INSTITUTIONAL INVESTORS AND FINANCIAL STATEMENT ANALYSIS

By

NICOLE YUNJEONG CHOI

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
College of Business

MAY 2009

To the faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of NICOLE YUNJEONG CHOI find it satisfactory and recommend that it be accepted.

Richard W. Sias, Ph. D., Chair

John R. Nofsinger, Ph. D.

Harry J. Turtle, Ph. D.

ACKNOWLEDGEMENT

First, I would like to extend heartfelt appreciation to my advisor, Richard Sias, without whose support and guidance I would not have been where I am now. He has the determination and insight of a true scholar. I hope I can be even half as good a mentor to my future students as he has been to me. I have been extremely fortunate to have him as my advisor during my Ph.D program and I look forward to working with him for many more years to come.

I would also like to thank John Nofsinger who has been tremendously supportive as my committee member and a Ph.D coordinator throughout my Ph.D years and during my job search process. I admire his creativity and the breadth of his curiosity a great deal. I am grateful to Harry Turtle, another member of my dissertation committee, who has encouraged me and helped me throughout the entire process with his precision and excellent writing skills.

I also thank all the other members of the Department of Finance, Insurance and Real Estate at Washington State University – Gene Lai, Michael McNamara, Donna Paul, Nathan Walcott, David Whidbee and Lily Xu – for their support and valuable advice. I especially thank Sandra Boyce for her administrative and motherly support for the six years that I have been at Washington State University.

I would not have been able to finish this dissertation and my Ph.D degree without the support from friends at the department and in Pullman. I acknowledge all the fellow Ph.D students in the department of Finance at Washington State University – Cherry, Heather, Kevin, Kainan, Chune, Sean, Chengping, Erin and Bela – for being great colleagues and friends. I would especially like to thank Athena, Sanatan and Abhi for always having been there to share good and bad times with me during the four years I have known them. They have been my colleagues, friends, family and everything I needed to go through tough times during the program.

I would also like to thank my beloved friends in Pullman, especially Jisung's and Hyuntaek's family for providing me with a family atmosphere and countless homemade meals I could never have dreamt of without them. Because of them, I have been less homesick than I might have been.

I am also grateful to the members of Hog Heaven Toastmasters club. Because of the superb communication training I got from the club, my presentations were more polished and I was far more comfortable in the social settings than I might have been otherwise. Beyond the skills that I learned from the club, I thank them also for the friendly environment they provided me.

Last but not least, I would like to acknowledge all the selfless and unconditional support I have received from my family. I have all the respect in the world for my parents. I can never express enough how much I am proud of them and how much I am thankful that I am their daughter. I would also like to thank my sister, Sunyong, and brother, Hyunsuk, for bearing with me for all those years.

INSTITUTIONAL INVESTORS AND FINANCIAL STATEMENT ANALYSIS

Abstract

By Nicole Yunjeong Choi, Ph.D.
Washington State University
May 2009

Chair: Richard W. Sias

My dissertation consists of two essays related to institutional investors and financial statement analysis. In the first paper, we examine whether institutional investors follow each other into and out of the same industries. Our empirical results reveal strong evidence of institutional industry herding. The cross-sectional correlation between the fraction of institutional traders buying an industry this quarter and the fraction buying last quarter, for example, averages 40%. Additional tests suggest that correlated signals primarily drive institutional industry herding. Our results also provide empirical support for 'style investing' models.

The second paper investigates the relation between changes in financial health, subsequent returns, and demand by individual and institutional investors to differentiate between the rational and irrational pricing explanation for why financial statement based analysis predicts the future returns. Recent studies show changes in financial health forecast future returns. Piotroski (2000, 2005) and Fama and French (2006) point out that there are two potential explanations for this predictability. First, a riskier firm (with a higher expected return) must have higher expected income growth to justify the same book-to-market ratio as a safer firm. Thus,

controlling for book-to-market ratios, firms with higher income growth should have higher returns and expectations are realized (on average). Alternatively, changes in financial health may predict future returns because market participants are slow to react to signals contained in financial statements, i.e., expectations are slowly revised over time. I investigate net trading of institutional investors to test whether investors' expectations are realized or revised. Consistent with the latter interpretation, improving financial health predicts both future returns and future demand by institutional investors.

TABLE OF CONTENTS

ACKNOWLEDGEMENT iii
ABSTRACT v
LIST OF TABLES x
LIST OF FIGURES xi
DEDICATION xii
CHAPTER ONE: GENERAL INTRODUCTION 1
CHAPTER TWO: INSTITUTIONAL INDUSTRY HERDING 3
1. Introduction 3
2. Background and data 9
2.1. Herding 9
2.2. Empirical tests of institutional stock herding 12
2.3. Data 12
3. Tests for institutional industry herding 14
3.1. Correlation between contemporaneous and lag institutional industry demand 14
3.2. Buy herds and sell herds 17
3.3. Value-weighted correlation and alternative industry definitions 18
3.4. Does stock herding drive industry herding? 19
3.5. The Lakonishok, Shleifer, and Vishny (1992) herding measure 22
4. Why do institutions industry herd? 24
4.1. Do underlying investors drive institutional industry herding? 24
4.2. Does industry momentum trading drive industry herding? 26
4.3. Herding and reputation 28

4.4. Industry herding and herding into size and book/market styles 30

4.5. Herding pre- and post-Electronic Data Gathering and Retrieval (EDGAR) service 34

4.6. Institutional industry demand and industry returns 36

5. Conclusions 40

REFERENCES 42

APPENDIX A: PROOFS 57

CHAPTER THREE: FINANCIAL STATEMENT ANALYSIS, FUTURE STOCK RETURNS AND DEMAND BY INSTITUTIONAL AND INDIVIDUAL INVESTORS 64

1. Introduction 64

2. Literature review 66

2.1. Under-reaction and use of financial statement analysis 66

2.2. Value-Growth Effect 68

3. Data 71

3.1. Institutional ownership data 71

3.2. Compustat/CRSP data 73

4. Replicating Piotroski (2000, 2005)'s results 78

4.1. Replicating Piotroski (2000) 78

4.1.1. Univariate analysis 79

4.1.2. Regression analysis 80

4.2. Replicating Piotroski (2005) 82

4.2.1. Univariate analysis 82

4.2.2. Regression analysis 84

5. Rational vs. irrational explanations for the explanatory power of the signal representing a firm's financial condition 86
5.1. Univariate analysis 87
5.2. Preliminary test 89
5.3. Regression analysis 90
6. Conclusion 93
REFERENCES 94

LIST OF TABLES

Tables for Chapter two

Table 1. Descriptive statistics 48

Table 2. Tests for herding 50

Table 3. Regression of weighted institutional industry demand on lag weighted institutional industry demand.. 52

Table 4. Tests for herding and momentum trading... 53

Table 5. Analysis by investor type.. 54

Table 6. Institutional industry herding into same size-BE/ME style stocks and different size-BE/ME style stocks.. 55

Table 7. Industry herding and subsequent returns 56

Tables for Chapter three

Table 1. Annual market adjusted returns to *f-score* portfolios (from 1976 to 1996)................. 100

Table 2. Annual returns to *f-score* portfolios for high book-to-market stocks (from 1976 to 1996)....... 101

Table 3. Regression of annual returns on other control variables and *f-scores* (1976-1996)..... 103

Table 4. Annual return to *f-score* portfolios (from 1972 to 2001).. 104

Table 5. Regressions of annual returns on f-scores and other control variables (1972-2001).... 107

Table 6. Annual returns and institutional ownership changes (1983-2005) 108

Table 7. Regression of institutional ownership change 110

LIST OF FIGURES

Figure 1 Cumulative returns and NID from *t*-12 to *t*+15 113

Figure 2 Cumulative returns and adjusted NID from *t*-12 to *t*+15 114

Dedication

I dedicate this dissertation to my dear family and especially to my late grandmother.

CHAPTER ONE: GENERAL INTRODUCTION

My dissertation consists of two essays related to the behavior of institutional investors and financial statement analysis. Institutional investors have long been known to be marginal investors who set prices and therefore a center of attention in asset pricing literature. In the first paper, we examine one of the popular behaviors of institutional investors: herding. Our goal in this paper is to test if institutional investors follow each other into and out of the same industry. This study contributes to two related literatures: institutional herding and style investing literature. Theoretical herding motives documented in the numerous literatures level should hold at industry level as much as or more than at stock level. Additionally, "style investing" literature argues a group of investors herd to a style and this behavior impacts returns. We find strong evidence of institutional investors herding behavior across industries and it is not a manifestation of stock herding. There are various reasons for why institutional investors follow each other and our results are most consistent with the correlated signals explanation.

The second paper investigates the relation between changes in financial health of the firm, its subsequent return, and demand by individual and institutional investors. There are two competing arguments about why financial statement analysis predicts future returns. Piotroski (2000, 2005) demonstrates a simple accounting based metric can successfully indentify the stocks with higher future returns from the stocks with low future profitability. Piotroski concludes this predictability comes from investors underreacting to information contained in financial statement analysis. On the other hand, Fama and French (2006) argue that financial statement analysis predicts future return because higher expected earnings firms should have higher expected returns. We attempt to disentangle two competing explanation for the return predictability of financial statement analysis. If the predictability comes from investors' underreaction, financial statement based metric will be correlated to the measures of investor demand. If financial statement analysis predicts the future returns because of risk based explanation, it will be independent of investor demand. We expect because institutional investors are more sophisticated than retail investors, institutional investors will be more likely than individual investors to exploit the information. We find strong relation between financial statement analysis and demand of institutional investors and our results support behavior-related explanation, rather than risk-based explanation for the predictability of financial statement analysis for the future returns.

CHAPTER TWO: INSTITUTIONAL INDUSTRY HERDING

"*The gains represent institutional herding, in which money managers chase each other into the hot performing areas regardless of the price they are paying…*" (Financial Times, July 5, 2004)

1. Introduction

The popular press often portrays institutional investors as driving prices from fundamental values and generating excess volatility as they herd to and from the latest 'fad.' Moreover, a rich theoretical literature suggests five additional reasons institutions may herd including underlying investors' flows, institutional positive feedback trading, attempting to preserve reputation by acting like other managers (reputational herding), inferring information from each others' trades (informational cascades), and following correlated signals (investigative herding). Although a growing empirical literature focuses on testing institutional herding in individual securities, the proposed reasons for institutional herding hold at least equally well at the industry level. If, for example, institutions are "piling in" to the technology industry, then an institution attempting to preserve their reputation may follow others into the technology industry. In addition, given institutional investors' dominant role in the market, institutional industry herding would likely impact industry valuations.[1]

The primary goal of this paper is to address this fundamental question: Do institutional investors herd across industries?[2] By moving beyond examining herding at the individual

[1] Institutional investors now dominate the ownership and trading of U.S. securities accounting for 63% of equity holdings in 2002 (NYSE factbook) and 70% to 96% of turnover (Schwartz and Shapiro, 1992; Jones and Lipson, 2003). See Chakravarty (2001), Boyer and Zheng (2004), Froot and Teo (2004), Sias, Starks, and Titman (2006), Kaniel, Saar, and Titman (2008), and Campbell, Ramadorai, and Schwartz (2007) for evidence that institutional investors are generally the price-setting marginal investors.
[2] Several previous studies (Lakonishok, Shleifer, and Vishny, 1992; Sharma, Easterwood, and Kumar, 2006) examine whether institutional investors herd at the individual stock level in some industries more than others, e.g., are institutions more likely to following each other from Microsoft to IBM than they are to follow each other from

security level, our study contributes to two related literatures. First, our results have direct implications for understanding why institutional investors herd and the potential price effects associated with such herding. Second, our study is closely related to the rapidly growing "style investing" literature. Barberis and Shleifer's (2003) groundbreaking model of style investing, for example, requires two key elements related to our study: (1) that a group of investors herd to and from styles, and (2) that these investors' herding impacts prices.[3] The growing empirical work on style investing (e.g., Teo and Woo, 2004; Barberis, Shleifer, and Wurgler, 2005; Froot and Teo, 2007) is also based on the proposition that a group of investors herd to a style and this behavior impacts returns.

Although most previous style investing studies focus on portfolios determined by market capitalization and book-to-market ratios, we focus on industry classifications because we believe institutions more often have signals regarding fundamental classifications such as industries than statistical classifications such as size and value/growth. Analysts, for example, are usually assigned on an industry basis. Institutional Investor's (the magazine) annual "All-America Research Team" analyst rankings, for instance, are by industry, e.g., Aerospace and Defense, Autos and Auto Parts, etc. Moreover, several studies suggest that industry information is impounded at different rates across securities within the same industry (e.g., Moskowitz and Grinblatt, 1999; Hou, 2007) and that investors may be able to infer information about a given firm based on information about other firms in the same industry (e.g., Lang and Lundholm, 1996). Last, many professional managers make industry/sector recommendations (e.g.,

---

Pacific Gas and Electric to Duke Energy? Our work, however, focuses on herding across industries, e.g., do institutional investors follow each other out of utilities and into technology stocks?

[3] In Barberis and Shleifer's (2003) model, an investment style (which, as the authors note, includes industry styles) experiences return momentum and reversals as a result of investors' style herding. The authors propose that institutions may be style investors (page 170), "…if we think of switchers as institutions chasing the best-performing style, then our model is consistent with evidence that demand shifts by institutions in particular influence security prices (Gompers and Metrick, 2001)."

overweight technology) just as they make individual security recommendations (e.g., overweight Microsoft). Although we find some anecdotal evidence of size or value/growth recommendations, such advice appears much less common.[4]

Our empirical results reveal strong evidence of institutional industry herding. The cross-sectional correlation between the fraction of institutional traders buying an industry this quarter and the fraction buying last quarter, for example, averages 40%. A number of robustness tests reveal that industry herding holds for alternative industry definitions and occurs both on the buy side (institutions following each other into the same industries) and the sell side (institutions following each other out of the same industries). Moreover, institutional investors' demand for a stock is a positive function of both their lag demand for that stock and their lag demand for other stocks in the same industry.

The balance of the paper focuses on understanding what drives institutional industry herding. Although these additional tests suggest institutional investors intentionally following each other into the same industries (as in informational cascades or reputational herding) likely plays some role in explaining the results, the aggregate evidence suggests that industry herding primarily arises from the manner in which information is incorporated into prices. Thus, the results are consistent with models (e.g., Froot, Scharfstein, and Stein, 1992; Hirshleifer, Subrahmanyam, and Titman, 1994) where informed investors receive signals at different times and, as a result, late informed investors follow early informed investors (i.e., herd) and information is incorporated into prices over time. Hirshleifer, Subrahmanyam, and Titman argue

[4] A search of marketwatch.com revealed sector/industry recommendations by Prudential, Lehman, Morgan Stanley, Credit Suisse, Wachovia, Goldman Sachs, Piper Jaffrey, Deutsche Bank, Bear Sterns, UBS, Bank of America, and Citi. Moreover, a Google search of "sector rotation" yielded over 200,000 hits. We find anecdotal evidence that managers occasionally make recommendations based on value/growth or size characteristics. A MarketWatch report (Turner, 2008), for example, notes "Portfolio strategists at Lehman Brothers on Monday said that they believe there is a tactical case for overweighting deep value companies."

this is reasonable because, "…in reality some investors, either fortuitously or owing to superior skill, acquire pertinent information before others." Similarly, Froot, Scharfstein, and Stein propose that even if investors attempt to acquire the same information, some will likely learn it before others.

We begin to examine what causes institutional industry herding by evaluating whether underlying investors' flows contribute to industry herding, e.g., retail investors moving funds from managers that focus on utility stocks and to managers that focus on healthcare stocks. We run two sets of tests to examine this explanation. First, following Dasgupta, Prat, and Verardo (2007), we exclude those institutional investors who are most subject to retail flows (mutual funds and independent advisors) from our analysis. Second, we examine changes in institutional investors' industry portfolio weights (that should not be impacted by underlying investors' flows) rather than changes in institutional investors' positions (that will be impacted by underlying investors' flows). Both tests suggest that institutional industry herding results from managers' decisions rather than underlying investors' flows.

Second, we investigate whether institutional investors' preference for industries with high lag returns might drive their herding as in the Barberis and Shleifer (2003) style investing model. Specifically, if institutional demand impacts returns and institutional investors industry momentum trade, then institutions chasing lag returns will also be chasing lag institutional industry demand. Although institutional investors tend to purchase (sell) industries that have done well (poorly) in the past, such momentum trading does not explain their herding: Institutional industry demand is largely independent of lag industry returns once controlling for lag institutional industry demand. Our results suggest institutions momentum trade industries because they herd and their lag demand is positively correlated with lag returns.

Third, we examine herding by investor type (banks, insurance companies, mutual funds, independent advisors, and unclassified investors) to test the reputational herding explanation. Following Sias (2004), we hypothesize that: (1) institutional investors concerned about their reputations are more likely to follow similarly classified institutions than differently classified institutions (e.g., mutual funds are more likely to follow other mutual funds than insurance companies), and (2) mutual funds and independent advisors will be more concerned about their reputations than other investors and therefore exhibit stronger herding propensities. We find mixed evidence for the reputational herding explanation. Four of the five investor groups are more likely to follow similarly classified institutions than differently classified institutions. We find little evidence, however, that mutual funds and independent advisors are more likely to herd than other institutional investors.

Fourth, we examine the relation between herding to similar size and book to market (henceforth, size-BE/ME) style stocks and industry herding to: (1) ensure that industry herding is unique from size-BE/ME style herding, (2) test whether industry signals may sometimes contain size-BE/ME components, and (3) help differentiate the correlated signals explanation from the informational cascades explanation. Specifically, we propose that size-BE/ME herding contributing to industry herding supports the correlated signals explanation over the informational cascades explanation because the informational cascades explanation would require that: (1) an investor infer both an industry signal and a size-BE/ME signal from previous investors' trades, and (2) be willing to ignore her own industry and/or size-BE/ME signals to follow the perceived industry signal *and* the perceived size-BE/ME signal of previous traders. Alternatively, the correlated signals explanation is consistent with size-BE/ME style herding contributing to industry herding if signals are sometimes related to size-BE/ME characteristics.

Institutions' correlated signals, for example, may suggest that although the banking industry is overvalued, small capitalization banks are more overvalued than large capitalization banks. Our results indicate that although industry herding is unique from size-BE/ME style herding, size-BE/ME style herding contributes to industry herding consistent with the correlated signals explanation (assuming industry signals sometimes contain an size-BE/ME component).

Fifth, we investigate whether institutional industry herding is stronger once institutions have easy electronic access to other institutions' positions. Specifically, institutions were required to file their position reports through the SEC's *Electronic Data Gathering and Retrieval* (EDGAR) system after 1996. If herding is primarily driven by institutions intentionally following each other into the same industries (as in informational cascades or reputational herding), then the level of herding should be much greater once institutions have easy access to much less noisy signals of other institutions' demand. Consistent with the hypothesis that reputational herding and/or informational cascades *contribute* to industry herding, we find that institutional herding increases slightly once institutions can easily view other institutions' lag trades. Nonetheless, consistent with the hypothesis that industry herding *primarily* arises from correlated signals, we find strong evidence of industry herding both prior to, and following, mandatory electronic filing and the increase in herding following mandatory electronic filing is relatively small.

Last, we investigate whether institutional industry herding drives prices from fundamentals as expected if: (1) herding does not fully result from the manner in which information is incorporated into prices (i.e., correlated signals) and (2) herding impacts prices. Our results reveal that institutional industry demand is strongly positively correlated with industry returns over the herding period, i.e., those industries institutions most heavily purchase over a given period average significantly higher returns over that period than those industries institutions sell. We only find weak evidence, however, that industries institutions herd to underperform those they herd out of in the year *following* the herding. The strong relation between institutional industry demand and same period industry returns and the weak relation between institutional industry demand and subsequent industry returns are consistent with the explanation that correlated signals primarily drive institutional industry herding.

In sum, the results suggest that whatever causes institutional investors to herd has an industry component and are consistent with the Barberis and Shleifer (2003) style investing model. Overall, the evidence is most consistent with the correlated signals explanation. Specifically, (1) the lack of strong evidence of industry return reversals following herding, (2) the small change in herding levels pre- and post-mandatory electronic filing, and (3) the relation between size-BE/ME herding and industry herding, all favor the correlated signals explanation over the alternatives.

The balance of the paper is organized as follows—we provide a brief review of related literature and discuss data in the next section. Section 3 presents our primary empirical tests while Section 4 focuses on the causes of institutional industry herding. The final section presents conclusions.

2. Background and data

2.1. Herding

Industry (stock) herding is defined as a group of investors following each other into and out of the same industry (stock) over *some period*.[5] Previous work proposes six reasons institutional investors may herd—underlying investors' flows, fads, momentum trading, reputational herding, informational cascades, and investigative herding. First, institutional investors may herd to industries because underlying investors shift toward those industries (see Frazzini and Lamont, 2008). For example, if retail investors' flows shift to technology funds both this quarter and last quarter (for whatever reason), then, as a group, mutual funds will herd to technology stocks.

The fads argument proposes that institutional investors may herd to industries simply because those industries become more popular. Friedman (1984), for example, notes the close-knit nature of the professional investment community, the importance of relative performance, and the asymmetry of incentives (i.e., the cost of poor relative performance is greater than the reward for superior performance), all suggest that institutional investors will herd to and from the latest fad.

Institutional investors' momentum trading could drive their herding. In the framework of the Barberis and Shleifer (2003) model, for example, style investors follow other style investors into and out of the same industries as they chase returns that are driven by the trades of previous style investors. If, for instance, institutions strongly buy the technology industry this quarter (for whatever reason) and their demand drives up the value of the technology industry this quarter, then other institutions chasing returns next quarter will follow these institutions into the technology industry.

[5] As noted by Sias (2004), herding is sometimes loosely defined as investors buying or selling the same industry (or security) at the 'same' time. Because trades occur sequentially, however, investors cannot buy or sell the same stock at the same time–hence, stock herding has a temporal component. Although it is possible for a group of investors to buy (or sell) the same industry at the same time (e.g., one institution buys Yahoo while another buys Google at the same time), we focus on industry herding over time.

Institutional investors may herd because they face a reputational cost from acting different from the herd, i.e., it is more costly to be alone and wrong than to be with the herd and wrong (see Scharfstein and Stein, 1990; Trueman, 1994; Zwiebel, 1995; Dasgupta, Prat, and Verardo, 2007). Value managers who did not purchase technology stocks in the late 1990s, for example, suffered large investor withdrawals (see Shell, 2001).

Informational cascades occur when investors ignore their own noisy signals and attempt to infer information from previous investors' trades (see Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992). Thus, these models require that investors receive valuation signals and trade sequentially.[6] At the firm level, these signals may occur sequentially and contain private information regarding future firm performance. Given many professional managers make industry/sector recommendations, they must also believe they have information (i.e., signals) regarding industry/sector valuation not yet reflected in prices. Moreover, because sector upgrades and downgrades do not occur simultaneously, managers must either receive or act on industry signals sequentially. Thus, for example, a manager who's industry signal indicates energy stocks are overvalued may nonetheless ignore the signal and increase his/her energy sector exposure if managers trading earlier increased their exposure to the energy sector.

Investigative herding results from investors following correlated signals at different times and, therefore, may reflect the process by which information is impounded into prices (see Froot, Scharfstein, and Stein, 1992; Hirshleifer, Subrahmanyam, and Titman, 1994). If, for example, an investor receives a private signal at time *t* that Google is undervalued and another investor

[6] Agents receive private signals sequentially in the classical informational cascade models, e.g., Bikhchandani, Hirshleifer, and Welch (1992). Later work demonstrates this assumption can be relaxed as long as agents act on signals in sequence. In the Chamley and Gale (1994) model, for example, agents may wait to act on information because they learn from watching the decisions of previous traders. In the Gul and Lundholm (1995) and Zhang (1997) models, agents act sequentially because their signal quality differs.

receives a private signal at time t+1 that Yahoo is undervalued, then investors will follow each other into technology stocks.

2.2. Empirical tests of institutional stock herding

Most early studies of institutional stock herding focus on the Lakonishok, Shleifer, and Vishny (1992) "herding measure" (see Section 3.5 for details). In general, these studies find statistically significant, but relatively weak, evidence of institutional investors herding in the average stock (e.g., Lakonishok, Shleifer, and Vishny, 1992; Grinblatt, Titman, and Wermers, 1995; Wermers, 1999; Wylie, 2005). A number of recent papers (Sias, 2004; Foster, Gallagher, and Looi, 2005; Dasgupta, Prat, and Verardo, 2007; Puckett and Yan, 2008), however, find strong evidence of institutional stock herding by directly examining whether cross-sectional variation in institutional demand for securities this quarter is related to cross-sectional variation in institutional demand for securities in the previous quarter(s).

2.3. Data

Data for this study come from three sources. We use Compustat data to compute book values and the Center for Research in Security Prices (CRSP) for return, market capitalization, and industry classification (SIC codes). Each institutional investor's holdings of each stock come from their quarterly 13(f) reports.[7] Our institutional ownership data span the first quarter of 1983 through the last quarter of 2005 for a total of 92 quarters. We include all ordinary (CRSP share code of 10 or 11) securities with adequate data.

---

[7] The data were purchased from Thomson Financial. All institutions with at least $100 million under management are required to report equity positions (greater than 10,000 shares or $200,000) to the SEC each quarter. Managers with stale reports (i.e., report date unequal to quarter-end date) are excluded for the quarter. The data are also cleaned of obvious reporting errors (e.g., lags in adjustment for stock splits).

