ESSAYS IN EMPIRICAL ASSET PRICING

by

WEIKE XU

A dissertation submitted to the

Graduate School - Newark

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Professor Yangru Wu

and approved by

____

____

____

____

Newark, New Jersey

October, 2016

ABSTRACT OF THE DISSERTATION

Essays in Empirical Asset Pricing

By WEIKE XU

Dissertation Director:

Professor Yangru Wu

This dissertation includes two essays. The first essay examines how changes in ownership breadth affect the profitability of 21 anomaly-based strategies. I find that the profitability of these strategies is weaker following a growth in ownership breadth in the prior quarter. The return pattern is primarily attributed to the insignificant returns in the short portfolios. In addition, reduction in short-sale constraints due to increase in the ownership breadth can explain the insignificant return in the short portfolio. The conclusions stay the same after controlling for the common risk factors including the Fama-French three factors and the momentum factor. My results are robust to different size groups, different portfolio weighting methods, an alternative measure of active institutional investors and cross-sectional regression tests. These findings indicate that active institutional investors improve market efficiency.

In the second essay, I examine how the relaxation of short-sale constraints affects the readability in financial disclosures using a natural experiment. From 2005 to 2007, the SEC implemented a pilot program in which one-third of the Russell 3000 stocks were randomly selected as pilot stocks and were exempted from short-sale price tests. I find that the readability of 10-K reports for the pilot stocks significantly decreases during the program period. Moreover, the relation between a reduction in short-sales constraint and annual report readability is not uniform in the cross-section. I find that the results are more pronounced for firms that are smaller, less profitable or riskier; for firms that have lower institutional ownership or analyst coverage; and for firms with worse corporate governance or corporate social responsibility. I conclude that Regulation SHO leads to lower readability in the context of financial disclosures.

ACKNOWLEDGEMENTS

I am indebted to my adviser, Yangru Wu, for his endless guidance, patience, and support of my academic endeavors. He introduced the field of asset pricing to me. I have learned so much from him over the years. He has become one of my closest mentors and collaborators over the past few years. Working with him has been a real inspiration and I look forward to many more collaborations with him in the future. I would also like to thank my other committee members, Jin-Mo Kim and Ankur Pareek, for the time and effort that they put into my dissertation. Their comments and encouragements have been invaluable. And finally, I would like to thank my outside committee member, Lin Peng, for pointing me in the right direction both in terms of research and my career.

My Parents, Guoyi Xu and Xiaoyan Huang, and my wife, Minxing Sun, have supported me every step of way. Additionally, I would not have made it through the PhD program without the help and support of my friends and colleagues: Philip Davies, Rose C. Liao, S. Abraham Ravid, Ivan Brick, Simi Kedia, Tavy Ronen, Frank McIntyre, Yuzhao Zhang, Yichuan Liu, Chong Xiao, Xianjue Wang, Xinjie Wang, and Ge Wu.

TABLE OF CONTENT

LIST OF TABLES

Chapter1: Change in Ownership Breadth and Anomaly Returns

(jointly with Yangru Wu)

1.1 Introduction

Asset pricing literature documents that firm characteristics have significant power to predict future stock returns. For instance, future stock returns are positively related to past 6-12 months' returns (Jegadeesh and Titman, 1993), and to book-to-market ratios (Fama and French, 1992). In addition, future stock returns are negatively related to past one month return (Jegadeesh, 1990; and Lehmann, 1990), and idiosyncratic volatility (Ang et al., 2006). These patterns are considered as anomalies because they are not explained by standard asset pricing models, such as the CAPM and the Fama and French three-factor model.

Institutional investors play a growing role in the US equity market. Whether institutional investors add value to correct market mispricing is a longstanding debate in the finance literature.[1] On one hand, Lewellen (2011) argues that institutional investors hold the market portfolio and fail to take advantage of well-known anomalies. Edelen, Ince and Kadlec (2015) find that institutional investors increase their holdings of overpriced stocks and decrease their holdings of underpriced stocks and therefore can be a possible source of anomalies. On the other hand, several papers find that institutional investors trade on anomalies. Green, Hand, and Soliman (2011) find that the growth of hedge fund investments can partly explain the demise of the accrual anomaly. Cheng et al.

[1] Other papers that study the relationship between institutional investors and anomalies: mutual funds and accruals (Ali et al., 2008), short-term trading and anomalies (Cremers and Pareek, 2014), transient institutional investors and PEAD (Ke and Ramalingegowda, 2005).

(2015) find that short-term reversal is temporally higher following declines in the number of active institutional investors. Hanson and Sunderam (2014) show that short sellers trade on well-known anomalies such as value and momentum.

This paper empirically tests how active institutional investors affect stock price efficiency. Specifically, we examine how changes in ownership breadth affect the profitability of 21 anomaly-based strategies. The ownership breadth is introduced by Chen, Hong, and Stein (2002) (hereafter CHS) who argue that the breadth of ownership, as defined by the number of institutions that long a stock, can be a proxy for how tightly the short-sale constraints bind. Decreases in breadth of ownership forecast low future returns as the short-sale constrains bind more tightly. We use changes in breadth ownership rather than the breadth itself is because the level of breadth ownership is a permanent firm characteristic (quarterly autocorrelation of 0.99), which is highly correlated with firm size (CHS, 2002).

There are two reasons why we study the trading behavior of active institutional investors. First, compared with passive institutional investors such as pension funds, active institutional investors have shorter investment horizons and different investment objectives. Therefore, they are more likely to exploit anomalies. Second, active institutional investors have grown significantly over the past two decades. According to the Thomson Reuters 13f database, active institutional investors take up 71% of total institutional trading at the end of 2013.[2] It is important to separately study the behavior of active institutional investors.

[2] We calculate the ratio by using the market value held by active institutional investors divided by the total market value held by all institutional investors. Following Abarbanell, Bushee and Raedy (2003), I define active institutional investors as investment companies and independent advisors.

This paper makes several contributions. First, we show evidence that active institutional investors add value to correct market mispricing. We find that the 21 anomalies documented in the literature are weaker for the stocks following the growth in the ownership breadth. Second, we study the role of institutional investors on a broad set of anomalies from the angle of short-sale constraints. We argue that movements of active institutional investors would affect the short-sale constraints, and therefore would affect the profitability of anomalies. Third, combining ownership breadth changes and well-known anomalies, we develop new trading strategies that outperform these pure anomaly-based strategies.

This paper has four findings. First, we find that profitability of anomaly-based strategies is lower following a growth in the ownership breadth in the prior quarter. Stocks whose changes in ownership breadth are in the top tertile (HIGH group) are expected to be less mispriced than those in the bottom tertile (LOW group), as short-sale constraints are binding less tightly for the stocks in the HIGH group and more capital can be devoted to these stocks to correct mispricing. We show that 15 out of 21 anomalies are significantly weaker in the HIGH group than in the LOW group. The combination strategy (the strategy that takes equal positions across all 21 anomalies) yields a risk-adjusted return of 76 bps per month in the LOW group, and 34 bps per month in the HIGH group. The difference in profitability between the LOW and the HIGH groups is 43 bps per month (t-statistc=5.75), which is statistically significant at the 1% level.

Second, we find that the lower profits of anomaly-based strategies in the HIGH group are primarily attributed to the insignificant abnormal returns in the short portfolios. For instance, the monthly profits for the combination strategy are 39 bps (t-statistics=4.85), 5 bps (t-statistics=0.59), and 34 bps (t-statistics=5.57) for the long, short and long-short portfolios in the HIGH group, respectively.

Third, the insignificant returns from the short side in the HIGH group are due to the relaxation of short-sale constraints. We find that returns from the short side are significantly higher for each anomaly in the HIGH group. In other words, shorting is less profitable for the stocks in the HIGH group. When averaged across all 21 anomalies, the risk-adjusted return of the short side in the HIGH group is positive and insignificant, supporting that stocks in the HIGH group are less mispriced. Two factors can account for this result: (1) reduction in short-sale constraints due to increase in ownership breadth; (2) stock picking skills of active institutional investors (they buy underpriced stocks and sell overpriced stocks). We show evidence to support the short-sale constraint hypothesis. We use two other short-sale constraints proxies (size and total institutional ownership, Nagel, 2005) to divide stocks into short-sale constrained and unconstrained stocks. We find that the return pattern that the anomalies are weaker in the HIGH group is more significant for short-sale constrained stocks and is insignificant for unconstrained stocks.

Finally, strategies that long the stocks with high-performing characteristics and highest changes in ownership breadth, and short the stocks with the low-performing characteristics and lowest changes in ownership breadth outperform pure anomaly-based strategies. For instance, compared with the value-weighted risk adjusted return of pure anomaly-based strategies, when averaged across all 21 anomalies, the new strategy increases payoffs by 53 bps per month.

We conduct a battery of robustness check. We demonstrate that the results are similar in cross-section regression tests and are not driven by small stocks. We show that the results are robust after controlling for the Fama-French three factors and the momentum factor. We provide evidence that the return pattern is hold by forming value-weighted portfolios and by independent sorting. Finally, our results are robust using an alternative measure of active institutional investors based on institutions' portfolio turnover (Gaspar, Massa and Matos, 2005 (hereafter GMM, 2005)) and an alternative measure of ownership breadth (Choi, Yan and Jin, 2011).

Closely related to our work is by Edelen, Ince and Kadlec (2015), who find that institutional investors tend to trade contrary to signals implied by anomalies at one-year horizon. Our study has several differences with theirs. First, we examine the quarterly changes in ownership breadth to allow institutional investors to make decisions based on timely information, which are consistent with the measures used in CHS (2002) and CJY (2011). Edelen, Ince and Kadlec (2015) study the changes in institutional ownership during previous six quarters. Second, we study the effect of quarterly changes in ownership breadth on 21 anomalies. We argue that stocks with decline in ownership breadth in the previous quarter are more mispriced than those with growth in ownership breadth in the prior quarter. We find that the risk-adjusted return difference between LOW and HIGH group are significantly positive for 15 out of 21 anomalies. Edelen, Ince and Kadlec (2015) find that the risk-adjusted return differences between the stocks with institutional buy and the stocks with institutional sell are significantly positive for 3 out of 7 anomalies. Finally, we focus on the trading behavior of active institutional investors who are more likely to exploit these anomalies.

The remainder of the paper is organized as follows. Section 1.2 describes the data and methodology. Section 1.3 presents returns of 21 anomaly-based strategies and the return pattern of ownership breadth changes. Section 1.4 reports sources of the return pattern. Section 1.5 reports the returns of new strategies. Section 1.6 conducts some robustness tests and Section 1.7 concludes.

1.2 Data and Methodology

Our data are from four sources. Accounting data come from the Annual and Quarter CRSP/Compustat Merged database. Monthly data on stock return, stock price, and shares outstanding are from Center for Research in Security Prices (CRSP) monthly files. Returns on Fama-French's three factors and the momentum factor are from Kenneth French's website. Finally, the institutional investor holdings data come from Thomson Reuters Institutional (13f) holdings S34 files.

The SEC requires all institutional investors with greater than $100 million of securities under management to report their quarter-end holdings within 45 days after each calendar quarter. All common stock positions greater than 10,000 shares or $200,000 must be disclosed. Thomson Reuters classifies institutions as five types: (1) banks, (2) insurance companies, (3) investment companies, (4) independent advisors, and (5) others. Unfortunately, the type code is not reliable after 1998, as Thomson Reuters improperly classified many institutions as endowment and others.[3] So we use the institutional investors' classification data from Brian Bushee's personal website.[4] Following Abarbanell, Bushee and Raedy (2003), we define active institutional investors as investment companies and independent investment advisors. This measure excludes

[3] According to *User's Guide to Thomson Reuters Mutual Fund and Investment Company Common Stock Holdings Databases* on WRDS, Thomson Reuters has no plan to fix the mapping errors.

[4] We thank Brian Bushee for providing the institutional investor classifications data at this website: http://acct.wharton.upenn.edu/faculty/bushee/IIclass.html.

passive institutions such as bank trusts, insurance companies, university and foundation endowments, and pension funds. We also use an alternative measure of active institutional investors based on institutions' portfolio turnover (GMM, 2005). Each quarter, institutions with above median average churn rate are classified as active institutions.[5] Following CHS (2002), we require institutions to hold at least one stock in both quarter *q* and quarter *q-1* to control for the growth effect.

Our sample consists of all NYSE/AMEX/Nasdaq common stocks (share codes 10 and 11) from January 1981 to March 2014. The sample begins in 1981 due to the availability of institutional investors' type data. We exclude ADRs, REITs, financials, closed-end funds, foreign shares, all firms in the financial sectors (SIC 6000-6999), and stocks with share prices less than one dollar at the beginning of portfolio formation date. Following Lewellen (2011), we reverse Thomson Reuters 13f's split adjustment using CRSP cumulative shares and price adjustment factor whenever there is a difference between the filing date and the reporting date. To ensure that the results are not driven by small stocks, following Fama and French (2008), we group all stocks into two subsamples based on market equity at the beginning of each month: all-but-tiny stocks (those larger than NYSE 20 percentile), and large stocks (those larger than NYSE 50 percentile).

Our tests include 21 well-documented anomalies. Following Kogan and Tian (2014), we separate them into 7 groups below (the detailed construction of these anomalies is described in Appendix 1):

(1) Prior returns: short-term reversal (STREV), long-term reversal (LTR), momentum (MOM);

[5] The detailed construction of churn rate is summarized in Appendix 1.

(2) Valuation: book-to-market (BM),[6] earnings-to-price (EP), sales-to-price (SP);

(3) Earnings: returns on assets (ROA), standardized unexpected earnings (SUE), sales growth (SG), growth profitability premium(GP);

(4) Distress: O-score (OS);

(5) Investment: investment-to-assets (IA), asset growth (AG), accruals (TOTA), net operating assets (NOA), investment-to-capital (IK), investment growth (IG);

(6) External financing: net stock issuance (NSI), long-term stock issuance (LSI);

(7) Others: turnover (TO), idiosyncratic return volatility (IVOL).

Financial statements are often released to the public after the fiscal end period. Following Fama and French (1992), we match the annual accounting data at fiscal year ending in calendar year *y-1* with the monthly variables at the end of June in calendar year *y*. Following Chen, Marx and Zhang (2014), we match the monthly return data with the quarterly accounting data using the months immediately after the most recent earnings announcement data (RDQ). For example, if a company announces the financial statements of the fourth quarter of year *y-1* on February 20th of year *y*, we use the returns data in March of year *y* to match with the fourth quarter financial information of year *y-1*. In addition, we merge the institutional investor's holdings data with CRSP, and CRSP/Compustat using CUSIP numbers and report dates. Specifically, the CUSIP number from institutional investor's holding data at the end of quarter *q-1* of report year is matched with NCUSIP from CRSP at calendar quarter *q*.

[6] We construct BM as book equity in the prior fiscal year divided by market equity in last month. See Lewellen (2014). Similarly, we use last month market equity as the denominator for EP, SP, and LEV. See Lewellen (2014).

1.3 Empirical Results: change in ownership breadth and anomaly returns

1.3.1 Returns of Pure Anomaly-based Strategies

This section examines the significance of the 21 anomaly-based strategies. Specifically, we sort all stocks into five quintile portfolios based on 21 firm characteristics. The five quintile portfolio breakpoints are determined by sorting 21 firm characteristics using NYSE firms. Then we compute the equal-weighted and value-weighted returns for each portfolio and average hedge portfolio returns and risk-adjusted returns. We consider the higher-performing quintile as the long portfolio, and the lower-performing quintile as the short portfolio. For example, quintile 5 in book-to-market ratio (BM) is the long portfolio; however, quintile 5 in the short-term reversal (STREV) is the short portfolio. In addition, we construct a combination strategy (COMB) by taking equal positions across all 21 anomalies.

The portfolio formation and rebalancing frequencies differ across strategies. Specifically, for variables from CRSP, and CRSP/Compustat annual merged data, we form portfolios at the beginning of each month and hold them for one month.[7] For variables from Compustat annual data, we form portfolios at the end of each June and hold them for one year.[8] For variables from Compustat quarterly data (ROA, OS, and SUE), we form portfolios at the month after the recent quarterly announcement, and hold them for three months.

[7] They include short-term reversal (STREV), momentum (MOM), long-term reversal (LTR), net stock issuance (NSI), long-term stock issuance (LSI), turnover ratio (TO), idiosyncratic return volatility (IVOL), book-to-market (BM), earnings-to-price (EP), and sales-to-price (SP).
[8] They include investment-to-assets (IA), assets growth (AG), accruals (TOTA), investment-to-capital (IK), net operating assets (NOA), investment growth (IG), sales growth (SG), and growth profitability (GP).

*** Table 1.1 ***

Table 1 reports the profits from all 21 anomaly-based strategies across different size groups from January 1981 to March 2014. The risk-adjusted return (α) is computed from the following regression:

$$R_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_t \, , \qquad (1)$$

where the R_t is the return spread between the long and the short portfolios at month *t*, MKT_t is the value-weighted market excess return at month t , SMB_t is return spread between small and large stocks at month *t*, HML_t is the return spread between high and low value stocks at month *t*. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

Panels A and B report the equal-weighted average hedge portfolio returns and risk-adjusted returns of all 21 anomalies for each size group. Consistent with prior literature, all anomaly-based strategies generate significantly positive average hedge portfolio returns except for idiosyncratic volatility (IVOL). Similarly, all anomaly-based strategies generate significantly positive risk-adjusted returns. In addition, profits of these strategies are highest in all stocks, and lowest in large stocks, implying that anomalies are strongest in tiny stocks. For instance, the combination strategy yields a monthly average return of 73 basis points (bps) (t-statistic=9.71) and a risk-adjusted return of 69 bps (t-statistic=15.85) among all stocks, while the strategy earns a mean return of 35 bps (t-statistic=3.60) and a risk-adjusted return of 27 bps (t-statistic=4.81) among large stocks.

Panels C and D report value-weighted average hedge portfolio returns and risk-adjusted returns of all 21 anomalies for each size group. Consistent with prior literature, anomalies are usually weak by forming value-weighted portfolios except for idiosyncratic volatility (IVOL). We find that some anomalies are statistically insignificant in Panels C and D. For instance, short term reversal (STREV), turnover (TO), sales growth (SG), and investment-to-capital (IK) are statistically insignificant in each size group for both average hedge portfolio returns and risk-adjusted returns. The combination strategy yields an average monthly return of 36 bps (t-statistic=3.86) and a risk-adjusted return of 28 bps (t-statistic=5.70) for all stocks. Overall, the results in Table 1 show that these firm characteristics have power to predict future stock returns.

