6-2019

Institutional management and institutional trading

Jingi HA

Singapore Management University, jingiha.2014@pbs.smu.edu.sg

INSTITUTIONAL MANAGEMENT AND INSTITUTIONAL TRADING

JinGi HA

SINGAPORE MANAGEMENT UNIVERSITY

2019

Institutional Management and Institutional Trading

JinGi HA

Submitted to Lee Kong Chian School of Business
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Business (Finance)

Dissertation Committee:

Jianfeng HU (Supervisor / Chiar)
Assistant Professor of Finance
Singapore Management University

Ekkehart BOEHMER
Professor of Finance
Singapore Management University

Roger LOH
Associate Professor of Finance
Singapore Management University

Gennaro BERNILE
Associate Professor of Finance
University of Miami

Singapore Management University
2019

I hereby declare that this PhD dissertation is my original work
and it has been written by me in its entirety.
I have duly acknowledged all the sources of information
which have been used in this dissertation.

This PhD dissertation has also not been submitted for any degree
in any university previously.

JinGi HA
03 June 2019

Abstract

This dissertation consists of three papers in mutual fund governance or market microstructure that analyze the causal effect of board independence on mutual fund performance or the trading behavior of institutional trading and informed trading.

Chapter I studies how board independence affects fund performance, in relation to investment experience of independent directors. Using the SEC amendment in 2001 as an exogenous shock, I find that board independence does not improve or damage fund performance on average. When a fund board has independent directors with investment experience, however, it boosts fund performance. I also find that a fund manager is less constrained and the management fee on a contract is more aligned with fund performance under such a fund board. My findings suggest that board independence is not always beneficial to mutual fund shareholders, but its effectiveness varies depending on independent directors' investment experience.

Chapter II estimates daily aggregate order flow of individual stocks from all institutional investors as well as for hedge funds and other institutions separately. This study is coauthored with my advisor, Prof. Jianfeng Hu. We achieve this by extrapolating the relation between quarterly institutional ownership in 13F filings, aggregate market order imbalance in TAQ, and a representative group of institutional investors' transaction data. We find that the estimated institutional order imbalance positively predicts stock return on the next day and outperforms other institutional order flow estimates. The institutional order flow from hedge funds creates smaller contemporaneous price pressure and generates greater and more persistent price impact than the order flow from all other institutions. We also find that hedge funds trade on well-known anomalies while the other institutions do not. Our findings suggest that the superior trading skills of institutional investors can be largely attributed to hedge funds.

Lastly, I propose a simple measure of informed trading based on the Kyle (1985) model in Chapter III. This study is also coauthored with my advisor, Prof. Jianfeng Hu. We first calculate implied order imbalance (IOI) as contemporaneous stock returns divided by low-frequency illiquidity measures. The implied informed trading (IIT) is the residual of IOI regressed on its components (returns and illiquidity). We find that IIT positively predicts short-term future stock returns without subsequent reversals in the cross-section between 1927 and 2016. This predictability is robust in subperiods, and strengthens in stocks with high information asymmetry and before corporate events. The predictability survives existing measures of informed trading including short selling activities, order imbalance, and institutional trading in recent periods. Finally, IIT has the same predictive ability in G10 equity markets.

Acknowledgements

I would like to express my special appreciation and thanks to my friend and supervisor, Dr. Jianfeng Hu. It has been my privilege to have you as my mentor for the last four years. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been invaluable. I would also like to thank my committee members, professor Ekkehart Boehmer, professor Roger Loh, professor Gennaro Bernile for serving as my committee members. I also want to thank you for letting my defense be an enjoyable moment, and for your insightful comments and suggestions.

My sincere thanks also goes to my family. Words can not express how grateful I am to my mother and father for all of the sacrifices that you've made on my behalf. Without your precious support, this PhD would not have been achievable. I would also like to thank to my beloved spouse. Your prayer for me was what sustained me thus far. Thank you for supporting me for everything, and especially I can't thank you enough for encouraging me throughout this experience. I am especially grateful to my mother-in-law for believing in my research achievement and for her emotional support.

Finally I thank my God. I have experienced Your guidance day by day. Guide my steps for my future and I always pray "not my will, but thine be done."

Contents

List of Figures

List of Tables

I. When is board independence beneficial to mutual fund shareholders? Evidence from the 2001 SEC amendment

A. Introduction

The fiduciary duty of independent directors in a fund board closely relates to the interest of mutual fund investors. Independent directors separately oversee and evaluate the performance of a mutual fund. Based on the oversight and evaluation, they vote in person to approve an advisory contract and its renewal on an annual basis. The terms of the contract independent directors bargain with a fund manager include the amount of management fee and investment practice restrictions which should be directly tied to fund performance. Along with its important role, regulators have attempted to raise the proportion of independent directors in fund boards to advocate and protect the interest of fund investors. The Securities and Exchange Commission (SEC), for example, mandates fund boards to increase the minimum proportion of independent directors from 40 percent to a majority in 2001. The Investment Company Institute (ICI) recommends each board has a two-thirds majority of independent directors (ICI, 1999). As a result, the number of fund complexes with at least 75 percent of board seats held by independent directors increases from 46 percent in 1996 to 84 percent in 2016 (ICI, 2017).

However, the oversight function performed by independent directors can be limited because they are in a difficult position to acquire necessary information for effective oversight or proper evaluation of mutual funds. As an outsider in the mutual fund they monitor, an independent director likely has no inner source which can provide full information for the directors' duties. Also, inside directors or fund managers are reluctant to share information disadvantageous to them with independent directors at the risk of hash monitoring (Harris and Raviv, 2008). In addition, perhaps more importantly, independent directors usually attend board meetings on a *quarterly* basis to monitor the management of *multiple* funds.[1] It is challenging for independent directors

[1]ICI (2017) reports the majority of investment companies have four scheduled board meetings in a year. Also, Ferris and Yan (2007) document that the average (median) number of funds overseen by an independent director is 18.54 (6.00) in 2002.

to detect management problems in a number of mutual funds in such a short time. So it is possible that independent directors may be hard to obtain material information regarding mutual fund management and represent the interest of mutual fund investors effectively. Consitent with this idea, prior studies have failed to identify significant relation between board independence and fund performance despite of the important role of independent directors as a watchdog in fund management.[2]

Investment experience can improve the ability of independent directors to evaluate and monitor mutual fund management by lowering their cost for information acquisition. Specifically, I focus on investment experience from working experience as a fund manager, a general partner, or an executive officer in an investment company and as a private investor. Independent directors with investment experience likely have knowledge on the nature of investment activities, skills to quickly process financial information, or connections to help them to gain information about fund management. Therefore, investment experience can bridge information gap between independent directors and insiders like inside directors or fund managers. By doing so, it can help independent directors overcome information challenge while evaluating the investment ability of fund managers or the benefit and cost of current investment practice restrictions. Investment experience is helpful especially in the mutual fund industry where the decision-making process in daily trading practices is arguably opaque. In the line of thinking, this paper examines the effect of board independence on fund performance, depending on investment experience of independent directors.

The most challenging issue in this study is that board composition is endogenously determined (Hermalin and Weisbach, 1998; Raheja, 2005). I employ an SEC amendment in 2001 as an "exogenous" regulation shock to address the endogeneity issues in board independence. The amendment requires mutual funds to increase the minimum proportion of independent directors in the board from 40 percent to a majority. I define a treatment (control) group of mutual funds as mutual funds in which independent directors did not (did) constitute a majority of their board of directors before the 2001 SEC amendment. As a result of the amendment, 179 mutual funds

[2]See Almazan, *et al.* (2003), Ferris and Yan (2007), Khorana, Serves, and Wedge (2007), Chen, Goldstein, and Jiang (2008), and Cremers, *et al.* (2009).

in the treatment group significantly increase the averaged proportion of independent directors in a fund board from 47.0 percent to 67.9 percent, while other mutual funds in the control group increases from 74.5 percent to 75.8 percent only. The sudden increase of treated mutual funds in board independence gives me confidence to estimate the effect of fund board independence which is largely free from endogeneity concerns. I compare fund performance of treated mutual funds with that of controlled mutual funds before and after the 2001 amendment, depending on investment experience of independent directors who exist in a fund board before 2001. With the difference-in-difference test, this paper answers whether and when board independence is beneficial to mutual fund investors.

My main finding is that board independence does not improve or damage fund performance on average. However, it significantly increases fund return by 1.98 percent per year at the same level of fund risk when a fund board has an independent director with investment experience. The conditional effect is economically significant, given the average annual return in the treatment group of mutual funds is 0.69 percent. Similarly, I also find that board independence does not on average affect the return of a fund portfolio as well as a return gap – the difference between the reported fund return and the return on a fund portfolio disclosed on the previous quarter. With investment-experienced independent directors, however, board independence makes the return gap significantly increased without notable changes in portfolio return. All findings are robust in a propensity score match analysis, suggesting that they are not driven by pre-existing heterogeneity between treatment and control groups of mutual funds. Note that the return gap can partially capture the trading ability of fund managers and that the return of a portfolio represent their stock-picking ability (Kacperczyk, Sialm, and Zheng, 2008). My findings suggest it is possible that independent directors with investment experience would provide investment environment for fund managers to exert better trading ability.