We begin by assigning each security (each quarter) to one of the 49 Fama and French (1997) industries (using updated definitions posted on Ken French's website). To ensure our results are not influenced by a change in a stock's SIC code, we do not allow stocks to change industry classifications over the herding or return evaluation period. If ABC, for example, is classified in industry 1 at the beginning of quarter t-1, but industry 2 at the beginning of quarter t, then the company is classified as in industry 1 when evaluating herding between quarters t-1 and t, but industry 2 when evaluating herding between quarters t and t+1.

We define institution n as purchasing industry k if the dollar value of the institution's position in the industry increased over the quarter. As pointed out by Grinblatt, Titman, and Wermers (1995), however, the dollar value of a manager's position will increase (decrease) if the industry had a positive (negative) return even if the investor does not trade. To eliminate such "passive momentum," we use the product of beginning of quarter prices and end of quarter shares held to compute the "dollar value" of end of quarter holdings for manager n.[8] Specifically, manager n is classified as a buyer in industry k if:

$$\sum_{i=1}^{I_{k,t}} P_{i,t-1}\left(Shares_{n,i,t} - Shares_{n,i,t-1}\right) > 0\,, \tag{1}$$

where $I_{k,t}$ is the number of securities in industry k in quarter t, $P_{i,t\text{-}1}$ is the price of security i ($i \in k$) at the beginning of quarter t, and $Shares_{n,i,t\text{-}1}$ and $Shares_{n,i,t}$ are the number of (split-adjusted) shares of security i held by manager n at the beginning and end of quarter t, respectively. Analogously, manager n is classified as an industry k seller if Eq. (1) is negative. We define institutional industry demand (henceforth "institutional demand") as the ratio of the number of

[8] Previous work (e.g., Badrinath and Wahal, 2002; Wermers, Yao, and Zhao, 2007) uses the product of end of quarter prices and beginning of quarter shares held to compute the "dollar value" of beginning of quarter holdings for manager *n*. We find qualitatively equivalent results using this approach. We report results based on beginning of quarter prices because there may be correlation between end of quarter prices and institutional demand.

institutional investors buying industry k in quarter t to the number of institutions trading industry k in quarter t:

$$\Delta_{k,t} = \frac{\#\,Institutional\ buyers\ of\ industry\ k\ in\ quarter\ t}{\#\,Institutional\ buyers\ of\ industry\ k\ in\ quarter\ t + \#\,Institutional\ sellers\ of\ industry\ k\ in\ quarter\ t}. \quad (2)$$

Panels A and B in Table 1 report the time-series mean of cross-sectional quarterly descriptive statistics. Panel A reports the average industry has 692 institutional traders each quarter ranging from a minimum of 150 to a maximum of 1,076. Institutional demand averages near 50% reflecting that, on average, institutional investors are as likely to be buyers as sellers. There is, however, substantial cross-sectional variation in institutional demand—on average, institutional buyers account for over 60% of institutional traders in the highest institutional demand industry and less than 40% of institutional traders in the lowest institutional demand industry. Panel B reports that, on average, industries contain 116 stocks, ranging from a minimum of six securities to a maximum of 609 securities. The largest industry, on average, accounts for 11.35% of the market portfolio. Industries also have high levels of concentration. On average, the single largest firm accounts for 32% of the industry's capitalization. Panel C reports time-series descriptive statistics for each of the 49 industries.

[Insert Table 1 about here]

3. Tests for institutional industry herding

3.1. Correlation between contemporaneous and lag institutional industry demand

Following Sias (2004), we test for institutional herding by computing the cross-sectional correlation between institutional investors' industry demand this quarter and last quarter. The intuition is straightforward—if institutional investors industry herd, then cross-sectional variation in institutional demand last quarter will predict cross-sectional variation in institutional demand this quarter. A given institutional investor following their own lag industry trading, however, will also induce positive correlation between institutional demand this quarter and last quarter. Positive correlation may arise, for example, if: (1) Fidelity Investments purchased the healthcare industry both this quarter and last, or (2) Fidelity Investments purchased the healthcare industry this quarter and other institutions purchased it last quarter. Sias (2004) demonstrates that the correlation between institutional demand this quarter and last can be directly partitioned into these two components. Specifically, the correlation can be written as the sum of the products of demeaned dummy variables (denoted $D_{n,k,t}$) that equal one if institution n buys industry k in quarter t and zero if institution n sells industry k (see Appendix A for proof):

$$\rho\left(\Delta_{k,t},\Delta_{k,t-1}\right)=\left[\frac{1}{(K)\sigma\left(\Delta_{k,t}\right)\sigma\left(\Delta_{k,t-1}\right)}\right]\sum_{k=1}^{K}\left[\sum_{n=1}^{N_{k,t}}\left(\frac{D_{n,k,t}-\overline{\Delta_{k,t}}}{N_{k,t}}\bullet\frac{D_{n,k,t-1}-\overline{\Delta_{k,t-1}}}{N_{k,t-1}}\right)\right]+$$

$$\left[\frac{1}{(K)\sigma\left(\Delta_{k,t}\right)\sigma\left(\Delta_{k,t-1}\right)}\right]\sum_{k=1}^{K}\left[\sum_{n=1}^{N_{k,t}}\sum_{m=1,m\neq n}^{N_{k,t-1}}\left(\frac{D_{n,k,t}-\overline{\Delta_{k,t}}}{N_{k,t}}\bullet\frac{D_{m,k,t-1}-\overline{\Delta_{k,t-1}}}{N_{k,t-1}}\right)\right], \tag{3}$$

where K is the number of industries (49 in our primary tests), $N_{k,t}$ is the number of institutions trading industry k in quarter t, and $\sigma(\Delta_{k,t})$ and $\overline{\Delta_{k,t}}$ are the cross-sectional standard deviation and average institutional demand in quarter t, respectively. The first term on the right-hand side of Eq. (3) is the portion of the correlation attributed to individual institutional investors following their own lag demand (i.e., investor n following her own lag demand for industry k) and the second term is the portion attributed to institutions following the lag demand of other institutional investors (i.e., investor n following investor m's lag demand for industry k).

Panel A in Table 2 reports the time-series average of the 90 cross-sectional correlation coefficients between institutional demand this quarter and last quarter [and associated t-statistics based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates; henceforth, Newey-West *t*-statistics]. Institutional investors' demand for an industry this quarter is strongly related to their demand last quarter—the cross-sectional correlation averages 40% and is statistically significant at the 1% level. The next two columns report the time-series averages of the portion of the correlation (and associated Newey-West *t*-statistics) due to institutional investors following their own lag industry demand [i.e., the first term in Eq. (3)] and the portion due to institutions following the lag demand of other institutional investors [i.e., the second term in Eq. (3)]. Both components are statistically significant at the 1% level. The evidence that institutional investors follow their own lag demand is consistent with the hypothesis that institutional investors spread their trades out over time to minimize the price impact of their trading consistent with Barclay and Warner (1993), Chakravarty (2001), and Sias (2004). The results also reveal that 92% of the average correlation (0.3743/0.4049) arises from institutional investors following other institutional investors into and out of the same industries, i.e., industry herding.

[Insert Table 2 about here]

To help gauge the economic significance of the results, we more closely examine the herding in those industries that contribute the most to the correlation. We begin by computing each industry's contribution to the cross-sectional correlation between institutional demand this quarter and last quarter, i.e., each industry's contribution to Eq. (3):

$$\textit{Industry } k\textit{'s contribution}_t = \left[\frac{1}{(K)\sigma(\Delta_{k,t})\sigma(\Delta_{k,t-1})}\right](\Delta_{k,t} - \overline{\Delta_{k,t}})(\Delta_{k,t-1} - \overline{\Delta_{k,t-1}}). \tag{4}$$

We then denote (each quarter) the 10 industries where the last two terms are both positive (i.e., institutions bought the industry more than average both this quarter and last) that contribute the most to the industry herding measure [i.e., with the largest Eq. (4)] as buy-herding industries. Analogously, we denote the top 10 industries where the last two terms are both negative as sell-herding industries. The top 10 buy-herding industries average 608 institutional traders in quarter t-1 of which 330 are buyers (54.32%) and 278 are sellers (45.68%). The following quarter (t), these buy-herding industries average 624 traders of which 338 are buyers (54.17%) and 285 are sellers (45.67%). Similarly, the top 10 sell-herding industries average 704 traders in quarter t-1 of which 379 are sellers (53.85%) and 325 are buyers (46.15%). In quarter t, the sell-herding industries average 710 traders of which 384 are sellers (54.18%) and 325 are buyers (45.82%).

3.2. Buy herds and sell herds

A number of previous studies (e.g., Grinblatt, Titman, and Wermers, 1995; Wermers, 1999; Wylie, 2005) of stock herding examine buy herding (institutions following each other into the same stock) versus sell herdings (institutions following each other out of the same stock). As pointed out by Brown, Wei, and Wermers (2007), for example, it is possible that institutional sell herding may be more limited than buy herding because many institutional investors cannot sell securities short.

To examine whether institutional investors are more likely to buy herd or sell herd industries, we partition Eq. (3) into those industries institutions bought in quarter t-1 ($\Delta_{k,t-1} > 0.5$) and those industries institutions sold in quarter t-1 ($\Delta_{k,t-1} < 0.5$) to compute the portion of the correlation arising from institutions following each other into the same industries (first row in Panel B of Table 2) and institutions following each other out of the same industries (second row in Panel B). The third row in Panel B reports the difference and associated Newey-West t-

statistics. The results reveal no evidence that industry buy herding differs meaningfully from industry sell herding.

3.3. Value-weighted correlation and alternative industry definitions

Table 1 reveals that the smallest industry accounts for, on average, 0.05% of the total market capitalization. Each industry, however, contributes equally in the calculation of the correlation between institutional demand this quarter and last. To ensure the correlations are not driven by the very smallest of industries, we compute and decompose the industry-weighted correlation, where each industry's weight is equal to their fraction of market capitalization at the beginning of quarter *t*-1 (see Appendix A for additional detail). Panel C in Table 2 reports the time-series average of the 90 cross-sectional industry-weighted correlation coefficients and associated Newey-West *t*-statistics. The results are nearly identical to the equal-weighted correlations—institutional industry demand is strongly correlated with lag institutional demand and is primarily driven by institutions following other institutions into and out of the same industries.

Although the 49 Fama and French (1997) industries are often used in academic studies, they serve as only one of a number of possible industry definitions. To examine the sensitivity of our results to finer industry definitions, we repeat the analysis in Panel A but define industries based on two digit SIC codes (on average, this results in 73 industries each quarter). Results, reported in the first row of Panel D, reveal strong, albeit slightly weaker correlation (averaging 24.65%) that is primarily driven by institutions following other institutions into the same industry. We next try coarser industry definitions—repeating the analysis with the additional industry definitions available on Ken French's website that classify firms into 5, 10, 12, 17, 30, and 38 industries. The results, presented in the bottom six rows of Panel D, are consistent with base case—strong evidence of institutional industry herding primarily driven by institutions following other institutions into the same industry.

3.4. Does stock herding drive industry herding?

Table 1 reveals that many industries are highly concentrated, e.g., the largest single stock in an industry accounts for, on average, 32% of the total industry capitalization. It is possible, therefore, that industry herding is simply a manifestation of stock herding. If institutional investors are herding to Microsoft and Microsoft accounts for nearly half the technology industry, then institutional investors are likely herding to the technology industry (as long as institutions' Microsoft purchases are not fully offset by sales of other technology stocks).

To examine whether industry herding is a manifestation of stock herding, we define an alternative measure of institutional industry demand as the capitalization-weighted average institutional demand for securities in each industry. We begin by defining the institutional demand for each stock *i* (in quarter *t*) as the number of institutions buying (i.e., increasing the split-adjusted number of shares they hold) the stock as a fraction of the number of institutions trading the stock:

$$\Delta_{i,t} = \frac{\#\,Institutional\ buyers\ of\ stock\ i\ in\ quarter\ t}{\#\,Institutional\ buyers\ of\ stock\ i\ in\ quarter\ t + \#\,Institutional\ sellers\ of\ stock\ i\ in\ quarter\ t}. \tag{5}$$

We then define the weighted institutional demand for industry *k* (henceforth, "weighted institutional demand" and denoted $\Delta^{*}_{k,t}$) as the market capitalization weighted average institutional demand across stocks in industry k (where $w_{i,t}$ is security i's capitalization weight in industry k at the beginning of quarter t):[9]

$$\Delta^*_{k,t} = \sum_{i=1}^{I_{k,t}} w_{i,t} \Delta_{i,t} \,. \tag{6}$$

Because the weighted institutional industry demand is a linear function of institutional demand for each security in that industry, we can directly decompose the cross-sectional correlation between weighted institutional demand this quarter and last quarter into four components: the portions that arise from following each other or themselves into the same stock and the portions that arise from following each other or themselves into different stocks in the same industry (see Appendix A for proof):

$$\rho\left(\Delta^*_{k,t}, \Delta^*_{k,t-1}\right) = \frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})} \sum_{k=1}^{K} \left(\sum_{i=1}^{I_{k,t}} \left(w_{i,t} w_{i,t-1} \left(\sum_{n=1}^{N_{i,t}} \left(\frac{D_{n,i,t} - \overline{\Delta^*_{k,t}}}{N_{i,t}} \bullet \frac{D_{n,i,t-1} - \overline{\Delta^*_{k,t-1}}}{N_{i,t-1}} \right) \right) \right) \right) +$$

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})} \sum_{k=1}^{K} \left(\sum_{i=1}^{I_{k,t}} \left(w_{i,t} w_{i,t-1} \left(\sum_{n=1}^{N_{i,t}} \sum_{m=1, m \neq n}^{N_{i,t-1}} \frac{D_{n,i,t} - \overline{\Delta^*_{k,t}}}{N_{i,t}} \bullet \frac{D_{m,i,t-1} - \overline{\Delta^*_{k,t-1}}}{N_{i,t-1}} \right) \right) \right) +$$

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})} \sum_{k=1}^{K} \left(\sum_{i=1}^{I_{k,t}} \sum_{j=1, j \neq i}^{I_{k,t-1}} \left(w_{i,t} w_{j,t-1} \left(\sum_{n=1}^{N_{i,t}} \frac{D_{n,i,t} - \overline{\Delta^*_{k,t}}}{N_{i,t}} \bullet \frac{D_{n,j,t-1} - \overline{\Delta^*_{k,t-1}}}{N_{j,t-1}} \right) \right) \right) +$$

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})} \sum_{k=1}^{K} \left(\sum_{i=1}^{I_{k,t}} \sum_{j=1, j \neq i}^{I_{k,t-1}} \left(w_{i,t} w_{j,t-1} \left(\sum_{n=1}^{N_{i,t}} \sum_{m=1, m \neq n}^{N_{j,t-1}} \frac{D_{n,i,t} - \overline{\Delta^*_{k,t}}}{N_{i,t}} \bullet \frac{D_{m,j,t-1} - \overline{\Delta^*_{k,t-1}}}{N_{j,t-1}} \right) \right) \right), \tag{7}$$

[9] We verify that this alternative measure of institutional industry demand is closely related to the number of institutions increasing their position in the industry divided by the number trading the industry [i.e., Eq. (2)]. Specifically, the cross-sectional correlation across the 49 industries between the measures given in Eq. (2) and Eq. (6) averages 81%.

where $N_{i,t}$ is the number of institutions trading security i in quarter t and $D_{n,i,t}$ is a dummy variable that equals one if institutional investor n increases her position in security i in quarter t and zero if the investor decreases her position in security i.

The first term on the right hand side of Eq. (7) is the portion of the correlation that arises from institutional investors following their own trades in the same stock (i.e., institution n following their own lag trades in security i) and the second term is the portion that arises from institutional investors following other institutions into the same stock (i.e., institution n following institution m's lag trades in security i). The third term is the portion of the correlation that arises from institutions following themselves into different stocks in the same industry (i.e., institution n's trades in security i following their lag trades in security j where both i and j are in industry k), while the last term is the portion that arises from institutions following other institutions into different stocks in the same industry (i.e., institution n's trades in security i following institution m's lag trades in security j where both i and j are in industry k).

As shown in the bottom right-hand cell in Table 3, the cross-sectional correlation between weighted institutional demand this quarter and last averages 57% (statistically significant at the 1% level). The four interior cells of Table 3 report the time-series average of each of the four components given in Eq. (7) and associated Newey-West t-statistics. The results reveal that all four components are statistically significant at the 1% level. The results in the top row are consistent with the hypothesis that institutional investors spread their trading out over time in both an individual security and in an industry to minimize the price impact of their trading. The results also reveal, consistent with the explanation that the combination of stock herding and high industry concentration contributes to industry herding, institutional investors following other institutional investors into the same stock accounts for the largest single component of the quarterly correlation (0.3235/0.5716). This result is consistent with recent evidence that institutional investors herd into and out of individual securities (Sias, 2004; Foster, Gallagher, and Looi, 2005; Dasgupta, Prat, and Verardo, 2007; Puckett and Yan, 2008).

[Insert Table 3 about here]

The figure shown in the center cell, accounting for 34% of the overall correlation (0.1942/0.5716) and statistically significant at the 1% level (*t*-statistic=11.10), however, is the key result reported in Table 3. Specifically, an institutional investor's demand for a stock this quarter is related not only to other institutions' demand for that stock last quarter, but also to other institutional investors' demand for different stocks in the same industry last quarter. In sum, although institutional investors herding into individual stocks contributes to institutional industry herding, industry herding is unique from stock herding.[10]

3.5. The Lakonishok, Shleifer, and Vishny (1992) herding measure

Most early investigations of institutional herding focus on the Lakonishok, Shleifer, and Vishny (1992) herding measure:

$$H_{k,t} = \left|\Delta_{k,t} - \overline{\Delta_{k,t}}\right| - AF_{k,t}\,, \tag{8}$$

where, as in Eq. (2), $\Delta_{k,t}$ is the ratio of the number of institutions buying industry k to the number trading industry k in quarter t (and $\overline{\Delta_{k,t}}$ is its cross-sectional average). The adjustment factor ($AF_{k,t}$) accounts for the fact that the expected value of the first term is positive regardless of institutional herding and is computed by assuming the number of institutional traders in industry

[10] As a robustness test, we also compute an industry-weighted, weighted institutional demand [i.e., Eq. (6)] correlation (analogous to Panel C in Table 2) and correlations based on the alternative industry definitions (analogous to Panel D in Table 2). Although specific results are not reported (to conserve space), with the exception of the extremely broad 5-industry classification, these alternative approaches yield qualitatively identical results.

k during quarter t follows a binomial distribution with the probability of buying set equal to $\overline{\Delta_{k,t}}$. This metric tests for herding by recognizing that if institutional investors follow each others' demand then institutional investors will primarily be buyers of industries they herd to and primarily be sellers of industries they herd from within that quarter.[11]

For our sample, the Lakonishok, Shleifer, and Vishny (1992) herding measure averages 1.39% across the 4,459 industry-quarter observations (91 quarters * 49 industries) and differs significantly from zero at the 1% level (t-statistic=34.66). Given the average institutional demand (i.e., $\Delta_{k,t}$) is approximately 50% (see Table 1), the average herding measure of 1.39% can be interpreted as meaning that if there were 100 institutional traders in a random industry-quarter, we would expect 51.39 on one side of the market (buyers or seller) and 48.61 on the other. Thus, consistent with previous work (e.g., Wermers, 1999; Sias, 2004), the measure reveals highly significant, albeit not particularly large, levels of institutional herding in the average industry-quarter.

The key to reconciling the 'strength' of the results between the Lakonishok, Shleifer, and Vishny (1992) and Sias (2004) herding tests is that the correlation focuses on whether those industries that had the greatest institutional demand (or supply) last quarter have the greatest demand (or supply) this quarter. In contrast, the Lakonishok, Shleifer, and Vishny measure evaluates the average herding across *every* industry *every* quarter. Thus, the correlation tests will reveal strong evidence of herding if institutions are strongly herding into three industries and strongly herding out of three other industries, but have net demand near zero for the remaining 43 industries. The Lakonishok, Shleifer, and Vishny measure will also capture such herding,

[11] Both the Lakonishok, Shleifer, and Vishny (1992) and Sias (2004) herding tests measure herding over time, i.e., whether institutions follow other institutions. The Lakonishok, Shleifer, and Vishny metric, however, indirectly captures the temporal nature of the herding by testing whether institutional investors *follow* other institutional investors within the same quarter.

although the *average* across all 49 industries will be relatively small.[12] In short, the results of the Lakonishok, Shleifer, and Vishny tests are fully consistent with our previous tests.

4. Why do institutions industry herd?

We next attempt to differentiate between the six proposed herding motives: underlying investors' flows, momentum trading, reputational herding, informational cascades, investigative herding, and fads.

4.1. Do underlying investors drive institutional industry herding?

Institutional industry herding could simply reflect underlying investors' flows. Frazzini and Lamont (2008) note, for example, that in 1999 retail investors added $37 billion to technology-oriented Janus Funds while adding only $16 billion to more conservative, and much larger, Fidelity funds. And by 2001, retail investors moved strongly out of Janus and into Fidelity. We take two approaches to testing whether underlying investors' flows can explain institutional industry herding. First, we repeat our empirical tests excluding those institutional investors most subject to retail flows. Specifically, Thomson Financial classifies institutions into five groups: banks, insurance companies, mutual funds (investment companies), independent investment advisors, and unclassified institutions.[13] Dasgupta, Prat, and Verardo (2007) argue

[12] Consider an extreme example: Assume that institutional investors are herding to three industries such that 70% of institutional traders are buyers both this quarter and last, and institutional investors are herding out of three industries such that 70% of institutional traders are sellers this quarter and last. In the remaining 43 industries, institutional traders are exactly 50% buyers and 50% sellers. Further assume the sample sizes are large enough that the adjustment factors in the Lakonishok, Shleifer, and Vishny (1992) measure are approximately zero. In such a case, the average Lakonishok, Shleifer, and Vishny metric is 0.024 (measure over either quarter, or both quarters together) while the cross-sectional correlation is one, i.e., the cross-sectional variation in last quarter's institutional demand perfectly explains the cross-sectional variation in this quarter's institutional demand.

[13] The classifications are inexact in that institutions file 13(f) reports in the aggregate and some institutions would qualify as more than one type. For example, mutual funds that also act as independent investment advisors are classified as mutual funds if more than 50% of their assets are in mutual funds and as independent investment

that mutual funds and independent investment advisors are most likely to be subject to the vagaries of retail investors. Thus, if institutional industry herding is primarily driven by underlying investor flows, our results should be substantially weaker when excluding mutual funds and independent investment advisors.

Panel E in Table 2 reports the industry herding analysis [i.e., Eq. (3)] when excluding mutual funds and independent advisors. The results reveal no evidence that institutional industry herding is driven by retail investors' flows. In fact, the point estimates are slightly larger when excluding mutual funds and independent advisors from the analysis (Panel E) than when including them (Panel A).

As a second test of whether underlying investors' flows explain institutional industry herding, we focus on changes in institutions' industry portfolio weights rather than industry positions (following Sias, 2004). The intuition is straightforward—although underlying investors' flows would impact whether a manager buys an industry, it should not impact the managers' industry portfolio weight.[14] Thus, we redefine whether an institution buys or sells an industry each quarter by examining changes in institutions' industry portfolio weights. Specifically, manager *n* is classified as a buyer of industry *k* if their end of quarter portfolio industry weight is greater than their beginning of quarter industry portfolio weight:

$$\frac{\sum_{i=1}^{N_{k,t}} P_{i,t-1} Shares_{n,i,t}}{\sum_{k=1}^{K} \sum_{i=1}^{N_{k,t}} P_{i,t-1} Shares_{n,i,t}} - \frac{\sum_{i=1}^{N_{k,t-1}} P_{i,t-1} Shares_{n,i,t-1}}{\sum_{k=1}^{K} \sum_{i=1}^{N_{k,t-1}} P_{i,t-1} Shares_{n,i,t-1}} > 0. \tag{9}$$

advisers otherwise. Thomson Financial began a different classification scheme at the end of 1998. Classifications from December 1998-2005 were based on additional classification data provided by Thomson Financial (details available on request).

[14] It is possible, however, that some large managers have different investment vehicles and therefore the manager may be affected by correlated flows, e.g., money flowing out of Fidelity's utility fund and into Fidelity's healthcare fund.

As before, we use beginning-of-quarter share prices at both the beginning and end of the quarter to ensure we capture changes in portfolio weights driven by trading rather than differences in industry returns. We then compute institutional investors' demand for industry *k* as the number of institutions increasing their industry *k* portfolio weight divided by the number of institutions changing their industry *k* portfolio weight [analogous to Eq. (2)].

Panel F of Table 2 reports the time-series average correlation between institutional demand (based on changes in portfolio weights) this quarter and last as well as the portion that arises from institutions following their own lag changes in industry portfolio weights and the portion that arises from following other institutions' lag changes in industry portfolio weights. The results, nearly identical to the previous analysis (reported in Panel A), reveal no evidence that underlying investors' flows drive institutional investors' industry herding.

4.2. Does industry momentum trading drive industry herding?

Institutions may herd because institutional demand last quarter is positively correlated with last quarter's industry returns and institutions, as a group, are attracted to industries with high lag returns and repelled from industries with low lag returns as in Barberis and Shleifer's (2003) style investing model. To investigate this possibility, we first test whether institutional investors momentum trade industries by estimating quarterly cross-sectional regressions of institutional industry demand [i.e., Eq. (2)] on industry returns over the previous quarter, six months, or year [following Fama and French (1997) industry returns are value-weighted]. For comparison, we also estimate quarterly cross-sectional regressions of institutional demand on lag institutional demand over the previous quarter, six months, or year. To directly compare coefficients in subsequent tests, we standardize (i.e., rescale to zero mean and unit variance, each quarter) both institutional industry demand and industry returns.