1.3.2 Change in ownership breadth and anomalies returns

This section tests how changes in ownership breadth affect profitability of 21 anomaly-based strategies using sequential sorting. First, we sort stocks into three tertile portfolios: low (T1), medium (T2), and high (T3), according to quarterly changes in ownership breadth. Stocks in the T1 portfolio (LOW group) are those with lowest changes in ownership breadth in the prior quarter. Stocks in the T3 portfolio (HIGH group) are those with highest changes in ownership breadth in the prior quarter. We expect that stocks in the LOW group are more likely to be mispriced than those in the HIGH group. Second, we sort stocks into five quintile portfolios based on the lagged 21 firm characteristics within each ownership breadth change group.[9] Third, for each

---

[9] For anomalies constructed from the Compustat annual database, we use the quarterly changes in the number of active institutional investors to form quarterly-balanced portfolios. We get similar results by

anomaly, we compute average hedge portfolio returns and risk-adjusted returns across the LOW and HIGH groups. Finally, for each anomaly, we calculate the mean difference of average hedge portfolio returns and risk-adjusted returns between LOW and HIGH groups. To obtain the difference of risk-adjusted returns, we compute the difference of hedge portfolio returns between the LOW and the HIGH groups each month, and then run the following regression:

$$R_{d,t} - R_{i,t} = \alpha_{dif} + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_{it}, \quad (2)$$

where $R_{d,t}$ is the hedge portfolio return in the LOW group at month *t*, $R_{i,t}$ is the hedge portfolio return in the HIGH group at month *t*. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

*** Table 1.2 ***

Table 2 reports the average hedge portfolio returns and risk-adjusted returns of all anomalies for the LOW and HIGH groups, and differences in profits between the LOW and HIGH groups. For the average hedge portfolio returns, 16 out of 21 anomalies are significantly weaker in the HIGH group than in the LOW group. On one hand, among the LOW group, except for idiosyncratic volatility (IVOL), all anomalies are statistically significant at the 1% level. On the other hand, among the HIGH group, some anomalies (short-term reversal (STREV), idiosyncratic volatility (IVOL), turnover (TO), book-to-

using the changes in the number of active institutional investors in June as the information for the whole year and form annual-balanced portfolios.

market (BM), earnings-price (EP), sales growth (SG), investment-to-capital (IK), and o-score (OS)) become insignificant. Moreover, the combination strategy yields a monthly mean return of 80 bps (t-statistic=8.85) in the LOW group, and 39 bps (t-statistic=4.36) in the HIGH group. The difference in profitability between the LOW and HIGH groups is 41 bps (t-statistic=6.11), which is statistically significant at the 1% level.

The results for the risk-adjusted returns are similar. We find that 15 out of 21 anomalies are significantly weaker in the HIGH group than in the LOW group. Among the LOW group, except for earnings-price (EP), all anomalies are statistically significant at the 1% level. Among the HIGH group, some anomalies such as short-term reversal (STREV), idiosyncratic volatility (IVOL), long-term reversal (LTR), book-to-market (BM), sales-price (SP), earnings-price (EP), sales growth (SG), investment-to-capital (IK), and accrual (TOTA) are insignificant. In addition, the combination strategy yields a risk-adjusted return of 76 bps per month (t-statistic=12.85) in the LOW group, and 34 bps per month (t-statistic=5.57) in the HIGH group, as compared to 69 bps per month (t-statistc=15.84) in Table 1. The difference in profitability between the LOW and HIGH groups is 43 bps per month (t-statistic=5.75), which is significant at the 1% level.

The evidence in Table 2 shows that profitability of anomaly-based strategies is significantly weaker following a growth in ownership breadth in the prior quarter. In addition, anomalies are not fully eliminated after the growth in ownership breadth in the previous quarter.

1.4 Sources of the Return Pattern

To determine sources of the return pattern, we examine the risk-adjusted returns to the long and short sides for each anomaly across the LOW and HIGH groups in Table 3.

*** Table 1.3 ***

1.4.1 Long side

Profits of the anomaly-based strategies can be attributed to the long sides. It is possible that the lower returns from the long sides of anomalies contribute to lower returns in the HIGH group. However, we do not find evidence to support it. The returns from the long side of anomalies are significantly higher in the HIGH group. The return difference between LOW and HIGH groups is statistically negative for each anomaly. For instance, the return from the long side of the combination strategy is 11 bps per month (t-statistic=0.87) in the LOW group, and 39 bps per month (t-statistic=4.85) in the HIGH group and the difference between the two groups being -28 bps per month (t-statistic=-2.12), which is statistically significant at the 5% level. The evidence show that the long sides of anomalies do not contribute to lower returns in the HIGH group.

However, we find that long sides contribute to profits of anomaly-based strategies in the HIGH group. 18 out of 21 anomalies earn statistically positive returns in the long side. For instance, the spread from the long side of the combination strategy is 39 bps per month (t-statistic=4.85). The positive spreads in the long side can partially explain that anomalies are not fully eliminated as the number of active institutional investors increase.

One explanation to the above results is that changes in ownership breadth are positively correlated with future stock return (CHS (2002)). We find that changes in ownership breadth indeed have a strong positive predict power for future returns. Specifically, we sort stocks into five quintile portfolios based on changes in ownership breadth and compute the raw excess return and the Fama-French alpha for the long-short portfolio in Table 4. The equal-weighted FF alpha is 79 bps per month (t-statistic=4.28); the value-weighted FF alpha is 49 bps per month (t-statistic=3.37). Another explanation is that the stock picking skills of active institutional investors. They buy underpriced stocks and sell overpriced stocks. We discuss these two explanations in Section C.

*** Table 1.4 ***

1.4.2 Short Side

Shorting plays a critical role in generating profits for anomaly-based strategies. Profits from anomaly-based strategies are mainly attributed to short leg portfolios. For instance, from Appendix 2 which describes returns to long and short portfolios for pure anomaly-based strategies, we can see that the short portfolio generates a risk-adjusted return of 44 bps per month, which contributes 64% of total profit of the combination strategy[10]. We expect that the return pattern, anomalies are weaker following a growth in ownership breadth, would be primarily attributed to higher return (less negative) from short sides of the anomalies.

From Table 3, we find that returns from the short side for each strategy are significantly higher in the HIGH group. In other words, in the high group, shorting is less profitable. More importantly, short sides earn insignificant abnormal returns in the HIGH

[10]We present the returns to long and short portfolio of unconditional strategies in Appendix 2.

group, implying that stocks are correctively priced. For example, for the short side, the combination strategy yields a risk-adjusted return of -65 bps per month (t-statistic=-4.40) for the LOW group, and 5 bps per month (t-statistic=0.59) for the HIGH group, as opposed to -44 bps per month (t-statistic=-4.03) in Appendix 2. For the high group, 19 out of 21 anomalies earn insignificant or positive abnormal returns for the short sides. In addition, risk-adjusted return differences between the LOW and HIGH groups are statistically negative for each anomaly. For instance, the difference between the HIGH and LOW groups for the combination strategy is -71 bps per month (t-statistic=-4.41), which is statistically significant at the 1% level. Therefore, among the HIGH group, the insignificant abnormal returns in the short portfolios lead to lower returns of anomaly-based strategies.

1.4.3 Short-sale Constraints and Managerial Skills

Table 3 shows that the higher returns from short sides contribute to lower profits of anomalies for the HIGH group. Two factors can explain this result: (1) relaxation in short-sale constraints due to increase in ownership breadth; and (2) stock picking skills of active institutional investors. Active institutional investors have ability to identify mispriced stocks.

On one hand, CHS (2002) argue that reduction in ownership breadth would tighten the short-sale constraints. Stocks from the short side of anomalies in the low group should be more overpriced as the short-sale constraints bind more tightly and shorting is more difficult. Therefore, the future returns of anomalies would be lower in the low group. On the other hand, active institutional investors may have skill to identify mispriced stocks. They buy the stocks in the short side of anomalies because the stocks are underpriced. For instance, a stock was overpriced and the price went down by 20% in January. Active institutional investors realized that the price of this stock was below its fundamental value and bought this stock in February and then the price went up. Therefore, the stocks with increase in the number of active institutional investors have higher returns than those with decrease in the number of active institutional investors. Unfortunately, we are unable to directly test the managerial skills hypothesis because 13f data do not capture the intra-quarter institutional trading activities and the exact execution time of institutional trading. In addition, it is difficult to isolate the effect of relaxing short-sale constraints and the impact of superior managerial skills of active institutional investors on anomaly returns. However, we find evidence to support the relaxation of short-sale constraints hypothesis. If the return pattern is driven by reducing short-sale constraints, we should expect that the results would be stronger for the short-sale constrained stocks and be weaker or even be wiped out for the short-sale unconstrained stocks.

We first use two proxies for short-sale constraints: size and total institutional ownership (Nagel, 2005) to divide stocks into short-sale constrained stocks and unconstrained stocks. Then we independent sort stocks into $3\times3\times5$ portfolios based on short-sale constraints proxies, changes in ownership breadth and anomalies and then we compute the Fama and French Alpha for each anomaly and each short-sale constraints proxy. Table 5 reports the results for size and total institutional ownership in Panels A and B, respectively. The findings in Table 5 support the short-sale constraints hypothesis. First, Panel A shows that the return pattern is strong for the small stocks and vanishes for large stocks, consistent with our prediction. For instance, for the combination strategy, the monthly profits difference between the LOW and HIGH groups is 55 bps (t-statistics=4.66) for small stocks and 13 bps (t-statistic=1.30) for large stocks. Second, Panel B shows that the return pattern is strong for the low institutional ownership stocks and is wiped out for the high institutional ownership stocks. For instance, for the combination strategy, the monthly risk-adjusted return difference between the LOW and HIGH groups is 100 bps (t-statistics=5.25) for the low institutional ownership stocks and 9 bps (t-statistic=0.81) for the high institutional ownership stocks.

*** Table 1.5 ***

1.4.4 Passive Institutional Investors

We conduct the same tests as in Table 2 using changes in the number of passive institutional investors. Table 6 reports the risk-adjusted returns of anomalies for the LOW and HIGH groups. We also report the return difference between the LOW and HIGH groups. The return pattern is weaker but still statistically significant. 11 out of the 21 anomalies are significantly weaker in the LOW group. For instance, for the combination strategy, the difference between the LOW and HIGH groups is 21 bps (t-statistic=3.07) (as compared to 43 bps (t-statistic=5.75) in Table 2). The evidence in Table 6 further supports the short-sale constraints hypothesis. In principle, ownership breadth should include all investors including active and passive institutional investors. Short-sale constraints also can be reduced by increasing in the number of passive institutional investors.

*** Table 1.6 ***

1.5 New strategies

We develop new trading strategies by combining changes in ownership breadth and 21 anomalies. Specifically, we long the stocks in the high-performing quintile (quintile 5) and the top ownership breadth changes tertile (tertile 3), and short the stocks in the low-performing quintile (quintile 1) and bottom ownership breadth changes tertile portfolio (tertile 1). As the SEC requires all institutional investors with greater than $100 million of securities under management to report their quarter-end holdings within 45 days after each calendar quarter, we left a two-month gap between the variables from institutional holding data and returns to ensure that these strategies are tradable. To directly compare the performance of strategies, we report the performance of pure anomaly based strategies in Table A2. We present both equal-weighted and value-weighted average hedge portfolio returns and risk-adjusted returns of the new strategies in Table 7.

*** Table 1.7 ***

We find that these strategies outperform pure anomaly-based strategies. In Panel A of Table 7, the combination strategy results in profits of 111 bps and 110 bps per month for the equal-weighted average hedge portfolio return and risk-adjusted return, respectively (as opposed to 69 and 64 bps in Table A2). In addition, the combination strategy returns 71 bps and 80 bps per month for the value-weighted average hedge portfolio return and risk-adjusted return, respectively (as compared to 33 and 27 bps in Table A2). When average across all 21 anomalies, the value-weighted risk-adjusted return of the new strategy increase its profit by 53 bps per month compared with that of the pure anomaly-based strategy.

1.6 Robustness Check

1.6.1 Regression Analysis

In this section, we examine quarterly changes in ownership breadth and anomalies in regression analysis. We run the following monthly cross-section regressions for each anomaly:

$$R_{i,t+1} = \alpha + \beta_1 A_{i,t} + \beta_2 D_{i,Num} + \beta_3 \left(D_{i,Num} \times A_{i,t}\right) + controls + \varepsilon_{i,t} \quad (5)$$

where $R_{i,t+1}$ is the return on stock i at month *t+1*, $A_{i,t}$ is the lagged firm characteristics (anomaly), $D_{i,Num}$ is a dummy variable that takes value 1 if stocks are in the lowest tertile for changes in ownership breadth and zero otherwise. In addition, we add other firm characteristics that have predicted power for future returns as control variables. They are: price momentum, Amihud (2002) illiquidity, market capitalization, and book-to-market ratio.

The parameter β_3 in the regressions captures the marginal effect of changes in breadth ownership on the relation between firm characteristics and future return. If a firm characteristic is negatively correlated with future stock returns, a negative sign of β_3

implies that the predict power of future stock return for this firm characteristic is stronger following the decline in the ownership breadth in the previous quarter. Similarly, If a firm characteristic is positively correlated with future stock returns, a positive sign of β_3 implies that the predict power of future stock return for this firm characteristic is stronger following the decline in the ownership breadth in the prior quarter.

We report the estimation of coefficients and their *t*-statistic of each anomaly in Table 8. Models (1) and (2) present the results without and with the control variables, respectively. The results from regression analysis are similar to those from portfolio sorts. For model (1), 12 out of 21 anomalies are stronger following a decline in ownership breadth in the prior quarter. For instance, β_3 of short-term reversal (STREV) is -3.23 (t-statistic=-7.94). As short-term reversal is negatively related to future stock return, significant negative β_3 implies that short-term reversal is stronger following a decline in ownership breadth. In addition, the results are robust after controlling for other firm characteristics that have predicted power for future returns.

*** Table 1.8 ***

1.6.2 Independent Sorting

We repeat the analysis in Table 2 using independent sorting. We independently sort stocks into 3×5 porfolios based on quarterly changes in ownership breadth and firm characteristics. We report the results of risk-adjusted return in Table 9. The findings are similar to those in Table 3. For instance, 13 out of 21 anomalies are weaker in the HIGH group than in the LOW group. The return difference between the LOW and HIGH group is 39 bps per month (as opposed to 43 bps per month in Table 2).

*** Table 1.9 ***

1.6.3 Value-weighted Portfolios

To test whether the results are robust by forming value-weighted portfolios, we repeat the tests in Table 2. Table 10 reports the value-weighted risk-adjusted returns of 21 conditional strategies based on changes in ownership breadth and anomalies. The results are robust. We find that 12 out of 21 anomalies are significantly weaker in the HIGH group than in the LOW group. Specifically, the monthly profits for the combination strategy is 51 bps (t-statistic=7.95) in the LOW group, and 15 bps (t-statistic=2.27) in the HIGH group. The profits difference between the LOW and HIGH groups is 36 bps (t-statistic=4.76), which is statistically significant at the 1% level.

*** Table 1.10 ***

1.6.4 Alternative Measures

We conduct the same analysis in Table 2 using an alternative measure of active institutional investors introduced by GMM (2005) and an alternative measure of ownership breadth documented by CJY (2011). I report the results of risk-adjusted returns in Table 11.

Gaspar, Massa and Matos (2005) develop a measure of active institutional investors using the institution's portfolio turnover (churn rate). Those with above median average churn rate are classified as active institutions in each quarter. The findings are similar to those in Table 2. For the GMM's measure of active institutional investor, we find that 13 out of 21 anomalies are weaker in the HIGH group than in the LOW group. The monthly profits difference between the LOW and HIGH groups is 39 bps (as opposed to 43 bps in Table 2). For the CYJ's measure of ownership breadth, 14 out of 21 anomalies are significantly weaker in the HIGH group than in the LOW group. The monthly difference between the LOW and HIGH groups is 40 bps (t-statistic=5.35), which is statistically significant at the 1% level.

*** Table 1.11 ***

1.6.5 Size

Anomalies are usually strongest among small stocks, and weakest among large stocks. To show that our results are not driven by small stocks, we repeat the analysis conducted in Table 2 using all-but-tiny stocks and large stocks. The results are presented in Table 12. For all-but-tiny stocks, 13 out of the 21 anomalies are weaker in the HIGH group than in the LOW group; for large stocks, 10 out of the 21 anomalies are weaker in the HIGH group than in the LOW group. In addition, among all-but-tiny stocks, the monthly profits difference between the LOW and HIGH groups is 30 bps (t-statistic=4.07), which is statistically significant at the 1% level. Among large stocks, the monthly returns between the LOW and HIGH groups is 29 bps (t-statistic=3.50), which is statistically significant at the 1% level.

*** Table 1.12 ***

1.6.6 The Carhart Four-factor Alpha

We also use the Carhart Four-factor model (Carhart, 1997) to calculate the risk-adjusted returns for each anomaly and then we reproduce the tests in Table 2. We run the following regression to obtain alpha for each anomaly:

$$R_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \varepsilon_t, \tag{3}$$

where UMD_t is the return spread between winner and loser stocks at month *t.* I run the following to obtain the alpha difference across LOW and HIGH group for each anomaly:

$$R_{d,t} - R_{i,t} = \alpha_{dif} + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \varepsilon_{it}, \tag{4}$$

Table 13 reports the Carhart four-factor alphas of anomalies for the LOW and HIGH groups. The results are still robust. We find 9 out of 21 anomalies are significantly weaker in the HIGH group. Difference in profitability between the LOW and HIGH groups for the combination strategy is 34 bps per month (t-statistics=4.35), as opposed to 43 bps per month (t-statistics=5.75) in Table 2. The weak results imply that the momentum factor can explain the return patterns of some anomalies. For instance, the return patterns for turnover (TO) and earnings-price ratio (EP) vanish after controlling the momentum factor.

*** Table 1.13 ***

1.7 Conclusion

In this paper, we find that the profitability of 21 anomaly-based strategies is significantly weaker following a growth in the ownership breadth in the prior quarter. The return pattern remains robust after controlling for common risk factors including the Fama-French three factors and the momentum factor. Our results are robust to different size groups, different portfolio weighting methods, an alternative measure of active institutional investors and cross-sectional regression tests.

The return pattern is primarily attributed to the insignificant abnormal returns from the short portfolios for these stocks with increase in ownership breadth in the prior quarter. We find that a reduction in short-sale constraints due to an increase in ownership breadth can explain the insignificant abnormal returns for the short side. As the ownership breadth grows, the short-sale constraints are binding less tightly and overpricing is easier to be corrected. In addition to the short-sale constraint hypothesis, the stock picking skills of active institutional investors may also explain the lower returns in the short portfolio. These findings indicate that active institutional investors improve market efficiency.

Our tests have several limitations. First, the measure of active institutional investors includes passive mutual funds. Second, it is difficult to disentangle two explanations for my results: relaxation in short-sale constraints and stock-picking skills of active institutional investors. These are left for future research

Appendix 1: Constructions of anomalies and variables from institutional investor holding

Prior Returns

Short-term reversal (STREV)

Jegadeesh (1990) and Lehmann (1990) find that prior one month return is negatively related to future return. Short-term reversal at month t is the monthly return at month t-1.

Momentum (MOM)

Jegadeesh and Titman (1993) first document that stocks with high returns over past 3 to 12 month have abnormally high average returns for the next 3 to 12 month. The result is confirmed by others (Fama and French (1996, 2008), Jegadeesh and Titman (2001)). Momentum at month t is the cumulative continuously compounded return from month t-7 to month t-2.

Long-term reversal (LTR)

DeBondt and Thaler (1985, 1987) first document that stocks with low returns over the past 3-5 years have abnormally high average returns. The long-term reversal can be explained by Fama-French 3 factors (Fama and French (1996)). The long-term reversal in month t is defined as the cumulated continuously compounded stock return from month t-60 to month t -13.

External Financing

Net stock issuance (NSI)

Stocks returns are lower after stock issuance ((Loughran and Ritter (1995)). Stocks with low net stock issues have abnormally high average returns (Fama and French (2008), Pontiand Woodgate (2008), Lewellen (2014)). The net stock issuance at month t-1 is computed by the split-adjusted shares outstanding at month t-1 divided by lagged 12 month split-adjusted shares outstanding. The split-adjusted shares outstanding are the product of common shares outstanding (CRSP item SHROUT) and the cumulative adjustment factor (CRSP item CFACSHR).