I look into this possibility by investigating the change of investment practice restrictions and the relation between contractual management fee and fund performance around the SEC amendment. I find that the treatment group of mutual funds decreases 1.40 out of eighteen investment restrictions in the grace period of the 2001 SEC amendment, while the control group decreases 0.56 restrictions only. Notably, the decrease of investment restrictions in the treatment group is mainly attributed to treated mutual funds with at least one independent director with investment experience. I also find weak evidence on that board independence causes management fee on a contract to be more positively associated with fund performance when a fund board has independent directors with investment experience. My findings suggest that independent directors with investment experience may provide fund managers better investment environment by relaxing investment restrictions and adjusting the level of contractual management fee in line with fund performance.

In addition, I explore endogenous concerns regarding investment experience of independent directors, defined as the proportion of independent directors with investment experience in a fund board who exist in 2000 before the SEC amendment. It would be possible that investment experience of independent directors who are newly hired or resigned after 2000 may affect the effect of board independence on fund performance. If it is the case, one may interpret my main findings as meaning that fund performance improves in response to a change in the proportion of independent directors with investment experience, rather than an increase in the proportion of independent directors. To address that concern, I compare the effect of board independence, conditional on investment experience of independent directors who (1) hold the position, (2) get hired, and (3) are resigned in the grace period of the 2001 amendment. I find that my main findings are mainly driven by independent directors who keep holding the director position in the period, while newly hired independent directors marginally influence the effectiveness of board independence and resigned independent directors have statistically no influence.

Also, it would be possible that independent directors without investment experience may affect the effect of board independence on fund performance. For example, they can provide material inside information to fund managers, using their working experience in non-finance industries (Dass, *et al.*, 2014). If it is the case, the interpretation of my findings is much complicated. Taking this concern into consideration, I compare the effectiveness of board independence, conditional on independent directors with (1) investment experience, (2) working experience in the finance industry except an investment company, and (3) working experience in non-finance industries. I find that independent directors with investment experience play a major role to enhance the effectiveness of board independence, while independent directors with other types of working experience marginally hurt.

This paper makes three contribution to literature. First, this paper provides evidence that board independence in the mutual fund industry is not always beneficial to mutual fund investors, but its effect varies depending on investment experience of independent directors. Specifically, board independence unconditionally does not increase or decrease fund performance. Having independent directors with investment experience, however, board independence boosts fund return by 1.98 percent per year at the qualitatively same level of fund risk. Furthermore, the evidence suggests that prior studies have failed to find significant relation between board independence and fund performance because the effect of board independence can be cancelled out on average.

Second, this paper provides an insight to mutual fund investors on how to exploit board information in fund investment. Mutual fund investors care little about board information despite of its importance. According to ICI (2006), only five percent of mutual fund holders consider board information to be very important for fund investment. This paper documents evidence that mutual fund investors can have the benefit of board information when they choose an independent fund board which has an independent director with investment experience.

To the best of my knowledge, this is the first paper to study the SEC amendment of 2001. The SEC requires mutual funds to increase the minimum proportion of independent directors in the board from 40 percent to a majority. The amendment is as useful as the Sarbanes-Oxley Act of 2002 to examine the effectiveness of board independence in the mutual fund industry because it makes an exogenous impact on the proportion of independent directors in a board of a non-compliant mutual fund. Specifically, The amendment causes 179 non-compliant mutual funds in my sample to increase the proportion of independent directors from 47.0 to 67.9 percent on average. In sharp contrast, compliant mutual funds increase from 74.5 to 75.8 percent only.

B. Empirical strategy

In this section, I will describe my identification strategy and difference-in-difference estimators based on the SEC amendment of 2001.

The simplest approach to examine the effect of board independence on fund shareholders' interest is to regress after-fee fund return on the proportion of independent directors in a fund board. However, board independence is endogenous in the mutual fund industry. For example, poor fund performance may induce fund shareholders to invite more independent directors into a fund board, as in Hermalin and Weisbach (1998). Also, a fund manager who has strong bargaining power due to his investment ability may strategically move into a fund with a smaller fraction of independent directors in a board, seeking higher fee, as in Tufano and Sevick (1997). Hence results from the regression estimation would be interpreted in various ways. To address the endogeneity concerns and clarify a causal effect, I use the 2001 SEC amendment as a source of exogenous changes in board independence.

In 1992, the Division of Investment Management re-examined the adequacy of the governance structure for investment companies. The Division concluded that the monitoring role of independent directors has served investors well at minimal cost. However, the Division recommended the SEC to increase the minimum proportion of independent directors on fund boards from forty percent to more than fifty percent, taking into account the increasingly significant responsibilities placed on independent directors, e.g., the approval of rule 12b-1 plans. In response, the SEC held a Roundtable in 1999 to discuss the role of independent investment company directors and particularly how to enhance their effectiveness.

After evaluating the ideas and suggestions offered by the Division and the Roundtable, the SEC adopted three amendments in 2001 to enhance director independence and its effectiveness. The Commission requires any mutual fund that relies upon certain exemptive rules (1) to raise the minimum percent of independent directors in board from 40 percent to a majority, (2) to let independent directors nominate and select other independent directors, (3) for independent directors to hire only counsel that does not have substantial tie to fund managers, if they hire counsel. I study the first amendment among three relevant amendments because it has a clear treatment group of mutual funds for which independent directors did not constitute a majority of the board before the 2001 amendment, while the other amendments influence entire mutual fund industry. The first amendment was announced on January 3, 2001 and its compliance date was July 1, 2002.

I begin by hand-collecting fund-level board information from mutual fund's Statement of Additional Information (SAI). The information contained in the SAI includes the name of fund directors, whether a director is independent, and their principal occupations during the past five years. With the SAI, I define a treatment (control) group of mutual funds in which independent directors did not (did) constitute a majority of their board of directors before the SEC amendment of 2001. I keep mutual funds which has fund performance information in CRSP Survival-Bias Free Mutual Fund Database (CRSP) after matching funds in the SAI and CRSP by a fund name. Finally, I retain mutual funds which have both board and fund information on December 2000 and July 2002. The producedure yields my sample with 179 mutual funds in the treatment group and $5,409$ mutual funds in the control group.

I also collect biographical information on fund directors to examine how the effectiveness of board independence varies with investment experience of independent directors. I identify the portion of independent directors with investment experience in a fund board, using the information from the SAI on directors' work history over the last five years. I consider an independent director to have investment experience when the director has working experience as a fund manager, a general partner, or an executive officer in an investment company or as a private investor. In total, I am able to obtain working-experience information on $4,333$ fund directors in the mutual fund industry on December 2000 and July 2002.

[Place Figure 1 about here]

Figure 1 presents the change in the proportion of independent directors on a mutual fund board before and after the SEC amendment of 2001. 179 treated mutual funds significantly increase the averaged proportion of independent directors in a fund board from 47.0 percent to 67.9 percent, while other mutual funds in the control group increase from 74.5 percent to 75.8 percent only.[3] The sudden increase of treated mutual funds in board independence gives me confidence to estimate the effect of fund board independence which is largely free from endogeneity concerns. Using the exogenous shock in board independence, I analyze the effect of board independence on fund performance, conditional on investment experience of independent directors who exist in a fund board before the amendment announcement.

My approach to test for the change of fund performance following the enhanced board independence is to compare the difference in a fund performance measure around the 2001 amendment between treated and controlled mutual funds and examine how the difference depends on investment experience of independent directors. The difference-in-difference methodology, however, would be vulnerable to pre-existing heterogeneity between treatment and control groups because the heterogeneity may affect the estimated impact of independent boards. Therefore, the ex-ante difference between two groups should be controlled to ensure I am correctly identifying the effect of board independence. I take two ways to mitigate the concern. First, I add control variables into my regression models to reduce their effect in my estimation. Second, I use the propensity-score matching procedure to compare treated funds with similar controlled funds.

I estimate the following regression specification to empirically examine how mutual funds react to the exogenous improvement in board independency:

$$\text{Fund Performance}_i = \alpha + \beta_1 \text{After}_i + \beta_2 \text{Treat}_i + \beta_3 \text{After}_i \times \text{Treat}_i + \gamma' \mathbf{X}_i + \epsilon_i, \tag{1}$$

where for each mutual fund i, I use four measures for Fund Performance including fund return, portfolio holding return, the return gap between fund return and portfolio holding return, idiosyncratic risk of fund return, as detailed in Section II.B; "After" denotes a dummy variable that is equal to one on July 2002 and zero on December 2000; "Treat" denotes a dummy variable that indicates

[3]A usual way for non-compliant mutual funds to increase the proportion of independent directors in a fund board is to replace its insider director with an independent director and maintain the size of a fund board. The replacement does not, at least in my sample, cause a significant change in the portion of investment-experienced independent directors on a fund board around the amendment. That is because treated mutual funds usually elect (remove) an independent (inside) director who has working experience in a non-finance industry, rather than in an investment company.

if a mutual fund is treated or not; **X** denotes a vector of control variables including investment experience of independent and inside directors, board size, director age, director compensation, ownership of independent and inside directors, institutional ownership, fund age, fund total net asset, and expense ratio on December 2000. β_3 is the coefficient of interest which captures the effect of enhanced board independence on fund performance of treated funds relative to controlled funds.

I also test how investment experience of independent directors contributes to the effectiveness of board independence by using the following regression model:

$$\begin{aligned} \text{Fund Performance}_i &= \alpha + \beta_1 \text{After}_i + \beta_2 \text{Treat}_i + \beta_3 \text{After}_i \times \text{Treat}_i \\ &+ \beta_4 \text{After}_i \times \text{IIE}_i + \beta_5 \text{Treat}_i \times \text{IIE}_i + \beta_6 \text{After}_i \times \text{Treat}_i \times \text{IIE}_i \\ &+ \gamma' \mathbf{X}_i + \epsilon_i, \end{aligned} \tag{2}$$

where for each mutual fund i, IE denotes the portion of independent directors with investment experience on a fund board and other variables are the same as above. In the regression model, β_6 is the coefficient of interest which corresponds to the conditional effect of enhanced board independence on fund performance in treated funds relative to controlled funds.