The first column of Table 4 reports that the cross-sectional correlation between institutional demand and lag quarterly institutional demand averages 40% consistent with Table 2.[15] [As before, all *t*-statistics are based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates.] The fourth and seventh columns reveal that institutional demand is also positively correlated with institutional demand measured over the previous six months or year. For the lag six month and lag annual industry demand, we redefine buyers and sellers based on changes in their holdings over the previous six months or year [analogous to Eq. (1)], respectively.[16] The second, fifth, and eighth columns in Table 4 also reveal, however, that institutional demand is positively correlated with industry returns over the previous quarter, six months, and year, respectively (all statistically significant at the 5% level or better). Thus, the results reveal that institutional investors momentum trade at the industry level consistent with the Barberis and Shleifer (2003) style investing model and evidence at the individual security level [see Sias (2007)].

[Insert Table 4 about here]

To test whether institutional industry momentum trading explains their industry herding, we include both lag institutional demand and lag industry returns in the quarterly regressions (the tildes indicate the variables are standardized):

$$\widetilde{\Delta}_{k,t} = \beta_{1,t}\widetilde{\Delta}_{k,t-1} + \beta_{2,t}\widetilde{R}_{k,t-1} + \varepsilon_{k,t}\,. \tag{10}$$

---

15 Because both variables are standardized and there is only one independent variable, the average coefficient is the average correlation.

16 For example, if an institutional investor made a large increase in their utilities holdings two quarters ago and a small decrease last quarter, the investor would be classified as a seller last quarter but a buyer over the lag six month period.

The average coefficients for the 90 cross-sectional regressions are reported in the third (lag quarter), sixth (lag six months), and last (lag year) columns of Table 4. Institutional momentum trading does not explain institutional industry herding, i.e., institutional demand remains positively related to lag institutional demand even after accounting for lag industry returns. In fact, the evidence suggests that institutional investors' industry momentum trading results from their herding—there is no evidence that institutional demand is related to lag industry returns once accounting for lag institutional demand.

4.3. Herding and reputation

Sias (2004) hypothesizes that if professional investors' reputational concerns drive their herding, then institutional investors should be more likely to follow similarly classified institutions than differently classified institutions. Sias also proposes, consistent with Dasgupta, Prat, and Verardo (2007), that mutual funds and independent advisors are most likely to experience investor flows as a result of changes in their reputation. Thus, if reputational concerns drive herding, then mutual funds and independent advisors should exhibit a greater herding propensity than other investor types.

Sias (2004) points out that analysis by investor type is complicated by the fact that the number of each type of institutional investor differs. As a result, a given investor type may contribute more to the herding measure [i.e., the second term in Eq. (3)] because there are many of those investors rather than because that investor type exhibits a greater herding propensity. Thus, we follow Sias and measure each investor types' propensity to engage in herding as their average (rather than total) contribution from following similarly classified institutions and their average contribution from following differently classified institutions. For a given quarter, the average same-type herding contribution for banks is given by the last portion of the second term in Eq. (3) limited to banks averaged over the 49 industries:

$$Average\ same-type\ herding\ contribution_t^{Banks} = \frac{1}{49}\sum_{k=1}^{49}\left[\sum_{b=1}^{B_{k,t}}\sum_{m=1,m\neq b,m\in B}^{B^*_{k,t-1}}\frac{\left(D_{b,k,t}-\overline{\Delta_{k,t}}\right)\left(D_{m,k,t-1}-\overline{\Delta_{k,t-1}}\right)}{B_{k,t}B^*_{k,t-1}}\right], \tag{11}$$

where $B_{k,t}$ is the number of banks trading industry k in quarter t and $B^*_{k,t-1}$ is the number of different banks trading industry k in quarter t-1. Similarly, the average different-type herding contribution for banks is given by the last portion of the second term in Eq. (3) limited to banks trading in quarter t and non-banks trading in quarter t-1 (averaged over the 49 industries):

$$Average\ differnt-type\ herding\ contribution_t^{Banks} = \frac{1}{49}\sum_{k=1}^{49}\left[\sum_{b=1}^{B_{k,t}}\sum_{m=1,m\notin B}^{N_{k,t-1}-B_{k,t-1}}\frac{\left(D_{b,k,t}-\overline{\Delta_{k,t}}\right)\left(D_{m,k,t-1}-\overline{\Delta_{k,t-1}}\right)}{B_{k,t}\left(N_{k,t-1}-B_{k,t-1}\right)}\right], \tag{12}$$

where $N_{k,t-1}$-$B_{k,t-1}$ is the number of non-banks trading industry k in quarter t-1. We compute analogous statistics for each of the other investor types. For completeness, we also compute the average contribution from following their own previous trades [i.e., the last portion of the first term in Eq. (3) limited to each investor type] and the average contribution from following other investors' trades regardless of trader type.

Table 5 reports the time-series average of the 90 estimates by investor type and associated Newey-West t-statistics. The first and second columns in Table 5 report the average contribution from following their own industry trades and the average contribution from following other investors' (regardless of classification) industry trades, respectively. The results reveal strong evidence of following their own trades and following other investors' trades for each investor type (statistically significant at the 1% level in all cases). The third and fourth columns report the average contribution from following similarly classified traders [i.e., Eq. (11)] and from following differently classified traders [i.e., Eq. (12)], respectively. The last column reports the difference between the third and fourth columns as a test of whether each investor type is more likely to follow similarly classified investors or differently classified investors.

[Insert Table 5 about here]

The results reveal mixed support for the reputational herding hypothesis. The results in the last three columns reveal that four of the five types are more likely to follow similarly classified institutions than differently classified institutions consistent with the reputational herding explanation. Independent advisors (who, as shown in Table 1, are the largest investor group), however, do not exhibit this pattern.[17] Moreover, inconsistent with the reputational herding explanation, mutual funds and independent advisors exhibit among the lowest herding propensities.

4.4. Industry herding and herding into size and book/market styles

Although Barberis and Shleifer (2003) note that style investing includes industry styles, most empirical work (e.g., Teo and Woo, 2004) focuses on styles defined by market capitalization and book-to-market ratios. Size-BE/ME styles are also often used in defining mutual fund classifications or manager strategies. In this section, we investigate the relation between industry herding and size-BE/ME style herding for three reasons. First, because industry membership is correlated with size-BE/ME styles (e.g., the technology industry primarily consists of low BE/ME growth stocks), it is possible that institutions industry herd because they herd to and from size-BE/ME styles rather than industry styles per se.

[17] One possible reason that independent managers do not follow each other more than other investors is that hedge funds (who are included in the set of independent advisors) recognize that 13(f) reports only reflect long positions that may be offset by unreported short positions. Therefore, 13(f) reports may be less informative regarding other independent investors' net positions.

Second, it is possible that institutional investors' industry signals may sometimes contain size-BE/ME components. We found a number of examples of analysts recommending securities within an industry based on size or valuation characteristics. For example, analysts at Fox-Pitt Kelton Cochran Caronia Waller (2008) argue investors should avoid small-cap bank stocks, "We expect third-quarter results in general will focus on credit-quality deterioration and capital adequacy. However, results will likely be bifurcated among regions and market cap…Bottom line, we believe the message coming out of the third quarter will be different than the prior three quarters for larger-caps, but will likely be similar or worse for the smaller-caps."

Third, we examine the relation between size-BE/ME herding and industry herding to help differentiate informational cascades from correlated signals. We propose that herding to similar size-BE/ME style stocks contributing to industry herding fits the correlated signals explanation better than the informational cascades explanation. Specifically, the correlated signals explanation is consistent with herding to similar size-BE/ME style securities contributing to industry herding if signals are sometimes related to size-BE/ME characteristics. If institutions agree with the analysts cited above, for example, institutions may herd out of small bank stocks more so than large bank stocks. Alternatively, the informational cascades explanation would require that an investor: (1) infer both an industry signal and a size-BE/ME signal from previous investors' trades, and (2) be willing to ignore her own industry and/or size-BE/ME signals to follow the perceived industry signal *and* the perceived size-BE/ME signal of previous traders. In the above example, for instance, informational cascades would require an institution who viewed banks as undervalued and small banks as more undervalued than large banks, to ignore both signals and follow the previous trader out banks and out of small banks more than large banks. And an investor who believed all banks were equally undervalued, would ignore her industry signal (and sell banks) and also sell small banks to a greater degree than large banks (despite believing all banks are equally undervalued). Thus, although the informational cascades argument is not necessarily inconsistent with size-BE/ME herding contributing to industry herding, the relation is more tenuous.

We begin to investigate the relation between industry herding and size-BE/ME herding by partitioning securities into six styles based on the median NYSE market equity breakpoint (big/small) and the 30th and 70th book to market NYSE percentile breakpoints (value/neutral/growth) following Fama and French (1993).[18] Because Eq. (7) can be decomposed to the stock level, we can investigate the relation between industry herding and size-BE/ME style herding by further partitioning the last term in Eq. (7) (i.e., the industry herding contribution) into managers following other managers into: (1) different, but same size-BE/ME style, stocks in the same industry, and (2) different style stocks in the same industry (see Appendix A for proof):

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})}\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right)=$$

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})}\sum_{k=1}^{K}\left(\sum_{i=1 i\in s}^{I_{k,t}}\sum_{j=1,j\neq i,j\in s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right)+$$

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})}\sum_{k=1}^{K}\left(\sum_{i=1,i\in s}^{I_{k,t}}\sum_{j=1,j\neq i,j\notin s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right),\quad (13)$$

where $i\in s$ indicates security i is in size-BE/ME style s.

[18] Following Fama and French (2006) book equity is computed as total assets (Compustat item #6) minus liabilities (#181) plus balance sheet deferred taxes and investment tax credits (#35) if available, minus preferred stock liquidating value (#10) if available, or redemption value (#56) if available, or carrying value (#130). Further following Fama and French, the book to market ratio is computed each year t based on market value at the end of December in year t and the book value for the fiscal year that ends in calendar year t. For the quarters ending in June, September, and December of year t, we use the book to market ratio from the end of year t-1. For the quarter ending in March, we use the book to market ratio from the end of year t-2.

The first column in Table 6 (identical to the middle cell in Table 3) reports the portion of the correlation attributed to institutional industry herding. The next two columns in the first row further partition the industry herding contribution into the portion that arises from institutions following other institutions into different, but same size-BE/ME style, stocks in the same industry [the first term on the right hand side of Eq. (13)] and the portion that arises from institution following other institutions into different size-BE/ME style stocks in the same industry [the last term in Eq. (13)]. All *t*-statistics in Table 6 are based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates.

[Insert Table 6 about here]

The results reveal that institutions following each other into and out of same size-BE/ME style stocks and different size-BE/ME style stocks both contribute to industry herding. Specifically, 65% (0.1260/0.1942) of the industry herding contribution [i.e., the last term in Eq. (7)] is due to following each other into same size-BE/ME style stocks and 35% (0.0683/0.1942) results from following each other into different size-BE/ME style stocks in the same industry. Both portions are statistically significant at the 1% level. The results demonstrate that industry herding is unique from size-BE/ME style herding.

Although the decomposition reveals that size-BE/ME style herding does not fully explain industry herding, it does not test whether size-BE/ME style herding contributes to industry herding. To examine this question, we compute the expected contribution by same and different style stocks by recognizing that if size-BE/ME herding does not contribute to industry herding, then manager n should be as likely to purchase (as opposed to sell) security i following manager m's purchase of security j ($i,j \in k$) whether securities i and j are in the same size-BE/ME styles or in different styles (see Appendix A for details). The second row in Table 6 reports the time-series average of expected contributions from following other managers into same and different size-BE/ME style stocks in the same industry under the null that managers are as likely to follow each other into and out of same size-BE/ME style stocks as different size-BE/ME style stocks. The last row reports the difference between the realized and expected contributions.

The results reveal that size-BE/ME style herding contributes to industry herding. Specifically, the realized contribution from following others into same size-BE/ME style stocks in the same industry accounts for 65% of the herding contribution (0.1260/0.1942) versus 39% (0.0765/0.1942) under the null hypothesis that institutional industry herding is independent of size-BE/ME style. The difference (0.0495=0.1260-0.0765) is statistically significant at the 1% level.[19] The results are consistent with the hypothesis that industry signals sometimes contain size-BE/ME components and provide support for the correlated signals explanation.

4.5. Herding pre- and post-Electronic Data Gathering and Retrieval (EDGAR) service

If institutional industry herding arises from institutional investors intentionally following each other into and out of the same industries (as in informational cascades or reputational herding), then institutions must somehow learn what industries other institutions are buying or selling. Noisy estimates of this information may arise from a number of sources. Given a positive relation between aggregate institutional demand for a security and same period security returns (e.g., Sias, Starks, and Titman, 2006) and a positive relation between aggregate institutional demand for an industry and same period industry returns (see Section 4.6), institutions may be able to garner some idea of whether other institutions are buying or selling from returns. Second, there is some evidence of word-of-mouth effects between institutions. Hong, Kubik, and Stein

[19] Because the first two rows of Table 6 are a simple partitioning of the last term in Eq. (7), the differences (reported in the last row) are exactly offsetting.

(2005), for example, find that a mutual fund manager is more likely to buy (sell) a stock if other managers in the same city are buying (selling) the same stock. Similarly, Cohen, Frazzini, and Malloy (2007) report that mutual fund managers who attended the same university tend to buy (or sell) the same stocks at the same time. Moreover, in a survey of institutional managers' purchases, Shiller and Pound (1989) report that half of their respondents claim "an investment professional" motivated their initial interest in the company.[20] Third, institutions may also gain information from interaction with broker-dealers or investor relations departments.

In 1996, however, the SEC began requiring institutions to file their 13(f) reports electronically through the SEC's EDGAR service.[21] Thus, in the last 40 quarters of the sample, institutional investors were able to easily access every other intuitional investors' previous quarter's trades.[22] If institutions intentionally following other institutions into the same industries is primarily responsible for industry herding, then the much less noisy signal available to all investors following mandatory EDGAR filing should result in much stronger levels of herding. Alternatively, if correlated signals primarily drive the results, then industry herding should be strong both prior to, and following, mandatory electronic filing.

Panel G of Table 2 reports the average correlation and its partitioned components for the post-EDGAR period (1996-2005, n=40 quarters) and the pre-EDGAR period (1983-1995, n=50 quarters). The results in the last column reveal the mean herding component averages 17% larger (0.4066/0.3484 – 1) in the post-EDGAR period. The last two rows in Panel G report a t-statistic from a difference in means test and a z-statistic from a Wilcoxon rank sum test that the herding

[20] There is also anecdotal evidence of word-of-mouth effects. In an interview with Ticker Magazine (2006), for example, Matthew Patsky of Winslow Green Growth Fund answers the question, "Can you explain your research process?" with "We consider ourselves bottomup stock pickers…We also have long-lasting relationships with other managers and we regularly share ideas."
[21] Managers were able to voluntarily file electronic 13(f) reports prior to this period.
[22] Institutions must file 13(f) reports within 45 days of quarter-end.

components are equal in the pre- and post-EDGAR periods. Although we cannot reject the hypothesis with the *t*-test for difference in means (*p*-value=0.11), the non-parametric Wilcoxon test rejects the hypothesis at the 5% level.

Consistent with the hypothesis that reputational herding and/or informational cascades sometimes contribute to industry herding, Panel G reveals that institutional industry herding is slightly greater in the post-EDGAR period. Nonetheless, consistent with the explanation that correlated signals primarily drive industry herding, the increase in the herding estimate is relatively small and there is strong herding both prior to, and following, mandatory EDGAR filing.

4.6. Institutional industry demand and industry returns

Investigative herding models propose that herding may result from institutions receiving, or acting on, correlated information at different times and therefore reflects the process by which information is incorporated into prices. In contrast, the alternative explanations suggest herding may drive prices from fundamentals—assuming, consistent with recent empirical work (e.g., Chakravarty, 2001; Froot and Teo, 2004; Sias, Starks, and Titman, 2006; Kaniel, Saar, and Titman, 2008; Campbell, Ramadorai, and Schwartz, 2007), that institutional investors are usually the price-setting marginal investor.

Recognize, however, that any relation between institutional demand and contemporaneous or subsequent security/industry prices does not necessarily imply institutional *herding* (i.e., institutions following other institutions) impacts prices but may simply reflect institutional demand shocks. Gompers and Metrick (2001), for example, propose that demand shocks associated with the growth in institutional assets under management and institutional investors' preference for large capitalization stocks may help explain the disappearance of the small firm premium in recent years.

Assuming institutional herding impacts returns, we can differentiate the correlated signals explanation from the alternatives by examining the relation between institutional demand, contemporaneous returns, and subsequent returns. If institutional industry herding reflects the manner that industry information is impounded into prices, then institutional demand should be positively correlated with contemporaneous industry returns and not inversely related to subsequent industry returns. In contrast, if herding does not always reflect the process by which information is incorporated into prices, then institutional demand should be positively related to contemporaneous industry returns and inversely related to subsequent industry returns.

We begin by computing, each quarter, each industry's contribution to the cross-sectional correlation between institutional demand this quarter and last quarter [i.e., Eq. (4)]. As before, we denote the 10 industries where the last two terms are both positive that contribute the most to the industry herding measure [i.e., with the largest Eq. (4)] as buy-herding industries and the top 10 industries where the last two terms are both negative as sell-herding industries. To compute buy- and sell-herd industry returns, each quarter, we calculate the average return across the 10 buy-herding industries and the 10 sell-herding industries. We then examine industry returns for the formation period (quarters -1 to 0) and up to three years following formation (quarters 1 to 12).

We use Jegadeesh and Titman's (1993) calendar time aggregation method to calculate returns each quarter from overlapping observations.[23] From the time-series of quarterly buy- or

[23] Because the portfolios are updated each quarter, evaluation periods longer than one quarter produce overlapping observations. Following Jegadeesh and Titman (1993), we aggregate results for each calendar quarter. Consider, for example, the first quarter of 1999 when evaluating the holding period for the two quarters following formation. The

sell-herd returns (as well as their difference), we estimate the abnormal return as the intercept from a time-series regression of the quarterly portfolio return on the Fama and French (1993) market, size, and value factors:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p\left(R_{m,t} - R_{f,t}\right) + \beta_{SMB} R_{SMB,t} + \beta_{HML} R_{HML,t} + \varepsilon_{p,t}, \tag{14}$$

where $R_{p,t}$ is the quarterly return on the buy-herd (or sell-herd or difference) portfolio, $R_{f,t}$ is the risk-free rate and $R_{m,t}$, $R_{SMB,t}$ and $R_{HML,t}$ are the Fama-French market, size, and value factor returns, respectively.[24]

The first two columns of Panel A in Table 7 report the average quarterly raw return from the buy- and sell-herding industry portfolios over the indicated period. The third column reports their difference and associated Newey-West *t*-statistic. The next three columns report the buy-herding portfolio, sell-herding portfolio, and difference portfolio (quarterly) alphas from Eq. (14).[25]

[Insert Table 7 about here]

The results reveal evidence consistent with the hypothesis that institutional industry demand impacts prices. In the two formation quarters, industries most heavily purchased by institutions outperform those most heavily sold by 2.73% per quarter (the difference in alphas is slightly larger).[26] In the four quarters immediately following formation, however, buy-herding

---

cross-sectional average return for the second quarter following the April-September of 1998 formation period is the first observation for the first quarter of 1999. The cross-sectional average return for the first quarter following the July-December 1998 formation period is the second observation for the first quarter of 1999. Averaging these two observations yields the average return during the first calendar quarter of 1999 over event quarters 1 and 2.

[24] Quarterly market, size, and value factor returns and the quarterly risk-free rate are calculated as compound monthly values (downloaded from Ken French's website).

[25] The *t*-statistics for the Fama-French alphas are based on time-series regressions of the Jegadeesh and Titman calendar aggregation returns and Newey and West (1987) standard errors.

[26] This is consistent with previous studies that show a positive relation between institutional demand (or subsets of institutional investors such as mutual funds) and individual security returns the same quarter including Grinblatt and Titman (1989, 1993), Grinblatt, Titman, and Wermers (1995), Jones, Lee, and Weis (1999), Nofsinger and Sias (1999), Wermers (1999, 2000), and Sias (2007).

industries underperform the sell-herding industries by 1.03% per quarter (marginally statistically significant at the 10% level).[27] Some of this difference, however, is due to differences in exposure to the Fama-French factors. Specifically, the difference in 3-factor alphas is -0.67% per quarter (over quarters 1 to 4), but not statistically significant at traditional levels. Although factor loadings are not reported (to reduce clutter), this largely arises from sell-herding industries' greater sensitivity to the value factor. In sum, although the results in the first row of Panel A reveal a strong positive relation between institutional industry demand and industry returns the same period, we only find weak evidence of a subsequent return reversal.

In an interesting study, Dasgupta, Prat, and Verardo (2007) find that securities persistently purchased by institutions (e.g., over the last four quarters) subsequently underperform those persistently sold by institutions. The authors interpret the apparent price correction as resulting from mispricing induced by long-term institutional herding. To investigate this possibility for industries, each quarter we partition the 49 industries into those that were purchased more than average (i.e., $(\Delta_{k,t} - \overline{\Delta_{k,t}}) > 0$) by institutions in each of the four previous quarters (t=0 to t=-3) and those that were sold more than average in each of the four previous quarters. The number of industries that meet these criteria ranges from 2 to 14 and averages 7.31 industries that institutions bought over each of the last four quarters and 7.94 industries that institutions sold over each of the last four quarters. We then repeat the analysis in the previous section based on these longer-term buy- and sell-herd industries.

[27] Although early work suggests that cross-sectional variation in institutional demand for individual securities is positively related to future returns (e.g., Nofsinger and Sias, 1999; Gompers and Metrick, 2001), recent work (e.g., San, 2007; Dasgupta, Prat, and Verardo, 2007) suggests an inverse relation between institutional demand and subsequent security returns in more recent periods. In untabulated results, we split the sample into two periods and find that although sell-herding industries subsequently outperform buy-herding industries in both the early (1983:12-1994:12) and late (1995:03-2005:12) periods, the difference is greater (-1.49% versus -0.56% per quarter over quarters 1 to 4) and statistically significant only in the early period. The Fama and French (1993) 3-factor alpha is also statistically significant in the early period.

The results, reported in Panel B of Table 7, reveal slightly stronger evidence that institutional industry herding sometimes drives prices from fundamental values. Specifically, those industries institutions purchased over the last four quarters subsequently underperform, on average, those industries institutions sold over the last four quarters. In the first year following formation, differences are statistically significant at the 10% and 5% levels for raw and abnormal returns, respectively.

In sum, the results in Table 7 reveal evidence consistent with the explanation that informational cascades, fads, and reputational herding may sometimes play a role in driving institutional industry herding. Because evidence of return reversals is weak, however, the analysis suggests that correlated signals primarily drive institutional industry herding.

5. Conclusions

Institutional investors follow each other into and out of the same industries (i.e., "industry herd"). Our results have implications for two related literatures. First, whatever factors drive institutional investors to herd appear to have an industry component. (Although, the primary factors that drive stock herding may differ from the primary factors that drive industry herding.) If, for example, some institutional investors herd in an attempt to preserve reputation, then our results are consistent with the hypothesis that managers attempt to preserve reputation by adjusting industry positions as well as stock positions. Analogously, if fads sometimes contribute to institutional herding, then there must be industry fads. If informational cascades contribute to industry herding, then institutions must, at least sometime, infer industry signals from each others' trades. And if following correlated signals cause institutional herding, then institutions' signals must have an industry component.

Second, our evidence is consistent with the growing style investing literature. Specifically, the Barberis and Shleifer (2003) style model requires a group of investors to style herd and that their herding impacts prices. Related empirical studies also contain these assumptions. Our results demonstrate that institutions herd to industry styles and are consistent with the explanation that such herding impacts prices.

Additional tests suggest a number of factors contribute to industry herding. Consistent with reputational herding, most institutions are more likely to follow similarly-classified institutions than differently-classified institutions. Inconsistent with the reputational explanation, however, we find no evidence that those investors who should be most concerned about their reputations (mutual funds and independent advisors) are more likely to herd than other investors. We also find that institutional investors momentum trade at the industry level. Institutional industry momentum trading, however, does not explain their herding—once accounting for lag industry demand, institutional industry demand is independent of lag industry returns.

In aggregate, our tests are most supportive of the correlated signals explanation. Specifically, three results support the explanation that correlated signals primarily drive institutional industry herding. First, evidence that size-BE/ME herding contributes to industry herding fits the correlated signals explanation better than the informational cascades explanation. Second, evidence of institutional herding is nearly as strong prior to mandatory electronic filing of ownership positions as following mandatory electronic filing. If the results are primarily driven by institutions intentionally following other institutions into the same industry (and not correlated signals), then, contrary to our empirical findings, the herding should be much weaker prior to electronic filing. Third, consistent with the correlated signals explanation, we find only weak evidence of subsequent industry return reversals.

References

Badrinath, S.G., Wahal, S., 2002. Momentum trading by institutions. Journal of Finance 57, 2449-2478.

Banerjee, A.V., 1992. A simple model of herd behavior. Quarterly Journal of Economics 107, 797-817.

Barberis, N., Shleifer, A., 2003. Style investing. Journal of Financial Economics 68, 161-199.