Long-term stock issuance (LSI)

Stocks with low composite issuance have abnormally high average returns, after controlling for other known predictors of returns (Daniel and Titman (2006), Fama and French (2008), Pontiand Woodgate (2008)). The long-term stock issuance at month t-1 is computed by the split-adjusted shares outstanding at month t-1 divided by lagged 36 month split-adjusted shares outstanding.

Valuation

Book-to market (BM)

Stocks with high book-to-market have abnormally high average returns (Fama and French (1992), Fama and French (1993)). Book-to-market is the book equity in the prior fiscal year divided by market equity at month t-1. The book equity is stockholders' equity, plus deferred taxes and investment credit (Compustat item TXDITC), and minus the book value of preferred stock. The book value of stockholder's equity is stockholder equity (Compustat item TEQ). If TEQ is missing, I use total common equity (Compustat item CEQ) plus preferred stock carrying value ((Compustat item UPSTKC)) to replace the missing TEQ. Otherwise, I use total assets (Compustat item AT) minus total liabilities (Compustat item LT). Depending on availability, I use the redemption (Compustat item PSTKRV), liquidation (Compustat item PSTKL), or par value (Compustat item UPSTK) (in that order) to estimate the book value of preferred stock.

Earning-to-price (EP)

Stocks with high earnings-to-price ratio have abnormally high average returns (Basu (1977), Basu (1983)). Fama and French (1992) argue that the effect seems to be subsumed by book-to-market ratio. Earnings-to-price is earnings before extraordinary items (Compustat item EBIT) divided by market equity at month t-1.

Sales-to-price (SP)

Stocks with high sales-to-price ratio have abnormally high average returns (Fama and French (1992) and Lakonishok, Shleifer, and Vishny (1994)). Sales-to-price (SP) is sales in the prior fiscal year divided by market equity at month t-1.

Earnings

Returns on assets (ROA)

Stocks with high return on assets have abnormally high average returns (Fama and French (2006), Chen, Novy-Marx, and Zhang (2010)). Return on assets is the ratio of income before extraordinary items (Compustat item IBQ) over total assets (Compustat item ATQ)

Standardized unexpected earnings (SUE)

Stocks with high standardized unexpected earnings have abnormally high average returns (Bernard and Thomas (1989)). Standardized unexpected earnings is the change in the most recently announced quarterly earnings per share (Compustat item EPSPIQ) from its announced value four quarters ago divided by the standard deviation of the change in quarterly earnings over the prior eight quarters.

Sales growth (SG)

Stocks with low past sales growth have abnormally high average returns (Lakonishok, Shleifer, and Vishny (1994)). Sales growth is the percent change in net sales over net turnover (Compustat item SALE).

Growth profitability premium (GP)

More profitability firms have higher returns than lower profitability firms, and the gross profit premium is the cleanest accounting measure of true economic profitability (Norvy-Marx (2010)). Growth profitability premium is the ratio of sales minus cost of goods sold (Compustat item COGS) over total assets.

Distress

O-score (OS)

Stocks with lower Ohlson score (lower probability of default) have abnormally high average returns (Grffin and Lemmon (2002)). O-score is calculated using model one in table four (Ohlson (1980))

Investment

Investment-to-assets (IA)

Stocks with low investment-to-assets ratios have abnormally high average returns (Lyandres, Sun, and Zhang (2008), Chen, Novy-Marx, and Zhang (2010)). Investment-to-assets (IA) is the annual change in property, plant, and equipment (Compustat item PPEGT) plus annual change in total inventories (Compustat item INVT) divided by prior fiscal year total assets (Compustat item AT).

Asset growth (AG)

Stocks with low asset growth have abnormally high average returns (Cooper, Gulen, and Schill (2008)). Asset growth (AG) is the percentage change in total assets (Compustat item AT).

Accruals (TOTA)

Stocks with low accruals have abnormally high average returns (Sloan(1996)). Accrual (TOTA) is change in current assets (Compustat item ACT) minus the change in cash and short-term investments (Compustat item CASH) minus the change in current liabilities (Compustat item LCT) plus the change in debt in current liabilities (Compustat item DLC) plus change in income taxes payable (Compustat item TXP) minus depreciation and amortization (Compustat item DP) scaled by total assets (Compustat item AT).

Net operating assets (NOA)

Stocks with low net operating assets have abnormally high average returns (Hirshleifer, Hou, Teoh, and Zhang (2004)). Net operating assets (NOA) is total operating assets minus total operating liabilities scaled by the average total assets (Compustat item AT) over the past two years. Specifically,

$$NOA_{t-1} = \frac{OA_{t-1} - OL_{t-1}}{(AT_{t-1} + AT_{t-2})/2}$$

$$OA_{t-1} = AT_{t-1} - CHE_{t-1}$$

$$OL_{t-1} = AT_{t-1} - DLC_{t-1} - DLTT_{t-1} - MIB_{t-1} - PSTK_{t-1} - CEQ_{t-1}$$

where *OA* is total operating assets, *OL* is total operating liabilities, *AT* is total assets, *CHE* is cash and short-term investments, *DLC* is debt in current liabilities, *DLTT* is long term debt, *MIB* is non-controlling interest, *PSTK* is preferred capital stock, and *CEQ* is common equity.

Investment-to-capital (IK)

Stocks with low investment-to-capital ratios have abnormally high average returns (Xing (2008)). Investment to capital is the ratio of capital expenditure (Compustat item CAPX) over property, plant, and equipment (Compustat item PPENT).

Investment growth (IG)

Stocks with low investment growth rates have abnormally high average returns (Xing (2008)). Investment growth is the percentage change in capital expenditure (Compustat item CAPX).

Others

Turnover (TO)

Stocks with low average turnover over past 3-12 months have abnormally high average returns (Lee and Swaminathan (2000)). Turnover ratio (TO) is the average monthly turnover from month t-12 to month t-3. The monthly turnover is the ratio of shares traded over shares outstanding. I followed Gao and Ritter (2010) to adjust the Nasdaq trading volume. Specifically, prior to February 1, 2001, I divide Nasdaq volume by 2. From February 1, 2001 to December 31, 2001, I divide Nasdaq volume by 1.8. For the years 2002 and 2003, I divide Nasdaq volume by 1.6. No adjustment is made from 2004 onwards.

Idiosyncratic volatility (IVOL)

Stocks with high idiosyncratic return volatility have abnormally low returns (Ang, Hodrick, Xing and Zhang (2006)). Idiosyncratic volatility (IVOL) is the standard deviation of residuals from a regression of daily excess returns on the Fama-French three factor model.

Illiquidity ratio (AMIHUD)

The expected returns on illiquid firms are higher than those on liquid firms (Amihud (2002)). The illiquidity ratio is defined as the absolute value of daily stock return scaled by the dollar trading volume of the stock on that day.

$$Amihud_{i,t} = \frac{1}{N_{i,t}} \sum_{\tau=1}^{N_{i,t}} \frac{\left|R_{i,t}\right|}{DVOL_{i,t}}$$

where $N_{i,t}$ is the number of trading days with positive trading volume for stock *i* in month *t*. $\left|R_{i,t}\right|$ is the absolute value of the return of stock *i* on day *t*, and $DVOL_{i,t}$ is the dollar trading volume of stock *i* on day.

Institutional Investors

Number of active institutional investors: The number of active institutional investors that hold a stock in a quarter. Following Abarbanell, Bushee and Raedy (2003), I define the active institutional investors as investment companies and independent investment advisors. Others are considered as passive institutional investors.

Value of active institutional investors: Market value of active investors in a given quarter q is calculated by price multiply by shares owned by the investors in the quarter.

Active institutional investors (Gaspar, Massa and Matos (2005))

Gaspar, Massa and Matos (2005) develop a measure of active institutional investors using the institution's portfolio turnover (churn rate). The churn rate of institution *f* in a given quarter *q* is computed as follows

$$CR_{f,q} = \frac{\sum_{i\in Q} |N_{i,f,q}P_{i,q} - N_{i,f,q-1}P_{i,q-1} - N_{i,f,q-1}\Delta P_{i,q}|}{\sum_{i\in Q} \frac{N_{i,f,q}P_{i,q} + N_{i,f,q-1}P_{i,q-1}}{2}}$$

where $N_{i,f,q}$ refers to the number of shares of stock *i* held by institution *f* in quarter q, $P_{i,q}$ refers to the stock price at the same time, Q represents the set of companies held by institution *f*. The average churn rate is calculated over the past four quarters to rank the institutions, and those with above median average churn rate are classified as active institutions in each quarter.

Appendix 2: Returns to long and short Sides of pure anomaly-based strategies

This table presents the returns to long and short sides of 21 pure anomaly-based strategies. We sort all stocks with respect to the 21 firm characteristics and construct 5 quintile portfolios. The combination strategy (COMB) is constructed by taking equal positions across the 21 anomalies. Abbreviations of the anomalies are as follows: short-term reversal (STREV), momentum (MOM), long-term reversal (REV), net stock issues (NSI), Long-term stock issuance (LSI), idiosyncratic volatility (IVOL), turnover (TO), book-to-market (BM), earnings-to-price (EP), sales-to-price (SP), growth profitability premium (GP), sales growth (SG), investment growth (IG), investment-to-capital (IK), investment-to-assets (IA), accruals (TOTA), asset growth (AG), net operating assets (NOA), return on assets (ROA), Ohlson score (OS), and standardized unexpected earnings (SUE).

	FF Alphas			
	Long		Short	
	ret	t-stat	ret	t-stat
STREV	0.09	0.52	-0.54	-5.22
MOM	0.42	4.44	-0.73	-3.73
REV	0.09	0.70	-0.14	-1.47
NSI	0.24	2.70	-0.54	-4.46
LSI	0.28	3.19	-0.48	-3.83
IVOL	0.10	1.29	-0.43	-2.94
TO	0.20	1.74	-0.58	-4.71
BM	0.22	1.27	-0.35	-4.75
EP	0.08	0.59	-0.30	-2.38
SP	0.09	0.57	-0.38	-3.95
GP	0.33	3.20	-0.36	-2.23
SG	0.07	0.48	-0.33	-2.69
IG	0.20	1.56	-0.26	-2.12
IK	0.17	1.46	-0.25	-1.92
IA	0.25	1.83	-0.63	-4.91
TOTA	0.08	0.61	-0.24	-2.05
AG	0.28	1.79	-0.52	-4.16
NOA	0.27	2.07	-0.51	-4.06
ROA	0.58	6.19	-0.58	-3.06
OS	0.44	4.63	-0.62	-3.69
SUE	0.93	8.03	-0.62	-4.65
COMB	0.25	2.42	-0.44	-4.03

Table 1.1: Pure anomaly-based strategies

This table presents the performance of 21 anomaly-based strategies for each size group. I sort all stocks based on the 21 firm characteristics and construct five quintile portfolios using NYSE breakpoints. The combination strategy (COMB) is constructed by taking equal positions across the 21 anomalies. All stocks consist of all NYSE/AMEX/Nasdaq common stocks; all-but-tiny stocks are those whose sizes are higher than NYSE 20 percentile; and large stocks are those whose sizes are higher than NYSE 50 percentile. Panels A and B report the equal-weighted average hedge portfolio returns, and Fama-French alphas of all anomalies for each size group. Panels C and D report results by forming value-weighted portfolios. Abbreviations of the variables are as follows: short-term reversal (STREV), momentum (MOM), long-term reversal (LTR), net stock issues (NSI), Long-term stock issuance (LSI), idiosyncratic volatility (IVOL), turnover (TO), book-to-market (BM), earnings-to-price (EP), sales-to-price (SP), growth profitability premium (GP), sales growth (SG), investment growth (IG), investment-to-capital (IK), investment-to-assets (IA), accruals (TOTA), asset growth (AG), net operating assets (NOA), return on assets (ROA), Ohlson score (OS), and standardized unexpected earnings (SUE). Definitions of variables are described in Appendix 1. Returns and alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

Panel A: Equal-weighted Average Hedged Portfolio Returns						
	All Stocks		All-but-tiny Stocks		Large Stocks	
	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.79	4.03	0.41	2.03	0.37	1.76
MOM	0.95	4.01	0.58	2.30	0.35	1.35
REV	0.40	2.71	0.30	2.12	0.26	1.60
NSI	0.77	4.23	0.51	2.66	0.44	2.42
LSI	0.75	4.03	0.50	2.72	0.45	2.77
IVOL	0.27	0.95	0.37	1.23	0.24	0.83
TO	0.66	2.80	0.27	1.13	0.13	0.54
BM	0.90	4.36	0.47	2.02	0.36	1.58
EP	0.68	3.31	0.63	2.53	0.58	2.30
SP	0.89	4.15	0.65	2.74	0.50	2.22
GP	0.67	5.26	0.51	3.72	0.40	2.73
SG	0.49	4.41	0.22	1.65	0.19	1.24
IG	0.50	6.04	0.35	3.81	0.34	3.20
IK	0.45	3.50	0.39	2.12	0.27	1.26
IA	0.86	7.90	0.47	4.17	0.40	3.24
TOTA	0.40	5.39	0.24	2.81	0.13	1.37
AG	0.88	6.94	0.50	4.08	0.35	2.36
NOA	0.66	4.49	0.40	2.78	0.38	2.78
ROA	1.01	4.93	0.71	3.57	0.22	1.11
OS	0.86	5.17	0.65	4.07	0.38	2.05
SUE	1.50	12.13	0.91	6.61	0.70	3.89
COMB	0.73	9.71	0.48	5.11	0.35	3.60

Panel B: Equal-weighted Fama-French Alphas						
	All Stocks		All-but-tiny Stocks		Large Stocks	
	α	t-stat	α	t-stat	α	t-stat
STREV	0.63	3.11	0.23	1.10	0.16	0.73
MOM	1.15	5.17	0.80	3.37	0.52	2.07
REV	0.23	1.79	0.08	0.69	0.07	0.49
NSI	0.78	6.57	0.47	3.99	0.35	2.84
LSI	0.76	6.05	0.49	4.20	0.39	3.32
IVOL	0.53	3.25	0.61	4.43	0.39	2.56
TO	0.78	5.71	0.39	3.24	0.25	1.82
BM	0.57	3.42	0.01	0.07	-0.05	-0.35
EP	0.38	3.00	0.26	1.78	0.21	1.37
SP	0.47	3.12	0.11	0.71	0.00	-0.03
GP	0.69	5.22	0.54	3.88	0.44	3.16
SG	0.40	3.97	0.10	0.98	0.06	0.50
IG	0.45	5.36	0.29	3.16	0.28	2.71
IK	0.42	4.88	0.32	3.49	0.19	1.83
IA	0.89	8.07	0.43	3.94	0.35	3.02
TOTA	0.32	4.37	0.19	2.17	0.10	1.11
AG	0.80	6.89	0.40	3.68	0.22	1.99
NOA	0.78	6.52	0.55	4.59	0.49	4.35
ROA	1.16	5.79	0.90	4.96	0.35	1.67
OS	1.05	6.36	0.85	5.52	0.47	2.47
SUE	1.55	12.98	0.97	7.36	0.76	4.19
COMB	0.69	15.84	0.41	8.41	0.27	4.80

Panel C: Value-weighted Average Hedged Portfolio Returns						
	All Stocks		All-but-tiny Stocks		Large Stocks	
	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.26	1.16	0.29	1.31	0.30	1.35
MOM	0.36	1.32	0.25	0.97	0.16	0.62
REV	0.33	1.93	0.33	1.94	0.33	1.82
NSI	0.35	2.04	0.30	1.77	0.25	1.48
LSI	0.45	2.97	0.42	2.85	0.37	2.73
IVOL	0.43	1.39	0.30	1.04	0.18	0.68
TO	0.08	0.32	-0.02	-0.07	-0.08	-0.28
BM	0.36	1.79	0.28	1.47	0.18	0.97
EP	0.46	2.15	0.42	1.97	0.35	1.64
SP	0.51	2.54	0.39	2.02	0.27	1.45
GP	0.41	2.99	0.38	2.61	0.35	2.35
SG	0.13	0.77	0.09	0.51	0.03	0.17
IG	0.32	2.60	0.31	2.29	0.23	1.66
IK	0.26	1.15	0.22	0.96	0.15	0.65
IA	0.38	2.91	0.33	2.45	0.30	2.15
TOTA	0.24	2.07	0.23	1.90	0.18	1.42
AG	0.32	2.12	0.28	1.68	0.21	1.19
NOA	0.48	4.27	0.45	4.14	0.48	4.17
ROA	0.55	2.50	0.46	2.16	0.35	1.54
OS	0.40	2.10	0.39	2.04	0.34	1.62
SUE	0.53	3.06	0.50	2.70	0.46	2.29
COMB	0.36	3.86	0.31	3.28	0.25	2.68

Panel D: Value-weighted Fama-French Alphas						
	All Stocks		All-but-tiny Stocks		Large Stocks	
	α	t-stat	α	t-stat	α	t-stat
STREV	0.04	0.16	0.10	0.42	0.10	0.44
MOM	0.56	2.17	0.39	1.56	0.25	0.99
REV	0.07	0.47	0.11	0.75	0.14	0.88
NSI	0.32	2.64	0.28	2.24	0.21	1.60
LSI	0.42	3.72	0.38	3.34	0.31	2.63
IVOL	0.65	4.05	0.48	3.06	0.34	2.09
TO	0.21	1.39	0.14	0.86	0.04	0.25
BM	-0.05	-0.32	-0.13	-0.99	-0.20	-1.44
EP	0.11	0.78	0.04	0.32	-0.02	-0.15
SP	-0.03	-0.20	-0.14	-1.04	-0.22	-1.59
GP	0.53	4.14	0.54	4.01	0.55	3.87
SG	-0.05	-0.40	-0.08	-0.61	-0.14	-0.99
IG	0.24	2.19	0.24	1.97	0.19	1.48
IK	0.18	1.43	0.14	1.05	0.07	0.49
IA	0.31	2.67	0.27	2.30	0.23	1.82
TOTA	0.22	1.96	0.25	2.10	0.21	1.66
AG	0.15	1.34	0.13	1.12	0.06	0.44
NOA	0.58	5.87	0.57	5.64	0.60	5.53
ROA	0.77	3.58	0.69	3.39	0.52	2.26
OS	0.61	3.45	0.59	3.21	0.43	2.06
SUE	0.59	3.90	0.59	3.44	0.55	2.85
COMB	0.28	5.70	0.24	4.72	0.18	3.28

Table 1.2: Change in ownership breadth and anomaly returns: sequential sorting

We first sort stocks into three tertile portfolios (LOW, MED and HIGH) based on changes in ownership breadth in the last quarter using NYSE breakpoints. Then we sort stocks into five quintile portfolios based on 21 firm characteristics within each breadth change group using NYSE breakpoints. The ownership breadth is the number of active institutional investors long a stock. To control for the growth of new institutional investors, we require each institutional investor to hold at least one stock in quarter *q* and *q-1*. The combination strategy (COMB) is constructed by taking equal positions across the 21 anomalies. We present the average hedge portfolio return and risk-adjusted return of each anomaly for the LOW and HIGH groups. In addition, we compute the risk-adjusted return differences between the LOW and HIGH groups (column “LOW-HIGH”) for each anomaly using the following regression:

$$R_{d,t} - R_{i,t} = \alpha_{dif} + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_{it}$$

where the $R_{d,t}$ is the hedge portfolio return in the LOW group at month *t*, $R_{i,t}$ is the hedge portfolio return in the HIGH group at month *t*, MKT_t is the value-weighted market excess return at month *t* , SMB_t is return spread between low and high stocks at month *t*, HML_t is the return spread between high and low value stocks at month *t*. The α_{dif} is the Fama-French alpha difference between the Low and High groups. Returns and alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