I estimate several versions of equation (1) and (2) in this paper, including fund fixed effects to remove potential time-invariant factors which may influence the effect of the regulation shock on fund performance. The regressor of "After" in the above regression models accounts for time fixed effects because my empirical setting focuses on a single time period around the SEC amendment in 2001. Lastly, standard errors for all specifications are robust to heteroskedasticity and clustered at the fund level.

C. Data

C.1. Sample selection

This study employs three data sources to construct main sample. The Statement of Additional Information (SAI) is a source of board information for a mutual fund such as director names, director ownership, director compensation, if a director is independent and their principal occupations during the past five years. The SAI is the part B of the registration statement which a mutual fund is required to file with SEC on an annual basis. Another source is the Semi-Annual Report (N-SAR) for fund information about what kind of investment practice restrictions a mutual fund imposes on (Item 70) and the management fee on a contract between a mutual fund and a fund manager (Item 48). The SEC requires a mutual fund to report N-SAR on a semi-annual basis. Both the SAI and the N-SAR are available electronically from SEC's EDGAR database. The third and last source is CRSP Survival-Bias Free Mutual Fund Database (CRSP). I gather fund information such as expense ratio, actual 12b-1 fee, actual management fee, institutional ownership, fund age, fund total net asset (TNA), fund object, fund return and fund portfolio holding.

I hand-collect the above information from the SAI and N-SAR either in the last report of each mutual fund on SEC EDGAR database before January 2, 2001 when the SEC amendment is announced or in the first report after July 1, 2002 when the amendment becomes effective. I also use fund information of mutual funds listed on CRSP between January 1997 and June 2006. I merge the three data sources in two steps: First, I match the SAI and N-SAR by Central Index Key (CIK) and use a fund name in N-SAR (Item 7) to create fund-level observations. Second, I merge the EDGAR-sourced sample with CRSP database by matching a fund name manually.[4] I combine different share classes under the same fund name into a single fund by computing the

[4]The SAI and N-SAR report at a Central Index Key (CIK) level. CIK is a 10-digit number used on the SEC to identify a corporation or an individual who have filed disclosure with the SEC. It is too noisy to construct sample at fund level based on CIK because it could be either at a fund level or a fund complex level. So, I use a fund name provided by N-SAR to create fund-level observations after matching the SAI and N-SAR by CIK. CRSP records data at a share class level with the name of a share class (fund_name) in the following format: [Registrant company name]:[Fund name];[Share class type], e.g., BlackRock Index Funds, Inc: BlackRock Small Cap Index Fund; Investor A Shares. To combine the SAI and N-SAR with CRSP, I refer a fund name to the second part of the share class name. In case of a fund name is missing, I use registrant company name instead.

asset-weighted average of the class-level variables. For comparison reason, I require that each fund observation has non-missing data on board composition and monthly return on both December 2000 and July 2002.

C.2. Summary statistics and variable description

This section conveys summary statistics and brief description for variables of interest. Table A1 lists the definitions of the variables in detail.

[Place Table 19 about here]

Table 19 presents summary statistics of variables used in this paper at the fund level. The sample consists of 5,564 unique funds from 452 fund families, covering 4,333 directors. The sample funds include 179 (5,385) mutual funds in a treatment (control) group which didn't constitute (constituted) a majority of independent director in a board before the SEC amendment of 2001. The sample funds managed $5.50 trillion total assets in 2000, 79.0 percent of all U.S. mutual funds[5].

(1) Board information

The first part of Table 19 starts with the proportion of independent directors. Treated fund boards has experienced its increase from 0.47 to 0.68 by the SEC amendment of 2001. The difference between before and after the amendment is substantially significant, i.e., 0.21 with a t-statistic of 21.76. Controlled mutual funds, however, increase the proportion of independent board directors by 0.01 with t-statistics of 6.44.

Next, Table 19 reports descriptive statistics for working experience of directors. Directors' investment experience is about whether a director has working experience as a professional or private investor. A professional investor indicates that the director has worked in an investment company, and a private investor is for a director who is described as a private investor in the biographical information of the SAI. 26 percent of independent directors in the treatment group have investment experience before the 2001 amendment, and the portion of investment-experienced

[5]Investment Company Institute (ICI) reports the amount of total mutual fund assets is $6.96 trillion in 2000.

independent directors marginally falls to 23 percent after the 2001 amendment. The control group also has similar portion of independent directors with investment experience, 20 percent, and has no change across the grace period of the amendment. Most dependent directors has investment experience, regardless of treatment and control groups, because dependent directors are usually a fund manager or an executive of the fund complex.

In addition, both treated and controlled funds have about twenty percent of independent directors who has working experience in a finance industry except an investment company and about half of independent directors who do not have working experience in finance industry. Note that treated mutual funds have the portion of independent directors from a non-finance industry increased from 42 to 48 percent (t-stat=2.03) and the portion of dependent directors from a non-finance industry decreased from 10 to 6 percent (t-stat=2.58). Considering the board size is not changed in the adopting period, the statistical change suggests that treated funds usually comply the amendment by replacing its insider directors with independent directors who have working experience in a non-finance industry.

Independent chairmanship on a fund board has rarely been changed for both treatment and control groups in the SEC amendment: 12 percent of treated funds and 18 percent of controlled funds hold independent chairmanship on the board. The dollar value of fund shares owned by its directors is \$0.04 million of for treated funds and \$0.08 million for controlled funds.[6] The averaged age of directors across a fund board is around 55 and 60 for treated and controlled funds, respectively. The amount of compensation for directors is \$0.01 million and \$0.07 million for treated and controlled funds. Note that the huge gap in compensation between treated and controlled groups is because controlled funds are, on average, five times bigger than treated funds in terms of total net asset (TNA). The size difference between two groups may cause potential bias in difference-in-difference estimation, so I address this concern by using the propensity-score

[6]Table 19 reports directors' ownership in a fund they oversee only after the 2001 amendment because mutual funds start to document the information in the SAI on February 2001. For the rest of this paper, I assume that the director ownership before the amendment is the same as after the amendment. My rationale for the assumption is based on the finding of Chen, Goldstein, and Jiang (2008): Director ownership is virtually identical in 2002 and 2003. The finding suggests that the ownership of directors is likely constant over time. For the reason, I believe the constant assumption in director ownership does not cause serious issues in my analyses.

match in Section I.D.1.

(2) Performance information

The second part of Table 19 is about performance information. I employ four measures for fund performance including three return-related measures and two risk-related measures. I require four years of fund return history to compute the four performance measures. The first measure is the cumulative risk-adjusted fund return with respect to Fama-French (2015) five factors. The second measure is the return gap between fund return and portfolio return. I first compute the difference between quarterly cumulative FF5-adjusted fund return and the quarterly buy-and-hold FF5-adjusted return on a portfolio that invests in the most recently disclosed stock positions. Then, I take an average of the quarterly difference over four years before and after the amendment, i.e., 1997:Q1-2000:Q4 and 2002:Q3-2005:Q2. The third measure is the FF5-adjusted return on a portfolio disclosed most recently. Similar to the return gap, I compute the quarterly portfolio return first and average it over four years before and after the amendment. The fourth and last measure is about idiosyncratic risk of monthly fund return, calculated as the standard deviation of the residual from a regression of monthly fund returns on the Fama-French (2015) five-factor model.

The mean value of fund performance measures is almost the same for treated and controlled mutual funds except that idiosyncratic risk for treated funds is higher than for controlled funds. For example, the mean value of cumulative FF5-adjusted return is 0.03 for both groups of funds before the amendment, and it drops together to close to 0.00 after the amendment. However, the time trends for the measures are mixed: fund return, portfolio return, and idiosyncratic risk decrease, but return gap increase over time. The time trends can be partially explained by market-wide events like the dot-com bubble for 1994-2000 and the mutual fund scandal of 2003. The regressor of "After" in the equation (1) and (2) accounts for changes in the level of the U.S. captial markets over time. Note that the treated group has a lower t value for the difference between before and after the SEC amendment, indicating that there is greater variation across treated mutual funds in the impact of the amendment on fund performance. The number of observations for fund performance measures is always less than board information variables because I impose a four-year minimum in estimating the performance measures. Also, the number of observations for the return gap and portfolio return is less than for fund return and idiosyncratic risk because some mutual funds like bond funds do not have portfolio information.

(3) Fund information

The third and last part of Table 19 is about fund information. First, five versions of investment restriction measures are reported. The total number of restrictions on investment in the treated funds is 8.14 before the amendment, but it significantly fallss to 6.74 with t-statistics of 2.50 after the amendment. The control group also shows a similar pattern, but it relaxs, on average, 0.56 restrictions only. Following Almazan, *et al.* (2004), investment restriction items which can affect fund performance are decomposed into three categories: Derivative (Item 70B to Item 70I), Leverage (Item 70O, Item 70Q, and Item 70R), Illiquidity (Item 70J). The number of restrictions in the three categories are all significantly reduced after the amendment, particularly in the treatment group of mutual funds. I also construct a constraint score, termed as C-Score, proposed by Almazan, *et al.* (2004) and observe the same time trend as the other restriction measures: Greater reduction of a constraint score for treated funds than for controlled funds after the amendment. In this paper, I will mainly use the total number of investment practice restrictions as a proxy for the level of constraints a fund manager faces. Other restriction measures are not heavily discussed in the paper, but their test results are qualitatively the same as the reported results.