Barberis, N., Shleifer, A., Wurgler, J., 2005. Comovement. Journal of Financial Economics 75, 283-317.

Barclay, M.J., Warner, J.B., 1993. Stealth trading and volatility: which trades move prices? Journal of Financial Economics 34, 281-305.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100, 992-1026.

Boyer, B., Zheng L., 2004. Who moves the market? A study of stock prices and sector cashflows. Unpublished working paper, Brigham Young University and University of Michigan.

Brown, N., Wei, K., Wermers, R., 2007. Analyst recommendations, mutual fund herding, and overreaction in stock prices. Unpublished working paper, University of Southern California, University of Texas – Dallas, University of Maryland.

Campbell, J., Ramadorai, T., Schwartz, A., 2007. Caught on tape: institutional trading, stock returns, and earnings announcements. Centre for Economic Policy Research Discussion paper No. 6390.

Chakravarty, S., 2001. Stealth trading: which traders' trades move stock prices? Journal of Financial Economics 61, 289-307.

Chamley, C., Gale, D., 1994. Information revelation and strategic delay in a model of investment. Econometrica 62, 1065-1085.

Cohen, L., Frazzini, A., Malloy, C., 2007. The small world of investing: Board connections and mutual fund returns. NBER Working Paper No. 13121.

Dasgupta, A., Prat, A., Verardo M., 2007. Institutional trade persistence and long-term equity returns. Centre for Economic Policy Research Discussion Paper No. 6374.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

Fama, E., French, K., 1997. Industry costs of equity. Journal of Financial Economics 43, 153-194.

Fama, E., French, K., 2006. Profitability, investment and average returns. Journal of Financial Economics 82, 491-518.

Foster, F.D., Gallagher, D.R., Looi, A., 2005. Institutional trading and share returns. Unpublished working paper, The University of New South Wales, Australia.

Fox-Pitt Kelton Cochran Caronia Waller, 2008. Outlook for small U.S. banks may worsen. Barron's Investors' Soapbox PM, October 7, 2008.

Frazzini, A., Lamont, O.A., 2008. Dumb money: mutual fund flows and the cross-section of stock returns. Journal of Financial Economics 88,299-322.

Friedman, B.M., 1984. A comment: stock prices and social dynamics. Brookings Papers on Economic Activity 2, 504-508.

Froot, K.A., Scharfstein, D.S., Stein J.C., 1992. Herd on the street: informational inefficiencies in a market with short-term speculation. Journal of Finance 47, 1461-1484.

Froot, K.A., Teo, M., 2004. Equity style returns and institutional investor flows. NBER Working Paper No. 10355.

Froot, K.A., Teo, M., 2007. Style investing and institutional investors. Journal of Financial and Quantitative Analysis, forthcoming.

Gompers, P., Metrick, A., 2001. Institutional investors and equity prices. Quarterly Journal of Economics, 116, 229-260.

Grinblatt, M., Titman, S., 1989. Portfolio performance evaluation: old issues and new insights. Review of Financial Studies 2, 393-422.

Grinblatt, M, Titman, S., 1993. Performance measurement without benchmarks: an examination of mutual fund returns. Journal of Business 66, 47-68.

Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior. American Economic Review 85, 1088-1105.

Gul, F., Lundholm, R., 1995. Endogenous timing and the clustering of agents' decisions. Journal of Political Economy 103, 1039-1066.

Hirshleifer, D., Subrahmanyam, A., Titman, S., 1994. Security analysis and trading patterns when some investors receive information before others. Journal of Finance 49, 1665-1698.

Hong, H., Kubik, J., Stein, J., 2005. Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. Journal of Finance 60, 2801-2824.

Hou, K., 2007. Industry information diffusion and the lead-lag effect in stock returns. Review of Financial Studies 20, 1113-1138.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65-91.

Jones, C., Lipson, M., 2003. Are retail orders different? Unpublished working paper, Columbia University and University of Georgia.

Jones, S., Lee, D., Weis, E., 1999. Herding and feedback trading by different types of institutions and the effects on stock prices. Unpublished working Paper, Indiana University–Indianapolis, Kennesaw State University and Merrill Lynch.

Kaniel, R., Saar, G., Titman, S., 2008. Individual investor trading and stock returns. Journal of Finance 63, 273-310.

Lakonishok, J., Shleifer, A., Vishny, R.W., 1992. The impact of institutional trading on stock prices. Journal of Financial Economics 32, 23-43.

Lang, M., Lundholm, R., 1996. The relation between security returns, firm earnings, and industry earnings. Contemporary Accounting Research 13, 607-629.

Moskowitz, T.J., Grinblatt, M., 1999. Do industries explain momentum? Journal of Finance 54, 1249-1290.

Newey, W.K., West, K.D., 1987. A simple positive semi-definitive, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703-708.

Nofsinger, J., Sias, R., 1999. Herding and feedback trading by institutional and individual investors. Journal of Finance 54, 2263-2295.

Puckett, A., Yan, X.S., 2008. Short-term institutional herding and its impact on stock prices. Unpublished working paper, University of Missouri.

San, G., 2007. Who gains more by trading–institutions or individuals? Unpublished working paper, Hebrew University of Jerusalem.

Scharfstein, D., Stein, J., 1990. Herd Behavior and investment. American Economic Review 80, 465-479.

Schwartz, R., Shapiro, J., 1992. The challenge of institutionalization for equity markets. In: Saunders, A. (Ed.), Recent developments in finance. Business One Irwin, Homewood, IL, 31-45.

Sharma, V., Easterwood, J., Kumar R., 2006. Institutional herding and the internet bubble. Unpublished working paper, University of Michigan-Dearborn, and Virginia Tech.

Shell, A., 2001. Tech teetotalers have the last laugh. USA Today, January 10, 2001.

Shiller, R., Pound, J., 1989. Survey evidence on diffusion of interest and information among investors. Journal of Economic Behavior and Organization 12, 47-66.

Sias, R.W., 2004. Institutional herding. Review of Financial Studies 17, 165-206.

Sias, R.W., 2007. Reconcilable differences: momentum trading by institutions. The Financial Review 42, 1-22.

Sias, R.W., Starks, L., Titman, S., 2006. Changes in institutional ownership and stock returns: assessment and methodology. Journal of Business 79, 2869-2910.

Teo, M., Woo, S., 2004. Style effects in the cross-section of stock returns. Journal of Financial Economics 74, 367-398.

Ticker Magazine, 2006. Green Returns. 123jump.com website: www.123jump.com/mutual-fund/Green-Returns/296/.

Trueman, B., 1994. Analyst forecasts and herding behavior. Review of Financial Studies 7, 97-124.

Turner, S., 2008. Lehman Brothers see tactical case for deep value overweight. Marketwatch, April 14, 2008.

Wermers, R., 1999. Mutual fund herding and the impact on stock prices. Journal of Finance 54, 581-622.

Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transactions costs, and expenses. Journal of Finance 55, 1655-1695.

Wermers, R., Yao, T., Zhao, J., 2007. The investment value of mutual fund portfolio disclosure. Unpublished working paper, University of Maryland, University of Iowa and University of Arizona.

Wylie, S., 2005. Fund manager herding: a test of the accuracy of empirical results using U.K. data. Journal of Business 78, 381-403.

Zwiebel, J., 1995. Corporate conservatism and relative compensation. Journal of Political Economy 103, 1-25.

Zhang, J., 1997. Strategic delay and the onset of investment cascades. RAND Journal of Economics 28, 188-205.

Table 1. Descriptive statistics

Stocks are classified each quarter (between March 1983 and December 2005) into one of 49 industries. Panel A reports the time-series average of the cross-sectional descriptive statistics for the number of institutional investors trading in each industry (overall and by type) and the ratio of the number of institutional buyers to institutional traders in each industry. Panel B reports the time-series average of the cross-sectional descriptive statistics for the number of firms in each industry, the fraction of market capitalization accounted for by each industry, and the fraction of industry capitalization accounted for by the largest firm in the industry. Panel C reports time-series descriptive statistics for each of the 49 industries including the average number of firms in the industry, the industry's market capitalization weight, and the average, time-series standard deviation, and first order autocorrelation of institutional demand for the industry.

	Panel A: Institutional investor statistics				
	Mean	Median	Minimum	Maximum	Std. Dev.
Number of institutions trading in an industry	692	748	150	1,076	270
Number of banks trading	134	153	32	177	44
Number of insurance companies trading	36	38	9	55	12
Number of mutual funds trading	42	45	12	60	13
Number of independent advisors trading	440	468	82	723	191
Number of unclassified institutions trading	40	42	7	64	15
#Buyers/(#Buyers + #Sellers)	50.04%	50.08%	39.76%	60.38%	4.08%

	Panel B: Industry statistics				
	Mean	Median	Minimum	Maximum	Std. Dev.
Number of firms in industry	116	77	6	609	118
Industry capitalization/Market capitalization	2.04%	1.18%	0.05%	11.35%	2.44%
Largest firm in industry/Industry capitalization	31.79%	26.99%	5.09%	80.19%	19.20%

	Panel C: Industry statistics by industry				
Industry	# of Firms	%Market Cap.	#Buyers/(#Buyers + #Sellers) Mean	Std. Dev.	Autocorrelation
Agriculture	18	0.08%	0.514	0.054	0.160
Food Products	80	2.12%	0.478	0.037	0.396
Candy & Soda	11	1.85%	0.465	0.040	0.323
Beer & Liquor	18	0.45%	0.473	0.045	0.430
Tobacco Products	6	1.37%	0.467	0.059	0.615
Recreation	49	0.43%	0.507	0.035	0.411
Entertainment	76	0.84%	0.498	0.047	0.160
Printing and Publishing	60	1.38%	0.484	0.034	0.281
Consumer Goods	103	4.46%	0.485	0.037	0.405
Apparel	67	0.46%	0.487	0.039	0.430
Healthcare	120	0.95%	0.508	0.039	0.465
Medical Equipment	156	1.39%	0.502	0.035	0.348
Pharmaceutical Products	203	6.63%	0.491	0.039	0.203
Chemicals	89	2.68%	0.489	0.034	0.432
Rubber and Plastic Products	49	0.22%	0.507	0.043	0.264
Textiles	36	0.17%	0.497	0.057	0.378
Construction Materials	125	1.54%	0.490	0.033	0.566

Table 1. Descriptive statistics (continued)

Industry	# of Firms	%Market Cap.	Panel C: Industry statistics by industry #Buyers/(#Buyers + #Sellers) Mean	Std. Dev.	Autocor-relation
Construction	63	0.30%	0.508	0.043	0.229
Steel Works Etc	71	0.89%	0.503	0.033	0.188
Fabricated Products	20	0.08%	0.523	0.054	0.355
Machinery	174	1.60%	0.500	0.032	0.364
Electrical Equipment	159	1.97%	0.505	0.029	0.240
Automobiles and Trucks	67	2.40%	0.488	0.047	0.139
Aircraft	24	1.03%	0.494	0.039	0.361
Shipbuilding, Railroad Equipment	8	0.14%	0.501	0.050	0.244
Defense	9	0.25%	0.492	0.042	0.277
Precious Metals	28	0.19%	0.520	0.060	0.162
Non-Metallic and Industrial Metal Mining	26	0.27%	0.501	0.060	0.397
Coal	9	0.07%	0.514	0.052	0.273
Petroleum and Natural Gas	226	6.52%	0.488	0.045	0.498
Utilities	167	5.71%	0.514	0.046	0.582
Communication	132	5.54%	0.504	0.052	0.243
Personal Services	60	0.40%	0.504	0.035	0.410
Business Services	313	2.21%	0.515	0.028	0.074
Computer Hardware	176	4.62%	0.488	0.030	0.192
Computer Software	290	3.51%	0.532	0.045	0.178
Electronic Equipment	262	3.41%	0.506	0.038	0.260
Measuring and Control Equipment	113	0.84%	0.502	0.046	0.320
Business Supplies	52	1.34%	0.486	0.036	0.151
Shipping Containers	24	0.73%	0.484	0.039	0.443
Transportation	111	1.35%	0.497	0.042	0.516
Wholesale	235	1.61%	0.511	0.031	0.075
Retail	274	5.58%	0.503	0.044	0.225
Restaraunts, Hotels, Motels	126	1.22%	0.493	0.032	0.185
Banking	455	5.22%	0.505	0.046	0.022
Insurance	157	3.55%	0.497	0.040	0.429
Real Estate	57	0.23%	0.509	0.054	0.382
Trading	465	9.63%	0.509	0.049	0.041
Almost Nothing	58	0.56%	0.515	0.047	0.362

Table 2. Tests for herding

The first column in Panel A reports the time-series average of 90 correlation coefficients between institutional industry demand this quarter and last quarter (from September 1983 to December 2005). Institutional industry demand is defined as the number of institutional investors buying the industry that quarter divided by the number of institutional investors trading the industry that quarter. The next two columns partition the correlation coefficient into the portion that results from institutional investors following their own lag industry demand and the portion that results from institutions following the lag industry demand of other institutional investors [see Eq. (3)]. In Panel B, the correlation is further partitioned into those industries institutions purchased in quarter *t*-1 (buy herding) and those industries institutions sold in quarter *t*-1 (sell herding). Panel C reports time-series average industry-weighted correlation (and its components). Panel D uses alternative industry definitions. Panel E excludes mutual funds and independent investment advisors from the analysis. In Panels A-E, an institution is defined as a buyer (seller) if the institution increases (decreases) their position in industry over the quarter. In Panel F an institution is defined as a buyer (seller) if the institution increases (decreases) their industry portfolio weight over the quarter. Panel G partitions the results in Panel A into the post-EDGAR period (*n*=40 quarters) and the pre-EDGAR period (*n*=50 quarters). In Panels A-F, *t*-statistics (reported in parentheses) are based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates. ** indicates statistical significance at the 1% level; * at the 5% level.

Table 2. Tests for herding (continued)

	Average correlation coefficient	Partitioned correlation coefficient: Institutions following their own lag industry demand	Partitioned correlation coefficient: Institutions following other institutions' lag industry demand
	Panel A: All institutions		
49 industries	0.4049	0.0307	0.3743
	(18.02)**	(16.32)**	(16.53)**
	Panel B: Buy herding versus sell herding		
Buy herding	0.2016	0.0157	0.1859
Sell herding	0.2034	0.0150	0.1884
Difference	-0.0018	0.0007	-0.0025
	(-0.10)	(0.45)	(-0.14)
	Panel C: All institutions – Industry-weighted correlation		
49 industries	0.4177	0.0238	0.3939
	(13.10)**	(16.10)**	(12.42)**
	Panel D: All institutions – Alternative industry definitions		
2-digit SIC code	0.2465	0.0218	0.2246
	(6.44)**	(0.79)	(7.80)**
38 industries	0.3475	0.0572	0.2903
	(12.38)**	(5.83)**	(9.49)**
30 industries	0.4135	0.0293	0.3842
	(14.56)**	(14.74)**	(13.55)**
17 industries	0.3627	0.0289	0.3338
	(10.07)**	(14.09)**	(9.27)**
12 industries	0.3930	0.0266	0.3664
	(9.83)**	(14.29)**	(9.06)**
10 industries	0.4073	0.0259	0.3814
	(9.97)**	(13.50)**	(9.22)**
5 industries	0.2811	0.0499	0.2312
	(4.47)**	(8.37)**	(3.59)**
	Panel E: Excludes mutual funds and independent advisors		
49 industries	0.4121	0.0326	0.3795
	(19.76)**	(17.77)**	(17.93)**
	Panel F: All institutions – Buyer if increased portfolio weight in industry		
49 industries	0.3687	0.0189	0.3498
	(16.24)**	(14.48)**	(15.58)**
	Panel G: Pre- and post-EDGAR electronic 13(f) filing		
Post-EDGAR (1996-2005)	0.4284	0.0217	0.4066
Pre-EDGAR (1983-1995)	0.3861	0.0378	0.3484
t-test for difference			1.65
Wilcoxon *z*-statistic			2.04*

Table 3. Regression of weighted institutional industry demand on lag weighted institutional industry demand

Institutional demand for security *i* is computed as the number of institutional investors buying security *i* in quarter *t* divided by the number of institutions trading security *i* in quarter *t*. Weighted institutional demand for industry *k* is computed as the cross-sectional weighted average (by beginning of quarter capitalization) demand for all securities in industry *k*. The bottom right-hand cell reports the time-series average of 90 correlation coefficients between weighted institutional industry demand this quarter and last quarter (from September 1983 to December 2005). This correlation is partitioned [see Eq. (7)], each quarter, into four components: (1) institutions following themselves into the same stock (top left-hand cell), (2) institutions following other institutions into the same stock (middle row, left-hand cell), (3) institutions following themselves into different stocks in the same industry (top row, middle cell), and (4) institutions following other institutions into different stocks in the same industry (middle row, middle cell). Summing across columns (last column) yields the totals for following themselves versus following other institutions. Summing across rows (last row) yields the totals for following into the same stock versus following into different stocks in the same industry. All *t*-statistics (reported in parentheses) are based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates. ** indicates statistical significance at the 1% level.

	Into the same stock	Into different stocks in the same industry	Total
Following themselves	0.0206	0.0333	0.0539
	(10.28)**	(5.64)**	(7.46)**
Following others	0.3235	0.1942	0.5177
	(14.99)**	(11.10)**	(23.23)**
Total	0.3441	0.2275	0.5716
	(16.41)**	(11.39)**	(27.32)**

Table 4. Tests for herding and momentum trading

Each column in this table reports the time-series average coefficient from 90 cross-sectional regressions of standardized institutional industry demand this quarter on: (1) standardized lag institutional industry demand over the previous quarter, six months, or year (first, fourth, and seventh columns), (2) standardized industry returns the previous quarter, six months, or year (second, fifth, and eighth columns), or (3) standardized industry returns and standardized institutional industry demand over the previous quarter, six months, or year (third, sixth, and last columns). Institutional industry demand is defined as the number of institutional investors increasing their position in the industry divided by the number of institutional investors trading the industry. All *t*-statistics (reported in parentheses) are based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates. ** indicates statistical significance at the 1% level; * at the 5% level.

	Measured over previous quarter			Measured over previous six months			Measured over previous year		
Lag institutional demand	0.4049 (18.02)**		0.4052 (17.29)**	0.3802 (17.98)**		0.3752 (17.23)**	0.3858 (18.41)**		0.3716 (17.83)**
Lag return		0.0590 (2.48)*	-0.0134 (-0.64)		0.0928 (3.14)**	0.0004 (0.02)		0.1221 (3.69)**	0.0356 (1.38)
Adjusted R^2	17.46%	2.70%	19.23%	15.21%	3.28%	16.91%	15.57%	4.69%	17.99%

Table 5. Analysis by investor type

Institutional demand for each industry quarter is computed as the ratio of the number of institutional buyers to the number of institutional traders. This table reports the average contribution to the correlation between institutional demand this quarter and last quarter by investor type. The first column reflects each investor's propensity to follow their own lag industry demand and the second column reflects each investor's propensity to follow other institutional investors into and out of the same industry. The third column reports the average contribution to the correlation from each investor type following similarly classified institutions, e.g., banks following other banks [see Eq. (11)]. The fourth column reports the average contribution to the correlation from each investor type following differently classified institutions, e.g., banks following insurance companies [see Eq. (12)]. The last column reports the difference between columns three and four. All *t*-statistics (reported in parentheses) are based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates. ** indicates statistical significance at the 1% level.

	Average contribution from following their own industry trades	Average contribution from following others' industry trades	Average contribution from following same type traders	Average contribution from following different type traders	Average "same contribution" less average "different contribution"
Banks	0.0232	0.0011	0.0023	0.0007	0.0016
	(26.64)**	(15.74)**	(13.01)**	(10.00)**	(10.52)**
Insurance	0.0226	0.0003	0.0019	0.0002	0.0017
companies	(15.24)**	(5.75)**	(6.27)**	(3.57)**	(5.34)**
Mutual funds	0.0326	0.0004	0.0011	0.0003	0.0008
	(21.41)**	(4.95)**	(4.36)**	(3.93)**	(2.90)**
Independent	0.0296	0.0004	0.0003	0.0005	-0.0001
advisors	(28.04)**	(11.84)**	(7.86)**	(10.32)**	(-3.42)**
Unclassified	0.0283	0.0007	0.0022	0.0006	0.0015
investors	(12.28)**	(8.00)**	(6.33)**	(6.53)**	(4.20)**

Table 6. Institutional industry herding into same size-BE/ME style stocks and different size-BE/ME style stocks

Institutional demand for security *i* is computed as the number of institutional investors buying security *i* in quarter *t* divided by the number of institutions trading security *i* in quarter *t*. Weighted institutional demand for industry *k* is computed as the cross-sectional weighted average (by beginning of quarter capitalization) demand for all securities in industry *k*. Each quarter we compute the correlation coefficient between weighted institutional industry demand this quarter and weighted institutional industry demand last quarter (from September 1983 to December 2005). The first column reports the portion of this correlation due to institutions following other institutions into different stocks in the same industry (this figure is identical to the middle row of the middle column in Table 3). The next two columns in the first row further partition the contribution into the portion attributed to institutions following others into (and out of) different stocks in the same industry within the same size-BE/ME style and into (and out of) different size-BE/ME style stocks in the same industry, respectively. The second row reports the time-series mean expected values computed under the null hypothesis that managers are as likely to follow other managers into and out of same size-BE/ME style stocks as different size-BE/ME style stocks (see Appendix A). The last row reports the mean difference between the realized and expected values. All *t*-statistics (reported in parentheses) are based on Newey and West (1987) standard errors computed from the time-series of coefficient estimates. ** indicates statistical significance at the 1% level.

	Into different stocks in the same industry	Same size-BE/ME style	Different size-BE/ME style
Realized contribution	0.1942	0.1260 (12.48)**	0.0683 (5.78)**
Expected contribution	0.1942	0.0765 (11.31)**	0.1178 (10.81)**
Realized - expected		0.0495 (6.88)**	-0.0495 (-6.88)**

Table 7. Industry herding and subsequent returns

This table reports the average quarterly raw and abnormal returns for buy-herding and sell-herding industries over the formation period and the post-formation period. Institutional industry demand is defined as the number of institutional investors increasing their position in the industry that quarter divided by the number of institutional investors trading the industry that quarter. In Panel A, the 49 industries are sorted, each quarter, into the top 10 buy-herding industries (those industries that institutions buy in both quarter t=0 and t=-1 that contribute the most to the cross-sectional correlation between demand this quarter and last) and the top 10 sell-herding industries (those industries that institutions sell in both quarter t=0 and t=-1 that contribute the most to the cross-sectional correlation between demand this quarter and last). In Panel B, the 49 industries are sorted, each quarter, into those with above average institutional demand (buy herds) in each of the four previous quarters (t=0 to t=-3) and those with below average institutional demand (sell herds) in each of the four previous quarters. The t-statistics (reported in parentheses) for raw industry returns are based on non-overlapping quarters following the calendar-aggregation method in Jegadeesh and Titman (1993) and Newey and West (1987) standard errors. The t-statistics for the alphas are based on time-series regressions of the Jegadeesh and Titman calendar aggregation returns on market, size, and value factors and Newey and West standard errors. ** indicates statistical significance at the 1% level; * at the 5% level.

	Raw industry returns			Fama-French 3-factor model alphas		
	Buy herds	Sell herds	Difference	Buy herds	Sell herds	Difference
	Panel A: Portfolios based on herding over quarters t=0 to t=-1					
Quarter -1 to 0	0.0477	0.0204	0.0273	0.0152	-0.0172	0.0324
			(4.50)**			(4.96)**
Quarter 1	0.0315	0.0384	-0.0069	-0.0002	0.0005	-0.0007
			(-1.11)			(-0.12)
Quarters 1 to 2	0.0321	0.0382	-0.0061	-0.0003	0.0002	-0.0005
			(-1.01)			(-0.09)
Quarters 1 to 4	0.0293	0.0396	-0.0103	-0.0042	0.0025	-0.0067
			(-1.94)			(-1.59)
Quarters 5 to 8	0.0319	0.0378	-0.0059	-0.0054	0.0010	-0.0064
			(-1.32)			(-1.62)
Quarters 9 to 12	0.0356	0.0381	-0.0026	-0.0026	0.0030	-0.0055
			(-0.56)			(-1.23)
	Panel B: Portfolios based on herding over quarters t=0 to t=-3					
Quarter -3 to 0	0.0498	0.0273	0.0224	0.0180	-0.0107	0.0286
			(3.29)**			(3.66)**
Quarter 1	0.0340	0.0430	-0.0090	-0.0006	0.0063	-0.0069
			(-1.33)			(-1.15)
Quarters 1 to 2	0.0304	0.0430	-0.0126	-0.0051	0.0065	-0.0116
			(-1.87)			(-2.14)*
Quarters 1 to 4	0.0298	0.0414	-0.0117	-0.0060	0.0051	-0.0110
			(-1.85)			(-2.28)*
Quarters 5 to 8	0.0292	0.0377	-0.0085	-0.0081	0.0012	-0.0093
			(-1.73)			(-2.00)*
Quarters 9 to 12	0.0336	0.0370	-0.0034	0.0002	0.0041	-0.0039
			(-0.67)			(-0.77)

Appendix A: Proofs

A. Proof of Eq. (3)

Eq. (2) defines institutional demand ($\Delta_{k,t}$) for industry k as the ratio of the number of institutions buying industry k in quarter t to the number of institutions buying or selling industry k in quarter t. Defining $D_{n,k\,t}$ as a dummy variable that equals one if institutional investor n increases her position in industry k in quarter t, and zero if the investor decreases her position in industry k, institutional demand can be written:

$$\Delta_{k,t} = \sum_{n=1}^{N_{k,t}} \frac{D_{n,k,t}}{N_{k,t}}, \tag{A1}$$

where $N_{k,\,t}$ is the number of institutions trading industry k in quarter t.