	Average Hedge Portfolio Returns						Fama-French Alphas					
	LOW		HIGH		LOW-HIGH		LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	1.08	4.74	0.12	0.57	0.96	**5.91**	0.89	3.81	-0.08	-0.36	0.97	**5.66**
MOM	0.74	2.47	0.65	2.65	0.09	0.34	1.00	3.92	0.80	3.35	0.21	1.10
REV	0.58	3.26	0.35	2.18	0.23	1.27	0.39	2.47	0.11	0.74	0.28	1.43
NSI	0.89	4.36	0.41	1.99	0.48	**3.18**	0.91	6.34	0.40	2.80	0.50	**3.28**
LSI	0.89	4.32	0.39	1.95	0.50	**3.14**	0.90	5.87	0.39	2.67	0.51	**3.07**
IVOL	0.45	1.39	-0.03	-0.09	0.48	**2.71**	0.77	3.74	0.20	1.31	0.57	**3.30**
TO	0.81	3.08	0.30	1.20	0.51	**2.64**	0.99	5.69	0.36	2.19	0.63	**2.95**
BM	1.03	4.06	0.40	1.61	0.63	**2.98**	0.57	2.85	0.04	0.17	0.53	**2.68**
EP	0.62	2.67	0.48	1.78	0.13	0.70	0.23	1.54	0.21	1.20	0.02	0.14
SP	1.10	4.52	0.48	1.88	0.62	**3.13**	0.58	3.55	0.04	0.18	0.55	**2.96**
GP	0.73	4.88	0.41	2.70	0.32	**2.14**	0.78	5.03	0.47	2.94	0.31	**2.11**
SG	0.69	4.72	0.10	0.73	0.59	**4.04**	0.57	4.26	0.00	0.00	0.57	**3.72**
IG	0.57	5.02	0.36	3.29	0.21	1.67	0.49	4.18	0.31	2.80	0.18	1.30
IK	0.56	3.28	0.18	0.97	0.38	**2.34**	0.51	4.00	0.10	0.81	0.41	**2.40**
IA	1.01	7.19	0.45	3.53	0.56	**3.93**	1.01	6.99	0.43	3.27	0.58	**3.76**
TOTA	0.48	4.52	0.18	1.73	0.30	**2.37**	0.40	3.72	0.12	1.14	0.29	**2.21**
AG	1.03	7.04	0.49	3.48	0.53	**3.57**	0.94	6.88	0.41	3.01	0.53	**3.36**
NOA	0.71	4.76	0.41	2.30	0.30	**2.38**	0.83	6.15	0.55	3.80	0.29	**2.22**
ROA	0.91	3.33	0.78	3.54	0.13	0.56	1.30	5.34	1.05	4.90	0.26	1.10
OS	0.80	3.48	0.26	1.06	0.55	1.95	1.16	5.46	0.58	2.47	0.59	**2.11**
SUE	1.20	6.15	1.01	5.97	0.18	0.80	1.41	7.99	1.12	6.78	0.29	1.24
COMB	0.80	8.85	0.39	4.36	0.41	**6.11**	0.76	12.85	0.34	5.57	0.43	**5.75**

Table 1.3: Returns from long and short sides

This table reports the risk-adjusted returns from long and short sides across the LOW and HIGH groups for each anomaly in Table 2. We report the risk-adjusted return of each anomaly for the LOW and HIGH groups. We also compute the return difference between LOW and HIGH groups (column "LOW-HIGH") for each anomaly. Alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

	Long Side						Short Side					
	LOW		HIGH		LOW-HIGH		LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.06	0.25	0.20	1.33	-0.14	-0.76	-0.84	-6.53	0.28	2.15	-1.12	-7.21
MOM	0.15	1.45	0.64	4.51	-0.49	-3.15	-0.86	-3.45	-0.16	-1.01	-0.69	-3.65
REV	0.01	0.08	0.27	2.13	-0.25	-1.33	-0.38	-2.51	0.16	1.65	-0.53	-3.32
NSI	0.17	1.51	0.40	4.45	-0.23	-2.22	-0.74	-4.39	0.00	-0.01	-0.74	-3.76
LSI	0.19	1.70	0.41	4.56	-0.22	-2.15	-0.71	-4.11	0.02	0.14	-0.73	-3.73
IVOL	0.05	0.61	0.29	3.22	-0.24	-2.53	-0.71	-3.66	0.09	0.73	-0.80	-4.02
TO	0.16	1.41	0.38	3.19	-0.22	-1.57	-0.83	-4.45	0.02	0.23	-0.85	-4.11
BM	0.03	0.13	0.31	1.92	-0.28	-1.38	-0.54	-5.05	0.27	2.31	-0.81	-4.58
EP	-0.16	-0.89	0.36	2.78	-0.52	-3.27	-0.39	-2.29	0.15	1.23	-0.54	-2.70
SP	0.03	0.15	0.21	1.39	-0.18	-1.03	-0.55	-4.17	0.18	1.47	-0.73	-3.96
GP	0.13	0.98	0.56	5.83	-0.43	-2.56	-0.66	-3.82	0.09	0.61	-0.75	-3.84
SG	-0.05	-0.32	0.14	1.22	-0.19	-1.11	-0.58	-3.53	0.14	1.35	-0.72	-3.85
IG	0.03	0.22	0.38	3.73	-0.35	-2.06	-0.46	-2.87	0.09	0.90	-0.55	-3.15
IK	-0.01	-0.08	0.24	2.17	-0.25	-1.61	-0.51	-3.11	0.14	1.36	-0.65	-3.52
IA	0.15	1.03	0.32	3.03	-0.17	-1.06	-0.89	-5.12	-0.09	-0.73	-0.80	-4.55
TOTA	-0.06	-0.37	0.32	2.68	-0.38	-1.92	-0.46	-3.28	0.14	1.48	-0.60	-3.64
AG	0.13	0.80	0.42	3.33	-0.29	-1.53	-0.80	-4.83	-0.03	-0.36	-0.76	-4.48
NOA	0.10	0.68	0.47	4.25	-0.37	-2.06	-0.74	-4.68	-0.07	-0.61	-0.67	-4.39
ROA	0.39	2.79	0.63	4.59	-0.23	-1.32	-0.91	-3.82	-0.42	-2.61	-0.49	-2.08
OS	0.28	1.70	0.53	3.98	-0.24	-1.16	-0.88	-3.96	-0.05	-0.28	-0.83	-3.62
SUE	0.54	3.34	0.79	6.19	-0.25	-1.28	-0.87	-4.74	-0.33	-2.39	-0.54	-2.49
COMB	0.11	0.87	0.39	4.85	-0.28	-2.12	-0.65	-4.40	0.05	0.59	-0.71	-4.41

Table 1.4: Returns to a strategy based on changes in ownership breadths

This table reports the performance of a strategy that based on changes in ownership breadths. The ownership breadth is the number of active institutional investors long a stock. We sort all stocks into 5 quintile portfolios based on quarterly changes in ownership breadths. Five quintle portfolio breakpoints are determined by sorting changes in ownershio breadth using NYSE firms. Then we long the stocks have highest changes in the number of active institutional investors (quintile 5) and short the stocks have lowest changes in the number of active institutional investors (quintile 1). Returns and alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

	Equal Weight		Value Weight	
	Ret	Alpha	Ret	Alpha
P5	1.04	0.33	0.75	0.19
	3.40	4.06	3.12	2.66
P1	0.40	-0.46	0.39	-0.30
	1.17	-3.19	1.49	-3.15
P5-P1	0.64	0.79	0.36	0.49
	3.46	4.28	2.52	3.37

Table 1.5: Return patterns for short-sale constrained and unconstrained stocks

This table repeats the tests conducted in Table 2 for short-sale constrained and unconstrained stocks. We use two short-sale constrain proxies: size and total institutional ownership. Stocks are sorted into 5 quintile portfolios based on 21 firm characteristics, and 3 tertile portfolios (low, medium and high groups) based on changes in ownership breadth over the previous one quarter, and tertile portfolios based on short-sale constraints proxies. Then we independently sort stocks into $3 \times 3 \times 5$ portfolios and we compute the Fama-French alphas for the Low and High groups for each short-sale constrain proxy. All portfolio breakpoints are determined by sorting firm characteristic using NYSE firms. Panels A and B report the results for size and institutional ownership, respectively. Alphas are reported as percentage per month. The t-statistics are computed using heteroskedasticity- consistent standard errors of White (1980).

Panel A: Triple Sorts on Size, Institutional Number Changes and Anomalies

	Small Stocks						Large Stocks					
	LOW		HIGH		LOW-HIGH		LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	1.05	4.95	-0.84	-2.32	1.90	5.55	0.34	1.24	-0.09	-0.32	0.42	2.08
MOM	1.54	5.83	1.76	4.98	-0.23	-0.66	0.49	1.66	1.27	4.35	-0.78	-3.23
REV	0.32	1.52	-0.06	-0.18	0.38	0.97	-0.03	-0.18	0.06	0.34	-0.10	-0.45
NSI	1.20	6.26	0.21	0.81	0.99	3.54	0.68	4.19	0.50	2.80	0.17	0.85
LSI	1.16	6.12	-0.05	-0.20	1.21	4.04	0.71	4.70	0.29	1.68	0.42	2.07
IVOL	1.16	5.13	0.44	1.68	0.75	2.64	1.49	5.98	0.46	1.46	1.03	3.07
TO	1.24	4.98	0.37	1.15	0.95	2.62	0.53	2.55	0.48	2.12	0.05	0.17
BM	0.44	1.85	0.17	0.64	0.25	0.74	0.24	1.31	-0.02	-0.09	0.26	1.09
EP	0.41	2.10	0.48	1.43	-0.10	-0.31	0.89	4.03	0.33	1.55	0.55	2.26
SP	0.75	3.23	0.41	1.49	0.32	1.03	0.33	1.70	0.10	0.42	0.23	0.93
GP	0.58	3.38	0.82	3.02	-0.24	-0.83	0.70	3.85	0.49	2.30	0.21	0.96
SG	0.69	3.68	-0.23	-0.83	0.91	2.72	0.48	2.67	0.02	0.10	0.46	2.21
IG	0.30	2.08	-0.13	-0.49	0.43	1.36	0.40	2.50	0.44	2.25	-0.04	-0.19
IK	0.54	3.14	-0.24	-0.86	0.78	2.55	0.73	3.99	0.34	1.73	0.39	1.72
IA	0.70	4.21	0.15	0.59	0.51	1.72	0.53	3.46	0.33	1.81	0.21	1.08
TOTA	0.38	2.49	-0.31	-1.05	0.67	2.11	0.22	1.44	0.54	3.16	-0.32	-1.53
AG	0.73	3.73	0.14	0.48	0.57	1.79	0.75	4.52	0.54	2.96	0.21	1.09
NOA	0.61	3.19	0.23	0.79	0.34	1.04	0.53	3.35	0.61	3.13	-0.08	-0.40
ROA	1.64	4.92	2.24	4.59	-0.47	-0.82	1.08	2.96	1.02	3.05	0.08	0.17
OS	1.67	4.87	0.52	1.22	1.13	2.13	1.08	3.00	1.35	3.65	-0.22	-0.48
SUE	2.00	6.26	1.59	3.28	0.56	1.01	0.70	3.07	1.09	3.42	-0.38	-0.99
COMB	0.87	10.19	0.29	2.80	0.55	4.66	0.59	7.24	0.46	5.36	0.13	1.30

Panel B: Triple Sorts on Institutional Ownership(IO), Institutional Number Changes and Anomalies												
	Low IO Stocks						High IO Stocks					
	LOW		HIGH		LOW-HIGH		LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	1.68	4.42	-0.75	-1.62	2.42	5.21	0.95	3.11	-0.27	-0.78	1.22	4.45
MOM	1.59	4.99	1.08	2.30	0.51	1.06	1.28	3.53	1.50	4.12	-0.22	-0.65
REV	0.97	1.60	1.03	1.29	-0.06	-0.06	-0.39	-0.97	0.52	0.98	-0.91	-1.44
NSI	1.51	4.33	0.30	0.73	1.21	3.48	0.90	4.85	0.42	1.98	0.48	2.11
LSI	1.36	3.52	0.27	0.60	1.09	2.60	0.84	4.08	0.60	2.69	0.24	0.85
IVOL	1.06	2.57	0.58	1.08	0.48	1.03	1.28	3.45	0.77	2.14	0.51	1.20
TO	1.57	1.95	0.22	0.26	1.35	1.88	0.88	2.89	0.14	0.45	0.74	1.83
BM	1.95	3.83	0.88	1.61	1.07	1.86	0.22	0.63	0.28	0.66	-0.06	-0.17
EP	0.77	1.08	0.23	0.38	0.54	0.57	0.75	1.25	0.57	1.25	0.18	0.29
SP	1.81	4.41	1.92	3.27	-0.11	-0.17	0.21	0.67	0.41	1.04	-0.20	-0.59
GP	1.25	3.43	1.59	4.14	-0.34	-0.65	0.95	3.65	0.47	1.74	0.48	1.90
SG	1.42	5.22	0.07	0.18	1.34	2.81	0.31	1.38	0.18	0.76	0.13	0.49
IG	0.75	1.60	0.24	0.62	0.51	0.92	0.53	2.48	0.77	3.35	-0.24	-0.93
IK	0.95	3.18	0.38	0.99	0.57	1.28	0.34	1.53	0.38	1.70	-0.04	-0.16
IA	1.97	6.26	0.35	0.93	1.62	3.29	0.57	2.98	0.65	2.39	-0.08	-0.31
TOTA	0.82	2.74	0.00	0.00	0.82	1.56	0.19	1.08	0.78	3.31	-0.59	-2.33
AG	2.51	7.15	0.51	1.20	2.01	3.54	0.69	3.55	1.03	3.54	-0.34	-1.21
NOA	1.16	3.12	0.45	1.33	0.71	1.39	0.41	2.10	0.48	1.95	-0.07	-0.26
ROA	1.56	1.19	1.75	1.05	-0.18	-0.10	0.46	0.53	0.08	0.05	0.38	0.19
OS	2.33	3.25	0.42	0.41	1.91	1.76	0.76	1.29	1.27	2.16	-0.51	-0.59
SUE	1.12	1.32	0.18	0.22	0.94	0.83	0.08	0.17	1.28	1.89	-1.20	-1.67
COMB	1.45	11.10	0.45	2.14	1.00	5.25	0.55	4.99	0.46	4.07	0.09	0.81

Table 1.6: Sequential sorting: passive institutional investors and anomalies

This table repeats the tests in Table 2 using changes in the number of passive institutional investors. The passive institutional investors include insurance companies, bank trusts, pension funds, university endowments and others. We report the risk-adjusted returns across all 21 anomalies for the LOW and HIGH groups. Alphas are reported as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

Panel A: Long-Short

	LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.42	1.94	0.17	0.83	0.26	1.74
MOM	1.13	4.75	0.85	3.61	0.28	1.70
REV	0.11	0.78	0.16	1.05	-0.05	-0.28
NSI	0.80	6.02	0.53	3.62	0.27	1.69
LSI	0.90	6.63	0.62	4.39	0.28	1.85
IVOL	0.91	5.71	0.65	4.26	0.27	1.68
TO	0.88	5.55	0.35	2.37	0.53	3.13
BM	0.15	0.99	0.22	1.41	-0.06	-0.43
EP	0.39	2.52	0.43	2.73	-0.04	-0.25
SP	0.21	1.33	0.22	1.30	-0.01	-0.07
GP	0.68	4.54	0.44	2.81	0.24	1.72
SG	0.42	3.36	0.03	0.28	0.39	2.70
IG	0.35	3.08	0.37	3.31	-0.02	-0.17
IK	0.51	3.98	0.29	2.38	0.22	1.40
IA	0.77	5.84	0.39	2.89	0.38	2.40
TOTA	0.30	2.88	0.19	1.70	0.11	0.92
AG	0.71	5.13	0.42	3.50	0.28	1.83
NOA	0.64	4.64	0.44	3.35	0.20	1.44
ROA	1.19	5.26	1.13	4.90	0.07	0.23
OS	1.11	5.98	0.76	3.31	0.34	1.24
SUE	1.44	8.62	1.00	4.50	0.44	1.69
COMB	0.67	11.13	0.46	7.89	0.21	3.07

Panel B: Long sides						
	LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat
STREV	-0.31	-1.79	-0.14	-1.08	-0.17	-1.08
MOM	0.08	0.73	0.23	1.56	-0.15	-0.99
REV	-0.21	-1.78	-0.10	-0.86	-0.11	-0.71
NSI	0.05	0.53	0.14	1.52	-0.09	-0.90
LSI	0.11	1.18	0.09	0.94	0.02	0.23
IVOL	0.05	0.51	0.16	1.61	-0.11	-1.04
TO	0.07	0.70	0.04	0.38	0.03	0.26
BM	-0.29	-1.84	-0.02	-0.15	-0.27	-1.90
EP	-0.23	-1.67	0.07	0.63	-0.31	-2.19
SP	-0.32	-2.27	-0.09	-0.77	-0.23	-1.77
GP	0.01	0.06	0.09	0.95	-0.08	-0.63
SG	-0.23	-2.07	-0.32	-3.10	0.09	0.59
IG	-0.14	-1.38	0.01	0.16	-0.15	-1.06
IK	-0.09	-0.86	-0.07	-0.66	-0.01	-0.08
IA	-0.09	-0.87	-0.14	-1.41	0.05	0.31
TOTA	-0.26	-2.14	-0.16	-1.66	-0.09	-0.61
AG	-0.11	-0.93	-0.05	-0.54	-0.05	-0.34
NOA	-0.06	-0.54	0.00	0.01	-0.06	-0.44
ROA	0.38	3.08	0.41	2.73	-0.03	-0.15
OS	0.37	2.73	0.21	1.33	0.16	0.87
SUE	0.53	3.88	0.55	3.20	-0.02	-0.08
COMB	-0.03	-0.39	0.04	0.63	-0.08	-0.72

Panel C: Short sides						
	LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat
STREV	-0.73	-6.03	-0.31	-2.45	-0.43	-3.08
MOM	-1.05	-5.33	-0.62	-4.18	-0.43	-2.76
REV	-0.32	-2.46	-0.25	-2.43	-0.06	-0.45
NSI	-0.75	-5.71	-0.39	-3.68	-0.36	-2.16
LSI	-0.78	-6.07	-0.53	-5.01	-0.25	-1.57
IVOL	-0.87	-6.25	-0.49	-4.52	-0.38	-2.30
TO	-0.81	-5.68	-0.31	-2.91	-0.50	-3.05
BM	-0.44	-4.43	-0.24	-2.23	-0.20	-1.36
EP	-0.62	-5.14	-0.35	-3.30	-0.27	-1.73
SP	-0.52	-4.65	-0.30	-2.74	-0.22	-1.46
GP	-0.67	-4.90	-0.35	-2.70	-0.32	-1.91
SG	-0.65	-4.87	-0.35	-3.48	-0.30	-1.95
IG	-0.49	-3.89	-0.36	-3.74	-0.13	-0.88
IK	-0.59	-4.70	-0.37	-3.98	-0.23	-1.55
IA	-0.86	-6.19	-0.53	-4.41	-0.33	-2.21
TOTA	-0.55	-4.96	-0.35	-3.82	-0.20	-1.50
AG	-0.82	-5.94	-0.48	-4.86	-0.34	-2.20
NOA	-0.70	-5.33	-0.44	-3.96	-0.26	-1.84
ROA	-0.81	-4.13	-0.71	-4.25	-0.09	-0.42
OS	-0.74	-4.34	-0.56	-3.10	-0.18	-0.75
SUE	-0.91	-6.01	-0.45	-2.90	-0.46	-2.18
COMB	-0.70	-6.36	-0.42	-5.26	-0.28	-2.16