Then, I report statistics for fund fees including expense ratio, 12b-1 fee, management fee, and contractual management fee. Fund fee variables are constructed in two steps: I first obatin the amount of fund fees at the end of every month. Then, I take an average of the fund fees over four years before and after the amendment. By doing so, I can observe the change in the level of fund fees following the enhancement of board independency in 2001. I require four years of fund return history when computing the fee variables. Expense ratio and its detailed items are not significantly changed in the treatment group after the SEC amendment, while the control group of mutual funds raise the expense ratio with more spending on 12b-1 plans and other operating cost. Note that management fee provided by CRSP can be offset by fee waivers and reimbursements and, therefore, it can be lower than management fee on a contract, depending on the reaction of fund investors to their past performance. To idenfity the role of independent directors in adjusting the level of management fee, I gather information on contractual management fee from N-SAR, termed as K-mgmt Fee. The management fee on a contract marginally increases in treated funds, while it shows no change in controlled funds.

Other variables regarding fund information are following. There is no significant change in institutional ownership for both groups of mutual funds. Institutional ownership is calculated as the total net asset of institutional class in a fund divided by the total net asset of the fund. The treated mutual funds are about two years younger than the controlled. Lastly, treated funds are well-diversified: treated funds compose roughly sixty percent of equity funds, thity percent of fixed income funds, and ten percent of index funds. Controlled mutual funds also have well-diversified composition. Also, the fund objective has not been changed around the SEC amendment of 2001.

D. Results

In Section I.D.1, I confirm the averaged effect of an exogenous enhancement in board independence on fund performance as well as how its effect varies with investment experience of independent directors on a fund board. In Section I.D.2 and I.D.3, I explore additional difference-in-difference analyses to assess what causes the estimated treatment effect. Speicifically, I examine the change in investment contraints a fund manager faces and the change in the association of contractual management fee with fund performance. In Section I.D.4, I conduct several robustness tests.

D.1. Baseline

[Place Table 2 about here]

Table 2 presents the main results from the regression model of (1) and (2). My primary purpose is to examine how the improvement of board independence affects fund performance, depending on investment experience of independent directors: On averge, treated mutual funds should improve fund performance relative to non-treated mutual funds after the amendment when they have investment-experience independent directors on a fund board. I first test if enhanced board independence affects cumulative return and idiosyncratic risk unconditionally. Column 'Unconditional' of Table 2 reports the estimated coefficients of estimating equation of (1) without control variables in Panel A and with control variable in Panel B.[7] Consistent with the prior literature, I do not find a significant relation between board independence and performance, regardless of adding the control variables. For example, the difference-in-difference (DiD) coefficients, β_3, are statistically insignificant in Panel A: 0.007 for cumulative return (t-stat=0.22) and 0.020 for idiosyncratic risk (t-stat=0.64).

The next columns contain my cental results. In Column 'Conditional', I allow the effect of board independence to depend on investment experience of independent directors by introducing interaction terms with their experience. Recall that the investment experience of independent directors is the portion of independent directors on a fund board who have working experience as a professional or private investor. Two important findings emerge from these estimates. First, the estimated coefficient on the interaction term, β_6, is positive and significant for cumulative return. The conditional DiD coefficient in Column 'Conditional' is 0.314 with t-statistics of 2.71, indicating that the improvement in board independence among treated funds causes fund return to increase 1.98% per year when treated funds have independent directors with investment experience.[8] Even after adding control variables, the conditional coefficient is still positive and different from zero at high level of statistical significance. Second, the estimates reveal that the improved board independency has a negative impact on fund return for certain funds which do not have experienced independent directors on a board. The DiD coefficient, β_3, in the equation (2) is -0.074 with a t-value of -2.51 in Panel A and -0.054 with a t-value of -1.79 in Panel B. The

[7] I also replicate Table 2, removing fund fixed effect or restricting my sample to mutual funds which survive for entire estimation period from 1997 to 2006. The results are qualitatively the same as Table 2.

[8] The economic significance is calculated as follows. The average portion of experienced independent directors is 0.26 on a treated fund board, and the estimation period of cumulative fund return is four years. Therefore, the enhanced board independence is, on average, associated with 1.98% higher annual return ($= \sqrt[4]{1 + (0.314 \times 0.26)} - 1$) among treated funds with experienced independent directors.

outcome mirrors recent theories and empirical research that infer the cost of an independent board when it is informationally constrained.[9] I do not find a significant relation between idiosyncratic risk and the improvement in fund board independency, regardless of the presence of control variables.

Table 2 provides evidence that the increased proportion of independent directors has a material impact on fund performance, conditional on their investment experience. Independent directors seem to boost fund return when they have investment experience, and they damage when they do not have the experience, but their positive and negative impact cancel out on average. One interpretation of the evidence is that the working experience as a professional or private investor helps them to overcome the inherent information disadvantage of independent directors as outsiders and, therefore, they can govern fund management effectively. Prior studies have failed to figure out the substantial effect of board independence. One potential reason is that they have examined the unconditional effect of board independence which is statistically indifferent from zero in my sample. Another possible reason can be due to my empirical identification strategy. I employ the SEC amendment of 2001 as an exogenous shock in the degree of board independence among treated funds. Without using such an exogenous shock, evidence documented in previous literature may be subject to potential bias rooted in endogeneity of board composition and fund performance.

[Place Figure 2 about here]

I graphically illustrate the conditional effect of board indepedence in Figure 2. It plots cumulative return for treated (controlled) funds with versus without independent directors who has investment experience on a board. Consistent with Table 2, the left plot of Figure 2 reflects the increasing return difference between treated funds with and without investment-experience independent directors: -0.50% in the grace period of the amendment from January 2001 to June 2002 (t-stat=0.75), -0.50% in the first year after the amendment (t-stat=0.38), 2.83% in the second year (t-stat=1.95), 4.55% in the third year (t-stat=2.65), and 9.24% in the fourth year (t-stat=3.79). The

[9] Adams and Ferreira (2007) and Harris and Raviv (2008) provide theoretical models to prove that a board controlled by insiders, rather than independent directors, can be optimal for firm value due to their information advantage. Duchin, Matsusaka, and Ozbas (2010) empirically prove independent directors are less effective in worse information environment.

return difference between controlled funds has qualitatively the same time trend with much smaller variation, ranging from -0.40% in the grace period to 2.64% in the fourth year. Overall, this figure shows how investment experience of independent directors makes an economically significant increase in cumulative return of treated mutual funds relative to controlled mutual funds after the amendment.

Next, I repeat my baseline tests on portfolio-based performance measures such as the return gap and portfolio return. I note that the return gap can partially capture the trading ability of fund managers, and the return of a portfolio represents their stock-picking ability (Kacperczyk, Sialm, and Zheng, 2008). In line with the thought, I aim to clarify which ability is most prominently improved by the enhanced independency.

[Place Table 3 about here]

Table 3 presents the outcome of re-estimating equation (1) and (2) using the portfolio-based measures as dependent variables. The regression specifications are the same as Table 2, but I only report coefficients of interest for brevity's sake. Looking across Panel A and Panel B and focusing on the three-way interaction term in Column 'Conditional', the estimates indicate that the return gap is significantly increased following the independency enhancement, if a fund board has investment-experience independent directors and the return gap drops otherwise. One message from the table is that the trading ability of a fund manager, captured by the return gap, is primarily improved or damaged by the enhanced independency, depending on the presence of experienced independent directors on a fund board. These results suggest it is possible that the quality of trading environment managed by a fund board would depend on investment experience of independent directors. I will explore the possibility using the number of investment practice restrictions in Section I.D.2 and the association of management fee on a contract with fund return in Section I.D.3.

[Place Table 4 about here]

I implement a difference-in-difference matching estimator to address unobservable heterogeneity between the treatment and control groups of mutual funds. Specifically, I match each treated fund to a set of control funds using all the control variables in Table 2. First, I run a logit regression of an indicator variable for whether a particular fund is classified as treated or controlled on my matching variable on December 2000. The estimated coefficients from the logit regression are used to estimate probabilities of treatment for each fund in the sample. Second, these probabilities are used to perform a nearest neighbor match. I match with replacement using a standard tolerance of 0.005 caliper and allowing for up to four unique matches per treated fund. As a consequence, I can reconstruct my sample with 99 treated funds and 366 matched control funds.

Table 4 shows the results from the matching sample. Panel A displays the descriptive statistics for the treatment and matched control samples. The number of treated funds are 99, less than the total number of treated funds in my sample, i.e., 179, because they are required not to have any missing variables in the propensity matching analysis. Also, the number of matched control funds drops from 396 (= $99 * 4$) to 366 because some treated funds do not have four unique matches within the standard tolerance of 0.005 caliper. Importantly, the descriptive statistics in Panel A indicate that, at least for the variables I match on, treated funds are not statistically differentiable from matched control funds. This is a clear indication that the matching scheme performs well.

Panel B of Table 4 displays the impact of board independence on fund performance, conditional on independent directors' investment experience. The first and second rows present the DiD matching coefficient of β_3 in the equation (1) and β_6 in the equation (2), respectively. The estimators produces quantitatively similar estimated coefficients of the average treatment effect. The results indicate that my key findings in Table 2 and Table 3 are robust and not driven by heterogeneity between treatment and control groups.