The cross-sectional correlation between institutional demand this quarter and last is given by:

$$\rho\left(\Delta_{k,t}, \Delta_{k,t-1}\right) = \frac{1}{\sqrt{\sum_{k=1}^{K} w_k \left(\Delta_{k,t} - \overline{\Delta_{k,t}}\right)^2}\sqrt{\sum_{k=1}^{K} w_k \left(\Delta_{k,t-1} - \overline{\Delta_{k,t-1}}\right)^2}} \sum_{k=1}^{K} w_k \left(\Delta_{k,t} - \overline{\Delta_{k,t}}\right)\left(\Delta_{k,t-1} - \overline{\Delta_{k,t-1}}\right), \tag{A2}$$

where w_k is one divided by the number of industries ($1/K$) for the equal-weighted correlations and the industry's market weight at the beginning of quarter t-1 for the value-weighted correlations. Analogously, $\overline{\Delta_{k,t}}$ is equal-weighted average institutional demand across industries for the equal-weighted correlations and the value-weighted average institutional demand across industries for the value-weighted correlations.

For ease of notation, define:

$$C_t = \sqrt{\sum_{k=1}^{K} w_k \left(\Delta_{k,t} - \overline{\Delta_{k,t}}\right)^2}\sqrt{\sum_{k=1}^{K} w_k \left(\Delta_{k,t-1} - \overline{\Delta_{k,t-1}}\right)^2}\ . \tag{A3}$$

Substituting (A1) and (A3) into (A2) yields:

$$\rho\left(\Delta_{k,t},\Delta_{k,t-1}\right)=\left[\frac{1}{C_t}\right]\sum_{k=1}^{K}w_k\left[\left(\sum_{n=1}^{N_{k,t}}\frac{D_{n,k,t}-\overline{\Delta_{k,t}}}{N_{k,t}}\right)\left(\sum_{n=1}^{N_{k,t-1}}\frac{D_{n,k,t-1}-\overline{\Delta_{k,t-1}}}{N_{k,t-1}}\right)\right]. \tag{A4}$$

This sum of products can be further partitioned into those that arise from investors following their own lag industry demand (i.e., investor *n*'s industry demand at times *t* and *t*-1) and those that arise from investors following the lag industry demand of other institutional investors (i.e., investor *n*'s demand at time *t* and investor *m*'s demand at time *t*-1), yielding:

$$\rho\left(\Delta_{k,t},\Delta_{k,t-1}\right)=\left[\frac{1}{C_t}\right]\sum_{k=1}^{K}w_k\left[\sum_{n=1}^{N_{k,t}}\left(\frac{D_{n,k,t}-\overline{\Delta_{k,t}}}{N_{k,t}}\bullet\frac{D_{n,k,t-1}-\overline{\Delta_{k,t-1}}}{N_{k,t-1}}\right)\right]+$$

$$\left[\frac{1}{C_t}\right]\sum_{k=1}^{K}w_k\left[\sum_{n=1}^{N_{k,t}}\sum_{m=1,m\neq n}^{N_{k,t-1}}\left(\frac{D_{n,k,t}-\overline{\Delta_{k,t}}}{N_{k,t}}\bullet\frac{D_{m,k,t-1}-\overline{\Delta_{k,t-1}}}{N_{k,t-1}}\right)\right]. \tag{A5}$$

B. Proof of Eqs. (7) and (13)

Eq. (5) defines institutional demand for security *i* ($\Delta_{i,t}$) as the ratio of the number of institutions buying security *i* in quarter *t* to the number of institutions buying or selling security *i* in quarter *t*. Defining $D_{n,i,t}$ as a dummy variable that equals one if institutional investor *n* increases her position in security *i* in quarter *t*, and zero if the investor decreases her position in security *i*, institutional demand for security *i* can be written:

$$\Delta_{i,t}=\sum_{n=1}^{N_{i,t}}\frac{D_{n,i,t}}{N_{i,t}}, \tag{A6}$$

where $N_{i,t}$ is the number of institutions trading security *i* in quarter *t*. We define the weighted institutional demand for industry *k* (denoted $\Delta^*_{k,t}$) as the market-capitalization weighted average institutional demand across the securities in industry *k* (where $w_{i,t}$ is security *i*'s capitalization weight within industry *k* at the beginning of quarter *t*):

$$\Delta^*_{k,t} = \sum_{i=1}^{I_{k,t}} w_{i,t}\Delta_{i,t} \ , \tag{A7}$$

where $I_{k,t}$ is the number of securities in industry k in quarter t. For ease of notation, define:

$$C^*_t = \sqrt{\sum_{k=1}^{K} w_k\left(\Delta^*_{k,t} - \overline{\Delta^*_{k,t}}\right)^2}\sqrt{\sum_{k=1}^{K} w_k\left(\Delta^*_{k,t-1} - \overline{\Delta^*_{k,t-1}}\right)^2} \ , \tag{A8}$$

where w_k is as defined above (in subsection A). The correlation between weighted institutional industry demand this quarter and last is given by:

$$\rho\left(\Delta^*_{k,t}, \Delta^*_{k,t-1}\right) = \frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\Delta^*_{k,t} - \overline{\Delta^*_{k,t}}\right)\left(\Delta^*_{k,t-1} - \overline{\Delta^*_{k,t-1}}\right). \tag{A9}$$

Substituting Eq. (A7) into (A9) yields:

$$\rho\left(\Delta^*_{k,t}, \Delta^*_{k,t-1}\right) = \frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}} w_{i,t}\Delta_{i,t} - \overline{\Delta^*_{k,t}}\right)\left(\sum_{i=1}^{I_{k,t-1}} w_{i,t-1}\Delta_{i,t-1} - \overline{\Delta^*_{k,t-1}}\right). \tag{A10}$$

Because the weights sum to one, Eq. (A10) can be written:

$$\rho\left(\Delta^*_{k,t}, \Delta^*_{k,t-1}\right) = \frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}} w_{i,t}\left(\Delta_{i,t} - \overline{\Delta^*_{k,t}}\right)\right)\left(\sum_{i=1}^{I_{k,t-1}} w_{i,t-1}\left(\Delta_{i,t-1} - \overline{\Delta^*_{k,t-1}}\right)\right). \tag{A11}$$

Substituting Eq. (A6) into Eq. (A11) yields:

$$\rho\left(\Delta^*_{k,t}, \Delta^*_{k,t-1}\right) = \frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}} w_{i,t}\left(\sum_{n=1}^{N_{i,t}} \frac{D_{n,i,t}}{N_{i,t}} - \overline{\Delta^*_{k,t}}\right)\right)\left(\sum_{i=1}^{I_{k,t-1}} w_{i,t-1}\left(\sum_{n=1}^{N_{i,t-1}} \frac{D_{n,i,t-1}}{N_{i,t-1}} - \overline{\Delta^*_{k,t-1}}\right)\right). \tag{A12}$$

Which can be written:

$$\rho\left(\Delta^*_{k,t}, \Delta^*_{k,t-1}\right) = \frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}} w_{i,t}\left(\sum_{n=1}^{N_{i,t}} \frac{D_{n,i,t} - \overline{\Delta^*_{k,t}}}{N_{i,t}}\right)\right)\left(\sum_{i=1}^{I_{k,t-1}} w_{i,t-1}\left(\sum_{n=1}^{N_{i,t-1}} \frac{D_{n,i,t-1} - \overline{\Delta^*_{k,t-1}}}{N_{i,t-1}}\right)\right). \tag{A13}$$

Eq. (A13) can be partitioned into those terms that represent trading in the same security this quarter and last (i.e., institutional trading in security i in both quarter t and quarter t-1) and trading in different securities in the same industry [i.e., institutional trading in security i in quarter t and security j ($i,j \in k$) in quarter t-1]:

$$\rho\left(\Delta^*_{k,t},\Delta^*_{k,t-1}\right)=\frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}}\left(w_{i,t}\left(\sum_{n=1}^{N_{i,t}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\right)\bullet w_{i,t-1}\left(\sum_{n=1}^{N_{i,t-1}}\frac{D_{n,i,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{i,t-1}}\right)\right)\right)+$$

$$\frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}\left(\sum_{n=1}^{N_{i,t}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\right)\bullet w_{j,t-1}\left(\sum_{n=1}^{N_{j,t-1}}\frac{D_{n,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right).\text{(A14)}$$

Each term in Eq. (A14) can be further partitioned into investors following their own lag trades (i.e., investor n at time t and t-1) and following other investors' lag trades (i.e., investor n at time t and investor m at time t-1) yielding the general form of Eq. (7):

$$\rho\left(\Delta^*_{k,t},\Delta^*_{k,t-1}\right)=\frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}}\left(w_{i,t}w_{i,t-1}\left(\sum_{n=1}^{N_{i,t}}\left(\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{n,i,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{i,t-1}}\right)\right)\right)\right)+$$

$$\frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}}\left(w_{i,t}w_{i,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{i,t-1}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{m,i,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{i,t-1}}\right)\right)\right)+$$

$$\frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{n,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right)+$$

$$\frac{1}{C^*_t}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right). \quad \text{(A15)}$$

The last term in (A15) represents institutions following other institutions into different stocks in the same industry. This term can be further partitioned into managers following other managers into same size-BE/ME style stocks ($i,j \in k$, $i,j \in s$) and into different style stocks in the same industry ($i,j \in k$, $i \in s, j \notin s$) yielding Eq. (13):

$$\frac{1}{C_t^*}\sum_{k=1}^{K} w_k\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta_{k,t}^*}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta_{k,t-1}^*}}{N_{j,t-1}}\right)\right)\right)=$$

$$\frac{1}{C_t^*}w_k\sum_{k=1}^{K}\left(\sum_{i=1 i\in s}^{I_{k,t}}\sum_{j=1,j\neq i,j\in s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta_{k,t}^*}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta_{k,t-1}^*}}{N_{j,t-1}}\right)\right)\right)+$$

$$\frac{1}{C_t^*}w_k\sum_{k=1}^{K}\left(\sum_{i=1,i\in s}^{I_{k,t}}\sum_{j=1,j\neq i,j\notin s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta_{k,t}^*}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta_{k,t-1}^*}}{N_{j,t-1}}\right)\right)\right).\ \text{(A16)}$$

C. Expected contributions from same- and different-style stocks

The last term in Eq. (7) [or Eq. (A15)] represents institutional investors following other institutions into and out of different stocks in the same industry (i.e., industry herding):

$$\frac{1}{(K)\sigma(\Delta_{k,t}^*)\sigma(\Delta_{k,t-1}^*)}\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta_{k,t}^*}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta_{k,t-1}^*}}{N_{j,t-1}}\right)\right)\right).\quad\text{(A17)}$$

Rearranging terms yields:

$$\frac{1}{(K)\sigma(\Delta_{k,t}^*)\sigma(\Delta_{k,t-1}^*)}\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\left(D_{n,i,t}-\overline{\Delta_{k,t}^*}\right)\left(D_{m,j,t-1}-\overline{\Delta_{k,t-1}^*}\right)\right)\right)\right)$$
.(A18)

If manager *n* follows manager *m* into (or out of) a different stock in the same industry then the product of the last two terms $\left(i.e., \left(D_{n,i,t}-\overline{\Delta_{k,t}^*}\right)\left(D_{m,j,t-1}-\overline{\Delta_{k,t-1}^*}\right)\right)$ is positive. Conversely, if manager *n* trades in the opposite direction of manager *m* (e.g., manager *n* purchases security *i* following manager *m*'s sale of security *j*), the last term is negative. Under the null hypothesis that managers are as likely to follow each other into and out of same style stocks as different style stocks in the same industry, the expected value of the product is the same regardless of whether stocks i and j are in the same size-BE/ME style ($i,j \in k$, $i,j \in s$) or in different size-BE/ME styles ($i,j \in k$, $i \in s$, $j \notin s$). As a result, the expected contribution of same- and different size-BE/ME style herding (under the null) is determined by the remaining terms in Eq. (A18). Specifically, the expected proportion of the herding contribution [i.e., the last term in Eq. (7)] attributed to same style stocks is given by the ratio of the expected contribution from same style terms ($i,j \in s$) to the expected contribution from all (i.e., same style and different style) terms:

$$\frac{\dfrac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})}\sum_{k=1}^{K}\left(\sum_{i=1,i\in s}^{I_{k,t}}\sum_{j=1,j\neq i,j\in s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}{\dfrac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})}\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}. \quad \text{(A19)}$$

Cancelling the first term yields:

$$Expected\ proportion\ attributed\ to\ same\ style\ stocks = \frac{\sum_{k=1}^{K}\left(\sum_{i=1,i\in s}^{I_{k,t}}\sum_{j=1,j\neq i,j\in s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}{\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}. \text{(A20)}$$

Analogously, the expected proportion of the herding contribution attributed to following other managers into and out of different style stocks is given by the ratio of the expected contribution from different style terms (i.e., $i \in s$, $j \notin s$) to the expected contribution from all (i.e., same style and different style) terms:

$$Expected\ proportion\ attributed\ to\ different\ style\ stocks = \frac{\sum_{k=1}^{K}\left(\sum_{i=1,i\in s}^{I_{k,t}}\sum_{j=1,j\neq i,j\notin s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}{\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}. \text{(A21)}$$

The last term in Eq. (7) (i.e., the contribution to the correlation attributed to institutions following other institutions into different stocks in the same industry) times Eq. (A20) yields the expected contribution (under the null hypothesis) to the correlation attributed to institutions following other institutions into and out of same size-BE/ME style stocks in the same industry:

$$Expected\ \ proportion\ \ of\ \ correlation\ \ attributed\ \ to\ herding\ \ in\ same\ style\ stocks_t =$$

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})}\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,\,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right)*$$

$$\frac{\sum_{k=1}^{K}\left(\sum_{i=1,i\in s}^{I_{k,t}}\sum_{j=1,\,j\neq i,\,j\in s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}{\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,\,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}. \quad \text{(A22)}$$

Similarly, the last term in Eq. (7) times Eq. (A21) yields the expected contribution (under the null hypothesis) to the correlation attributed to institutions following other institutions into and out of different size-BE/ME style stocks in the same industry:

$$Expected\ \ proportion\ \ of\ \ correlation\ attributed\ \ to\ herding\ in\ different\ \ style\ stocks_t =$$

$$\frac{1}{(K)\sigma(\Delta^*_{k,t})\sigma(\Delta^*_{k,t-1})}\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,\,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{D_{n,i,t}-\overline{\Delta^*_{k,t}}}{N_{i,t}}\bullet\frac{D_{m,j,t-1}-\overline{\Delta^*_{k,t-1}}}{N_{j,t-1}}\right)\right)\right)*$$

$$\frac{\sum_{k=1}^{K}\left(\sum_{i=1,i\in s}^{I_{k,t}}\sum_{j=1,\,j\neq i,\,j\notin s}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}{\sum_{k=1}^{K}\left(\sum_{i=1}^{I_{k,t}}\sum_{j=1,\,j\neq i}^{I_{k,t-1}}\left(w_{i,t}w_{j,t-1}\left(\sum_{n=1}^{N_{i,t}}\sum_{m=1,m\neq n}^{N_{j,t-1}}\frac{1}{N_{i,t}}\frac{1}{N_{j,t-1}}\right)\right)\right)}. \quad \text{(A23)}$$

CHAPTER THREE: FINANCIAL STATEMENT ANALYSIS, FUTURE STOCK RETURNS AND DEMAND BY INSTITUTIONAL AND INDIVIDUAL INVESTORS

1. Introduction

In a set of clever studies, Piotroski (2000, 2005) demonstrates that a set of nine (collectively denoted *f-score*) simple indicator variables (e.g., an increase in return on assets) garnered from financial statements can successfully identify future 'winners' and 'losers.' Piotroski argues that these return patterns arise because market participants underreact to information contained in financial statements. As a result, smart investors can garner abnormal returns by exploiting the subsequent revision (and related price corrections) of the market's biased expectations. Moreover, Piotroski argues that return patterns across value stocks (i.e., *f-score* successfully predicts which value stocks will outperform) and glamour stocks (i.e., *f-score* successfully predicts which growth stocks will underperform) supports the argument that the value premium arises from investor overreaction rather than compensation for fundamental risk.

Fama and French (2006) confirm that Piotroski's (2000, 2005) *f-score* forecasts future stock returns. They note, however, that under clean surplus accounting (i.e., changes in book value reflect earnings less dividend payments) the market equity to book equity ratio can be written:

$$\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau})/(1+r)^{\tau}}{B_t} \qquad (1)$$

where M is the market value of the firm's equity, B is the book equity value of the firm, r is the required rate of return, and Y is the firm's income. Fama and French point out that controlling for M_t/B_t and changes in book value of equity, measured relative to current book equity ($dB_{t+\tau}/B_t$), more profitable firms ($Y_{t+\tau}/B_t$) have higher expected returns (r). That is, holding changes in book values constant, a riskier firm (with a higher required rate of return) will have to have higher income to generate the same *B*/*M* ratio. Fama and French argue that *f-score* proxies for expected income and therefore is expected to positively vary with future stock returns. As a result, the relation between *f-score* and future returns is consistent with both rational asset pricing and Piotroski's irrational asset pricing explanation.

This paper investigates the relation between changes in financial health (*f-score*), subsequent returns, and demand by institutional investors to differentiate between the rational and irrational pricing explanation for why *f-score* predicts returns. The key difference in the rational and irrational pricing explanations is *investor behavior*. Under the irrational pricing explanation, investors are surprised when high *f-score* companies (especially value stocks with high *f-scores*) do well and low *f-score* companies (especially growth stocks with low *f-scores*) do poorly. As a result, subsequent shifts in investor demand drive high *f-score* company's prices higher and low *f-score* company's prices lower. In contrast, under the rational pricing explanation, expectations are realized (on average) and expectations are not systematically revised. As a result, there should be no systematic shift in investor demand.

Because there is a buyer for every seller, demand by 'market participants' equals supply. I overcome this issue by focusing on net demand by institutional (rather than individual) investors. Specifically, evidence suggests institutional investors are the marginal investors who set prices (e.g., Chakravarty, 2001; Froot and Teo, 2004; Sias, Starks and Titman, 2006; Kaniel, Saar, and Titman, 2008; Campbell, Ramadorai, and Schwartz, 2007). Moreover, because evidence suggests institutional investors are more sophisticated than retail investors (e.g., Hribar, Jenkins, and Wang, 2004; Bartov, Radhakrishnan, and Krinsky, 2000; Collins, Gong, and Hribar, 2003; Amihud and Li, 2002; Ke and Petroni, 2004), I expect that if *f-score* predicts returns as a result of biased expectations, institutional investors will be more likely than individual investors to exploit the information.

In sum, I propose that if rational asset pricing fully drives the relation, then there should be no relation between f-score and subsequent institutional demand. Alternatively, if investors' slow reaction to information contributes to the relation between f-score and the future returns, then f-score should forecast institutional demand as well as future returns. Consistent with the irrational pricing explanation, I find that *f-score* predicts future demand by institutional investors.

The balance of the paper is organized as follow. In Section 2, I introduce related literature and Section 3 explains data and the financial statement analysis used throughout the study. Section 4 presents results for the main tests and Section 5 concludes.

2. **Literature review**

2.1. Under-reaction and use of financial statement analysis

Hong and Stein (1999) propose a theory documenting underreaction and overreaction of market participants. The authors posit that the slow diffusion of information leads to underreaction in the short run, but it is likely to lead to overreaction in the long run. Barberis, Shleifer and Vishny (1998) construct a model to demonstrate that the market underreacts to continuation of good news or bad news and overreact to single significant news. Daniel, Hirshleifer, and Subrahmanyam (1998) argue overconfident investors overweight private signals while underweighting public signals. These theoretical works show that, on realizing the existence of mispricing, market participants adjust their expectations slowly leading to a reversal in returns (growth stocks' return declining and value stocks' return increasing).

Empirical works demonstrate the market underreacts to corporate news events, resulting in post-event return drift over long horizons. Ritter (1991) documents market's slow reaction to initial public offerings, while Loughran, and Ritter (1995) and Spiess and Affleck-Graves (1995) find similar results using seasoned equity offerings. Ikenberry, Lakonishok, and Vermaelen (1994) report the market underreacts to open market share repurchases. The authors argue information embedded in repurchases announcement is neglected at the initial stage and the market is slow to react to news, generating four year abnormal returns of 12.1% after the announcement. Ikenberry and Ramnath (2002) examine the case of stock splits as an example of "self-selected" corporate news event. The authors confirm previous findings (cf. Ikenberry, Rankine, and Stice (1996) and Desai and Jain (1997)) by reporting positive return drift after split announcements. It is the general agreement in this area of studies that abnormal long-run return patterns are generated by the market slowly reacting to a corporate event correlated with changes in firm's fundamentals such as future operating performance.

A number of studies also show that financial statement analyses possess predictive power for subsequent returns because market participants underreact to information conveyed in various measures of a firm's economic condition. Among the earliest works exploiting the predictive power of a financial statement analysis are Ou and Penman (1989) and Holthausen and Larcker (1992). In an attempt to find the metrics that are easier to calculate and implement, Lev and Thiagarajan (1993) identify 12 fundamental variables analysts described as useful and related to quality of earnings. Abarbanell and Bushee (1997, 1998) verify these 12 signals are predictive of subsequent earnings change and can be used to predict future returns.

Several recent studies employ financial statement based analyses to further investigate the cause of the well-known book-to-market effect. Piotroski (2000, 2005) shows simple accounting based measures can predict future return patterns among the broader population of stocks (Piotroski 2005), as well as value stocks (Piotroski 2000). The author argues the heuristic representing changes in financial health, called *f-score*, is successful at recognizing future return winners among high book-to-market (B/M) ratio stocks and losers among low B/M ratio stocks because these are the stocks in which mispricing is more dominant than in their counterparts. Piotroski claims investors underreact to recent financial improvement of value stocks and worsening financial condition of glamour stocks (denoted as "contrarian firms") because they are the groups of stocks with historical expectations implied by book-to-market ratio, contrary to expected future outcome implied by *f-score*. Additionally, Mohanram (2005) builds an index denoted as GSCORE to proxy for signals based on the three categories, growth, conservatism, and naïve extrapolation, in an attempt to separate winners and losers among low book-to-market firms. The author finds growth firms with strong growth aspect outperform growth firms with weak future growth potential. This study rules out the risk based explanation as a cause of the book-to-market effect by proving high risk stocks earn lower returns. Griffin and Lemmon (2002) adapt a metric proposed by Ohlson (1980), named O-score to proxy for a firm's bankruptcy risk. The authors show high level of bankruptcy risk measured by O-score leads to the poor future returns. Specifically, the authors demonstrate growth firms with higher distress risk earn lower return. All these studies agree on the source of the book-to-market anomaly; that is, investors over-extrapolate past performance and mispricing arises. These studies attribute the success of the suggested investment strategies to the fact that subsets of stocks are mispriced and the financial statement based analyses can help identify mispriced securities.

2.2. Value-Growth Effect

There are two explanations for the well-documented value premium: (1) value stocks are fundamentally riskier or (2) markets over-extrapolate past performance and overvalue growth stocks and undervalue value stocks (mispricing). Lakonishok, Shleifer, and Vishny (1994) argue investors are overly optimistic about the low book-to-market stocks, or growth stocks, based on the stock's promising past growth rate. Investors then believe the trend will continue and buy growth stocks excessively, causing growth stocks to be overvalued. For the same reason, value stocks, or high book-to-market ratio stocks, are underpriced in the market because investors ignore the stocks which have performed poorly in the past. Return patterns of the two groups of stocks eventually reverse when the market realizes the true valuation, resulting in value premium. Joining the argument, LaPorta (1996) and LaPorta, Lakonishok, Shleifer, and Vishny (1997) report positive returns for value stocks (or stocks with low expected growth rates) and negative returns for glamour stocks (or high expected growth rate stocks), following earnings announcement. These results show investors' over-extrapolation of past information reverses at some point, making contrarian strategies (buying past losers and selling past winners) profitable in the long run.

Competing argument to the overextrapolation is a risk based explanation. Fama and French (1992, 1995) argue value stocks have higher risk of financial distress, and therefore require higher returns. Also Chen and Zhang (1998) demonstrate high book-to-market stocks have high leverage, a measure of a firm's fundamental risk. As a whole, the related works in risk based explanation contend investors take high risk with high book-to-market stocks, and as a result are compensated with high returns. Fama and French (2006) incorporate profitability and investment effect to further examine higher returns on the value portfolio. In line with Haugen and Baker (1996), the authors argue higher profitability is linked to high expected returns, after controlling for the book-to-market effect and expected investment. Also consistent with Fairfield, Whisenant, and Yohn (2003), higher investment rates lead to lower expected returns, after controlling for the book-to-market effect and the profitability.

A growing body of literature documents the role of institutional investors in this anomaly. Ali, Hwang, and Trombley (2003) find the book-to-market effect is more evident in a setting with the greater arbitrage risk. As a proxy for the arbitrage risk, the authors use the idiosyncratic return volatility, transaction costs, and investor sophistication. The study uses the level of institutional ownership as a measure for the transaction cost and finds a negative relation between institutional ownership and the book-to-market effect. Nagel (2005) documents underperformance of growth stocks are intensified in the stocks with low institutional ownership level. Phalippou (2007) proposes that individual investors, not institutions, drive value premium by showing stocks held by institutional investors do not have significant value premium. The author explains the reason as mispricing and a lack of arbitrage.