Table 1.7: New strategies

This table presents performance of new strategies combining 21 firm characteristics and changes in ownership breadth. We first sort stocks into three tertile portfolios (LOW, MED and HIGH) based on changes in ownership breadth in the prior one quarter. Then we sort stocks into five quintile portfolios based on 21 firm characteristics within each breadth change group. All portfolio breakpoints are determined by sorting firm characteristic using NYSE firms. The combination strategy (COMB) is constructed by taking equal positions across all 21 anomalies. The new strategies are long the stocks in the high-performing quintile (quintile 5) and in the highest changes in breadth ownership tertile (tertile 3), and short the stocks in the low-performing quintile (quintile 1) and in the lowest changes in breadth ownership tertile (tertile 1). We report the equal-weighted and value-weighted average hedge portfolio return and risk-adjusted return for each anomaly. Returns and alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

	Equal Weight				Value Weight			
	Mean Return		FF Alpha		Mean Return		FF Alpha	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	1.12	4.98	1.04	4.90	0.82	3.52	0.74	3.31
MOM	1.20	3.39	1.49	4.48	0.77	2.04	1.12	3.20
REV	0.75	3.82	0.64	3.40	0.66	3.19	0.55	2.81
NSI	1.08	4.43	1.14	6.21	0.78	3.79	0.88	5.34
LSI	1.04	4.30	1.12	5.85	0.84	4.28	0.95	5.70
IVOL	0.65	1.82	1.00	4.29	0.99	2.93	1.32	6.23
TO	1.03	3.45	1.21	5.83	0.56	2.01	0.81	4.31
BM	1.15	5.54	0.84	4.92	0.71	3.53	0.44	3.00
EP	0.97	3.66	0.75	3.92	1.00	4.13	0.80	4.53
SP	1.16	5.19	0.77	4.31	0.80	4.10	0.40	2.79
GP	1.06	4.94	1.21	5.65	0.87	4.64	1.08	5.91
SG	0.73	3.69	0.72	3.64	0.64	2.98	0.57	3.07
IG	0.85	4.68	0.86	4.56	0.79	4.16	0.81	4.22
IK	0.74	3.27	0.75	3.92	0.65	2.57	0.74	4.19
IA	1.15	6.09	1.22	6.42	0.75	3.99	0.80	4.44
TOTA	0.71	3.99	0.72	3.83	0.69	3.75	0.75	3.88
AG	1.16	5.86	1.17	5.71	0.87	4.29	0.84	4.70
NOA	1.02	4.70	1.22	6.00	0.80	4.48	1.01	5.66
ROA	1.13	3.54	1.54	5.14	1.05	3.26	1.55	5.04
OS	0.93	3.04	1.41	5.04	0.70	2.49	1.05	3.94
SUE	1.39	5.27	1.66	7.33	0.90	3.29	1.17	4.78
COMB	1.00	5.82	1.04	6.67	0.79	5.05	0.85	6.36

Table 1.8: Cross-sectional regression analysis

This table examines the marginal effect of changes in ownership breadth on the relation between each anomaly and future return. We run the following monthly Fama-Macbeth regression in model 1:

$$R_{i,t+1} = \alpha + \beta_1 A_{i,t} + \beta_2 D_{i,Num} + \beta_3 (D_{i,Num} \times A_{i,t}) + \varepsilon_{i,t}$$

where $R_{i,t+1}$ is the return on stock i at month $t+1$, $A_{i,t}$is the lagged firm characteristic (anomaly), $D_{i,Num}$ is a dummy variable that takes value 1 if change in the ownership breadth is in the lowest tertile and zero otherwise. In addition, We add other firm characteristics as the controls in the model 2. The control variables are: the logarithm of Amihud (2002) illiquidity, market capitalization, book-to-market ratio and price momentum.

	Reg (1)						Reg (2)					
	β_1	t-stat	β_2	t-stat	β_3	t-stat	β_1	t-stat	β_2	t-stat	β_3	t-stat
STREV	-0.32	-0.55	-0.43	-3.71	-3.23	**-7.94**	-0.90	-1.68	-0.28	-4.41	-3.07	**-7.98**
MOM	0.86	3.49	-0.32	-3.47	0.16	0.61	1.15	5.39	-0.34	-5.39	0.18	0.87
REV	-0.18	-2.13	-0.23	-2.09	-0.12	-1.60	-0.03	-0.35	-0.20	-3.27	-0.12	-1.58
NSI	-1.08	-2.50	-0.34	-2.95	-1.96	**-5.59**	-0.97	-3.16	-0.23	-3.97	-1.56	**-4.52**
LSI	-0.74	-3.14	-0.24	-2.24	-0.75	**-4.14**	-0.82	-4.96	-0.22	-3.85	-0.61	**-3.52**
IVOL	-4.26	-2.88	-0.29	-2.71	-0.09	-0.10	-5.41	-4.74	-0.28	-2.91	0.23	0.25
TO	-5.11	-2.31	-0.25	-2.10	-5.36	**-4.12**	-5.30	-2.80	-0.14	-1.72	-3.95	**-3.22**
BM	0.16	1.75	-0.28	-2.32	0.30	**4.68**	0.27	2.91	-0.17	-2.43	0.17	**2.83**
EP	0.71	1.12	-0.39	-2.88	-0.15	-0.34	0.28	0.66	-0.31	-4.43	-0.08	-0.23
SP	0.20	2.74	-0.45	-3.85	0.13	**3.05**	0.12	1.89	-0.29	-5.00	0.08	1.92
GP	0.77	4.03	-0.43	-2.78	0.12	0.67	0.91	4.63	-0.33	-3.84	0.14	0.82
SG	-0.53	-2.96	-0.27	-2.28	-0.61	**-3.54**	-0.25	-1.60	-0.21	-3.71	-0.55	**-3.31**
IG	-0.22	-3.55	-0.33	-2.77	-0.08	-1.20	-0.13	-2.33	-0.27	-4.56	-0.09	-1.46
IK	-0.15	-2.02	-0.55	-3.62	-0.14	**-2.44**	-0.09	-1.44	-0.42	-4.35	-0.10	-1.91
IA	-0.14	-3.87	-0.87	-5.27	-0.18	**-4.72**	-0.10	-3.26	-0.74	-6.25	-0.18	**-4.97**
TOTA	-2.18	-4.50	-0.34	-2.89	-0.28	-0.53	-1.79	-4.03	-0.29	-4.70	-0.01	-0.02
AG	-1.14	-5.99	-0.29	-2.41	-0.71	**-3.68**	-0.83	-5.07	-0.22	-3.66	-0.64	**-3.44**
NOA	-0.79	-2.55	-0.03	-0.16	-0.49	**-2.55**	-0.93	-4.23	0.10	0.79	-0.63	**-3.47**
ROA	9.35	3.54	-0.14	-1.19	-1.71	-0.65	13.69	5.40	-0.19	-2.18	-2.27	-0.87
OS	-0.16	-3.00	-0.18	-1.48	-0.09	**-2.03**	-0.27	-5.85	-0.31	-3.38	-0.12	**-2.20**
SUE	0.38	8.17	-0.10	-0.87	0.02	0.31	0.39	8.69	-0.20	-2.38	0.04	0.59

Table 1.9: Change in ownership breadth and anomaly returns: independent sorting

We repeat the analysis in Table 2 using independent sorting. We sort stocks into three tertile portfolios (low, medium and high) based on changes in ownership breadth in the prior one quarter and five quintile portfolios based on 21 firm characteristics. Then we independently sort stocks into 3×5 portfolios. All portfolio breakpoints are determined by sorting firm characteristic using NYSE firms. We compute risk-adjusted of each anomaly for the LOW and HIGH groups. We also compute the return difference between the LOW and HIGH groups (column "LOW-HIGH") for each anomaly. Returns and alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

	Average Hedge Portfolio Returns						Fama-French Alphas					
	LOW		HIGH		LOW-HIGH		LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.93	4.18	-0.07	-0.29	1.00	**6.29**	0.77	3.42	0.08	0.33	0.69	**3.38**
MOM	1.05	4.44	0.91	3.20	0.14	0.80	0.71	2.76	1.05	4.12	-0.34	-1.62
REV	0.38	2.42	0.05	0.35	0.32	1.73	0.07	0.42	0.32	1.91	-0.25	-1.36
NSI	0.95	6.59	0.34	2.49	0.61	**4.04**	1.05	7.37	0.27	1.91	0.79	**4.81**
LSI	0.91	6.07	0.31	2.18	0.59	**3.60**	1.08	7.25	0.11	0.76	0.97	**5.30**
IVOL	0.78	3.94	0.25	1.51	0.53	**3.29**	0.89	4.43	0.13	0.70	0.75	**3.22**
TO	0.95	5.51	0.44	2.70	0.51	**2.38**	1.09	6.32	0.06	0.40	1.03	**5.16**
BM	0.64	3.32	0.13	0.49	0.51	**2.36**	0.42	1.98	-0.01	-0.07	0.43	**2.19**
EP	0.31	2.08	0.22	1.23	0.09	0.55	0.28	1.68	0.05	0.36	0.23	1.19
SP	0.60	3.59	0.06	0.25	0.54	**2.91**	0.36	1.82	0.03	0.19	0.32	1.72
GP	0.64	4.31	0.45	2.83	0.19	1.32	0.73	4.67	0.28	1.82	0.44	**2.70**
SG	0.48	3.86	-0.03	-0.23	0.51	**3.58**	0.40	3.10	0.03	0.27	0.37	**2.43**
IG	0.44	4.09	0.29	2.59	0.15	1.11	0.41	3.50	0.31	2.71	0.11	0.75
IK	0.58	4.80	0.16	1.28	0.41	**2.61**	0.63	4.75	0.29	2.26	0.34	**2.06**
IA	0.80	6.41	0.33	2.57	0.46	**3.30**	0.75	5.99	0.37	2.69	0.38	**2.50**
TOTA	0.26	2.49	0.22	1.99	0.04	0.33	0.22	1.97	0.28	2.50	-0.06	-0.47
AG	0.67	5.26	0.35	2.65	0.32	**2.20**	0.68	5.58	0.35	2.60	0.33	**2.28**
NOA	0.58	4.60	0.44	3.13	0.14	1.11	0.67	4.87	0.47	3.27	0.19	1.29
ROA	1.43	5.89	0.99	3.78	0.44	1.71	1.34	5.59	0.96	3.64	0.39	1.42
OS	1.05	4.81	0.62	2.74	0.43	1.62	1.20	4.89	0.56	2.50	0.64	**2.32**
SUE	1.44	7.41	1.07	5.91	0.38	1.51	1.01	6.20	1.06	6.60	-0.05	-0.25
COMB	0.72	12.71	0.33	5.27	0.39	**5.42**	0.67	11.00	0.32	5.25	0.36	**4.80**

Table 1.10: Change in ownership breadth and anomaly returns: value-weighted portfolios

We repeat the analysis in Table 2 by forming value-weighted portfolios for each anomaly. We first sort stocks into three tertile portfolios (low, medium and high) based on changes in ownership breadth in the prior one quarter. Then we sort stocks into five quintile portfolios based on 21 firm characteristics within each breadth change group. All portfolio breakpoints are determined by sorting firm characteristic using NYSE firms. We compute the Fama French alpha of each anomaly for the LOW and HIGH groups. We also compute the alpha difference between the LOW and HIGH groups (column "LOW-HIGH") for each anomaly. Alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity- consistent standard errors of White (1980).

	LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.06	0.24	0.14	0.61	-0.09	-0.52
MOM	0.45	1.64	0.27	1.05	0.18	0.84
REV	0.07	0.39	0.06	0.38	0.01	0.04
NSI	0.61	4.00	0.20	1.28	0.41	**2.25**
LSI	0.73	4.72	0.31	2.22	0.43	**2.45**
IVOL	1.16	6.13	0.21	1.16	0.96	**5.06**
TO	0.57	2.97	0.00	0.01	0.56	**2.63**
BM	0.02	0.10	-0.12	-0.86	0.14	0.69
EP	0.48	2.49	-0.07	-0.41	0.55	**2.59**
SP	0.15	0.74	-0.15	-0.92	0.30	1.35
GP	0.69	4.25	0.34	2.26	0.36	**2.32**
SG	0.37	2.36	-0.23	-1.49	0.59	**3.62**
IG	0.59	4.07	0.11	0.70	0.48	**2.54**
IK	0.63	4.17	0.02	0.14	0.61	**3.55**
IA	0.55	3.66	0.23	1.63	0.32	1.92
TOTA	0.21	1.41	0.18	1.32	0.03	0.20
AG	0.60	4.58	0.06	0.42	0.54	**3.41**
NOA	0.60	4.65	0.47	3.72	0.13	0.78
ROA	1.24	4.75	0.81	3.16	0.43	1.41
OS	0.77	3.42	0.30	1.21	0.47	1.58
SUE	0.80	3.73	0.28	1.44	0.52	1.93
COMB	0.51	7.95	0.15	2.27	0.36	**4.76**

Table 1.11: Change in ownership breadth and anomaly returns: alternative measures

We repeat analysis in Table 2 using an alternative measure of active institutional investors (Gaspar, Massa and Matos, 2005 (GMM)) and an alternative measure of changes in ownership breadth (Choi, Jin and Yan, 2011(CJY)). GMM (2005) defined active institutional investors using the institution's portfolio turnover (churn rate). The churn rate of institution *f* in a given quarter *q* is computed as follows

$$CR_{f,q} = \frac{\sum_{i \in Q} |N_{i,f,q}P_{i,q} - N_{i,f,q-1}P_{i,q-1} - N_{i,f,q-1}\Delta P_{i,q}|}{\sum_{i \in Q} \frac{N_{i,f,q}P_{i,q} + N_{i,f,q-1}P_{i,q-1}}{2}}$$

where $N_{i,f,q}$ refers to the number of shares of stock *i* held by institution *f* in quarter q, $P_{i,q}$ refers to the stock price at the same time, Q represents the set of companies held by institution *f*. The average churn rate is calculated over the past four quarters to rank the institutions, and those with above median average churn rate are classified as active institutions in each quarter. CJY (2011) introduce changes in market value of a stock as an alternative measure of changes in ownership breadth. We present the Fama and French alpha of each anomaly for the LOW and HIGH group for these two measures. We also compute the return difference between the LOW and HIGH groups. Alphas are expressed as percentage per month. The t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980).

	GMM's Measure						CJY's Measure					
	LOW		HIGH		LOW-HIGH		LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.77	3.28	0.04	0.17	0.74	**4.37**	0.73	3.40	0.07	0.31	0.67	**3.39**
MOM	0.99	3.84	0.98	4.37	0.01	0.06	0.89	3.44	0.82	3.75	0.07	0.34
REV	0.39	2.37	0.00	0.00	0.39	**2.13**	0.24	1.51	0.33	2.02	-0.09	-0.49
NSI	1.01	7.05	0.58	4.33	0.43	**2.66**	0.97	6.78	0.30	2.08	0.67	**3.96**
LSI	0.94	6.11	0.51	3.54	0.43	**2.58**	1.08	7.12	0.15	1.02	0.93	**5.09**
IVOL	0.77	3.59	0.48	3.23	0.29	1.66	0.90	4.50	0.07	0.39	0.83	**3.57**
TO	0.97	5.49	0.35	2.25	0.62	**3.02**	1.10	6.35	0.10	0.63	1.00	**5.16**
BM	0.49	2.48	0.03	0.15	0.46	**2.35**	0.43	1.99	-0.03	-0.18	0.46	**2.21**
EP	0.27	1.78	0.37	2.36	-0.11	-0.65	0.22	1.31	0.10	0.64	0.12	0.60
SP	0.48	3.17	0.15	0.81	0.33	**2.10**	0.39	1.98	-0.03	-0.14	0.42	**2.19**
GP	0.78	5.00	0.72	4.82	0.07	0.45	0.82	5.33	0.32	2.03	0.50	**3.10**
SG	0.67	4.83	-0.09	-0.77	0.76	**4.86**	0.40	2.97	0.00	0.00	0.40	**2.60**
IG	0.54	4.45	0.32	3.18	0.22	1.64	0.36	3.05	0.28	2.61	0.08	0.62
IK	0.61	4.74	0.08	0.71	0.53	**3.15**	0.53	4.18	0.20	1.61	0.33	**2.11**
IA	1.05	7.20	0.37	2.74	0.68	**4.51**	0.84	6.09	0.40	2.76	0.45	**2.73**
TOTA	0.42	3.74	0.11	1.16	0.31	**2.30**	0.22	1.91	0.20	1.96	0.02	0.13
AG	0.96	6.26	0.27	2.04	0.68	**4.04**	0.80	6.43	0.35	2.55	0.44	**2.86**
NOA	0.97	6.49	0.47	3.64	0.50	**3.67**	0.87	6.16	0.55	3.77	0.31	**2.08**
ROA	1.19	4.95	1.48	6.61	-0.29	-1.29	1.33	5.65	0.97	4.33	0.36	1.46
OS	1.28	5.76	1.02	4.74	0.25	1.03	1.33	6.27	0.62	3.31	0.71	**3.02**
SUE	1.13	6.70	1.35	7.51	-0.21	-0.91	1.10	6.30	1.14	7.21	-0.04	-0.20
COMB	0.76	12.83	0.41	6.98	0.35	**4.94**	0.70	10.71	0.30	5.09	0.40	**5.35**

Table 1.12: Change in ownership breadth and anomaly returns: different size

We repeat analysis in Table 2 for different size of stocks: all-but-tiny stocks and large stocks. All-but-tiny stocks are those whose sizes are higher than NYSE 20 percentile; and large stocks are those whose sizes are higher than NYSE 50 percentile. For each size group, we first sort stocks into three tertile portfolios (low, medium and high) based on quarterly changes in ownership breadth. Then we sort stocks into five quintile portfolios based on 21 firm characteristics within each breadth change group. All portfolio breakpoints are determined by sorting firm characteristics using NYSE firms. For each size group, we report the risk-adjusted return of each anomaly for the LOW and HIGH group. We also compute the return difference between the LOW and HIGH groups. Returns are expressed in percentage per month. The t-statistics are computed using heteroskedasticity- consistent standard errors of White (1980).