D.2. Investment practice restriction

In this section, I examine whether board independence enhanced by the SEC amendment in 2001 reduces restrictions on investment activities, depending on investment experience of independent directors. A fund board of directors can use investment restrictions to fill the gap of their monitoring ability by prohibiting certain types of investment practices like investing in options or futures, purchasing restricted securities, borrowing of money, purchasing securities on margin, or short-selling (Almazan, *et al.*, 2004). Imposing such investment restrictions, however, may cause investment advisors to lose some chances of timely investment. Therefore, as the monitoring ability of independent directors is improved, they are less likely to bind investment advisors by restrictions. Independent directors with investment experience can be better informed and can more effectively monitor fund management because they are likely to have investment-specific knowledge, skills, and connections. Accordingly, when an independent board has independent directors with investment experience, the fund board is less likely to constrict the trading practices of a fund manager.

[Place Figure 3 about here]

Figure 3 illustrates the change of investment practice restrictions around the SEC amendment. I count the number of investment restrictions for a fund at the end of each year and take an average across four groups of mutual funds: a treatment (control) group of mutual funds with and without investment-experience independent directors. The figure depicts the time trend in the number of investment restriction items every year during pre-amendment period (1997-2000) and post amendment period (2002-2005). Table 19 reports that the treatment group of mutual funds decreases 1.40 out of eighteen investment restrictions in the grace period of the 2001 SEC amendment, while the control group decreases 0.56 restrictions only. Figure 3 shows the decrease of investment restrictions in the treatment group is mainly attributed to treated mutual funds with at least one independent director with investment experience. Specifically, mutual funds with experienced independent directors in the treatment group reduce the number of restricted investment activities by 2.74, while other treated mutual funds reduce it by 0.20. Similar to treated funds without experienced independent directors, controlled funds with and without experienced independent directors drop it by 0.86 and 0.26, respectively.

[Place Table 5 about here]

Table 5 presents the effects of the SEC amendment in 2001 on investment practice restrictions, conditional on investment experience of independent directors. I examine it by replacing fund performance measures with the number of restricted investment activities on the left-hand side of (1) and (2). Column 1 of Table 5 indicates that the overall effect of improved board independence is to ease constraints on fund managers, consistent with Almazan, *et al.* (2004). The estimated coefficient is -0.911 with a t-value of -2.17. Notably, the result of Column 3 indicates that the DiD coefficient, β_3, is insignificant but the conditional DiD coefficient, β_6, is negative and statistically significant. The estimates on the DiD terms in Column 3 are 0.784 for β_3 (t-stat=1.36) and -6.674 for β_6 (t-stat=-4.30), indicating that the negative effect is largely attributed to treated funds which has investment-experience independent directors on a fund board. After adding control variables, Column 2 and 4 of Table 5 also indicate that the fund board reduces the number of investment practice restrictions only when it has independent directors with investment experience, which is consistent with Figure 3. This confirms that independent directors with investment experience may provide fund managers better investment environment by relaxing investment restrictions.

D.3. Management fee on a contract

Next, I examine whether the association of contractual management fee with fund performance is changed by enhanced board independence with respect to investment experience of independent directors. Without market friction, all mutual funds are supposed to generate zero expected after-fee risk-adjusted returns in equilibrium. Otherwise, there would be positive (negative) cash flow for funds with positive (negative) expected after-fee risk-adjusted returns (Berk and Green, 2004). Experienced independent directors are better informed, so they can properly evaluate funds' alpha and fairly reflect into management fee. The board action made by independent directors can be directly measured in management fee on a contract. Therefore, I expect contractual management fee to be alingned more, or at least flatter, with fund performance when a fund board has independent directors with investment experience.

I first measure the relation between after-fee risk-adjusted fund return and management fee on a contract for four years before and after the SEC amendment in 2001 by using the following regression equation,

$$\text{FF5-Adjusted Fund Return}_{i,m} = \alpha_i + \beta_i \text{ Management Fee on a Contract}_{i,m} + \epsilon_{i,m}. \tag{3}$$

For example, I regress the risk-adjusted fund return in the first half of 1999 on contractual management fee reported in the last Semi-Annual Report (SAR) of 1998. I focus on the coefficient of interest, β_i, which should become more positive, or at least flatter, after the amendment for better-informed funds.

[Place Table 6 about here]

Table 6 presents the effects of the SEC amendment in 2001 on the association of contractual management fee with fund return, conditional on investment experience of independent directors. I examine it by using regression specification of (1) and (2) after replacing the dependent variable with β_i of equation (3). Column 1 and 2 of Table 6 indicates that the overall effect of improved board independence is to exacerbate the relation between fund performance and contractual management fee. The estimated coefficient is -0.087 with a t-value of -1.88. Notably, the result of Column 3 and 4 also shows that the DiD coefficient, β_3, is negative and significant at 10 percent significance, indicating the negative relation is largely attributed to treated funds which do not has investment-experience independent directors on a fund board. In sharp contrast, the conditional DiD coefficient, β_6, on Column 3 and 4 is positive. The estimates on the DiD terms in Column 3 are 0.144 (t-stat=0.69). After adding control variables, the conditional DiD terms in Column 4 is positive and even significant. These results suggests that the fund board sets the management fee on a contract properly reflecting the ability of a fund manager only when it has independent directors with investment experience.

D.4. Robustness tests

In this section, I report the results of several robustness exercises. First, I explore endogenous concerns regarding investment experience of independent directors, defined as the proportion of independent directors with investment experience in a fund board who exist in 2000 before the SEC amendment. It would be possible that investment experience of independent directors who are newly hired or resigned after 2000 may affect the effect of board independence on fund performance. If it is the case, one may interpret my main findings as meaning that fund performance improves in response to a change in the proportion of independent directors with investment experience, rather than an increase in the proportion of independent directors. To address that concern, I compare the effect of board independence, conditional on investment experience of independent directors who (1) hold the position, (2) get hired, and (3) are resigned in the grace period of the 2001 amendment.

[Place Table 7 about here]

The results are reported in Table 7. I define three variables first: the faction of independent directors on a fund board who who (1) hold the position, termed as 'Existing Indep Inv Exp', (2) get hired, termed as 'New Indep Inv Exp', and (3) are resigned, termed as 'Resigned Indep Inv Exp', in the grace period of the 2001 amendment. I then estimate my baseline model, allowing the treated effect to differ among these three groups of independent directors. The focal estimates indicate that the effect of the amendment is concentrated on independent directors who exist before the SEC amendment. For this group of independent directors, the estimated DiD coefficients for fund return and the return gap are positive and significant, regardless of control variables. This is not the case for independent directors who are resigned during the grace period. The estimated coefficient for newly elected independent directors is positive and significant for fund return only in Panel B. Thus, the test results suggest that my main findings are mainly driven by independent directors who keep holding the director position in the period, while newly hired independent directors marginally influence the effectiveness of board independence and resigned independent directors have statistically no influence.

Next, it would be possible that independent directors without investment experience may affect the effect of board independence on fund performance. For example, they can provide material inside information to fund managers, using their working experience in non-finance industries

(Dass, *et al.*, 2014). If it is the case, the interpretation of my findings is much complicated. Taking this concern into consideration, I compare the effectiveness of board independence, conditional on independent directors with (1) investment experience, (2) working experience in the finance industry except an investment company, and (3) working experience in non-finance industries.

[Place Table 8 about here]

The comparison results are reported in Table 8. Similar to the previous robustness test, I first construct three additional variables: the fraction of independent directors with with (1) investment experience, termed as 'Indep Inv Exp', (2) working experience in the finance industry except an investment company, termed as 'Indep Fin-But-Inv Exp', and (3) working experience in non-finance industries, termed as 'Indep Non-Fin Exp'. Then, I conduct the baseline analysis based on the three groups of independent directors. The point estimates indicate that the effect of improved board independence exclusively results from independent directors with investment experience. For this group of independent directors, the estimated DiD coefficients, β_6, for fund return and the return gap are positive and significant in both Panel A (Without control variable) and Panel B (With control variable). For the other groups of independent directors, the conditional DiD estimates are always negative for fund return and the return gap. Therefore, the results suggest that independent directors with investment experience play a major role to enhance the effectiveness of board independence, while independent directors with other types of working experience marginally hurt.

E. Conclusion

Independent directors play a vital role in fund management. Despite of the importance, prior studies have failed to find significant relation between board independence and fund performance so far. The SEC amendment of 2001 requires mutual funds to raise the minimum proportion of independent directors on a fund board from 40 percent to a majority. Taking advantage of the largely exogenous shock in board independency, I mitigate the endogeneity issues that has disturbed previous attempts to estimate the effect of board independence. My main finding is that the role of independent directors matters in fund performance, but the direction of their effect varies with their investment experience. Independent directors with investment experience can be better informed because they are likely to have knowledge on the nature of investment activities, skills to process financial information, and connections to provide full information for fund management. Consistent with the notion, I find that improved board independency is associated with significantly better fund performance when a fund board has independent directors with investment experience, while it is negatively associated with fund performance when a fund board does not have. These findings suggest that the failure of privious studies to find the impact of board independence on fund performance may result from the failure to separate the positive and negative effect of board independence.