Although the studies mentioned above use the level of institutional ownership to examine the role of institutional investors in value/growth effects, Jiang (2007) uses change in institutional investor holdings to argue that institutional investors are responsible for driving the mispricing effect. Specifically, the author utilizes the intangible return concept developed in Daniel and Titman (2006) and proposes institutions herd to positive intangible returns and out of negative intangible returns. The author furthermore shows that the book-to-market effect is greater in high institutional herding stocks than is in low institutional herding stocks.

These findings are important in the context of this study as I explore if the information embedded in the financial-statement-based metrics influence institutional versus individual demand and supply for securities, which could in turn explain the return mechanism generating the value premium. That is, this theory is typically framed as the value premium arising from “investors” slowly updating their priors on value stocks with improving fundamentals and growth stocks with declining fundamentals. As noted above if institutional and individual investors update (and act on) their priors at different speeds then financial statement analysis may predict returns because financial statement analysis may predict institutional versus individual investor demand.

3. **Data**

3.1. Institutional ownership data

This study uses quarterly institutional investor holdings data for the period between March 1983 and December 2006, garnered from 13F reports and purchased from Thomson Financial. Institutions with $100 million or more under management are required to disclose their equity holding of 10,000 shares or $200,000 in value to Security and Exchange Commission within 45 days of the end of each calendar quarter.

Five different measures are calculated to proxy for the demand of institutional investors. Net institutional demand (NID) is the net change in fractional ownership of institutional investors in stock *i* over period *t*:

$$\text{Net institutional demand}_{i,\,t} = \frac{\#\text{ of shares held by institutions}_{i,\,t}}{\#\text{ shares outstanding}_{i,\,t}} - \frac{\#\text{ of shares held by institutions}_{i,\,t-1}}{\#\text{ shares outstanding}_{i,\,t}} \quad (2)$$

where number of shares are split-adjusted.

Numerous studies (e.g., Nofsinger and Sias, 1999; Wermers 1999; Grinblatt, Titman, and Wermers, 1995) find a relation between the demand of institutional investors and the lag returns and the size of a firm plays a critical role. For instance, as Sias (2007) points out, it is more common for a larger stock to go from fractional change in institutional ownership of 50% to 60% than for a smaller stock to go from fractional change of 20% to 40%, in both of which cases NID is 10%. At the same time, it is more common for a larger stock to go from fractional ownership of 50% to 30% than for a smaller stock to go from 30% to 10%, in both of which cases NID is -10%. As a result, extreme net institutional ownership measures (either positive or negative) are likely to be dominated by large capitalization stocks. Therefore, it is important to account for the size of a firm in measuring the demand of institutional investors. I calculate Adjusted NID by subtracting the average NID for the stocks at the same capitalization at the same time.

$$\text{Adjusted net institutional demand}_{i,\,t} = \text{NID}_{i,\,t} - \overline{\text{NID}_{c,\,t}} \tag{3}$$

where NID is net institutional demand, defined in Eq. (2), and $\overline{\text{NID}_{c,t}}$ is average net institutional demand of the firms at the same capitalization decile at the beginning of quarter t. It measures an abnormal fractional change of institutional demand by alleviating the firm size effect and allows for comparison across different capitalization stocks. Another measure of relative institutional demand is the percentage net institutional demand, and is defined as:

$$\text{Percentage net institutional demand}_{i,\,t} = \frac{\text{NID}_{i,\,t}}{\overline{\text{NID}_{c,\,0}}} \tag{4}$$

where $\overline{\text{NID}_{c,\,0}}$ is average institutional demand of the firms at the same capitalization decile at time 0. Net institutional demand (NID) for the firms at the same capitalization, instead of NID of that specific firm, at time 0 is used to scale a firm's NID at time *t* because some firms have very small (or often times 0) NID at time 0. Adjusted percentage net institutional demand is calculated as:

$$\text{Adjusted percentage net institutional demand}_{i,\,t} = \text{P\_NID}_{i,\,t} - \overline{\text{P\_NID}_{c,\,t}} \tag{5}$$

where $P\_NID_{i,t}$ is percentage net institutional demand for a firm *i* at time *t* and $\overline{P\_NID_{c,t}}$ is average percentage net institutional demand for the firms at the same capitalization declie at the same time.

Because previous work (e.g., Sias, Starks, and Titman, 2006) demonstrates that returns are most strongly related to other measures of institutional demand than net institutional demand, I use the ratio of the number of institutional buyers to number of institutional traders as the other metric to measure institutional investor's trading. I define institutional investors as buyers if the institutions increase fractional ownership in stock *i* over holding period *t* and seller if they decrease their fractional ownership, where fractional ownership is defined as the number of shares owned by an institution divided by number of shares outstanding for each stock. The number of shares outstanding and number of shares held by institutions are adjusted for stock split. Following Sias (2004), buyratio is defined as:

$$\text{Buyratio}_{i,t} = \frac{\#\text{ institutions buying}_{i,t}}{\#\text{ institutions trading}_{i,t}} \tag{6}$$

3.2. **Compustat/CRSP data**

Stock prices and return data are from Center for Research in Security Prices (CRSP) monthly data and accounting related variables are extracted from annual Compustat database. The sample includes only ordinary shares (i.e. securities with CRSP share codes 10 or 11). At the end of each fiscal year ending, I calculate the book value of equity as the book value of total assets (Compustat item #6) minus liabilities (Compustat item #181) plus balance sheet deferred taxes and investment tax credit (Compustat item #32), if available, minus the book value of preferred stocks (liquidating value (Compustat item #10), redemption value (Compustat item #56) or carrying value (Compustat Item #130), in order of availability). The book-to-market (B/M) ratio is the book value of equity divided by the market value of equity at the end of each fiscal year ending. Firms with the negative book-to-market ratio and financial companies are excluded. Financial statement based metric, *f-score*, is also calculated at the end of each fiscal year ending.

3.2.1. Financial statement analysis-based signal: *f-score*

For the financial statement analysis of a firm, I focus on Piotroski (2000, 2005)'s *f-score*. *F-score* is an aggregate measure for a firm's financial health based on nine financial performance signals from the three areas: profitability, financial leverage/liquidity, and the operating efficiency. Each binary variable takes a value of one if the signal implies good financial performance and zero otherwise and *f-score* is the sum of nine binary variables listed at the next three subsections.

3.2.1.1. Components of *f-score* representing profitability

Piotroski (2000, 2005) uses four ratios to measure how well a firm generates profit to fund its operation. ROA is net income before extraordinary items (Compustat item #18) divided by total assets at the beginning of each year. A binary variable representing ROA takes a value of one if ROA is positive, and zero otherwise. Difference between ROA this year and last (dROA) is also used to gauge trend in a firm's profitability. Positive trend in return shows future earnings for a firm are promising, sending a "good" signal for a firm's profitability. The indicator variable corresponding to a change in ROA is assigned one if the change is positive and zero otherwise.

The binary variable for cash flow from operations (CFO) equals one if a firm's CFO is positive and zero otherwise[28]. ACCRUAL, calculated as income before extraordinary items minus CFO, is included to account for the quality of a firm's earnings. Ohlson (1999) and Barth, Beaver, Hand and Landsman (1999) report accruals have different predictive power from cash flow component of earnings. Also, Sloan (1995) points out accruals, or noncash portion of earnings, are less likely to persist than cash flow portion, implying positive accrual is a negative signal for a firm's future performance. The corresponding indicator variable equals one if ACCRUAL is negative, or CFO is greater than net income, and zero otherwise.

3.2.1.2. Components of *f-score* representing the leverage/liquidity

Piotroski (2000, 2005) uses the ratio of current asset (Compustat item # 4) to current liabilities (Compustat item #5) to incorporate into the aggregate measure a firm's ability to meet its short-term debt obligation. As Piotroski points out, a high value of current ratios can also represent an insufficient use of short term assets for some types of businesses. However, overall, a high ratio is viewed as positive signs for a firm's financial health, adding value to the aggregate measure. Change in the ratio (dLQ) is used to capture the improvement of liquidity and a dummy variable for liquidity measure is assigned one if the ratio is improved from the last term, or the difference is positive.

[28] A method to calculate cash flow from operation (CFO) depends on whether a firm files the statement of cash flows or statement of working capital. If the company reports statement of cash flow, CFO is the net cash flow from operating activities (Compustat item # 308). If the company files the statement of working capital, CFO is calculated as funds from operations minus other changes in working capital (Compustat Item #236). Funds from operation is the sum of the earnings before income and taxes (EBIT, Compustat Item #18), deferred taxes (Compustat item #50) and equity's share of depreciation expense, where equity's share of depreciation expense is defined as depreciation expenses × {market capitalization/ (total assets – book value of equity + market capitalization)}.In all other cases, CFO is funds from operations plus other changes in working capital. If CFO is positive for a firm, the binary variable takes a value of one.

Interpretation of leverage measures is also twofold. The higher the leverage of a firm is, the more a firm has a downward risk. As Harris and Raviv (1990) and Jensen and Meckling (1976) show, however, debt can be used to monitor management, reducing the agency cost. Piotroski (2000, 2005) considers use of debt as a bad signal in a firm's financial situation and uses two measures to represent a firm's leverage. Change in the leverage ratio (long term debt (Compustat item #9 plus #44) divided by total assets at year end) over the year is employed to capture the level of a firm's external financing. Since a decrease in the leverage ratio is a positive sign to a firm's financial health, the binary variable takes the value of one if the change in the leverage ratio (dLEVER) is negative.

Not only is the use of debt a signal against a firm's financial health, but a new issuance of equity can also be considered as demonstrating that a firm needs additional external financing. If sales of common equity and preferred stock (Compustat item #108) from a firm's statement of cash flow are positive, the indicator value equals zero and one if the company does not issue any new common stocks and preferred stocks over a year.

3.2.1.3. Components of *f-score* representing the operating efficiency

Gross margin ratio and asset turnover ratio are used to gauge how efficient a firm operates. Gross margin is calculated as 1-(cost of goods sold (Compustat item #41) / sales (Compustat Item #12)). An increase in gross margin indicates a firm's better control over its production cost and inventory management, and/or an increase in sales price, therefore giving a positive signal for a firm's financial condition. The binary variable equals one if the change in gross margin ratio (dGM) from last year to this year is positive, and zero otherwise.

Asset turnover ratio is defined as sales divided by average total assets and represents a firm's efficiency at utilizing assets to generate sales. Improvement, or a positive change, in asset turnover ratio shows the company's productivity level has been increased over the respective period and sends a good signal regarding a firm's financial condition. The indicator variable takes the value of one if the change in turnover ratio from last year to this is positive, and zero otherwise.

3.2.1.4. Aggregating nine binary variables to compute *f-score*

To calculate the final signal to proxy for an overall change in firm's financial health, Piotroski (2000, 2005) adds all nine binary variables demonstrated at the last three sections. A designated binary variable is equal to one if a signal from the area it represents indicates improvement and zero if the signal demonstrates deterioration of a firm's financial condition. *F-score* ranges from zero to nine, with zero corresponding to the firms with the greatest deal of deterioration in their financial condition among the sample and nine to the firms with the biggest improvement on their financial health.

Piotroski (2000, 2005) argues nine variables used to construct *f-score* are not chosen to represent the optimal measures for the overall progress or weakening of a target firm's financial condition. Piotroski (2005) stresses that "this approach (*f-score*) represents a "step-back" to a simple, firm-specific analysis using absolute benchmarks to classify trends in financial condition … However, despite appearing "ad hoc", these ratios are intuitive, easy-to-construct and commonly used in financial statement analysis" (p.15). This study takes the author's view that the purpose of *f-score* method is not to be exclusive sets of measures, but to present one of various sets of statistics to gauge an overall change in a firm's financial health, with ease of implementation and interpretation. I extensively use the metric throughout the study to represent the development on a financial condition and predict the future performance of a firm.

4. **Replicating Piotroski (2000, 2005)'s results**

4.1. Replicating Piotroski (2000)

In this section, I attempt to replicate the analysis in Piotroski (2000) to ensure the financial statement analysis presented as *f-score* does in fact forecast future returns in high book-to-market stock portfolios. Piotroski (2000) shows the metric constructed using the financial statement entries can predict the future returns among the high book to market stocks. The author claims the high book-to-market stocks provide a good environment for testing accounting based heuristics because other pieces of information, such as analyst recommendation and voluntarily disclosure, are often not available or not reliable for the high book-to-market or "financially distressed" firms.

The author finds separating strong high book-to-market stocks from the weak ones generate positive abnormal returns and attributes the result to the market's inefficiency of incorporating the recent information into the price. The high book-to-market firms, or the value firms, with the strong recent improvement on their financial situation generate positive abnormal returns because the market is surprised when those firms perform well, unlike the expectation of the market participants. The author argues the result is inconsistent with Fama and French (1992)'s risk based explanation to the phenomenon because in this study the healthier firms with high scores in financial statement based metric generate higher returns. Instead, the author concludes the success in the strategy of buying financially strong value stocks come from the market's initial underreaction to the historic information. I follow Piotroski (2000) to see if the strategy of "separating winners from the losers" in high book-to-market portfolios can be repeated.

4.1.1. Univariate analysis

First to get a glimpse at the return scheme, I present the returns for the stocks with the strong financial health and the stocks with the weak (or deteriorating) financial condition. As a proxy for the change in financial condition (improvement or worsening), I follow Piotroski (2000) and use *f-score*, as explained in the previous sections. The firms with *f-score* of 4 or greater is categorized as strong *f-score* firms, or the firms with the positive improvement in their financial health and the firms with the score less than 4 are labeled weak financial condition firms. Table 1 presents the returns for the returns for the strong *f-score* firms and weak *f-score* firms, the difference between the groups for each year within the sample (from 1976 to 1996).

[Table 1 about here]

The annual market adjusted returns computed from 5th month after the portfolio formation are used. The table clearly shows the firms with strong financial conditions garner higher returns for the every year in the sample (from the year 1976 to 1996) except for the four years. This test gives a good idea for the predictive power of the accounting based heuristic for the future returns. To test this predictability further, I calculate mean, median, and the various percentiles of the annual raw, and market adjusted returns for each *f-score* portfolio and two *f-score* (high and low) groups.

[Table 2 about here]

Table 2 presents the average of raw and market adjusted returns for each *f-score* portfolio, and high and low *f-score* portfolio, as well as the entire sample (presented at the top row). The average return for the stocks at the highest book-to-market quintile from 1976 to 1996 is 23.99% and the median is 12.32%. The returns by *f-score* demonstrate the same pattern as shown in the previous section. Mean, Median, and 10^{th}, 25^{th}, 75^{th} and 90^{th} percentile returns increase monotonically as the firms' financial situation signals improve. The results also reveal the strategy of buying the stocks with *f-score* higher than 6 (High group) and selling the stocks with *f-score* less than 4 (Low group) would generate an average raw return of 22.51%. Examining the returns difference between two groups for the mean (difference in means test) and median (Wilcoxon rank test) confirms the difference in returns between the groups with highest improvement in the firms' financial condition and the groups with the most deteriorating financial situation is significant.

The test with the market adjusted returns demonstrates similar results. Average market adjusted return for the entire sample is 5.4% with median return of -5.14%. As is with the raw returns, mean, median, and the various percentile market adjusted returns by each *f-score* portfolio show the increasing patterns and the difference between the high *f-score* and low *f-score* groups are significantly different (difference in mean test statistic is 4.81 and Wilcoxon rank Z statistic is 5.94). All these results confirm the predictability of the financial statement based metric.

4.1.2. Regression analysis

Piotroski (2000) suggests a few variables that might have the correlation with the accounting based signal and/or the future returns. The author states the underlying motivation of the momentum effect is the same as the underreaction to the historical information on a firm's financial situation. He also cites Sloan (1996) and Loughran and Ritter (1995) as the evidences of the level of accrual and the recent equity offering, respectively, having predictive powers for the future returns. I run the following regression model of annual raw and market adjusted returns on the explanatory variables mentioned in Piotroski:

$$RET_{t+5,t+16} = \alpha + \beta_{SZ}\log(SZ) + \beta_{BM}\log(BM) + \beta_{mom}\text{MOMRET} + \beta_{EQ}\text{EQOFF} + \beta_{acc}\text{ACCRUAL} + \beta_{FS} fscore + \varepsilon \ . \quad (7)$$

Twelve month buy and hold raw and market adjusted return are measured starting at the 5th month after the accounting based signal is computed. Log (SZ) is the log value of a firm's market capitalization, and log (BM) is the log value of the book-to-market ratio of the firm, measured at the end of the previous fiscal year. MOMRET is a 6 month holding return prior to the portfolio formation period and EQOFF is a binary variable which takes value of 1 if a firm issued a new equity in the respective fiscal year and 0 otherwise. ACCRUAL is net income minus cash flow from operation, scaled by total assets at the beginning of the fiscal year. Firm's book-to-market categories are determined based on the previous year's book-to-market ratios and highest quintile book-to-market firms (high book-to-market firms) are retained for the test.

[Table 3 about here]

When the market adjusted return is regressed on the primary explanatory variables (size, and the book-to-market ratio), the pooled regression result reveals the financial statement based signal is strongly positively related to future market adjusted returns. Increase in one unit of *f-score* would result in the increase of the market adjusted return by 2.62% on average. When other possible explanatory variables (momentum, equity offerings, and accruals) are added to the model, the significance of *f-score* remains strong at the significance level of 1% (*t*-statistic of 5.42). Average coefficients from 21 annual regressions show similar results. The predictability of *f-score* stays significant both when the size and the book-to-market factor are controlled for and the other additional explanatory variables (momentum, equity offering, and accruals) are included (with *t*-statistics of 5.89 and 2.82, respectively).

I confirm with univiriate and regression analysis Piotroski (2000)'s findings that the metric derived from simple nine accounting-related variables can predict the future returns among high book-to-market stocks. The results stay strong after possible variables that may affect the future returns are controlled for.

4.2. Replicating Piotroski (2005)

4.2.1. Univariate analysis

In this section, I attempt to repeat Piotroski (2005) to confirm financial statement analysis is predictive of future returns not only in value stock portfolio, but in the entire sample as well. Piotroski (2005) reports a signal constructed using nine financial statement related variables has a power to predict subsequent returns. The author calculates one year buy and hold returns starting the 5^{th} month after the signal (*f-score*) is calculated and shows one year raw, and market adjusted buy and hold returns increase monotonically as *f-score* increases. I first closely follow this method to see whether the monotonic pattern on the returns can be regenerated. *f-score* and the book-to-market ratio for each firm is calculated as explained at the Section 3.2, at the end of each firm's fiscal year using annual financial statement data and updated every year. I follow Piotroski (2005) for computing the returns; I start return compounding at the 5^{th} month after the firm's fiscal year ends. Market adjusted return is a raw return minus one year buy-hold CRSP value weighted market index return over the same period. Final sample consists of 100,778 firm-years with adequate returns and accounting data from 1972 to 2001. Table 4 presents the results for return analyses.

[Table 4 about here]

Consistent with Piotroski (2000, 2005) and Fama and French (2006), raw and market adjusted returns show monotonic patterns. In the case of raw returns, higher *f-scores* represent higher one year future return in average and percentiles presented (10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles). When *f-scores* are categorized into three groups, Low (*f-score*=<3), Median (4=<*f-score*=<6) and High (*f-score*>=7), mean and percentile returns increase as *f-score* moves to a higher group. Differences in returns between high and low *f-score*, presented at the last row, confirm there is a statistically significant difference in returns between high and low *f-score* groups. Average raw return for High *f-score* group is 20.02% and for Low group, it is 8.34%, with a difference statistically significant at 1% level. The only exception for the monotonic pattern in returns is at the 90^{th} percentiles, possibly due to outliers at this category not behaving as other firms do in terms of returns and other characteristics. Market adjusted returns present the same pattern. High *f-score* group outperforms Low group by 11.26% on average annually. Tests of differences in mean and median with the *t*-test and signed rank Wilcoxon test, respectively, prove *f-score* has a predictive power for future return, at least at a univariate analysis.

This result has broad implications on the improvement at the trading strategy based on fundamentals of the firms. A strategy of buying stocks with a high *f-score* level and selling stocks with a low *f-score* level generates a market adjusted return of 11.2% when the overall market adjusted return for the entire sample is 2.64%. More importantly, as Piotroski (2005) stresses, although long-short strategies yield significant returns, the benefit of the strategy does not just pertain to the selling side of the trading. Short sales constraint is an apparent issue in the market as several studies suggest (for example, see Almazan, Brown, and Carlson (2004) and Thaler and Lamont (2003)). Therefore, a trading strategy relying heavily on the availability of short sales cannot have practical implication. Buying stocks with high *f-score*s only can generate 20.02% raw return, and 7.3% market adjusted returns, both of which are greater than average corresponding returns for the market portfolio. Profits from *f-score*-based strategy do not come only from the markets with short sales allowed, but from more general circumstances as well, because *f-score* is able to select winners and the winner groups make significantly larger returns than the overall market.

4.2.2. Regression analysis

The univariate analysis gives a general idea about the return patterns by *f-score* but it does not incorporate possible effect of the other control variables which might have some explanatory powers for the future returns. I run the multiple regression models to see if after controlling for the other possible explanatory variables, *f-score* would still have the predictive power of one year buy-hold future returns.

$$RET_{t+5,t+16} = \alpha + \beta_{SZ}\ SZ + \beta_{BM}\,\mathrm{BM} + \beta_{mom}\mathrm{MOMRET} + \beta_{FS} fscore + \varepsilon \qquad (8)$$

$$RET_{t+5,t+16} = \alpha + \beta_{SZ}\ SZ + \beta_{BM}\,\mathrm{BM} + \beta_{mom}\mathrm{MOMRET} + \beta_{hs}\mathrm{H}score + \beta_{ls} Lscore + \varepsilon \qquad (9)$$

The dependent variable in both the model (8) and (9) represents raw (or market adjusted or size adjusted) holding returns computed starting the 5th month after the fiscal year ends (or equivalently after the financial statement based metric is calculated). SZ, BM, and MOMRET is a decile assignment (from 0 to 9) for a firm's market capitalization, book-to-market ratio, and six month holding returns prior to portfolio formation, respectively. *F-score* is calculated as explained in the previous sections. *Hscore* and *Lscore* in the model (9) are dummy variables for the high and low *f-score* groups. Market adjusted returns are computed as raw returns minus CRSP market index returns and size adjusted returns are raw returns minus CRSP corresponding size portfolio returns.

[Table 5 about here]

Table 5 shows the results of the two regression models. The explanatory variables are regressed on annual raw, market adjusted, and the size adjusted holding returns. The coefficients for the variable *f-score* are significant at 1% level for all three different measures for the returns. Additionally, when the dummy variables for the firms with strong and weak financial improvement are used, the regression results remain the same. The firms with high level of financial improvement generate significantly positive raw, market adjusted and size adjusted returns and the firms experiencing worsening of the financial health show negative figures in all three categories of returns.

These results are very much in line with Piotroski (2000, 2005) and Fama and French (2006) that financial statement based metric can explain the future returns. Additionally, the signs for the other explanatory variables are consistent with the literature documenting the common phenomenon of the market. In all three tests using different return measures, the variable relating to the size of the firm is negatively related to the returns, which agrees with the well-known "small firm effect". The regression results also confirm the high book-to-market stocks (or the value stocks) generate higher return on average (book-to-market anomaly) and the stocks with high past returns generate average higher future returns (momentum effect).

Overall, I confirm in Section 4 that the results from Piostroki (2000, 2005) can be replicated and a set of nine simple indicators can indeed forecast future returns in the entire sample, as well as value portfolio in different sample periods. Fama and French (2006) also confirm Piotroski's result, although the authors propose the risk-based explanation as a reason for the predictive power of financial statement analysis. In the next section, I attempt to disentangle two competing arguments (Piotroski (2000, 2005)'s investor behavior related and Fama and French (2006)'s risk based) for the financial statement analysis' predictive power of the future returns.

5. Rational vs. irrational explanations for the explanatory power of the signal representing a firm's financial condition

In this section, I differentiate the two explanations suggested in the previous section. Contrary to Fama and French (2006) in which the authors argue the profitability is as expected and the firms earn higher risk for compensation for higher risk, Piotroski (2000, 2005)'s argument is related to investor demand. Piotroski attributes benefits of the trading strategies based on *f-score* to the fact that investors are slow to react to a signal representing the improvement or worsening of a firm's financial situation.

If financial statement analysis predicts the future returns because "investors" slowly react to the information regarding a firm's financial condition, it implies that subset of investors who recognize this opportunity earlier than others will trade to exploit the information. Because literature concerning the behavior of institutional investors proposes institutional investors are more sophisticated than individual investors (e.g., Hribar, Jenkins, and Wang, 2004; Bartov, Radhakrishnan, and Krinsky, 2000; Collins, Gong, and Hribar, 2003; Amihud and Li, 2002; Ke and Petroni, 2004), and institutional investors are price setting marginal investors (Froot and Teo, 2004; Sias, Starks and Titman , 2006) I expect that institutional investors will be the one who exploit the information embedded in *f-score*, prior to individual investors.

As a result, institutional investors will buy the stocks with improvement of financial situation, or high *f-score* stocks and sell the stocks with worsening financial situation, or low *f-score* stocks. Given there is a buyer for every seller, net demand by institutions must be offset by net supply by individual investors. Thus, individual investors are expected to take the opposite side of the trading to institutions and buy low *f-score* stocks and sell high *f-score* stocks.