	All-but-tiny Stocks						Large Stocks					
	LOW		HIGH		LOW-HIGH		LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat	ret	t-stat
STREV	0.51	2.28	0.03	0.13	0.48	3.24	0.15	0.68	0.08	0.35	0.06	0.37
MOM	0.66	2.66	0.51	2.16	0.15	0.83	0.52	2.03	0.36	1.33	0.16	0.78
REV	-0.04	-0.27	0.17	1.11	-0.21	-1.30	0.14	0.85	0.14	0.87	0.00	0.00
NSI	0.71	5.17	0.33	2.09	0.38	2.33	0.60	3.92	0.27	1.69	0.32	1.69
LSI	0.75	5.46	0.29	1.88	0.47	2.72	0.68	4.35	0.25	1.68	0.43	2.33
IVOL	0.97	5.64	0.16	0.98	0.81	4.58	0.82	4.66	0.06	0.30	0.76	3.72
TO	0.68	4.16	0.18	1.17	0.50	2.64	0.59	3.09	-0.03	-0.15	0.62	2.72
BM	0.15	0.92	-0.08	-0.45	0.23	1.40	0.09	0.52	-0.17	-1.00	0.26	1.29
EP	0.32	2.16	0.15	0.84	0.17	0.93	0.39	2.17	0.00	-0.01	0.39	1.77
SP	0.33	1.98	-0.01	-0.05	0.34	1.94	0.21	1.23	-0.10	-0.49	0.31	1.45
GP	0.73	4.59	0.27	1.55	0.47	2.90	0.63	3.67	0.27	1.57	0.36	2.09
SG	0.24	1.71	-0.03	-0.26	0.28	1.86	0.33	2.12	-0.16	-1.04	0.49	2.81
IG	0.27	2.38	0.24	1.97	0.03	0.25	0.29	2.27	0.23	1.56	0.06	0.40
IK	0.57	4.55	0.10	0.82	0.47	3.32	0.50	3.31	-0.13	-0.92	0.63	3.57
IA	0.68	4.96	0.30	2.16	0.38	2.78	0.56	3.82	0.15	1.06	0.40	2.54
TOTA	0.21	1.76	0.16	1.34	0.06	0.38	0.06	0.45	0.17	1.20	-0.10	-0.65
AG	0.53	4.14	0.26	1.96	0.27	2.00	0.56	4.06	0.07	0.48	0.49	3.01
NOA	0.67	5.09	0.42	2.92	0.25	1.92	0.53	3.82	0.39	2.69	0.13	0.78
ROA	1.00	4.29	0.89	3.67	0.11	0.45	0.44	1.48	0.46	1.79	-0.01	-0.05
OS	0.92	4.29	0.42	1.82	0.51	1.84	0.20	0.67	0.25	1.07	-0.05	-0.17
SUE	1.06	6.00	0.87	4.41	0.19	0.83	0.56	2.66	0.56	2.69	0.00	0.01
COMB	0.54	8.25	0.24	3.89	0.30	4.07	0.42	5.60	0.13	1.79	0.29	3.50

Table 1.13: Change in ownership breadth and anomaly returns: the Carhart's four-factor alpha

This table repeat the tests in Table 2 using Carhart's four-factor alpha. We first sort stocks into 3 tertile portfolios (LOW, MED and HIGH) based on changes in ownership breadth in the prior one quarter. Then we sort stocks into 5 quintile portfolios based on 21 firm characteristics within each breadth change group. We report the risk-adjusted return of each anomaly for the LOW and HIGH groups. The t-statistics are computed using heteroskedasticity- consistent standard errors of White (1980).

	LOW		HIGH		LOW-HIGH	
	ret	t-stat	ret	t-stat	ret	t-stat
STREV	1.13	4.17	0.11	0.50	1.02	5.25
MOM	0.15	0.70	0.09	0.58	0.06	0.27
REV	0.36	2.23	0.11	0.73	0.25	1.19
NSI	0.74	4.75	0.52	3.58	0.23	1.44
LSI	0.70	4.17	0.47	3.09	0.24	1.35
IVOL	0.28	1.39	0.09	0.59	0.19	1.12
TO	0.61	3.77	0.40	2.31	0.21	1.11
BM	1.14	5.55	0.49	2.67	0.65	2.68
EP	0.43	2.23	0.47	2.86	-0.04	-0.20
SP	0.95	4.88	0.36	1.77	0.59	2.49
GP	0.73	4.47	0.46	2.88	0.27	1.80
SG	0.51	3.60	-0.01	-0.07	0.52	3.00
IG	0.40	3.26	0.25	2.25	0.15	0.93
IK	0.29	2.22	0.04	0.30	0.25	1.38
IA	0.86	5.90	0.36	2.67	0.50	3.01
TOTA	0.46	3.77	0.08	0.72	0.38	2.72
AG	0.90	6.54	0.34	2.35	0.56	3.21
NOA	0.80	5.54	0.50	3.39	0.30	2.15
ROA	0.89	3.40	0.81	3.31	0.08	0.29
OS	0.96	3.86	0.45	1.31	0.51	1.09
SUE	1.10	4.92	0.94	5.51	0.16	0.62
COMB	0.67	11.88	0.33	5.49	0.34	4.35

Chapter 2: Short Selling and Readability in Financial Disclosures

(Jointly with Minxing Sun)

2.1 Introduction

An important channel for corporate managers to communicate investors and analysts about a firm's financial disclosure is the form 10-K. Readability of 10-K reports is one aspect of textual analysis aiming to measure how effectively managers convey valuation relevant information to investors and analysts (Loughran and McDonald, 2014, 2016). In accounting and finance literature, many papers have studied the link between annual report readability and firm performance. Li (2008) find that firms with lower reported earnings tend to have annual reports that are more difficult to read using the Fog Index and the number of words contained in the annual report. Biddle, Hilary, and Verdi (2009) hypothesize that firms with high annual report readability are associated with greater capital investment efficiency. Lehavy, Li, and Merkley (2011) argue that firms with more readable annual reports have lower analyst dispersion and greater earnings forecast accuracy. Loughran and McDonald (2014) propose 10-K document size as a simple proxy to measure readability that outperforming the Fog index. They find that lower readability of 10-Ks (larger file size) is associated with high return volatility, earnings forecast errors, and earnings forecast dispersion [1].

[1] Miller (2010) finds that small investors significantly trade fewer shares of companies that have low annual report readability. Lawrence (2013) documents that retail investors invest more in firms whose annual reports have better readability. Hwang and Kim (2016) find that reductions in readability of annual reports of equity closed-end investment companies (CEFs) reduce firms' value.

In this paper, we examine how the relaxation of short-sale constraints affects the readability in financial disclosures. In practice, it is difficult to test the impact of a reduction in short-sale constraints on readability for at least two reasons. First, proxies of short-sale constraints (eg. short interests, ownership breadth .etc) are noisy and endogenously determined. Second, the observed relation between short-sale constraint proxies and annual report readability suffers endogeneity problem. For instance, high level of short interests could cause or result from low annual report readability.

To overcome this endogeneity issue, we test the casual effect of short-sale constraints on annual report readability using a natural experiment, Regulation SHO. Traditionally, the tick test and Nasdaq's bid price test imposed constraints on short selling. In July 2004, the SEC initiated a pilot program under Rule 202T of Regulation SHO to remove short-sale price tests for pilot stocks. The SEC randomly selected 986 pilot stocks from the Russell 3000 index. During May, 2 2004 and August 6, 2007, pilot stocks were exempted from short-sale price tests and significantly reduced short-sale constraints as opposed to nonpilot stocks (SEC, 2007; Diether, Lee, and Werner, 2009. Prior studies show evidence to support that short-selling activities increases significantly for pilot stocks compared to nonpilot stocks (e.g., Boehmer et al. 2008; Diether, Lee, and Werner, 2009; Grullon, Michenaud, and Weston, 2015). In addition, the SEC eliminated short-sale price tests for all exchange-listed stocks including nonpilot stocks on July 2007. Since Regulation SHO is an exogenous shock to short-sale constraints and has specific beginning and ending date, it provides us an ideal and clear setting to test the causal effect of the relaxation of short-sale constraints on annual report readability using difference-in-differences (DiD) analysis.

We begin by confirming whether treatment group is randomly selected by comparing firm characteristics of pilot and non-pilot firms one year before the announcement of the pilot program. Following Loughran and Mcdonald (2014), we use 10-K document file size as a proxy for readability. An annual report with a larger 10-K file size is considered as less readable. We find that pilot and non-pilot stocks have similar readability of 10-Ks (*log(file size)*), firm size, operations and financial complexity, analysts' coverage and institutional ownership before the Regulation SHO. Next, we examine readability difference between pilot and non-pilot firms for three time horizons: pre-, during- and post-Regulation SHO periods. The baseline results imply a significantly negative correlation between a reduction in short-sale constraints and 10-K readability. Pilot firms significantly increase 10-K document file sizes (decrease 10-K readability) as opposed to non-pilot stocks during the program period.

To further verify the relation between a reduction in short-sale constraints and annual report readability, we run multivariate regression analysis. We find that the readability of 10-Ks for the pilot stocks is 9.4% lower than that for the non-pilot firms during the Regulation SHO period. After controlling for other determinants of annual report readability, a reduction in short-sale constraints due to the Regulation SHO still leads to a 6.9% lower 10-K readability for pilot stocks compared to nonpilot stocks. We argue that corporate managers have incentives to maintain the stock price, especially when the short-sale constraints are loose, because managers' compensation is positively related to stock price. To prevent short selling, pilot firms, whose short-sale constraints are significantly reduced due to the Regulation SHO, obscure valuation-relevant information by increasing 10-K file sizes and therefore raise analysis costs for short sellers. In addition, the SEC eliminated short-sale price tests for all exchange-listed stocks on July 6, 2007. This setting provides us an alternative approach to further confirm the negative correlation between reduction in short-sale constraints and annual report readability. According to our difference-in-difference analysis, nonpilot firms, whose short-sale constraints are significantly reduced after the Regulation SHO, increase 10-K file size (decrease 10-K readability) by 6.2% compared to pilot stocks.

The relation between the relaxation of short-sales constraint and annual report readability is not uniform in the cross-section. First, firms with good earnings or low risk have no incentive to obscure valuation relevant information because they have good financial strength. Second, better corporate governance companies and firms that have higher institutional ownership or analysts' coverage are less likely to manipulate annual report readability because corporate governance, institutional investors and analysts have monitoring effects on firms. We examine several cross-sectional analyses using firm size, ROA, return volatility, institutional ownership, analysts' coverage, corporate social responsibility and corporate governance index. We find that readability is significantly decreased for firms that are smaller, less profitable or riskier; for firms that have lower institutional ownership or analyst coverage; and for firms with worse corporate governance or corporate social responsibility.

Our study makes several contributions to the literature. First, our study adds to the literature on the effect of short selling on corporate decisions[2]. The closely related paper is Fang, Huang and Karpoff (2015), who find that short selling reduces earnings

[2] See He and Tian, 2014; Gilchrist, Himmelberg, and Huberman, 2005; Grullon, Michenaud, and Weston, 2015

management and helps detect frauds using the Regulation SHO. We show that short selling activities have significant effect on annual report readability which is another prospect of financial statement quality. Second, we identify a new determinant of annual report readability, short-sales constraints, adding to the accounting and finance literature.[3]

The rest of the paper is organized as follows. Section 2.2 discusses the related literature. Section 2.3 describes sample selection and reports summary statistics. Section 2.4 presents the main results. Section 2.5 concludes.

2.2 Related Literature

2.2.1 Short-Sale Price Tests in U.S. Equity Markets

Short-sale price tests were initially introduced in the U.S. equity markets in the 1930s to avoid bear raids by short sellers in declining markets. The NYSE adopted an uptick rule in 1935, which was replaced in 1938 by a stricter SEC rule, Rule 10a-1, also known as the "tick test." The latter rule mandates that a short sale can only occur at a price above the most recently traded price (plus tick) or at the most recently traded price if that price exceeds the last different price (zero-plus tick). In 1994, the National Association of Securities Dealers (NASD) adopted its own price test (the "bid test") under Rule 3350. Rule 3350 requires that a short sale occur at a price one penny above the bid price if the bid is a downtick from the previous bid.

On June 23, 2004, the SEC adopted Regulation SHO to provide a new regulatory framework for short-selling in the U. S. stock markets. The Regulation SHO removed the

[3] for a review, see Loughran and McDonald (2016)

tick test for a group of randomly selected stocks from the Russell 3000 index in order to evaluate the effectiveness and necessity of short-selling restrictions. On July 28, 2004, 986 stocks were selected as the pilot stocks. The pilot stocks were exempt from the tick test staring from May 2, 2005. The temporary suspension expired on July 6, 2007 when the SEC permanently suspended the tick test for all the publicly-traded U.S. companies. The permanent suspension of the tick test drew criticisms from firms and former regulators, including former SEC chairman Christopher Cox. The criticism intensified with the financial crisis in 2008-2009 due to the concern that financial stocks may be subject to market manipulations via short-selling. On February 24, 2010, the SEC reinstated the uptick rule, but only under the circumstance when a security's price drops by 10% or more from the last day's closing price.

2.2.2 Readability in Financial Disclosures

Research on the readability of accounting narratives has a long history (Jones and Shoemaker, 1994). Much of the earlier work on readability suffered from small sample sizes or problematic methodologies. For instance, Lewis, Parker, Pound, and Sutcliffe (1986) provide an analysis of various readability measures (i.e., Fog and Flesch Indexes) for financial reports using only nine Australian firms over a four year period. Li (2008) is the first paper that examines the link between annual report readability and firm performance. He measures the readability of annual reports (i.e., Form 10-Ks) using the Fog Index and the number of words contained in the annual report. The Fog Index is a function of two variables: average sentence length (in words) and complex words (defined as the percentage of words with more than two syllables). A Fog Index value of 16 implies that the reader needs sixteen years of education—essentially a college degree—to comprehend the text on a first reading. Li (2008) finds that firms with lower reported earnings tend to have annual reports that are harder to read (i.e., high Fog Index values or high word counts). Li also finds that companies with more readable annual reports have higher earnings persistence.

Following Li (2008), other researchers have used the Fog Index as a measure of annual report readability. Biddle, Hilary, and Verdi [2009] find that firms with high reporting quality (using the Fog Index and two other variables) are associated with greater capital investment efficiency. Guay et al. (2015) find that companies with less readable annual reports (based on six different readability measures including the Fog Index) tend to mitigate this negative readability effect by issuing more managerial forecasts of earnings per share, sales, and cash flows. Miller (2010) finds that small investors trade significantly fewer shares of firms with high Fog Index values and word counts (i.e., less readable annual reports) around the 10-K filing date. Less readable annual reports should be harder to process, especially for less sophisticated investors. Focusing on the link between readability and analyst coverage, Lehavy, Li, and Merkley (2011) find that more readable annual reports, as measured by the Fog Index, have lower analyst dispersion and greater earnings forecast accuracy. They find that 10-K readability is related to how many analysts cover a stock. Firms with higher Fog Index values, after controlling for company characteristics, have more analysts covering the stock. The readability of analyst reports is also associated with investor behavior. De Franco, Hope, Vyas, and Zhou (2015) find that more readable analyst reports are associated with significantly higher trading volume over a three-day window surrounding the analyst report date. Furthermore, Hwang and Kim (2016) find that reductions in readability of annual reports of equity closed-end investment companies (CEFs) reduce firms' value.

Loughran and McDonald (2014) empirically demonstrate that the Fog Index is a poorly specified readability measure when applied to business documents. They find that the Fog Index is not significant in explaining analyst dispersion or earnings surprises and propose that the natural log of gross 10-K file size is a simple readability measure. They find that firms with bigger 10-K file sizes are significantly linked with larger subsequent stock return volatility, analyst dispersion, and absolute earnings surprises.

2.3 Data Description

2.3.1 Sample Selection

Our sample is constructed based on the Russell 3000 index in June 2004. On July 28, SEC announced a list of 986 pilot stocks that would trade without being subject to any price tests during the event (Regulation SHO) period. We excluded stocks that were not previously subject to price tests (i.e., not listed on NYSE, Amex, or NASDAQ-NM) and stocks that went public or had spin-offs after April 30, 2004. Then we sort the stocks based on their daily dollar volume computed over the June 2003 to May 2004 period. Among the 2,952 stocks, we identify 986 pilot stocks and 1966 nonpilot stocks.

We obtain the SEC annual filing data from WRDS SEC Readability and Sentiment database. This database contains the detailed information about firms' SEC filing since 1994, for instance, filing date, file size, number of words, etc. Following

Loughran and McDonald (2014), we include all 10-K filings, 10-K 405, 10KSB and 10KSB40 filings. We require that the firm have a Compustat Permeant ID match, be ordinary common stock, have at least 2,000 words in the 10-K, and have a gap of at least 180 days between two filings. To construct our control variables, we collect financial statement information from CRSP/Compustat Merged, stock returns from CRSP, institutional holdings from Thomson Reuters 13-F, analyst coverage data from IBES, corporate events information from Thomson Reuters SDC Platinum M&A and Global New Issues databases, and litigation data from Securities Class Action Clearinghouse.

Our sample period covers 75 months including firms whose fiscal year ending dates are between May 1, 2002 and June 30, 2004 for the *pre-event* period, between May 1, 2005 and June 30, 2007 for the *during-event* period and between May 1, 2008 and June 30, 2010 for the *post-event* period. We classify May 1, 2005 to June 30, 2007 as *during-event* period is because the Reg SHO program effectively ran from May 2, 2005 to July 6, 2007. In our sample, we excludes financial firms (SIC 6000-6999) and utilities firms (SIC 4900-4949) because disclosure requirements are significantly different for these highly regulated industries. We require that firms have data to calculate firm characteristics over the entire sample period, resulting in a final sample of 724 pilot stocks and 1,482 non-pilot stocks.

2.3.2 Key variables

Following Loughran and McDonald (2014), we measure annual report readability of firms using the natural logarithm of 10-K report size. Loughran and Mcdonald (2014) show that traditional readability measures like the Fog Index are poorly specified when used to evaluate financial documents. They argue that the file size of the 10-K is a good proxy for document readability and is better gauge how effectively managers convey valuation-relevant information to investors and analysts.

Following Li (2008), we control for a set of firm characteristics that determine the annual report readability. Our control variables include size (the natural logarithm of the market value of equity at the end of the fiscal year), firm age (the natural logarithm of firm age since a firm's first appearance in CRSP monthly return file), special items (special items to asset ratio), stock return volatility, earnings volatility, business complexity (the natural logarithm of the number of business and geographic segments), financial complexity (the natural logarithm of the number of non-missing items in Compustat) and corporate events (SEO and MA dummy variables). We also include other controls variables affect the annual report readability: profitability (ROA), institutional ownership, analysts' coverage, the natural logarithm of prior one year 10-K filing size, and a litigation dummy variable. The detailed descriptions of all key variables are in Appendix A.

2.3.3 Summary Statistics

Table 2.1 reports the summary statistics of all key variables from the whole sample. All variables are winsorized at their 1% and 99% levels to minimize the effect of outliers. On average, a firm has 10-K document size of 1.83 megabyte, the market value of 4639.5 million, BM ratio of 0.55, age of 21.6 years, number of business segment of 2.21, number of geographic segment of 2.74, number of analysts of 5.43, institutional ownership of 0.67, return volatility of 0.13, earning volatility of 0.06, number of non-missing items of 357.45, special item ratio of -0.02, ROA of 0.00.

*** Table 2.1 ***

2.4 Empirical Results

2.4.1 Univariate difference-in-difference Analysis

The Regulation SHO is a natural experiment to study the causal effects of short selling on the annual report readability because the selection of pilot and non-pilot stocks was random and the costs of short selling were significantly reduced for pilot stocks. Therefore, the difference-in-differences (DiD) method is appropriate to study the effects of short selling on the annual report readability.

To verify the selection of pilot stocks was random, we compare the firm characteristics of pilot and non-pilot firms one year before the announcement of the pilot program (July, 2004). Table 2.2 presents the summary statistics and mean differences of firm characteristics between the treatment (pilot) and control (non-pilot) groups. We report t-statistics of the two-sample t-tests and z-statistics of the Wilcoxon Ranksum tests. We find that the treatment group has similar firm characteristics with the control group. The results in Table 2.2 further support that Regulation SHO is a natural experiment that is suitable for testing the effects of removing short-sale constraints on stocks.