The governance environment in the mutual fund industry has been changing from early 2000s when the amendment becomes effective. Independent directors can hire independent staffs who can closely observe the management of mutual funds and directly report to the directors on compliance matters as well as independent legal counsels which can provide adequate legal advice on the resolution of conflicts between fund shareholders and a fund manager. Those newly adopted rule may lessen the concern regarding the information constraint of independent directors. Nonetheless, most independent directors should still monitor multiple funds at quarterly meetings and a fund manager is frequently in a position to have a monopoly over information about fund operation. Therefore, although my findings are supposed to be interpreted with a caution, they still have significant implication on the role of board independence even in the recent period.

REFERENCES

[1] Adams, R. B., and D. Ferreira, 2007, A theory of friendly boards, Journal of Finance 62, 217-250.

[2] Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman, 2004, Why constrain your mutual fund manager?, Journal of Financial Economics 73, 289-321.

[3] Berk, Jonathan B, and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, Journal of Political Economy 112, 1269-1295.

[4] Chen, Q., I. Goldstein, and W. Jiang, 2008, Directors' ownership in the us mutual fund industry, Journal of Finance 63, 2629-2677.

[5] Cremers, M., J. Driessen, P. Maenhout, and D. Weinbaum, 2009, Does skin in the game matter? Director incentives and governance in the mutual fund industry, Journal of Financial and Quantitative Analysis 44, 1345-1373.

[6] Dass, Nishant, Omesh Kini, Vikram Nanda, Bunyamin Onal, and Jun Wang, 2014, Board expertise: Do directors from related industries help bridge the information gap?, The Review of Financial Studies 27, 1533-1592.

[7] Duchin, R., J. G. Matsusaka, and O. Ozbas, 2010, When are outside directors effective?, Journal of Financial Economics 96, 195-214.

[8] Ferris, Stephen P., and Xuemin Yan, 2007, Do independent directors and chairmen matter? The role of boards of directors in mutual fund governance, Journal of Corporate Finance 13, 392-420.

[9] Harris, M., and A. Raviv, 2008, A theory of board control and size, Review of Financial Studies 21, 1797-1832.

[10] Hermalin, Benjamin E., and Michael S. Weisbach, 1998, Endogenously chosen boards of directors and their monitoring of the ceo, The American Economic Review 88, 96-118.

[11] Investment Company Institute, 1999, Understanding the role of mutual fund directors, Washington, DC.

[12] Investment Company Institute, 2006, Understanding investor preferences for mutual fund information, Washington, DC.

[13] Investment Company Institute, 2017, Investment Company Fact Book: A review of trends and activities in the investment company industry, Washington, DC.

[14] Kacperczyk, M., C. Sialm, and L. Zheng, 2008, Unobserved actions of mutual funds, Review of Financial Studies 21, 2379-2416.

[15] Khorana, A., H. Servaes, and L. Wedge, 2007, Portfolio manager ownership and fund performance, Journal of Financial Economics 85, 179-204.

[16] Raheja, C. G., 2005, Determinants of board size and composition: A theory of corporate boards, Journal of Financial and Quantitative Analysis 40, 283-306.

[17] Tufano, P., and M. Sevick, 1997, Board structure and fee-setting in the us mutual fund industry, Journal of Financial Economics 46, 321-355.

Figure 1. **Proportion of Independent Fund Directors Before and After the SEC amendment of 2001** This figure illustrates the change in the proportion of independent directors on a fund board before and after the SEC amendment of 2001, including 95 percent confidence intervals on the top. The sample includes all mutual funds listed on SEC's EDGAR database on December 2000, termed as "Before", and July 2002, termed as "After", which have information on the CRSP Mutual Fund Database. This figure shows the amendment significantly increases the proportion of independent directors in treated mutual funds relative to controlled mutual funds.

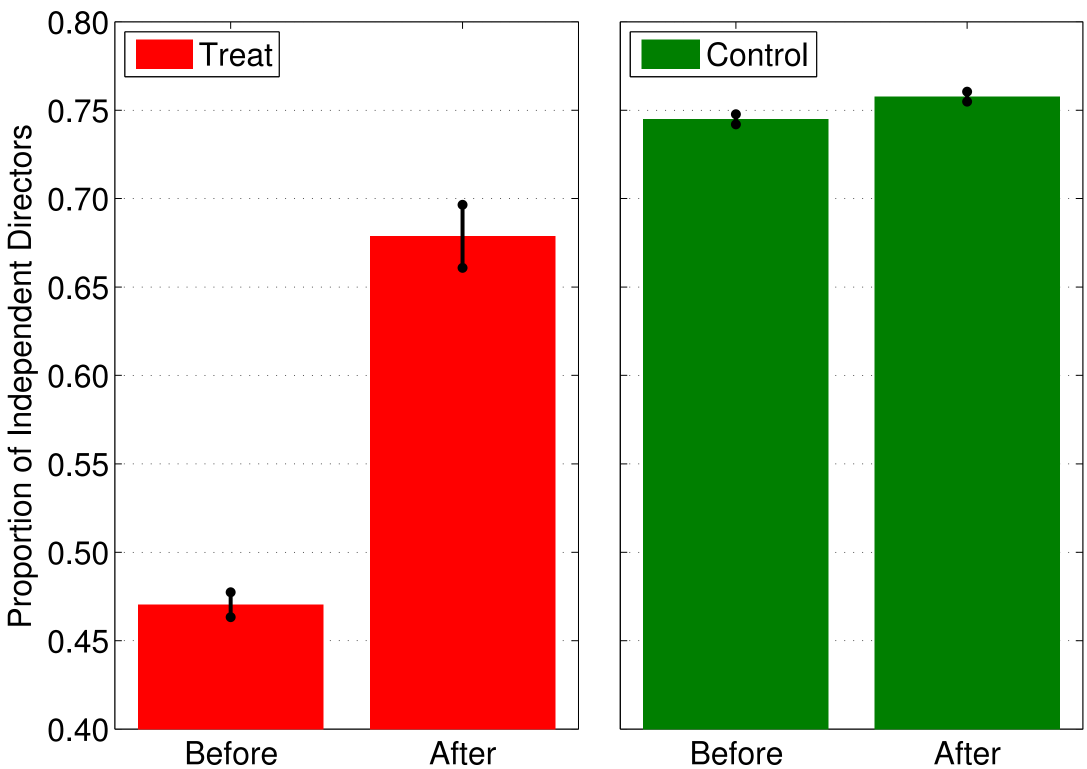

Figure 2. **Fund Cumulative Return Over Time** These figures plot fund cumulative return from January 2001 to June 2006 for four groups of mutual funds with 95 percent confidence intervals around them: The left (right) figure displays cumulative return for treated (controlled) mutual funds with versus without investment-experience independent directors on a board. The time period between January 2001 and June 2002 is the grace period of the SEC amendment. The sample includes all mutual funds listed on SEC's EDGAR database on December 2000 and July 2002 which have information on the CRSP Mutual Fund Database. This figure shows how investment experience of independent directors makes an economically meaningful increase in cumulative return of treated mutual funds relative to controlled mutual funds after the amendment.

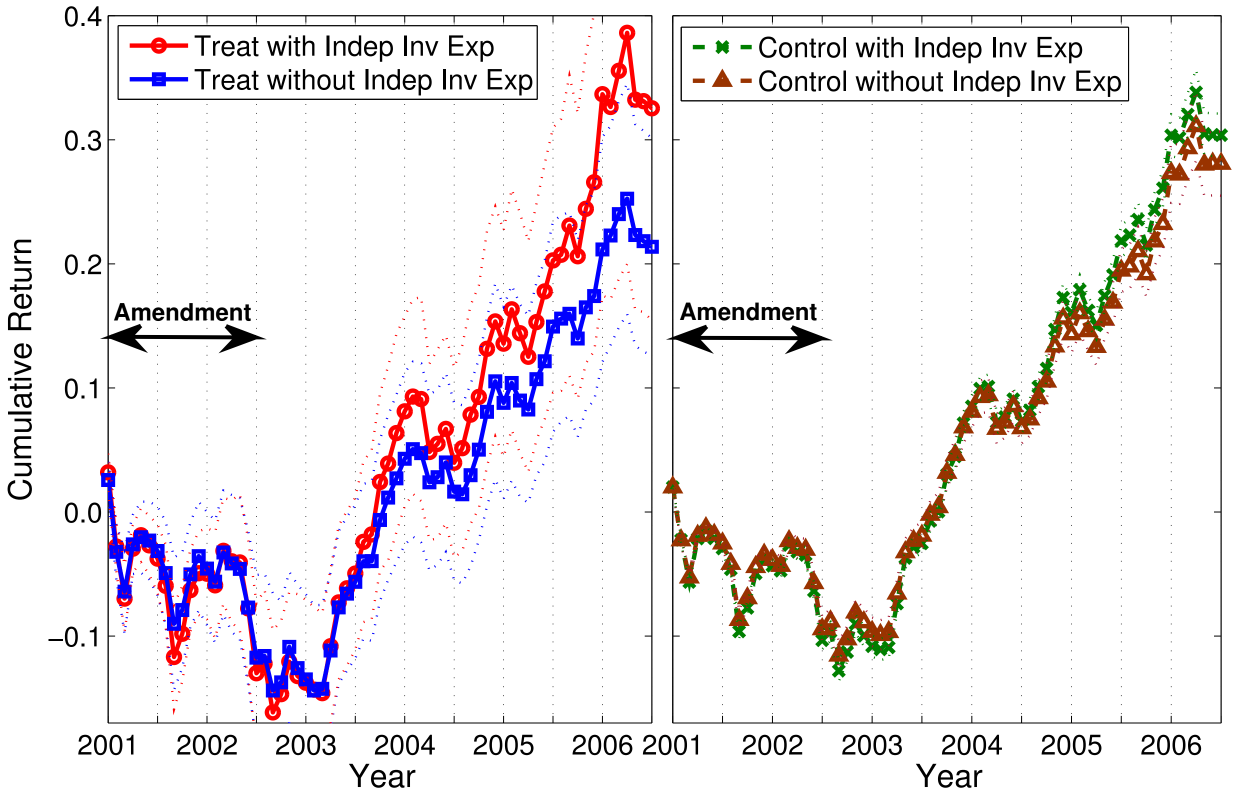

Figure 3. **Investment Practice Restriction Over Time** These figures depict the time trend in the number of investment restriction items. With 95 percent confidence intervals, I plot the average number of restriction items for 1997-2005 for treated (controlled) mutual funds with and without investment-experience independent director on a board. The time period between January 2001 and June 2002 is the grace period of the SEC amendment. The sample includes all mutual funds listed on SEC's EDGAR database on December 2000 and July 2002 which have information on the CRSP Mutual Fund Database.