5.1. Univariate analysis

In this section, I attempt to see if there is any trend for institutional demand variables as financial statement based metric increases, or financial situation of underlying firms improves. I examine net institutional demand, adjusted net institutional demand, percentage net institutional demand, adjusted percentage net institutional demand and buyratio, as defined in Section 3.1 at each *f-score* portfolios over a year starting the seventh month after the financial statement releases and the returns for the same period for a period of 1983 to 2005. I calculate annual returns over two time frames (1) starting the 5^{th} month after the portfolio formation (to match Piotroski (2000, 2005) and (2) starting the 7^{th} month after the formation so that the returns match the quarterly institutional ownership data. For example, if a firm's fiscal year ends in December 2002, financial statement based variables are collected in December 2002 and returns are calculated from May 2003 to April 2004 (t+5 to t+16) for a purpose of replicating Piotroski's results and from end of June 2003 to June 2004 (t+7 to t+18) to match returns with institutional investor demand variables. This allows me to match the quarterly institutional ownership data with the return data (e.g., I can evaluate institutional ownership changes from the end of June 2003 to the end of June 2004, but not from the end of April 2003 to the end of April 2004).

F-score is calculated at the end of each fiscal year ending. *F-score* and investor related variables are matched in a manner that investors have two quarters between when a firm's fiscal year ends and when investors start trading. That is, if a firm's fiscal year ends at March, for example, investors who start trading at the beginning of September, are able to exploit the information from the firm's annual financial statement released at March[29]. This method ensures financial statement information is available in public when an investor's investment horizon begins. For the simplicity, I exclude the firms whose fiscal year endings are not aligned with calendar quarter ending. The results of institutional holding measures test and quarterly returns are presented at Table 6.

[Table 6 about here]

Panel A in Table 6 presents the raw and market adjusted returns for the sample including institutional trading data (1983-2005). Four different measures of the annual returns (raw and market adjusted return at t+5 to t+16 and at t+7 to t+18) confirm the earlier conclusion that the returns increase as the strength of the firms' financial health increases in the various sample periods. All four returns demonstrate monotonically increasing pattern as *f-score* increases. The average differences between the groups with high *f-score* and the groups with low *f-score* are positive for all four return measures and are significant at 10% or better level (t-statistics with 3.62 for raw returns (t+5, t+16), 1.73 for market adjusted return (t+5, t+16), 2.15 for raw returns (t+7, t+18), and 9.40 for adjusted returns for (t+7, t+18)).

[29] Firms have statutory period of 90 days for their annual report filings and 45 days for quarterly filings. Stice (1991) and Griffin report majority of the firms submit their filing a few days before or on the statutory due date.

Panel B show institutional investor demand variables for each *f-score* portfolio, as well as the whole sample (presented at the first row of each panel). All the measures have tendency to increase as *f-score* increases. Almost all the cases presented in the table show a monotonic pattern of subsequent institutional demand by *f-score*. Difference in means test confirms there is a significant difference between low and high level of *f-score* groups for institutional demand measures (*t*-statistics for difference in means test are 5.13, 4.17, 3.75, 5.14, and 3.62 for net institutional demand (NID), adjusted NID, percentage NID, adjusted percentage NID, and buyratio, respectively).

On the whole, the results show institutional investors slowly react to the information contained in the signal for a company's financial health. As a result, they buy the securities with higher financial improvement (high *f-score*) more than the securities with deteriorating financial condition (low *f-score*) over a year, giving *f-score* predictive power for the subsequent institutional demand, where institutional demand is measured by various metrics. The result by the univariate analysis in this section supports Piotroski (2000, 2005)'s argument of investors' slow reaction to new information diffused on the market.

5.2. Preliminary test

Before testing any formal relation between returns, institutional demand, and the signal representing a firm's financial health, I first calculate average cumulative returns and institutional demand measures from *t*-12 to *t*+15 to see if there is any systematic trend in the three variables in interest between the financially healthiest groups and the groups with the most deteriorating financial situation. Institutional demand measures used in this preliminary test are net institutional demand and adjusted net institutional demand, defined as Eq. (2) and (3). Figure 1 shows the cumulated raw returns and net institutional demand for high and low *f-score* groups and figure 2 presents the cumulated raw returns and adjusted net institutional demand for the two groups.

[Figure 1 and 2 about here]

The figures present some remarkable results. First, there is a distinct difference between returns between two groups. Average return for the groups with high *f-score* is positive and the returns for the low *f-score* group are negative and the trend persists until about *t*+9 months. Institutional demand measures present the same pattern. Both net institutional demand and adjusted net institutional demand demonstrate distinction between the groups with high and low level of financial health. Especially Adjusted net institutional demand for the high *f-score* group is positive for almost entire test period (*t*-12 to *t*+15) and reveals there is a considerable difference between the high and low *f-score* groups.

5.3. Regression analysis

In this section, I repeat the regression models from Piotroski (2000, 2005), but using variables representing institutional demand, instead of returns, as the dependent variable to ensure the relation between institutional demand and *f-score* is not driven by the relation between institutional demand and other variables.

Although the univariate test and preliminary test performed at the previous section suggest *f-score* has a predictive power, it fails to rule out the other explanations for monotonic patterns in institutional demand variables by *f-score*s. For example, high *f-score* portfolios could result in high subsequent institutional demand if institutional investors prefer growth stocks (low book-to-market ratio) and growth stocks have high *f-score*s on average.

To further test the relation between *f-score* and institutional holdings variables, I run multiple regressions with the control variables known to have some explanatory power for institutional investment pattern and the ones included in Piotroski (2000, 2005)'s studies. To examine whether *f-score* predicts institutional investors' demand, I run the following models:

$$INS_{t+7,t+18} = \alpha + \beta_{SZ}\ \log(SZ) + \beta_{BM}\log(\text{BM}) + \beta_{mom}\text{MOMRET} + \beta_{FS}fscore + \varepsilon \qquad (10)$$

$$INS_{t+7,t+18} = \alpha + \beta_{SZ}\log(SZ) + \beta_{BM}\log(BM) + \beta_{mom}\text{MOMRET} + \beta_{EQ}\text{EQOFF} + \beta_{acc}\text{ACCRUAL} + \beta_{FS}fscore + \varepsilon \qquad (11)$$

$$INS_{t+7,t+18} = \alpha + \beta_{SZ}\log(SZ) + \beta_{BM}\log(BM) + \beta_{mom}\text{MOMDEC} + \beta_{EQ}\text{EQOFF} + \beta_{acc}\text{ACCDEC} + \beta_{FS}fscore + \varepsilon \qquad (12)$$

$$INS_{t+7,t+18} = \alpha + \beta_{SZ}\text{SZDEC} + \beta_{BM}\ \text{BMDEC} + \beta_{mom}\text{MOMDEC} + \beta_{FS}fscore + \varepsilon \qquad (13)$$

where INS is a variable representing change in institutional ownership (net institutional demand, adjusted net institutional demand, percentage net institutional demand, and adjusted percentage net institutional demand) on each panel) from *t*+7 to *t*+18, and SZDEC, BMDEC, MOMDEC , ACCDEC are decile assignments (from 0 to 9) to size, book-to-market ratios, and prior 6 month holding return, and accruals. EQOFF is a binary variable which takes value of 1 if a firm issued a new equity in the respective fiscal year and 0 otherwise. ACCRUAL is net income minus cash flow from operation, scaled by total assets at the beginning of the fiscal year. *F-score* is calculated at the end of each fiscal year ends. Time series average coefficients of the cross sectional regressions, run each fiscal year from 1983 to 2006, are reported at Table 7 and *t*-statistics are from time series standard error.

[Table 7 about here]

Four different measures of institutional investors' trading are used as the dependent variable and the results are presented at each panel of the Table 7. In the regression (10), *f-score,* log value of a firm's market capitalization, log value of book-to-market ratio are used as the explanatory variables, and in (11) additional explanatory variables representing recent equity offering and accrual are used. Regression (12) is the same as the regression (1), except the momentum measure is included as a decile assignment (from 0 to 9), rather than as the continuous variable. Regression (13) measures size, book-to-market ratio, and momentum as the decile assignments.

Regression result reveals with all five different measures of institutional investors' demand, *f-score* remains significant at 5% or better level when other possible explanatory variables are controlled for. The average regression coefficient for the variable *f-score* is significant for all five regressions. The signal for a firm's financial condition has the weakest relation to adjusted net institutional demand for the regression (10), although *f-score* variable is still significant at 1% level (with *t*-statistics of 2.16).

Regression results confirm a set of nine indicator variables representing a firm's financial health predicts the future returns because investors are slow to react to information embedded in the metric. As suggested in the introduction, if financial statement analysis predicts future return because information is slowly impounded into the metric and expectations are revised, there needs to be significant relation between investor behavior and *f-score*. If as Fama and French (2006) suggest, there is positive relation between accounting based metric and future return because riskier firms yield higher return, then there need not be a significant relation between investor behavior variables and financial statement analysis metric. The results support Piotroski

(2000, 2005)'s investor behavior' based explanation, rather than Fama and French (2006)'s risk based explanation for *f-score*'s predictive power for the future returns.

6. **Conclusion**

This study closely follows Piotroski (2000, 2005) and uses simple accounting based signals to proxy for financial situation for a firm and show the financial statement based analysis has predictive power for the future returns. The results are consistent with Fama and French (2006) and Piotroski (2000, 2005) that firms with the financial improvement outperform firms with worsening financial situation in various settings and the result is robust throughout the different samples (entire stock portfolio as well as value stock portfolio) in different sample periods.

I attempt to differentiate Piotroski (2000, 2005)'s investor behavior related explanation and Fama and French (2005)'s risk-based explanation by focusing on institutional investor demand. If information is slowly impounded into the metric representing a firm's financial condition, a group of investors who can detect the opportunity earlier than others will act on and exploit the information. Because institutional investors are known to be more sophisticated than retail investors, I conjecture institutional investors, rather than individual investors, exploit the information imbedded in the financial statement based metric, *f-score*.

I show institutional investors' demand is significantly related to the metric, *f-score* after controlling for other known factors, such as book-to-market ratio, size and momentum. Institutional investors purchase the firms with high level of financial condition more than firms with lower level of financial health. Therefore, I conclude the positive relation between the metric representing a firm's financial situation and future return is at least partially driven by information slowly impounded to the signals.

References

Abarbanell, J., Bushee, B., 1997. Fundamental analysis, future earnings, and stock prices. Journal of Accounting Research 35, 1-24.

Abarbanell, J., Bushee, B., 1998. Abnormal stock returns to a fundamental analysis strategy. Accounting Review 73, 19-45.

Ali, A., Hwang, L., Trombley, M., 2003. Arbitrage risk and the book-to-market anomaly. Journal of Financial Economics 69, 355-373.

Amihud Y., Li, K., 2002. The declining content of dividend announcements and the effects of institutional investors. FIN working paper No. 02-061.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307-343.

Bartrov, E., Radhakrishnan, S., Krinsky, I., 2000. Investor sophistication and patterns in stock returns after earnings announcement. Accounting Review 75, 43-63.

Bennett, J., Sias, R., Starks, L., 2003. Greener pastures and the impact of dynamic institutional preferences 16, 1203-1238.

Campbell, J., Ramadorai, T., Schwartz, A., 2007. Caught on tape: institutional trading, stock returns and earnings announcements. Unpublished working paper, Harvard University and University of Oxford.

Chakravarty, S., 2001. Stealth trading: which traders' traders move stock prices? Journal of Financial Economics 61, 289-307.

Chen, N., Zhang, F., 1998. Risk and return of value stocks. Journal of Business 71, 501-535.

Collins, D., Gong, G., Hribar, P., 2003. Investor sophistication and the mispricing of accruals. Review of Accounting Studies 8, 251-276.

Daniel, K., Titman, S., 2006. Market reactions to tangible and intangible information. Journal of Finance 61, 1605-1643.

Dasgupta, A., Prat, A., Verardo, M, Institutional trade persistence and long-term equity returns. Unpublished working paper, London school of economics.

Desai, H., Jain, P., 1997. Long-run common stock returns following stock splits and reverse splits. Journal of Business 70, 409-433.

Fairfield, P., Whisenant, J., Yohn, T., 2003. Accrued earnings and growth: implications for future profitability and market mispricing. Accounting Review 78, 353-371.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 48, 417-465.

Fama, E., French, K., 1995. Size and book-to-market factors in earnings and returns. Journal of Finance 50, 131-155.

Fama, E., French, K., 2006. Profitability, investment and average returns. Journal of Financial Economics 82, 491-518.

Frazzini,A., Lamont, O.A., 2008.Dumb money: mutual fund flows and the cross-section of stock returns. Journal of Financial Economics 88, 299-322.

Froot, K.A., Teo, M., 2004. Equity style returns and institutional investor flows. NBER Working Paper No. 10355.

Gompers, P., Metrick, 2001. Institutional investors and equity prices. Quarterly Journal of Economics 116, 229-260.

Griffin, J., Lemmon, M., 2002. Book-to-market equity, distress risk, and stock returns. Journal of Finance 57, 2317-2336.

Griffin, P., 2003. Got information? Investor response to form 10-K and form 10-Q EDGAR filings. Review of Accounting Studies 8, 433-460.

Haugen, R.A., Baker, N.L., 1996. Commonality in the determinants of expected stock returns. Journal of Financial Economics 41, 401-439.

Holthausen, R., Larcker, D., 1992. The prediction of stock returns using financial statement information. Journal of Accounting and Economics 15, 373-411.

Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143-2184.

Hribar, P., Jenkins, N.T., Wang, J., 2004. Institutional investors and accounting restatements.

Unpublished working paper, University of Iowa, Vanderbilt University, and Singapore management university.

Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. Market underreaction to open market share repurchases. Journal of Financial Economics 39, 181-208.

Ikenberry, D., Ramnath, S., 2002. Underreaction to self-selected news events: the case of stock splits. Review of Financial Studies 15, 489-526.

Ikenberry, D., Rankine, G., Stice, E., 1996. What do stock splits really signal? Journal of Financial and Quantitative Analysis 31, 357-375.

Jiang, H., 2007. Institutional investors, intangible information and the book-to-market effect. Working paper, RSM Erasmus University, the Netherlands.

Kaniel, R., Saar, G., Titman, S., 2008. Individual investor trading and stock returns. Journal of Finance 63, 273-310.

Ke, B., Petroni, K., 2004. How informed are actively trading institutional investors? Evidence from their trading behavior before a break in string of consecutive earnings increases. Journal of Accounting Research 42, 895-927.

Lakonishok, J., Shleifer A., Vishny R.W., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541-1578.

La Porta, R., 1996. Expectations and cross-section of stock returns. Journal of Finance 51, 1715-1742.

La Porta, R., Lakonishok, J., Shleifer A., Vishny R.W., 1997. Good news for value stocks: further evidence on market efficiency. Journal of Finance 52, 859-874.

Lev, B., Thiagarajan, R., 1993. Fundamental information analysis. Journal of Accounting Research 31, 190-215.

Loughran, T., Ritter, J.R., 1995. The new issues puzzle. Journal of Finance 50, 373-387.

Mohanram, L., 2005. Separating winners from losers among low book-to-market stocks using financial statement analysis. Review of Accounting Studies 10, 133-170.

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

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

Ou, J.A., Penman, S. H., 1989. Financial statement analysis and the prediction of stock returns. Journal of Accounting and Economics 11, 295-329.

Phalippou, L., 2007. Institutional ownership and the value premium. Unpublished working paper, University of Amsterdam.

Piotroski, J.D., 2000. Value investing: the use of historical financial statement information to separate winners from losers. Journal of Accounting Research 38,supplement, 1-41.

Piotroski, J.D., 2005. Further evidence on the relation between historical changes in financial condition, future stock returns and the value/glamour effect. Working paper, University of Chicago.

Ritter, J.R., 1991. The long-run performance of initial public offerings. Journal of Finance 46, 3-27.

Sias, R., 2004. Institutional herding. Review of Financial Studies 17, 165-206.

Sias, R., 2007. Reconcilable differences: Momentum trading by institutions. Financial Review 42, 1-22

Sias, R., Starks, L., Titman, S., 2006. Changes in institutional ownership and stock returns: assessment and methodology. Journal of Business 79, 2869-2910.

Spiess, D.K., Affleck-Graves, J., 1995. Underperformance in long-run stock returns following seasoned equity offering. Journal of Financial Economics 38, 243-267.

Stice, E., 1991. The market reaction to 10-K and 10-Q filings and to subsequent the Wall Street Journal earnings announcements. Accounting Review 66, 42-55.

Table 1. Annual market adjusted returns to *f-score* portfolios (from 1976 to 1996)

Table 1 presents annual holding return to the *f-score* portfolios by fiscal year from 1976 to 1996 for high book-to-market ratio firms. Firms' book-to-market categories are determined based on previous year's book-to-market ratios and highest quintile book-to-market firms are included in the sample. Strong *f-score* portfolios include the firms with *f-scores* greater than 4 and *f-scores* less than or equal to 4 are categorized into weak *f-score* portfolios. Annual market adjusted returns are calculated as raw return minus CRSP value weighted market index return measured from the beginning of the fifth month after a firm's fiscal year end. T-statistics are based on the time series standard error.

Year	Strong *f-score*	Weak *f-score*	Strong-Weak	Number of Observation
1976	0.3519	0.3513	0.0006	484
1977	0.1714	0.1710	0.0005	653
1978	-0.0354	-0.0596	0.0242	610
1979	0.1501	0.0654	0.0846	653
1980	0.1749	0.0235	0.1514	622
1981	0.2538	0.1437	0.1101	689
1982	0.2698	0.1963	0.0735	515
1983	0.0869	-0.1618	0.2487	318
1984	-0.0656	-0.1815	0.1159	959
1985	0.0681	-0.0981	0.1662	525
1986	0.1146	0.0508	0.0638	611
1987	0.0097	-0.0626	0.0723	1,081
1988	-0.0523	-0.1709	0.1185	755
1989	-0.0985	-0.0569	-0.0416	808
1990	0.1853	0.1187	0.0666	1,259
1991	0.2454	0.1507	0.0947	604
1992	0.2667	0.2703	-0.0036	683
1993	0.0270	0.0356	-0.0087	670
1994	-0.0278	-0.0037	-0.0241	1,118
1995	-0.0238	-0.1931	0.1693	912
1996	-0.0229	-0.0726	0.0497	997
Total				15,526
Average	0.0976	0.0246	0.0730	
(*t*-stat)	(3.36)	(0.74)	(4.52)	

Table 2. Annual returns to *f-score* portfolios for high book-to-market stocks (from 1976 to 1996)

Table 2 reports average and percentile 12 month holding returns for the samples from the period of 1976 to 1996. Firms' book-to-market categories are determined based on previous year's book-to-market ratios and highest quintile book-to-market firms are included in the sample. The first row shows the average for all the firms in the sample and the next nine rows document the returns to each *f-score* portfolio. *F-score* is calculated annually using nine variables representing three areas: profitability, liquidity/leverage and operating efficiency. Twelve-month holding return is calculated from the fifth months after formation (*t*+5 to *t*+16). Panel A shows one year raw return, and Panel B present market adjusted return. Market adjusted return is raw return minus CRSP value weighted market index return. Firms with *f-score* 0-3 are categorized into the portfolio Low, 4-6 into Med and 7-9 into High portfolio. T-statistics for the difference between Low and high portfolios are from difference in means test (for Mean). For median, the statistics is Wilcoxon size ranked *Z* statistics.

Panel A: Raw annual return to *f-score* portfolios

	Mean	10%	25%	Median	75%	90%	Number of Observation
All firms	0.2399	-0.3658	-0.1278	0.1232	0.4362	0.8693	15,526
f-score portfolio							
0	0.0968	-0.6379	-0.2407	0.0313	0.3667	0.9149	58
1	0.0859	-0.6000	-0.3019	0.0098	0.3564	0.7856	323
2	0.1480	-0.5435	-0.2572	0.0235	0.3889	0.8295	1,012
3	0.1813	-0.5000	-0.2381	0.0490	0.4082	0.8889	1,931
4	0.2289	-0.4000	-0.1667	0.1000	0.4259	0.8898	2,678
5	0.2554	-0.3421	-0.1105	0.1364	0.4479	0.8543	3,075
6	0.2676	-0.2848	-0.0789	0.1558	0.4444	0.8351	2,900
7	0.2706	-0.2651	-0.0694	0.1526	0.4452	0.8465	2,185
8	0.2984	-0.2315	-0.0489	0.1628	0.4655	0.9091	1,088
9	0.3689	-0.3064	-0.0733	0.1605	0.4689	1.0223	276
Low	0.0876	-0.6087	-0.2982	0.0104	0.3564	0.7857	381
High	0.3126	-0.2500	-0.0531	0.1617	0.4669	0.9468	1,364
High-Low	0.2251	0.3587	0.2451	0.1513	0.1105	0.1611	
(t-stat)	(5.07)			(6.87)			

Panel B: Market adjusted annual return to *f-score* strategy

	Mean	10%	25%	Median	75%	90%	Number of Observation
All firms	0.0540	-0.5443	-0.3026	-0.0514	0.2465	0.6693	15,526
f-score portfolio							
0	-0.0712	-0.8608	-0.3717	-0.1499	0.2756	0.7234	58
1	-0.0837	-0.7931	-0.4700	-0.1512	0.1797	0.6727	323
2	-0.0323	-0.7276	-0.4298	-0.1462	0.1900	0.6649	1,012
3	-0.0047	-0.6612	-0.4060	-0.1230	0.1986	0.6932	1,931
4	0.0446	-0.5786	-0.3312	-0.0734	0.2427	0.6902	2,678
5	0.0671	-0.5249	-0.2877	-0.0410	0.2594	0.6605	3,075
6	0.0831	-0.4534	-0.2515	-0.0191	0.2608	0.6388	2,900
7	0.0774	-0.4563	-0.2533	-0.0297	0.2472	0.6530	2,185
8	0.1125	-0.4144	-0.2268	-0.0073	0.3023	0.7068	1,088
9	0.1922	-0.4804	-0.2257	-0.0032	0.2818	0.8514	276
Low	-0.0818	-0.7968	-0.4638	-0.1512	0.1864	0.7006	381
High	0.1286	-0.4337	-0.2264	-0.0067	0.2939	0.7402	1,364
High-Low	0.2104	0.3631	0.2374	0.1445	0.1075	0.0396	
(t-stat)	(4.81)			(5.94)			

Table 3. Regression of annual returns on other control variables and *f-scores* (1976-1996)

This table presents regression results for the following model:

$$RET_{t+5,t+16} = \alpha + \beta_{SZ}\log(SZ) + \beta_{BM}\log(BM) + \beta_{mom}\text{MOMRET} + \beta_{EQ}\text{EQOFF} + \beta_{acc}\text{ACCRUAL} + \beta_{FS} fscore + \varepsilon$$

Panel A documents regression result from pooled regression and panel B shows time series average of the coefficients from the 21 annual regressions with the t-statistics (in the parentheses) from time series standard error. RET is a raw (or market adjusted) one year return starting the seventh month after the fiscal year end. SZ and BM are a firm's market capitalization and book-to-market ratio, respectively, measured at the end of the fiscal year. MOMRET is a 6 month holding return prior to the portfolio formation period and EQOFF is a binary variable which takes value of 1 if a firm issued a new equity in the respective fiscal year and 0 otherwise. ACCRUAL is net income minus cash flow from operation, scaled by total assets at the beginning of the fiscal year. Firms' book-to-market categories are determined based on previous year's book-to-market ratios and highest quintile book-to-market firms are included in the sample.

	Intercept	log(SZ)	log(BM)	MOMRET	EQOFF	ACCRUAL	*f-score*
Panel A: Pooled regression							
(1)	0.1496	-0.0164	0.0851				0.0262
	(1.92)	(-4.00)	(5.04)				(7.53)
(2)	0.1669	-0.0160	0.0788	0.0104	0.0080	-0.0092	0.0203
	(2.07)	(-3.83)	(4.62)	(4.85)	(0.58)	(-3.90)	(5.42)
Panel B: Time series average coefficients 21 annual regressions							
(1)	0.2898	-0.0196	-0.0260				0.0248
	(1.75)	(-2.32)	(-1.01)				(5.89)
(2)	0.0905	-0.0146	-0.0396	0.0422	0.0109	-0.0070	0.0141
	(0.58)	(-1.88)	(-1.81)	(11.46)	(0.63)	(-2.42)	(2.82)

Table 4. Annual return to *f-score* portfolios (from 1972 to 2001)

Table 4 reports average and percentile 12 month holding returns for the samples from the period of 1972 to 2001. The first row shows the average for all the firms in the sample and the next nine rows document the returns to each *f-score* portfolio. *F-score* is calculated annually using nine variables representing three areas: profitability, liquidity/leverage and operating efficiency. Twelve-month holding return is calculated from the fifth months after formation (t+5 to t+16). Panel A shows one year raw return, and Panel B presents market adjusted return. Market adjusted return is raw return minus CRSP value weighted market index return. Firms with *f-score* 0-3 are categorized into the portfolio Low, 4-6 into Med and 7-9 into High portfolio. T-statistics for the difference between Low and high portfolios are from difference in means test (for Mean). For median, the statistics is Wilcoxon size ranked Z statistics.