*** Table 2.2 ***

Table 2.3 reports the univariate DiD results for three time periods: *pre-event* period, *during-event* period and *post-event* period using 10-K document file size. We find several interesting patterns. First, among the *pre-event* period, the treatment group on average has similar 10-K file size with the control group. The mean *log(file size)* is -0.09 for the treatment group, is -0.06 for the control group and the difference between two groups being -0.03. The *t*-statistic of the two-sample *t-test* for the difference in means is -1.25 and the z-statistic of the Wilcoxon Ranksum test is -1.06, both insignificant. Second, during the event period, the treatment group on average has significantly larger 10-K file size than the control group, suggesting that the treatment group has a lower readability of 10-Ks than the control group. Specifically, the mean *log(file size)* is 0.47 for the treatment group and is 0.41 for the control group. The difference between the treatment and control groups is 0.06 with a t-statistic of 2.66 and a z-statistic of 1.97, implying that the treatment group has a 6.2% higher 10-K file size as opposed to the control group. Third, for the period after the event, the mean 10-K file size is 0.75 for both treatment and control groups. The two-sample t-test and the Wilcoxon Ranksum test show that the mean differences between the treatment and control groups are both insignificant.

*** Table 2.3 ***

The univariate DiD results show that despite pilot stocks have similar readability in financial disclosures with the non-pilot firms before and after the Regulation SHO program period, pilot firms significantly reduce readability of 10-Ks during the program period.

2.4.2 Multivariate difference-in-differences Analysis

In this section, we extend our tests using multivariate regressions. We summarize the results in Table 2.4. Following Fang, Huang, and Karpoff (2015), we run the following OLS regression:

$$Log\left(file\ size_{i,t}\right) = \alpha + \beta_1 * pilot_i + \beta_2 * pilot_i * during_t + during_t + controls + \varepsilon_{i,t} \qquad (1)$$

where $Log\left(file\ size_{i,t}\right)$ is the natural logarithm of 10-K document file size for firm i at year t. $Pilot_i$ is a dummy variable that equals one if a stock is selected as a pilot stock in Regulation SHO's pilot program and zero otherwise. $During_t$ is a dummy variable that equals one if the end of a firm's fiscal year t falls between May 1, 2005 and June 30, 2007 and zero otherwise.

*** Table 2.4 ***

The regression results of equation (1) are reported in column (1) in Table 4. The coefficient of interest is β_2, which captures the causal effect of short selling on annual report readability. The coefficient of $pilot_i * during_t$, β_2 is 0.091 and significant at the 1% level, implying that 10-K file sizes of pilot firms are 9.1% higher than those of non-pilot firms during the Regulation SHO program period, consistent with the results in baseline difference-in-differences tests. The coefficient on $Pilot_i$ is insignificant, suggesting that pilot and non-pilot firms exhibit similar 10-K file sizes before the pilot program.

In column (2), we augment equation (1) by including control variables previously shown to determine the annual report readability (Li, 2008) and other additional control variables: size, book-to-market ratio, firm age, special items to asset ratio, stock return and earnings volatility, business complexity, financial complexity, analysts coverage, institutional ownership, prior one year 10-K file size and corporate events (SEO, MA and litigation dummy variables). We find that β_2 reduces to 0.062 and significant at the 5% level after controlling for firm characteristics.

In column (3), we repeat the analysis in column (1) by adding industry (Fama and French 48 industries) fixed effects and fiscal year fixed effects. The variable $During_t$ is omitted because it is perfectly correlated with the fiscal year fixed effects. The standard errors are clustered by firm. The coefficient of $pilot_i * during_t$, β_2 increases to 0.94 and significant at the 1% level. In addition, we extend equation (1) by including controls, industry and year fixed effects in column (4). We find that β_2 is 0.069 and significant at the 5% level, suggesting that 10-K file sizes of pilot firms are 6.9% higher than those of non-pilot firms during the Regulation SHO program period. In other words, reduction in short selling costs causes pilot firms reduces readability of 10-K reports by 6.9%. Consistent with the univariate DiD analysis, the multivariate regression results show that pilot stocks significantly reduce readability of 10-Ks during the program period. We argue that corporate managers have incentives to maintain the stock price, especially when the short-sale constraints are loose, because managers' compensation is positively correlated with stock price. To prevent short selling, pilot firms, whose short-sale constraints are significantly reduced due to Regulation SHO, obscure valuation-relevant information by increasing 10-K file sizes and therefore raise analysis costs for short sellers.

The SEC eliminated short-sale price tests for all exchange-listed stocks on July 6, 2007 (Securities Exchange Act of 1934 Release No. 34-55970, July 3, 2007). This setting provides us an alternative approach to test the relation of short selling and annual report readability. We can test whether the pattern of short selling and annual report readability reverse during the *post-event* period. We run DiD tests using the same group of pilot and non-pilot firms and retain the sample from May, 2005 to June 2010 in Table 5. The regression is as follows:

$$Log\left(file\ size_{i,t}\right) = \alpha + \beta_1 * nonpilot_i + \beta_2 * nonpilot_i * post_t + controls + \varepsilon_{i,t} \quad (2)$$

where $nonpilot_i$ is a dummy variable that equals one if a stock is not selected as a pilot stock in Regulation SHO's pilot program and zero otherwise. $Post_t$ is a dummy variable that equals one if the end of a firm's fiscal year *t* falls between May 1, 2008 and June 30, 2010 and zero otherwise. *Industry* and *Year* are the industry fixed effects (Fama and French 48 industries) and fiscal year fixed effects, respectively. The results are reported in Table 2.5

*** Table 2.5 ***

Because non-pilot firms significantly reduced short selling costs among the *post-event* period, we expect that 10-K readability of non-pilot stocks would decrease compare with that of pilot stocks. In other words, the 10-K document size of non-pilot firms would increase as opposed to that of pilot firms and the coefficient on $nonpilot_i * post_t$ (β_2) would be positive. In column (1), β_2 is 0.062 and significant at the 5% level, consistent with our prediction. The result is similar after controlling for firm characteristics in column (2). Specifically, β_2 is 0.049 and also significantly at the 5% level, implying that non-pilot stocks increase 10-K document sizes by 4.8% compared with pilot firms. The results in Table 2.5 further confirm that reductions in short selling costs significantly reduce annual report readability of firms.

2.4.3 Cross-sectional analysis based on firm characteristics

The relation between a reduction in short-sales constraint and annual report readability is not uniform in the cross-section. Corporate managers' decision on annual report readability can be affected by various firm characteristics including size, return volatility, ROA, institutional ownership, analysts' coverage, corporate social responsibility and corporate governance. First, firms with good earnings or low risk have no incentive to obscure valuation relevant information because they have good financial strength. Second, better corporate governance companies and firms that have higher institutional ownership or analysts' coverage are less likely to manipulate annual report readability because corporate governance, institutional investors and analysts have monitoring effects on firms. We expect that the effect of a reduction in short-sale constraints on annual report readability would be significant only for firms that are smaller, less profitable or riskier; for firms that have lower institutional ownership or analyst coverage; and for firms with worse corporate governance or corporate social responsibility.

For each firm characteristic, we separate the whole sample into two groups: high and low. Then we run regressions in equation (1) for each group. The results are reported in Table 2.6. Panel A presents the results for firm size, analysts' coverage, institutional ownership and ROA. Panel B reports the findings for return volatility, corporate social responsibility and corporate governance.

*** Table 2.6 ***

The findings in Table 2.6 are consistent with our prediction. In Panel A, for size, the coefficient of $pilot_i * during_t$ is insignificant in large size group, but β_2 is 0.108 and significantly at the 5% level among small size group. For analysts' coverage, β_2 is positively significantly at the 5% level among low analyst coverage group, whereas β_2 is insignificant for high analyst coverage group. In addition, the slope of $pilot_i * during_t$ is 0.127 and significantly at the 1% level for low institutional ownership stocks, suggesting that readability of pilot firms are 12.7% lower than that of nonpilot firms. For firms with low institutional ownership, β_2 is insignificant. For ROA, β_2 is also positively significant at the 5 % level in high ROA firms, whereas it is insignificant for low ROA stocks.

We find similar results in Panel B. For return volatility, β_2 is positively significant at the 5% level in high risk firms, whereas it is insignificant for low risk stocks. Moreover, for corporate social responsibility and corporate governance, β_2 is significantly positive for the bad groups but insignificant for the good groups. To sum up, the evidence in Table 2.6 show that the effect of a reduction in short-sale constraints on annual report readability is significant only for firms that are smaller, less profitable or riskier; for firms that have lower institutional ownership or analyst coverage; and for firms with worse corporate governance or corporate social responsibility.

2.4.4 Robustness Check

In this subsection, we perform a placebo test for our DiD analysis reported in Table 3 to strengthen our causal argument. We address the concern that our identification tests mainly rely on the Regulation SHO that took place in 2004. Unobservable shocks which occurred prior to 2004 but are unrelated to Regulation SHO could have driven results. We conduct a placebo test by taking the true set of pilot and non-pilot firms identified by Regulation SHO but artificially picking a "pseudo-event" year when we assume a regulatory shock reduced short selling costs. We assume that the Regulation SHO is effective from May, 2002 to June, 2004. We repeat the difference-in-differences tests in Table 2.3 using the same set of pilot and non-pilot stocks. The results are presented in Table 2.7. The coefficient estimates on *pilot*during* is negative and insignificant.

*** Table 2.6 ***

2.5 Conclusion

This paper documents a causal relation between change in short-sale constraints and annual report readability. To establish causality, we use exogenous variation in short-sale constraints generated by a quasi-natural experiment, Regulation SHO, which randomly selects a group of stocks from the Russell 3000 index into a pilot program and removes short selling price tests. We find that the readability of 10-K reports for the pilot stocks significantly decreases during the program period. Furthermore, the negative relation between the short-sale constraints and annual report readability are more pronounced for firms that are smaller, less profitable or riskier; for firms that have lower institutional ownership or analyst coverage; and for firms with worse corporate governance or corporate social responsibility. Our results have important implications to users of financial statements, for instance, analysts and investors.

Appendix A: Definition of variables

Annual report readability measure:

10-K file size: Loughran and Mcdonald (2014) argue that file size of a 10-K is a good proxy for readability. Larger 10-K file size of a firm is less readable. The readability measure is defined as the natural logarithm of 10-K document filing size in fiscal year t.

Experiment related variables:

Pilot: A dummy variable that equals one if a stock is selected as a pilot stock in Regulation SHO's pilot program and zero otherwise.

During: A dummy variable that equals one if the end of a firm's fiscal year *t* falls between May 1, 2005 and June 30, 2007 and zero otherwise.

Post: A dummy variable that equals one if the end of a firm's fiscal year *t* falls between May 1, 2008 and June 30, 2010 and zero otherwise.

Firm characteristics:

Prior one year 10-K file size: The 10-K file sizes in fiscal year t are positively correlated with the 10-K file sizes in fiscal year t-1. The prior one year readability measure is defined as the natural logarithm of 10-K document filing size in fiscal year t-1.

Firm size: Larger firms have more complex 10-K reports. The size is defined as the natural logarithm of the market equity of firms at the end of fiscal year t.

Firm age: Older firms have more readable annual reports because there is less information asymmetry and less information uncertainty for these firms. The firm age is the number of years since a firm's first appearance in the CRSP monthly stock return file. We use the natural logarithm of the firm age in the regressions.

Special items (SI): Firms with a significant amount of special items are more likely to experience some unusual events. Firms with lower special items have more complex 10-K reports. SI is defined as the amount of special items scaled by book value of assets.

Volatility of business: Firms with higher volatility of business environment have more complex 10-K reports. To capture the volatility of business, we use two measures: stock return volatility (Ret_vol, measured as the standard deviation of the monthly stock returns in the prior year) and earnings volatility (Earn_vol, measured as the standard deviation of the operating earnings during the prior five fiscal years).

Complexity of operations: Firms with more complex operations are more likely to have complex 10-K reports. We use the number of business segments (NBSEG) and the number of geographic segments (NGSEG) to capture the operation complexity of firms.

Financial Complexity: Firms with more complex financial situations are more likely to have complicated 10-K reports. We use the logarithm of the number of non-missing items in Compustat as a proxy for financial complexity (NITEMS). Firms are more financially complex if they need to report more items in annual reports.

Corporate events: Unusual corporate events may require extra and more detailed disclosures, so firms with corporate events have more complex 10-K reports. We consider three events: seasoned equity offerings (SEO), merger and acquisition (MA) and litigation (LIT). The dummy variable SEO is equal to 1 if for a year in which a company has a common equity offering in the secondary market according to the SDC Global New Issues database and 0 otherwise. The dummy variable MA is set to 1 if for a year in which a company is an acquirer based on the SDC Platinum M&A database and 0 otherwise. The dummy variable LIT is equal to 1 if for a year in which a company appears in Securities Class Action Clearinghouse database at Stanford University and 0 otherwise.

Analyst coverage: Firms that are covered by more analysts are more likely to have complex 10-K reports. The analyst coverage is defined as the logarithm of the number of analysts following a stock from IBES database.

Profitability (ROA): Firms that earn higher profits have more readable 10-Ks. ROA is defined as the income before extraordinary items divided by lagged total assets.

Institutional ownership (IO): Firms with higher institutional ownership have more complex 10-K reports. The Institutional ownership is captured from Thomson Reuters 13-F database.

CSR score: A company's CSR score is the number of strengths minus the number of concerns from MSCI ESG KLD STATS.

Governance index: The governance index is introduced by Gompers, Ishii, and Metrick (2003). We apply the governance scores in 2006 for the data after 2006 because the index only covers from 1990 to 2006.

Table 2.1: Summary statistics

This table reports the summary statistics of firm characteristics of the treatment (pilot) and control (non-pilot) groups measured based on 2004 Russell 3000 index firms. The sample consists of 75 months including firms whose fiscal year ending dates are between May 1, 2002 and June 30, 2004 for the *pre-event* period, between May 1, 2005 and June 30, 2007 for the *during-event* period, and between May 1, 2008 and June 30, 2010 for the *post-event* period. We require firms that have data available to calculate firm characteristics and 10-k filing size over time. Definitions of variables are listed in the Appendix A.

Variables	N	Mean	SD
10-K file size (in megabytes)	11,880	1.83	1.49
Log (file size)	11,880	0.34	0.74
Lagged file size	11,492	1.53	1.19
Firm Size (in millions)	11,879	4.64	12.92
Age	11,880	21.60	17.67
NBSEG	11,880	2.21	1.58
NGSEG	11,880	2.74	2.17
# of analysts	10,679	5.43	4.99
RET_VOL	11,867	0.13	0.08
EARN_VOL	11,002	0.06	0.07
ROA	11,747	0.00	0.21
IO	11,880	0.67	0.23
Litigation	11,880	0.03	0.17
MA	11,880	0.37	0.48
SEO	11,880	0.06	0.24
NITEMS	11,880	357.45	27.61
SI	11,743	-0.02	0.07

Table 2.2: Firm characteristics before announcement of Regulation SHO

This table compares firm characteristics of the treatment (pilot) and control (non-pilot) groups one year before the announcement of the Regulation SHO (July 2004). Definitions of variables are listed in Appendix A. We report the t-statistics of the two-sample t-test and z-statistics of Wilcoxon Ranksum test for the difference between the treatment and control groups. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Treatment			Control			Difference	
	N	Mean	SD	N	Mean	SD	T-stat	Wilcoxon
Log(file size)	686	0.03	0.70	1,379	0.04	0.69	0.12	0.02
Log(lagged file size)	657	-0.18	0.78	1,303	-0.11	0.76	-1.87*	1.61
Log(firm size)	686	6.97	1.40	1,379	6.89	1.38	1.27	-1.40
Log(firm age)	686	2.60	0.85	1,379	2.54	0.86	1.63	-1.71*
Log(nbseg)	686	1.04	0.42	1,379	1.04	0.44	-0.34	-0.02
Log(ngseg)	686	1.13	0.45	1,379	1.15	0.47	-0.72	0.44
Log(# of analysts)	618	1.46	0.97	1,233	1.45	0.93	0.26	-0.32
RET_VOL	684	0.13	0.08	1,377	0.14	0.08	-1.36	1.04
EARN_VOL	602	0.06	0.07	1,185	0.07	0.09	-3.28***	2.45***
ROA	672	0.02	0.16	1,333	0.01	0.19	1.66*	-1.14
IO	686	0.62	0.21	1,379	0.61	0.23	1.27	-0.80
Litigation	686	0.03	0.18	1,379	0.03	0.17	0.20	-0.20
MA	686	0.40	0.49	1,379	0.35	0.48	1.99**	-1.99**
SEO	686	0.09	0.28	1,379	0.10	0.29	-0.50	0.50
Log(NITEMS)	686	5.80	0.04	1,379	5.80	0.04	0.93	-1.22
Log(SI)	680	-0.01	0.04	1,362	-0.01	0.04	0.02	0.01

Table 2.3: Univariate difference-in-differences (DiD) tests

This table presents the results of the univariate difference-in-differences (DiD) test on how the exogenous shock to short selling costs, Regulation SHO, affects the readability of firms' 10-K filings. We report the summary statistics of the natural logarithm of firms'10-K filing size for treatment (pilot) and control (non-pilot) groups under three time periods: pre-event (fiscal year ending date is between May 2002 and June 2004), during-event (fiscal year ending date is between May 2005 and June 2007), and post-event (fiscal year ending date is between May 2008 and June 2010) and differences in mean. ***, **, and * indicate significance at the 1%, 5%, and 10% levels using two-tailed tests.