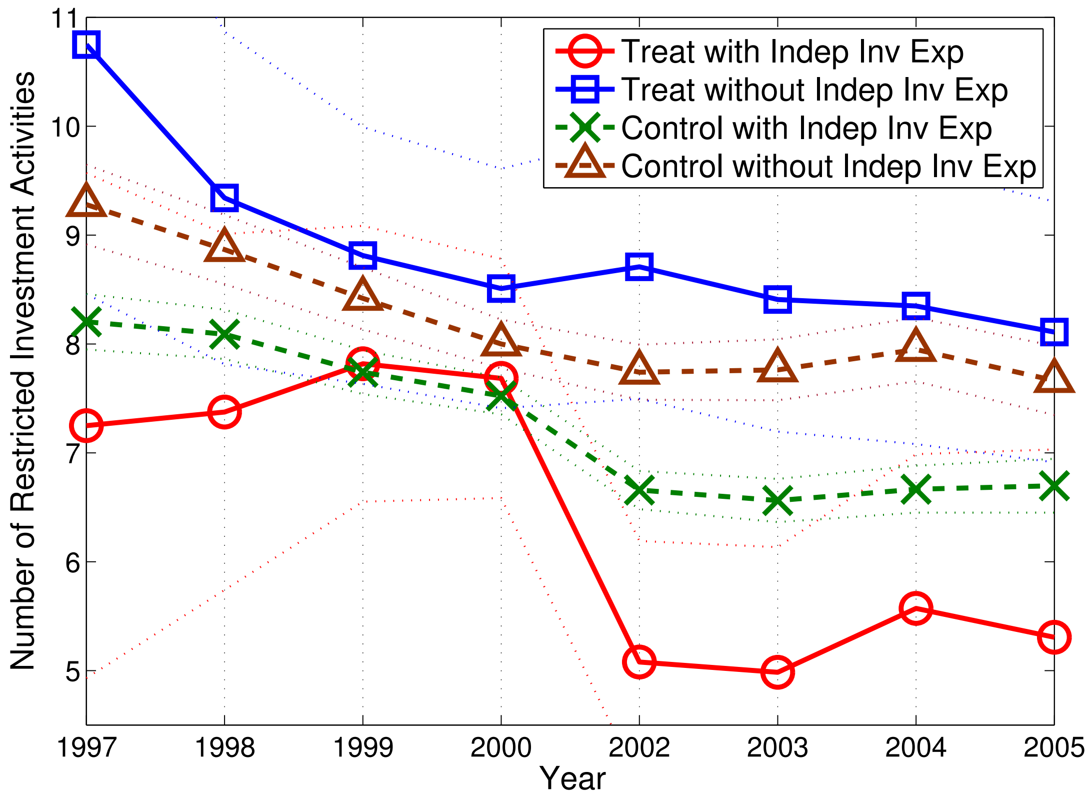

Table 1. Summary Statistics

This table presents summary statistics for treatment and control groups of mutual funds on December 2000 before the SEC amendment of 2001. The sample includes all mutual funds listed on SEC's EDGAR database on December 2000, termed as "Before", and July 2002, termed as "After", which have information on the CRSP Mutual Fund Database. Statistics for treated (controlled) funds are reported in the top (bottom) row for each variable. All variables are defined in Table A1. Stardard deviation and values at 25, 50, 70 percentiles are statistics before the amendment. t-statistic of difference in mean before and after the amendment is from a non-pair test assuming unequal variances.

	Mean									
Variable	#	Before	#	After	Diff	t	Std Dev	P25	P50	P75
(1) *Board Information*										
Indep Dir Proportion	179	0.47	179	0.68	0.21	(21.76)	0.05	0.43	0.50	0.50
	5,385	0.74	5,385	0.76	0.01	(6.44)	0.11	0.67	0.75	0.80
Dir Inv Experience										
- Indep Director	179	0.26	179	0.23	-0.03	(-1.08)	0.30	0.00	0.33	0.50
	5,385	0.20	5,385	0.20	0.00	(0.28)	0.19	0.00	0.20	0.33
- Dep Director	179	0.88	179	0.84	-0.04	(-1.46)	0.17	0.67	1.00	1.00
	5,385	0.83	5,385	0.86	0.04	(6.14)	0.31	0.67	1.00	1.00
Dir Fin-but-Inv Exp										
- Indep Director	179	0.21	179	0.23	0.03	(1.03)	0.24	0.00	0.00	0.50
	5,385	0.23	5,385	0.24	0.00	(0.98)	0.21	0.09	0.20	0.33
- Dep Director	179	0.02	179	0.02	0.00	(0.15)	0.08	0.00	0.00	0.00
	5,385	0.03	5,385	0.02	-0.01	(-6.92)	0.13	0.00	0.00	0.00
Dir Non-fin Exp										
- Indep Director	179	0.42	179	0.48	0.06	(2.03)	0.32	0.00	0.50	0.67
	5,385	0.50	5,385	0.51	0.01	(2.08)	0.25	0.33	0.50	0.67
- Dep Director	179	0.10	179	0.06	-0.04	(-2.58)	0.16	0.00	0.00	0.25
	5,385	0.07	5,385	0.07	-0.01	(-2.56)	0.20	0.00	0.00	0.00
Indep Chair	179	0.12	179	0.12	0.00	(0.00)	0.33	0.00	0.00	0.00
	5,385	0.18	5,385	0.18	-0.01	(-0.83)	0.39	0.00	0.00	0.00
Board Size (#)	179	5.46	179	5.46	0.00	(0.00)	1.83	4.00	6.00	6.00
	5,385	8.06	5,385	8.10	0.03	(0.62)	2.92	6.00	8.00	10.0
Indep Ownership (M$)	-	-	130	0.04	-	-	0.04	0.01	0.03	0.10
	-	-	4,673	0.08	-	-	0.03	0.06	0.08	0.10
Dep Ownership (M$)	-	-	130	0.08	-	-	0.12	0.03	0.10	0.10
	-	-	4,673	0.08	-	-	0.04	0.09	0.10	0.10
Director Age (Yr)	179	55.4	179	56.7	1.31	(1.96)	6.44	51.2	55.4	61.0
	5,385	60.6	5,385	61.1	0.50	(5.73)	4.67	58.1	60.9	63.9
Compensation (M$)	139	0.01	155	0.02	0.00	(1.74)	0.02	0.01	0.01	0.01
	4,676	0.07	4,741	0.09	0.02	(12.35)	0.06	0.02	0.06	0.10
(2) *Performance Information*										
FF5-Adj Cum Return	105	0.03	147	0.00	-0.03	(-1.18)	0.23	-0.05	0.02	0.09
	3,876	0.03	4,462	-0.01	-0.04	(-9.10)	0.19	-0.05	0.01	0.08
FF5-Adj Return Gap	70	0.02	98	0.02	0.01	(2.71)	0.02	0.01	0.01	0.02
	1,790	0.02	2,291	0.03	0.01	(10.46)	0.03	0.00	0.01	0.02
FF5-Adj Port Return	70	-0.01	98	-0.02	-0.01	(-4.10)	0.02	-0.02	-0.01	0.01
	1,790	-0.01	2,291	-0.02	-0.01	(-15.68)	0.03	-0.01	-0.01	0.00
FF5 Idio Risk	105	0.50	147	0.34	-0.16	(-3.68)	0.39	0.22	0.44	0.68
	3,876	0.41	4,462	0.28	-0.13	(-17.17)	0.39	0.14	0.27	0.58

(Continued)