Panel A: Raw annual return to *f-score* portfolios

	Mean	10%	25%	Median	75%	90%	Number of Observation
All firms	0.1514	-0.4951	-0.2251	0.0464	0.3606	0.7941	118,897
f-score portfolio							
0	0.0337	-0.7031	-0.5000	-0.1430	0.2258	0.8125	361
1	0.0476	-0.7214	-0.5100	-0.1608	0.2632	0.8571	2,932
2	0.0713	-0.6731	-0.4286	-0.0957	0.2958	0.8596	8,176
3	0.0985	-0.6210	-0.3521	-0.0303	0.3227	0.8436	14,740
4	0.1368	-0.5172	-0.2561	0.0239	0.3515	0.8128	21,316
5	0.1583	-0.4491	-0.1952	0.0588	0.3538	0.7678	24,301
6	0.1835	-0.3916	-0.1528	0.0873	0.3810	0.7778	22,369
7	0.1944	-0.3511	-0.1278	0.0984	0.3825	0.7761	16,086
8	0.2082	-0.3415	-0.1130	0.1083	0.4000	0.7840	7,342
9	0.2273	-0.3306	-0.1077	0.1085	0.3895	0.8333	1,274
Low	0.0834	-0.6560	-0.4000	-0.0652	0.3104	0.8478	26,209
Med	0.1599	-0.4545	-0.2000	0.0582	0.3634	0.7835	67,986
High	0.2002	-0.3474	-0.1226	0.1021	0.3881	0.7802	24,702
High-Low	0.1167	0.3086	0.2774	0.1673	0.0777	-0.0676	
(t-stat)	(16.88)			(42.47)			

Panel B: Market adjusted annual return to *f-score* portfolios

	Mean	10%	25%	Median	75%	90%	Number of Observation
All firms	0.0264	-0.5952	-0.3324	-0.0671	0.2245	0.6399	118,897
f-score portfolio							
0	-0.1172	-0.8687	-0.6192	-0.2848	0.0934	0.6071	361
1	-0.0719	-0.8289	-0.6000	-0.2654	0.1344	0.7139	2,932
2	-0.0513	-0.7728	-0.5328	-0.2124	0.1653	0.7106	8,176
3	-0.0247	-0.7174	-0.4567	-0.1486	0.1816	0.6802	14,740
4	0.0142	-0.6213	-0.3597	-0.0874	0.2202	0.6515	21,316
5	0.0326	-0.5538	-0.3021	-0.0550	0.2189	0.6102	24,301
6	0.0567	-0.4958	-0.2653	-0.0294	0.2417	0.6179	22,369
7	0.0672	-0.4607	-0.2416	-0.0191	0.2468	0.6282	16,086
8	0.0821	-0.4429	-0.2334	-0.0051	0.2682	0.6372	7,342
9	0.0959	-0.4553	-0.2278	-0.0004	0.2611	0.6741	1,274
Low	-0.0395	-0.7531	-0.4995	-0.1793	0.1723	0.6946	26,209
Med	0.0348	-0.5582	-0.3064	-0.0555	0.2273	0.6263	67,986
High	0.0731	-0.4567	-0.2384	-0.0146	0.2536	0.6342	24,702
High-Low	0.1126	0.2964	0.2611	0.1647	0.0813	-0.0604	
(t-stat)	(16.59)			(42.10)			

Panel C: Size adjusted annual return to *f-score* portfolios

	Mean	10%	25%	Median	75%	90%	Number of Observation
All firms	-0.0022	-0.6154	-0.3514	-0.0836	0.1989	0.5905	115,819
f-score portfolio							
0	-0.1040	-0.8319	-0.5908	-0.2591	0.1210	0.6400	359
1	-0.0704	-0.8303	-0.5788	-0.2539	0.1515	0.6975	2,892
2	-0.0661	-0.7792	-0.5296	-0.2117	0.1584	0.6611	8,048
3	-0.0501	-0.7348	-0.4664	-0.1610	0.1651	0.6278	14,439
4	-0.0150	-0.6442	-0.3797	-0.1020	0.1923	0.6057	20,822
5	0.0028	-0.5727	-0.3209	-0.0727	0.1921	0.5591	23,620
6	0.0252	-0.5240	-0.2844	-0.0469	0.2177	0.5677	21,740
7	0.0329	-0.4938	-0.2678	-0.0383	0.2178	0.5708	15,605
8	0.0471	-0.4905	-0.2591	-0.0231	0.2365	0.5887	7,081
9	0.0718	-0.4890	-0.2460	-0.0104	0.2293	0.6374	1,213
Low	-0.0581	-0.7625	-0.5031	-0.1841	0.1615	0.6419	25,738
Med	0.0046	-0.5819	-0.3271	-0.0728	0.2009	0.5746	66,182
High	0.0390	-0.4927	-0.2649	-0.0322	0.2234	0.5823	23,899
High-Low	0.0972	0.2698	0.2382	0.1519	0.0618	-0.0596	
(t-stat)	(14.35)			(37.36)			

Table 5. Regressions of annual returns on f-scores and other control variables (1972-2001)

This table presents average coefficients from 30 annual regressions for the following model:

$$RET_{t+5,t+16} = \alpha + \beta_{SZ}\ SZ + \beta_{BM}\,\mathrm{BM} + \beta_{mom}\mathrm{MOMRET} + \beta_{FS} fscore + \varepsilon \qquad (1)$$

$$RET_{t+5,t+16} = \alpha + \beta_{SZ}\ SZ + \beta_{BM}\,\mathrm{BM} + \beta_{mom}\mathrm{MOMRET} + \beta_{hs}\mathrm{H}score + \beta_{ls} Lscore + \varepsilon \qquad (2)$$

Where RET is annual raw (or market adjusted or size adjusted returns) holding returns measured from the seventh months after a firm's fiscal year ends, SZ, BM, and MOMRET is a decile assignment (from 0 to 9) for a firm's market capitalization, book-to-market ratio, and six month holding returns prior to portfolio formation, respectively. *F-score* is calculated at the end of each fiscal year end using nine variables representing three areas: profitability, liquidity/leverage and operating efficiency. *Hscore* is a binary variable which takes one for a firm whose *f-score* is between 7 and 9 and *Lscore* is an indicator for the firms with f-score ranging from 1 to 3.

	Intercept	SZ	BM	MOM	*f-score*	*L-score*	*H-score*
			Panel A: Annual raw returns				
(1)	0.0350	-0.0077	0.0079	0.0102	0.0146		
	(0.57)	(-2.03)	(2.19)	(3.80)	(3.79)		
(2)	0.1090	-0.0075	0.0081	0.0106		-0.0454	0.0252
	(2.25)	(-1.97)	(2.22)	(3.95)		(-3.35)	(3.30)
			Panel B: Annual market adjusted returns				
(1)	-0.0894	-0.0074	0.0078	0.0108	0.0145		
	(-1.76)	(-1.91)	(2.19)	(4.24)	(3.77)		
(2)	-0.0159	-0.0072	0.0080	0.0112		-0.0452	0.0252
	(-0.43)	(-1.86)	(2.22)	(4.40)		(-3.32)	(3.40)
			Panel C: Annual size adjusted returns				
(1)	-0.1313	-0.0028	0.0082	0.0113	0.0136		
	(-3.77)	(-1.18)	(2.31)	(4.77)	(3.40)		
(2)	-0.0623	-0.0026	0.0084	0.0117		-0.0427	0.0235
	(-3.16)	(-1.11)	(2.34)	(4.94)		(-3.05)	(3.07)

Table 6. Annual returns and institutional ownership changes (1983-2005)

Table 6 shows 12 months holding returns and institutional ownership change for the sample from the period of 1983 to 2005. The first row shows the average figures for all the firms in the sample and the next nine rows document the returns to each *f-score* portfolio. Next three rows presents the corresponding figures for f-score groups (Low for *f-scores* 0-3, Med for *f-scores* 4-6 and High for *f-scores* higher than 6). Last two rows document difference between high and low groups and the t-statistics are from difference in means test. *F-score* is calculated annually using nine variables representing three areas: profitability, liquidity/leverage and operating efficiency. Adjusted returns are calculated as raw return minus CRSP value weighted market index return for the corresponding periods. NID is change in fractional institutional ownership measured over 12 month period from the seventh month after a firm's fiscal year ends. Adj. NID is NID minus average NID for the similar size firms for the same period. P_NID is calculated as NID divided by fractional institutional ownership at portfolio formation. Adj. P_NID is measured by subtracting average P_NID for the similar size firms from P_NID for the same time period. Buyratio is number of the buyers divided by number of the traders.

Panel A: Annual returns to *f-score* portfolio

	Ret (*t*+5,*t*+16)	Adjret (*t*+5,*t*+16)	Ret (*t*+7,*t*+18)	Adjret (*t*+7,*t*+18)
All firms	0.1492	0.0223	0.1496	-0.0192
f-score portfolio				
0	0.0861	-0.0738	0.1206	-0.1141
1	0.2249	0.1020	0.2009	-0.0909
2	0.1112	-0.0020	0.1472	-0.0656
3	0.1347	0.0214	0.1409	-0.0506
4	0.1311	0.0113	0.1352	-0.0298
5	0.1382	0.0101	0.1368	-0.0273
6	0.1563	0.0242	0.1500	0.0023
7	0.1756	0.0390	0.1688	0.0093
8	0.1865	0.0448	0.1877	0.0294
9	0.2145	0.0713	0.1735	0.0137
Low	0.1364	0.0214	0.1487	-0.0602
Med	0.1421	0.0152	0.1407	-0.0181
High	0.1806	0.0422	0.1745	0.0153
H-L	0.0442	0.0208	0.0257	0.0754
t-stat	(3.62)	(1.73)	(2.15)	(9.40)

Panel B: Annual institutional ownership change to *f-score* portfolio

	NID	Adj.NID	P_NID	Adj. p_NID	Buyratio
All firms	0.0135	0.0019	0.0447	0.0053	0.5354
f-score portfolio					
0	-0.0022	-0.0154	-0.0026	-0.0488	0.5289
1	-0.0015	-0.0111	0.0029	-0.0495	0.5157
2	0.0085	-0.0022	0.0366	-0.0093	0.5230
3	0.0099	-0.0008	0.0393	-0.0037	0.5275
4	0.0143	0.0028	0.0456	0.0063	0.5309
5	0.0137	0.0021	0.0432	0.0061	0.5359
6	0.0148	0.0025	0.0462	0.0090	0.5423
7	0.0157	0.0038	0.0513	0.0139	0.5415
8	0.0187	0.0059	0.0583	0.0184	0.5448
9	0.0139	0.0025	0.0602	0.0200	0.5353
Low	0.0081	-0.0025	0.0341	-0.0108	0.5250
Med	0.0142	0.0025	0.0450	0.0071	0.5365
High	0.0165	0.0043	0.0537	0.0155	0.5421
H-L	0.0083	0.0068	0.0196	0.0263	0.0171
t-stat	(5.13)	(4.17)	(3.75)	(5.14)	(3.62)

Table 7. Regression of institutional ownership change

Table 7 documents the regression results from the following models:

$$INS_{t+7,t+18} = \alpha + \beta_{SZ} \log(SZ) + \beta_{BM}\log(\text{BM}) + \beta_{mom}\text{MOMRET} + \beta_{FS} fscore + \varepsilon \quad (1)$$
$$INS_{t+7,t+18} = \alpha + \beta_{SZ} \log(SZ) + \beta_{BM} \log(BM) + \beta_{mom}\text{MOMRET} + \beta_{EQ}\text{EQOFF} + \beta_{acc}\text{ACCRUAL} + \beta_{FS} fscore + \varepsilon \quad (2)$$
$$INS_{t+7,t+18} = \alpha + \beta_{SZ} \log(SZ) + \beta_{BM} \log(BM) + \beta_{mom}\text{MOMDEC} + \beta_{EQ}\text{EQOFF} + \beta_{acc}\text{ACCDEC} + \beta_{FS} fscore + \varepsilon \quad (3)$$
$$INS_{t+7,t+18} = \alpha + \beta_{SZ}\text{SZDEC} + \beta_{BM}\ \text{BMDEC} + \beta_{mom}\text{MOMDEC} + \beta_{FS} fscore + \varepsilon \quad (4)$$

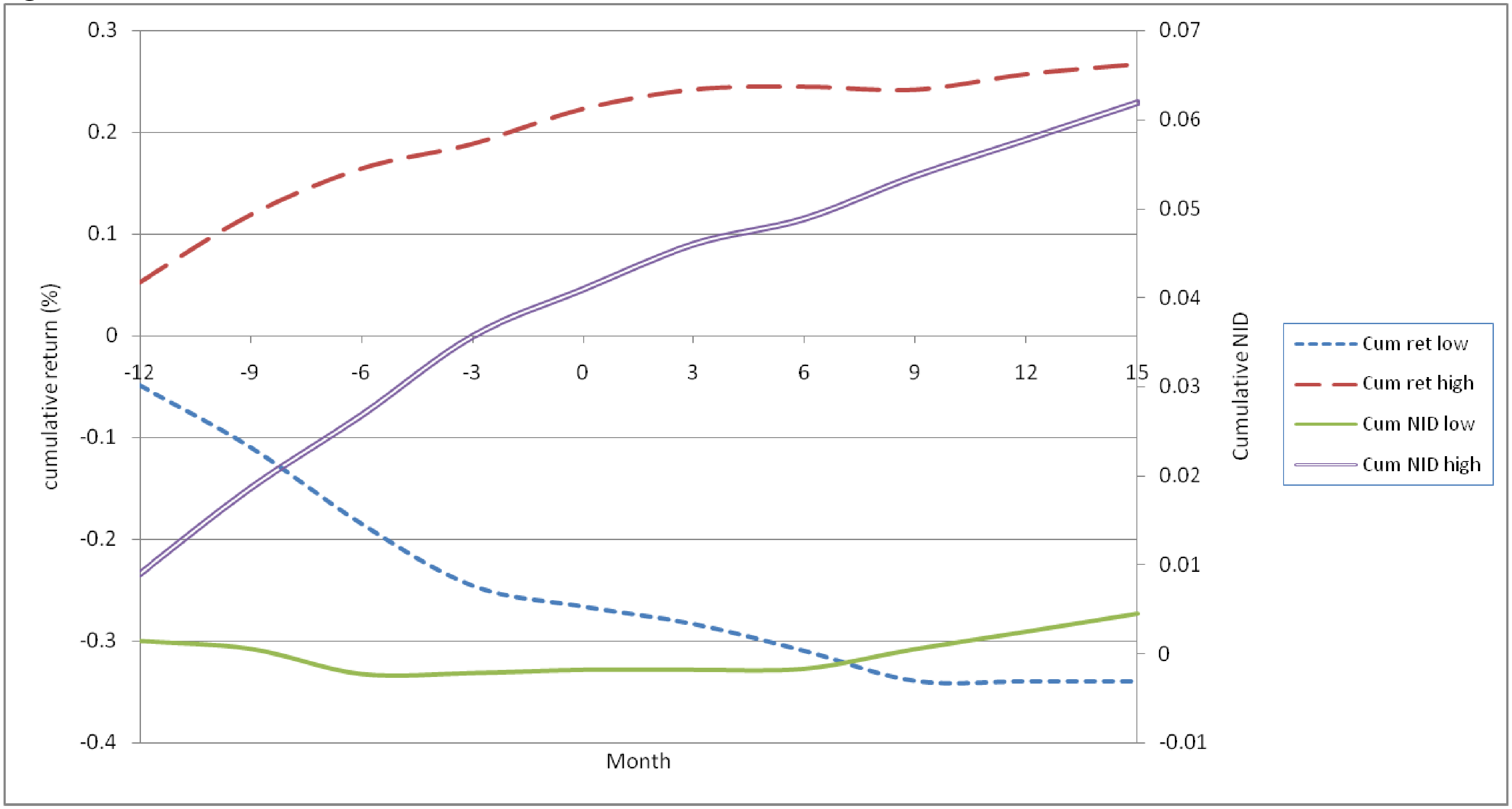
where INS is a variable representing change in institutional ownership (NID, Adj. NID, P_NID, Adj. P_NID and BuyRatio) on each panel) from *t*+7 to *t*+18, and SZDEC, BMDEC, MOMDEC , ACCDEC are decile assignments(from 0 to 9) to size, book-to-market ratios, and prior 6 month holding return, and accrual. EQOFF is a binary variable which takes value of 1 if a firm issued a new equity in the respective fiscal year and 0 otherwise. ACCRUAL is net income minus cash flow from operation, scaled by total assets at the beginning of the fiscal year. *F-score* is calculated at the end of each fiscal year end using nine variables representing three areas: profitability, liquidity/leverage and operating efficiency. BuyRatio is calculated as number of buyers of a firm divided by number of traders over the period from *t*+7 to *t*+18.

	Intercept	log(SZ)	SZdec	log(BM)	BMdec	MOM	MOMdec	EQOFF	ACCRUAL	ACCdec	*f-score*
					Panel A: NID						
(1)	0.0180	-0.0006		-0.0012		0.0350					0.0013
	(1.39)	(-0.94)		(-0.62)		(6.44)					(2.94)
(2)	0.0193	-0.0007		-0.0001		0.0350		-0.0058	-0.0065		0.0016
	(1.50)	(-1.07)		(-0.08)		(6.64)		(-2.42)	(-1.18)		(3.01)
(3)	0.0159	-0.0011		-0.0003			0.0037	-0.0060		-0.0053	0.0015
	(1.21)	(-1.82)		(-0.16)			(6.27)	(-2.36)		(-2.92)	(3.13)
(4)	-0.0051		-0.0005		-0.0003		0.0038				0.0013
	(-1.18)		(-1.30)		(-0.75)		(6.04)				(3.21)

	Intercept	log(SZ)	SZdec	log(BM)	BMdec	MOM	MOMdec	EQOFF	ACCRUAL	ACCdec	*f-score*
Panel B: Adj. NID											
(1)	-0.0217	0.0008		0.0000		0.0339					0.0010
	(-3.18)	(2.38)		(-0.01)		(6.23)					(2.16)
(2)	-0.0206	0.0008		0.0008		0.0339		-0.0046	-0.0084		0.0012
	(-3.18)	(2.23)		(0.46)		(6.42)		(-1.98)	(-1.48)		(2.29)
(3)	-0.0222	0.0003		0.0007			0.0037	-0.0049		-0.0008	0.0012
	(-4.24)	(0.91)		(0.41)			(6.24)	(-1.96)		(-3.20)	(2.29)
(4)	-0.0215		0.0002		-0.0001		0.0037				0.0011
	(-8.04)		(0.79)		(-0.15)		(5.96)				(2.36)
Panel C: P_NID											
(1)	0.2131	-0.0101		-0.0083		0.1336					0.0041
	(3.62)	(-3.66)		(-1.60)		(7.44)					(3.06)
(2)	0.2187	-0.0104		-0.0033		0.1342		-0.0250	-0.0144		0.0055
	(3.85)	(-3.86)		(-0.55)		(7.80)		(-2.00)	(-0.73)		(3.37)
(3)	0.2157	-0.0125		-0.0030			0.0139	-0.0026		-0.0008	0.0054
	(3.80)	(-4.51)		(-0.52)			(7.52)	(-2.02)		(-2.82)	(3.53)
(4)	0.0069		-0.0075		-0.0018		0.0139				0.0044
	(0.49)		(-3.65)		(-1.19)		(6.95)				(3.37)

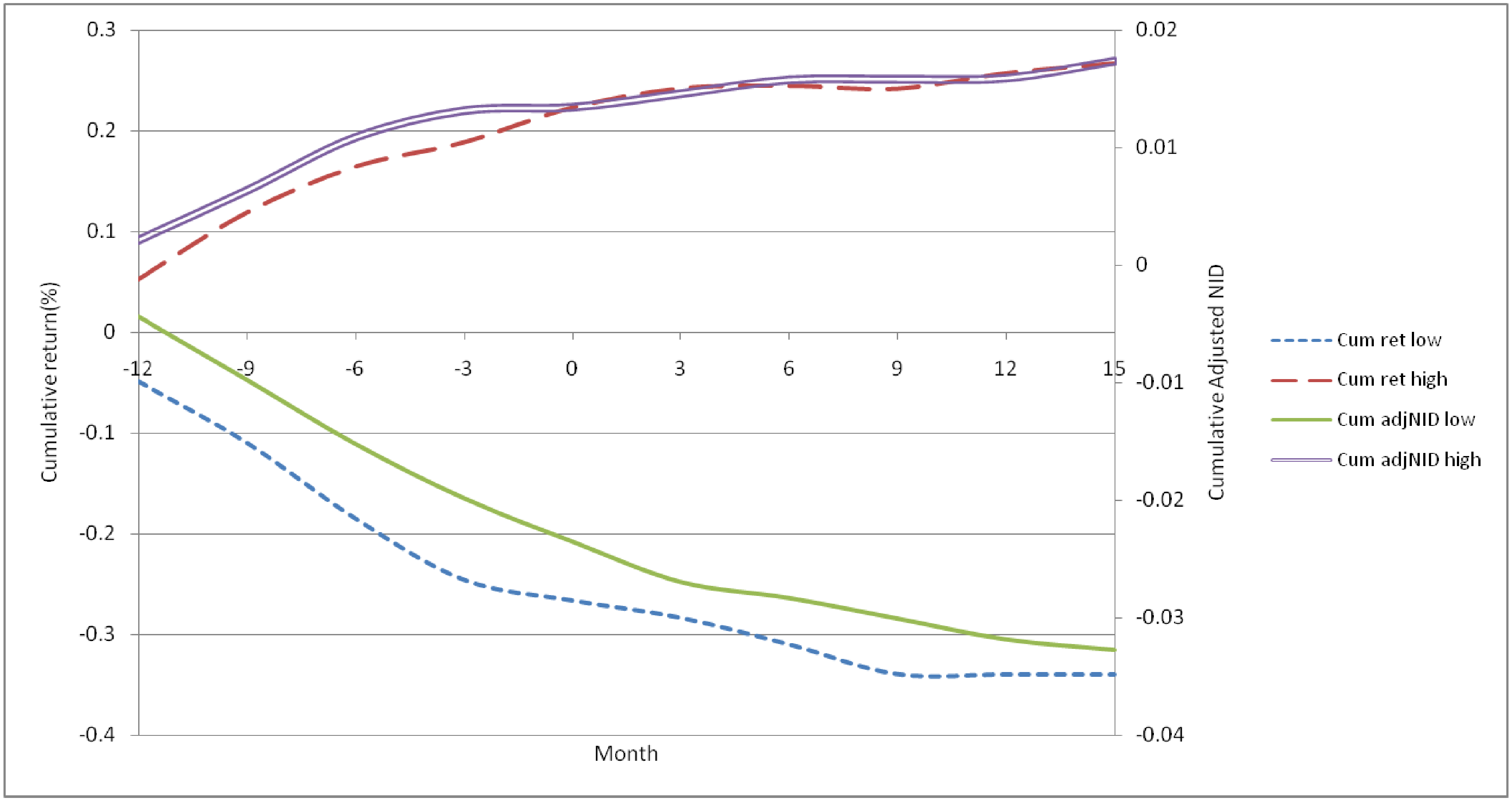
	Intercept	log(SZ)	SZdec	log(BM)	BMdec	MOM	MOMdec	EQOFF	ACCRUAL	ACCdec	*f-score*
					Panel D: Adj. P_NID						
(1)	-0.0699	0.0023		-0.0057		0.1309					0.0042
	(-1.62)	(1.10)		(-1.19)		(7.12)					(3.25)
(2)	-0.0638	0.0019		-0.0011		0.1314		-0.0232	-0.0164		0.0054
	(-1.68)	(1.03)		(-0.19)		(7.47)		(-1.89)	(-0.83)		(3.52)
(3)	-0.0651	-0.0002		-0.0009			0.0138	-0.0250		-0.0028	0.0052
	(-2.21)	(-0.09)		(-0.15)			(7.48)	(-1.90)		(-2.74)	(3.67)
(4)	-0.0772		0.0001		-0.0009		0.0137				0.0043
	(-6.81)		(0.07)		(-0.58)		(6.83)				(3.43)
					Panel E: Buyratio						
(1)	0.7433	-0.0121		-0.0109		0.0894					0.0032
	(10.94)	(-3.67)		(-3.54)		(10.00)					(3.18)
(2)	0.7470	-0.0122		-0.0083		0.0893		-0.0170	-0.0137		0.0041
	(10.81)	(-3.69)		(-2.30)		(10.54)		(-3.42)	(-2.34)		(4.17)
(3)	0.7256	-0.0131		-0.0084			0.0097	-0.0178		-0.0007	0.0039
	(10.07)	(-4.03)		(-2.32)			(9.38)	(-3.50)		(-1.81)	(3.78)
(4)	0.5184		-0.0066		-0.0023		0.0097				0.0031
	(29.80)		(-3.49)		(-2.77)		(8.76)				(2.95)

Figure 1 Cumulative returns and NID from *t*-12 to *t*+15



The figure present average cumulated raw returns and net institutional demand from t-12 to t+15. Dashed line represents the average cumulated return for the groups with high *f-score*, and dotted line represents the average cumulated return for the low *f-score* groups. Double solid line represents cumulated net institutional demand for high *f-score* groups and single solid line presents cumulated net institutional demand for low *f-score* group.

Figure 2 Cumulative returns and adjusted NID from *t*-12 to *t*+15



The figure present average cumulated raw returns and adjusted net institutional demand from *t*-12 to *t*+15. Dashed line represents the average cumulated return for the groups with high *f-score*, and dotted line represents the average cumulated return for the low *f-score* groups. Double solid line represents cumulated adjusted net institutional demand for high *f-score* groups and single solid line presents cumulated adjusted net institutional demand for low *f-score* group.