	Treatment Group (Pilot=1)		Control Group (Pilot=0)		Treatment-Control		
	N	Mean	N	Mean	Differences	T-stat	Wilcoxon
Pre-event	1,397	-0.09	2,845	-0.06	-0.03	-1.25	-1.06
During-event	1,137	0.47	2,310	0.41	0.06	**2.66**	**1.97**
Post-event	988	0.75	1,930	0.75	0.00	0.16	0.11

Table 2.4: Multivariate difference-in-differences (DiD) tests

This table reports the multivariate difference-in-differences tests on how reductions in short-sale constraints affect annual report readability using OLS regressions. Column (1) reports the results in the following regression:

$$Log\left(file\ size_{i,t}\right) = \alpha + \beta_1 * pilot_i + \beta_2 * pilot_i * during_t + during_t + \varepsilon_{i,t}$$

where $Log\left(file\ size_{i,t}\right)$ is the natural logarithm of 10-K document file size for firm *i* at year *t*. $Pilot_i$ is a dummy variable that equals one if a stock is selected as a pilot stock in Regulation SHO's pilot program and zero otherwise. $During_t$ is a dummy variable that equals one if the end of a firm's fiscal year t falls between May 2005 and June 2007 and zero otherwise. We augment the model by including the control variables in column (2), by adding industry (Fama and French 48 industries) and year fixed effects in column (3) and by further including the control variables, and industry and year fixed effects in column (4). We omit *during* in both columns (3) and (4) to avoid multicollinearity. The sample includes pre-event (fiscal year ending date is between May 2002 and June 2004) and during-event (fiscal year ending date is between May 2005 and June 2007). Variable definitions are provided in the Appendix A. Standard errors clustered by firms are displayed in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Log(file size)			
Variables	(1)	(2)	(3)	(4)
SHO	-0.031	-0.014	-0.022	-0.019
	(0.025)	(0.022)	(0.031)	(0.022)
*SHO*During*	**0.091*** **	**0.062** **	**0.094*** **	**0.069** **
	(0.033)	**(0.028)**	**(0.033)**	**(0.028)**
During	0.456***	0.089***		
	(0.019)	(0.027)		
Log(lagged file size)		0.540***		0.533***
		(0.013)		(0.013)
Log(firm size)		0.060***		0.057***
		(0.007)		(0.007)
Log(numbseg)		0.063***		0.042**
		(0.018)		(0.018)
Log(numgseg)		-0.004		0.008
		(0.015)		(0.017)
Log(# of analysts)		-0.004		0.002
		(0.010)		(0.011)
IO		0.020		0.039
		(0.040)		(0.040)
Earn_vol		-0.089		0.011
		(0.111)		(0.116)
Log(SI)		-0.075		0.007
		(0.132)		(0.136)
Ret_vol		0.600***		0.621***
		(0.131)		(0.134)
ROA		-0.265***		-0.282***
		(0.063)		(0.068)
Log (firm age)		-0.024**		-0.025**
		(0.012)		(0.013)
Log (# of non-missing items)		0.389*		0.681***
		(0.207)		(0.217)
Litigation		0.035		0.037
		(0.038)		(0.038)
SEO		0.003		-0.005
		(0.031)		(0.031)
MA		0.053***		0.051***
		(0.015)		(0.015)
Constant	-0.054***	-2.629**	-0.364***	-4.447***
	(0.014)	(1.186)	(0.118)	(1.253)
Observations	7,689	6,017	7,689	6,017
R-squared	0.109	0.467	0.152	0.477
Industry FE	NO	NO	YES	YES
Year FE	NO	NO	YES	YES

Table 2.5: Multivariate difference-in-differences (DiD) test: reverse of SHO

This table test whether the pattern of short selling and annual report readability reverse during the *post-event* period. The SEC eliminated short-sale price tests for all exchange-listed stocks on July 6, 2007. We run DiD tests using the same group of pilot and non-pilot firms and retain the sample from during-event (fiscal year ending date is between May 2005 and June 2007) period to post-event (fiscal year ending date is between May 2008 and June 2010) period. The regression in column (1) is as follows:

$$Log(file\ size_{i,t}) = \alpha + \beta_1 * nonpilot_i + \beta_2 * nonpilot_i * post_t + Industry_j + Year_t + \varepsilon_{i,t}$$

where $Log(file\ size_{i,t})$ is the natural logarithm of 10-K document file size for firm *i* at year *t*. $nonpilot_i$ is a dummy variable that equals one if a stock is not selected as a pilot stock in Regulation SHO's pilot program and zero otherwise. $Post_t$ is a dummy variable that equals one if the end of a firm's fiscal year *t* falls between May 2008 and June 2010 and zero otherwise. We omit *post* to avoid multicollinearity. *Industry* and *Year* are the industry fixed effects (Fama and French 48 industries) and fiscal year fixed effects, respectively. Variable definitions are provided in the Appendix A. Standard errors clustered by firms are displayed in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Log(file size)	
Variables	(1)	(2)
Nonpilot	-0.067**	-0.043**
	(0.028)	(0.018)
*Nonpilot*Post*	**0.062****	**0.049****
	(0.030)	**(0.023)**
Log(lag_file size)		0.606***
		(0.014)
Log(firm size)		0.076***
		(0.006)
Log(numbseg)		0.046***
		(0.016)
Log(numgseg)		0.022
		(0.015)
Log(# of analysts)		-0.020**
		(0.010)
IO		0.013
		(0.035)
Earn_vol		0.069
		(0.111)
Log(SI)		0.015
		(0.032)
Ret_vol		0.034
		(0.096)
ROA		-0.295***
		(0.053)
Log (firm age)		-0.001
		(0.011)
Log (# of non-missing items)		0.578***
		(0.210)
Litigation		0.108***
		(0.039)
SEO		0.026
		(0.029)
MA		0.046***
		(0.014)
Constant	0.194***	-3.824***
	(0.039)	(1.228)
Observations	6,365	5,388
R-squared	0.136	0.536
Industry FE	YES	YES
Year FE	YES	YES

Table 2.6: Multivariate difference-in-differences (DiD) test: subgroups

We repeat the difference-in-differences tests in Table 3 across different subgroups. Panel A reports the regression results for the small and large firms, the low and high analysts' coverage firms the bad and good earnings firms, and the low and high institutional ownership firms. Panel B reports the results for the low and high risk firms, the low and high CSR score firms and the bad and good corporate governance firms. Variable definitions are provided in the Appendix A. Standard errors clustered by firms are displayed in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Panel A: Cross-sectional regression analysis: size, analysts' coverage, IO, ROA								
	LARGE SIZE	SMALL SIZE	HIGH Analysts	LOW Analysts	High IO	Low IO	High ROA	Low ROA
pilot	-0.020	-0.018	-0.001	-0.014	0.023	-0.071**	-0.020	-0.028
	(0.031)	(0.032)	(0.030)	(0.030)	(0.031)	(0.033)	(0.032)	(0.032)
*pilot*During*	**0.051**	**0.108****	**0.020**	**0.101****	**0.021**	**0.127*****	**0.062**	**0.082****
	(0.037)	**(0.042)**	**(0.037)**	**(0.040)**	**(0.039)**	**(0.042)**	**(0.041)**	**(0.040)**
Log(lagfsize)	0.540***	0.528***	0.521***	0.558***	0.508***	0.547***	0.539***	0.511***
	(0.018)	(0.019)	(0.017)	(0.017)	(0.018)	(0.019)	(0.018)	(0.018)
Log(firm size)			0.061***	0.044***	0.070***	0.047***	0.057***	0.066***
			(0.008)	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)
Log(numbseg)	0.029	0.080***	0.019	0.087***	0.034	0.042*	0.025	0.051*
	(0.023)	(0.029)	(0.022)	(0.026)	(0.026)	(0.026)	(0.024)	(0.027)
Log(numgseg)	0.032	0.008	-0.011	0.064**	0.000	0.022	-0.003	-0.004
	(0.024)	(0.024)	(0.023)	(0.026)	(0.025)	(0.025)	(0.024)	(0.024)
Log(# of analysts)	0.051***	0.013			-0.016	0.011	0.005	-0.002
	(0.014)	(0.014)			(0.015)	(0.015)	(0.016)	(0.014)
IO	0.001	0.054	0.064	0.044				
	(0.063)	(0.056)	(0.060)	(0.050)				
Earn_vol	-0.223	0.096	-0.100	0.081	-0.132	0.148	0.113	0.110
	(0.219)	(0.134)	(0.154)	(0.155)	(0.201)	(0.140)	(0.213)	(0.130)
Log(SI)	-0.013	-0.077	-0.037	0.006	-0.084	0.064	-0.602	-0.202*
	(0.259)	(0.166)	(0.193)	(0.184)	(0.214)	(0.172)	(0.624)	(0.122)
Ret_vol	0.312	0.601***	0.551***	0.565***	1.061***	0.328*	0.888***	0.378**
	(0.231)	(0.175)	(0.192)	(0.174)	(0.217)	(0.174)	(0.255)	(0.161)
ROA	-0.358***	-0.149*	-0.291***	-0.193**	-0.374***	-0.210***		
	(0.138)	(0.078)	(0.092)	(0.077)	(0.121)	(0.079)		
Log (firm age)	0.016	-0.054**	-0.025	-0.038**	-0.025	-0.018	-0.019	-0.031*
	(0.016)	(0.022)	(0.017)	(0.017)	(0.018)	(0.019)	(0.018)	(0.018)
Log (# of non-missing items)	0.682**	0.661*	0.867***	0.412	0.608*	0.706**	1.107***	0.072
	(0.291)	(0.342)	(0.294)	(0.292)	(0.319)	(0.303)	(0.324)	(0.300)
Litigation	0.064	0.026	0.055	0.024	0.080	-0.015	0.036	0.037
	(0.050)	(0.061)	(0.047)	(0.063)	(0.057)	(0.047)	(0.062)	(0.047)
SEO	0.017	-0.038	-0.011	0.007	0.028	-0.044	0.019	-0.014
	(0.049)	(0.040)	(0.040)	(0.045)	(0.042)	(0.045)	(0.048)	(0.040)
MA	0.051**	0.069***	0.032*	0.054**	0.039*	0.066***	0.081***	0.018
	(0.020)	(0.024)	(0.019)	(0.023)	(0.021)	(0.023)	(0.022)	(0.021)
Observations	3,189	2,829	3,360	3,242	3,229	2,788	3,158	2,859
R-squared	0.470	0.470	0.463	0.501	0.455	0.516	0.500	0.447
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Panel B: Cross-sectional regression analysis: return volatility, CSR, governance						
	HIGH RET_VOL	LOW RET_VOL	HIGH CSR	LOW CSR	Good Gov	Bad GOV
pilot	-0.015	-0.022	-0.002	-0.041	-0.011	-0.023
	(0.031)	(0.032)	(0.030)	(0.034)	(0.029)	(0.034)
*pilot*During*	**0.083****	**0.054**	**0.022**	**0.127*****	**0.052**	**0.092****
	(0.040)	**(0.040)**	**(0.038)**	**(0.042)**	**(0.036)**	**(0.044)**
Log(lagfsize)	0.514***	0.540***	0.519***	0.545***	0.511***	0.542***
	(0.019)	(0.017)	(0.016)	(0.020)	(0.017)	(0.019)
Log(firm size)	0.049***	0.070***	0.042***	0.072***	0.056***	0.074***
	(0.011)	(0.010)	(0.010)	(0.011)	(0.009)	(0.012)
Log(numbseg)	0.035	0.037	0.069***	0.007	0.040	0.045*
	(0.027)	(0.024)	(0.024)	(0.028)	(0.025)	(0.027)
Log(numgseg)	0.004	0.012	0.015	0.009	0.016	0.009
	(0.023)	(0.026)	(0.023)	(0.027)	(0.022)	(0.029)
Log(# of analysts)	-0.003	0.000	0.008	0.005	-0.007	0.020
	(0.014)	(0.016)	(0.014)	(0.017)	(0.013)	(0.017)
IO	0.008	0.056	0.096*	-0.056	0.017	0.101
	(0.053)	(0.061)	(0.053)	(0.064)	(0.048)	(0.079)
Earn_vol	0.020	0.262	0.126	-0.141	-0.073	0.373
	(0.120)	(0.307)	(0.165)	(0.164)	(0.127)	(0.271)
Log(SI)	-0.179	0.488*	0.139	-0.140	-0.054	0.484**
	(0.160)	(0.275)	(0.213)	(0.184)	(0.169)	(0.239)
Ret_vol			0.625***	0.661***	0.645***	0.470*
			(0.185)	(0.205)	(0.156)	(0.271)
ROA	-0.173**	-0.763***	-0.278***	-0.267**	-0.196***	-0.793***
	(0.073)	(0.160)	(0.095)	(0.104)	(0.074)	(0.165)
Log (firm age)	-0.020	-0.030*	-0.008	-0.043**	-0.039**	-0.027
	(0.020)	(0.017)	(0.018)	(0.019)	(0.019)	(0.018)
Log (# of non-missing items)	0.578*	0.741**	0.771***	0.525	0.107	1.616***
	(0.328)	(0.289)	(0.294)	(0.331)	(0.289)	(0.317)
Litigation	0.034	0.057	0.101*	-0.048	0.054	0.016
	(0.044)	(0.073)	(0.056)	(0.051)	(0.051)	(0.058)
SEO	-0.043	0.047	0.013	-0.031	-0.023	0.020
	(0.036)	(0.057)	(0.042)	(0.046)	(0.034)	(0.067)
MA	0.058***	0.042**	0.043**	0.059**	0.074***	0.017
	(0.022)	(0.021)	(0.021)	(0.024)	(0.021)	(0.023)
Observations	2,902	3,115	3,508	2,509	3,715	2,302
R-squared	0.442	0.511	0.453	0.523	0.457	0.522
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Table 2.7: The placebo tests

This table reports the placebo tests results. We assume that the Regulation SHO is effective from May, 2002 to June, 2004. We repeat the difference-in-differences tests in Table 3 using the same set of pilot and non-pilot stocks. Variable definitions are provided in the Appendix A. Standard errors clustered by firms are displayed in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

	Log(file size)	
Variables	(1)	(2)
SHO	-0.015	-0.012
	(0.025)	(0.021)
*SHO*During*	-0.017	-0.023
	(0.029)	(0.031)
Log(lagged file size)		0.478***
		(0.017)
Log(size)		0.033***
		(0.008)
Log(numbseg)		0.061***
		(0.021)
Log(numgseg)		0.004
		(0.021)
Log(# of analysts)		0.007
		(0.044)
IO		0.047***
		(0.012)
Earn_vol		0.027
		(0.121)
Log(SI)		0.085
		(0.127)
Ret_vol		0.368***
		(0.112)
ROA		-0.390***
		(0.073)
Log (firm age)		-0.019
		(0.014)
Log (# of non-missing items)		0.449**
		(0.214)
Litigation		0.045
		(0.047)
SEO		-0.019
		(0.031)
MA		0.089***
		(0.017)
Constant	12.589***	3.691***
	(0.043)	(1.211)
Observations	7,929	5,481
R-squared	0.258	0.442
Industry FE	YES	YES
Year FE	YES	YES

BIBLIOGRAPHY

Abarbanell, Jeffery S., Brian J. Bushee, and Jana Smith Raedy, 2003, Institutional investor preferences and price pressure: The case of corporate spin-offs, *Journal of Business* 76, 233-261.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Asquith, Paul, Parag Pathak, and Jay Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243-276.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis, *Journal of Finance* 32, 663-682.

Basu, Sanjoy, 1983, The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 129-156.

Bernard, Victor L., and Jacob K. Thomas, 1989, Post-earnings-announcement drift: delayed price response or risk premium, *Journal of Accounting Research* 27:1-36.

Biddle, Gary C., Gilles Hilary, and Rodrigo S. Verdi, 2009, How does financial reporting quality relate to investment efficiency?. *Journal of Accounting and Economics* 48, 112-131.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Chan, Kalok, and Wai-Ming Fong, 2000, Trade size, order imbalance, and the volatility–volume relation, *Journal of Financial Economics* 57, 247-273.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2010, An alternative three-factor model, Unpublished paper.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein, 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171-205.

Cheng, Si, Allaudeen Hameed, Avanidhar Subrahmanyam, and Sheridan Titman, 2015, Short-term reversals: The effects of institutional exits and past returns, *Journal of Financial and Quantitative Analysis,* Forthcoming.

Choi, James J., Li Jin, and Hongjun Yan, 2012, What does stock ownership breadth measure, *Review of Finance* 17, 1239-1278.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609-1651.

Cremers, Martijn, and Ankur Pareek, 2015, Short-term trading and stock return anomalies: Momentum, reversal, and share issuance, *Review of Finance* 19,1649-1701.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605-1643.

De Bondt, Werner F.M., and Richard H. Thaler, 1985, Does the stock market overreact, *Journal of Finance 40*, 793-805.

De Bondt, Werner F.M., and Richard H. Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557-581.

Diether, Karl B., Kuan-Hui Lee, and Ingrid M. Werner, 2009, It's SHO time! Short-sale price tests and market quality, *Journal of Finance* 64, 37–73.

Edelen, Roger M., Ozgur S. Ince, and Gregory B. Kadlec, 2014, Institutional investors and stock return anomalies, *Journal of Financial Economic* 119, 472-488.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene F., and Kenneth R. French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491-518.

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653-1678.

Fang, Vivian W., Allen H. Huang, and Jonathan M. Karpoff, 2015, Short selling and earnings management: A controlled experiment, *Journal of Finance* 71, 1251-1294.

Gao, Xiaohui, and Jay R. Ritter, 2010, The marketing of seasoned equity offerings, *Journal of Financial Economics* 97, 33-52.

Green, Jeremiah, John R.M. Hand, and Mark T. Soliman, 2011, Going, going, gone? The apparent demise of the accruals anomaly, *Management Science* 57, 797-816.

Guay, Wayne R., Delphine Samuels, and Daniel J. Taylor, 2015, Guiding through the fog: Financial statement complexity and voluntary disclosure, Working paper

Gilchrist, Simon, Charles P. Himmelberg, and Gur Huberman, 2005, Do stock price bubbles influence corporate investment?. *Journal of Monetary Economics* 52, 805-827.

Grullon, Gustavo, Sebastien Michenaud, and James Weston, 2015, The real effects of short-selling constraints, *Review of Financial Studies* 28, 1737-1767.

Hanson, Samuel G., and Adi Sunderam, 2014, The growth and limits of arbitrage: Evidence from short interest, *Review of Financial Studies* 27, 1238-1286.

He, Jie, and Xuan Tian., 2015, SHO time for innovation: The real effects of short sellers, *Kelley School of Business Research Paper*.

Hwang, Byoung-Hyoun, and Hugh Hoikwang Kim, It Pays to Write Well, *Journal of Financial Economic*, Forthcoming.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297-331.

Huang, Allen, Reuven Lehavy, Amy Zang, and Rong Zheng, 2015, Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach, *Ross School of Business Paper* 1229.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

Jones, Charles M., Gautam Kaul, and Marc L. Lipson, 1994, Transactions, volume, and volatility, *Review of Financial Studies* 7, 631-651.

Jones, Michael John, and Paul A. Shoemaker, 1994, Accounting narratives: A review of empirical studies of content and readability, J*ournal of Accounting Literature* 13, 142.

Karpoff, Jonathan M., and Xiaoxia Lou, 2010, Short sellers and financial misconduct, *Journal of Finance* 65, 1879-1913.

Kogan, Leonid, and Mary Tian, 2012, Firm characteristics and empirical factor models: a data-mining experiment, working paper.

Lawrence, Alastair, 2013, Individual investors and financial disclosure, *Journal of Accounting and Economics* 56, 130-147.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

Lee, Charles, and Bhaskaran Swaminathan, 2000, Price momentum and trading volume, *Journal of Finance* 55, 2017-2069.

Lewis, N. R., Parker L.D., Pound G.D. and Sutcliffe Pet, 1986, Accounting report readability: The use of readability techniques, *Accounting and Business Research* 16, 199-213.

Lewellen, Jonathan, 2011, Institutional investors and the limits of arbitrage, *Journal of Financial Economics* 102, 62-80.

Lewellen, Jonathan, 2014, The cross section of expected stock returns, *Critical Finance Review* 4, 1-44.

Li, Feng, 2008. Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and economics* 45, 221-247.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23-51.

Loughran, Tim, and Bill McDonald, 2014, Measuring readability in financial disclosures, *Journal of Finance* 69, 1643-1671.

Loughran, Tim, and Bill McDonald, 2016, Textual analysis in accounting and finance: A survey, *Journal of Accounting Research* 54, 1187-1230.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2008, The new issues puzzle: testing the investment-based explanation, *Review of Financial Studies* 21, 2825-2855.

Miller, Edward M., 1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151-1168.

Miller, Brian P.,2010, The effects of reporting complexity on small and large investor trading, *The Accounting Review* 85, 2107-2143.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277-309.

Novy-Marx, Robert, 2013, The other side of value: Good growth and the gross profitability premium, *Journal of Financial Economics*108, 1 - 28.

Ohlson, James A, 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109-131.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921-945.

Securities and Exchange Commission, 2007, Economic analysis of the short sale price restrictions under the regulation SHO pilot, Office of Economic Analysis.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52: 35-55.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review 71*, 289-315.

Titman, Sheridan, Kuo-Chiang Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677-700.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

Xing, Yuhang, 2008, Interpreting the value effect through the Q-theory: An empirical investigation, *Review of Financial Studies* 21, 1767-1795.

You, Haifeng, and Xiao-jun Zhang, 2009, Financial reporting complexity and investor underreaction to 10-K information, *Review of Accounting Studies* 14, 559-586.