Table 19 – *Continued*

Variable	#	Mean Before	#	Mean After	Diff	*t*	Std Dev	P25	P50	P75
(3) *Fund Information*										
Restriction (#)	161	8.14	163	6.74	-1.40	(-2.50)	4.93	3.00	8.00	13.0
	4,815	7.67	4,965	7.11	-0.56	(-5.75)	4.75	3.00	7.00	12.0
- Derivative (#)	161	4.20	163	3.43	-0.78	(-2.11)	3.37	1.00	5.00	8.00
	4,817	4.06	4,970	3.79	-0.28	(-4.42)	3.06	1.00	4.00	8.00
- Leverage (#)	161	1.72	163	1.45	-0.27	(-2.80)	0.75	1.00	2.00	2.00
	4,815	1.82	4,966	1.75	-0.07	(-4.70)	0.79	1.00	2.00	2.00
- Illiquidity (#)	161	0.26	163	0.18	-0.08	(-1.66)	0.44	0.00	0.00	1.00
	4,817	0.15	4,970	0.14	-0.01	(-1.34)	0.35	0.00	0.00	0.00
- C-Score	161	0.45	163	0.37	-0.09	(-2.67)	0.28	0.22	0.47	0.57
	4,815	0.42	4,966	0.40	-0.02	(-4.53)	0.25	0.24	0.39	0.56
Expense Ratio (%)	105	1.18	147	1.32	0.14	(1.43)	0.68	0.72	1.05	1.46
	3,872	1.04	4,460	1.09	0.05	(2.64)	0.64	0.65	0.95	1.34
- 12b-1 Fee (%)	105	0.09	147	0.11	0.02	(0.71)	0.17	0.00	0.00	0.08
	3,872	0.18	4,460	0.20	0.02	(2.86)	0.25	0.00	0.05	0.27
- Mgmt Fee (%)	105	0.58	147	0.48	-0.10	(-0.77)	0.66	0.40	0.70	0.94
	3,869	0.54	4,460	0.54	0.00	(0.55)	0.38	0.36	0.52	0.75
- K-Mgmt Fee (%)	90	0.71	125	0.77	0.07	(1.63)	0.28	0.50	0.75	0.83
	2,911	0.56	3,477	0.56	0.00	(-0.48)	0.32	0.38	0.53	0.75
Inst Ownership	179	0.26	179	0.27	0.00	(0.02)	0.43	0.00	0.00	0.65
	5,385	0.26	5,385	0.26	0.00	(-0.47)	0.42	0.00	0.00	0.69
Fund Age (Yr)	179	6.35	179	8.33	1.98	(2.55)	7.38	3.00	4.00	8.00
	5,385	8.43	5,385	10.35	1.92	(11.95)	8.40	3.00	6.00	11.0
Fund TNA (B$)	179	0.24	179	0.34	0.11	(0.75)	0.89	0.02	0.05	0.16
	5,385	1.01	5,385	0.96	-0.05	(-0.63)	3.92	0.04	0.15	0.56
Int'l Equity Fund	179	0.09	179	0.09	0.00	(0.00)	0.29	0.00	0.00	0.00
	5,385	0.12	5,385	0.12	0.00	(0.00)	0.32	0.00	0.00	0.00
Dom Equity Fund	179	0.53	179	0.53	0.00	(0.00)	0.50	0.00	1.00	1.00
	5,385	0.36	5,385	0.36	0.00	(0.00)	0.48	0.00	0.00	1.00
Fix Income Fund	179	0.30	179	0.30	0.00	(0.00)	0.46	0.00	0.00	1.00
	5,385	0.43	5,385	0.43	0.00	(0.00)	0.50	0.00	0.00	1.00
Mixed Fund	179	0.07	179	0.07	0.00	(0.00)	0.26	0.00	0.00	0.00
	5,385	0.07	5,385	0.07	0.00	(0.00)	0.26	0.00	0.00	0.00
Index Fund	179	0.08	179	0.08	0.00	(0.00)	0.28	0.00	0.00	0.00
	5,385	0.04	5,385	0.04	0.00	(0.00)	0.20	0.00	0.00	0.00

Table 2. Fund Performance on Board Independence

This table presents the effect of the SEC amendment in 2001 on fund performance, conditional on investment experience of independent directors, using the regression specifications of (1) and (2). The sample includes all mutual funds listed on SEC's EDGAR database on December 2000 and July 2002, which have information on the CRSP Mutual Fund Database. All variables are defined in Table A1. Panel A and Panel B report the coefficient estimates from the regression models without and with control variables, respectively. Standard errors are reported in parentheses below the estimated coefficient and robust to heteroskedasticity and clustered at fund level. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Without control variable

	Unconditional		Conditional	
	Cum Return	Idio Risk	Cum Return	Idio Risk
Treat × After	0.007	0.020	-0.074**	0.055
	(0.22)	(0.64)	(-2.51)	(1.37)
Treat	-0.000	0.117***	0.016*	0.101***
	(-0.00)	(7.27)	(1.86)	(5.20)
After	-0.035***	-0.081***	-0.038***	-0.084***
	(-7.62)	(-14.02)	(-5.92)	(-10.36)
Treat × After × Indep Inv Exp			0.314***	-0.104
			(2.71)	(-0.79)
Treat × Indep Inv Exp			-0.060**	-0.174**
			(-2.18)	(-2.48)
After × Indep Inv Exp			0.013	-0.015
			(0.50)	(-0.51)
Indep Inv Exp			0.009*	0.233***
			(1.86)	(21.18)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.012	0.031	0.014	0.085
Num of Observations	8,590	8,590	8,590	8,590

Table 2 – *Continued*

Panel B. With control variable	Unconditional		Conditional	
	Cum Return	Idio Risk	Cum Return	Idio Risk
Treat × After	0.033	-0.016	-0.054*	0.001
	(0.92)	(-0.44)	(-1.79)	(0.02)
Treat	-0.017**	0.014	-0.007	-0.015
	(-2.30)	(0.68)	(-0.73)	(-0.65)
After	-0.030**	-0.108***	-0.040***	-0.113***
	(-2.57)	(-12.36)	(-3.21)	(-10.83)
Treat × After × Indep Inv Exp			0.392**	-0.085
			(2.35)	(-0.42)
Treat × Indep Inv Exp			-0.050	0.139
			(-1.30)	(1.18)
After × Indep Inv Exp			0.054*	0.022
			(1.75)	(0.68)
Indep Inv Exp	0.004	0.064***	0.003	0.055***
	(0.55)	(3.86)	(0.34)	(3.44)
Dep Inv Exp	0.004	0.059***	0.004	0.061***
	(1.32)	(9.98)	(1.35)	(10.17)
Indep Chair	0.002	0.007	0.002	0.008
	(0.81)	(1.17)	(0.78)	(1.36)
Board Size	-0.000	0.001	-0.000	0.001
	(-0.89)	(1.07)	(-0.71)	(0.97)
Dir Age	0.000	0.001	0.000	0.001
	(0.98)	(0.92)	(0.82)	(0.91)
Compensation	-0.013	-0.123	-0.028	-0.131
	(-0.18)	(-1.47)	(-0.39)	(-1.56)
Indep Ownership	-0.115***	-0.139*	-0.111***	-0.134*
	(-2.72)	(-1.92)	(-2.60)	(-1.85)
Dep Ownership	0.061*	-0.027	0.059	-0.028
	(1.71)	(-0.48)	(1.57)	(-0.50)
Inst Ownership	0.005*	0.010	0.005**	0.011
	(1.93)	(1.51)	(1.97)	(1.56)
Fund Age	-0.003	0.013***	-0.003	0.013***
	(-0.59)	(2.84)	(-0.57)	(2.91)
Fund TNA	0.013***	0.012***	0.013***	0.012***
	(4.11)	(4.74)	(4.09)	(4.71)
Expense Ratio	0.003	0.034***	0.002	0.034***
	(1.14)	(4.00)	(1.08)	(3.98)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.016	0.123	0.021	0.123
Num of Observations	6,808	6,808	6,808	6,808

Table 3. Portfolio-Related Fund Performance and Board Independence

This table presents the effect of the SEC amendment in 2001 on portfolio-related performance, conditional on investment experience of independent directors, using the regression specifications of (1) and (2). The sample includes all mutual funds listed on SEC's EDGAR database on December 2000 and July 2002, which have information on the CRSP Mutual Fund Database. All variables are defined in Table A1. For brevity, I report difference-in-difference estimators only. Panel A and Panel B report the coefficient estimates from the regression models without and with control variables, respectively. Standard errors are reported in parentheses below the estimated coefficient and robust to heteroskedasticity and clustered at fund level. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

	Unconditional		Conditional	
	Ret Gap	Port Ret	Ret Gap	Port Ret
Treat × After	0.001	-0.002	-0.006	-0.003
	(0.23)	(-0.53)	(-1.59)	(-0.62)
Treat	0.005***	-0.003***	0.006***	-0.004***
	(4.84)	(-2.99)	(4.71)	(-2.69)
After	0.014***	-0.016***	0.014***	-0.015***
	(16.80)	(-19.99)	(12.61)	(-14.42)
Treat × After × Indep Inv Exp			0.028**	0.003
			(2.29)	(0.20)
Treat × Indep Inv Exp			-0.020***	0.011**
			(-5.26)	(2.48)
After × Indep Inv Exp			-0.003	-0.001
			(-0.93)	(-0.29)
Indep Inv Exp			0.016***	-0.010***
			(13.47)	(-9.90)
Control Variable	No	No	No	No
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.088	0.122	0.118	0.134
Num of Observations	4,249	4,249	4,249	4,249

Panel B. With control variable

	Unconditional		Conditional	
	Ret Gap	Port Ret	Ret Gap	Port Ret
Treat × After	-0.002	0.003	-0.011**	0.002
	(-0.55)	(0.63)	(-2.32)	(0.42)
Treat	-0.002	-0.000	-0.001	-0.001
	(-1.30)	(-0.14)	(-0.43)	(-0.67)
After	0.010***	-0.013***	0.010***	-0.013***
	(4.57)	(-7.52)	(4.27)	(-6.81)
Treat × After × Indep Inv Exp			0.032**	0.002
			(2.56)	(0.14)
Treat × Indep Inv Exp			-0.003	0.003
			(-0.64)	(0.59)
After × Indep Inv Exp			0.002	-0.002
			(0.41)	(-0.39)
Indep Inv Exp	0.005***	-0.003*	0.005***	-0.003*
	(2.98)	(-1.72)	(2.73)	(-1.67)
Control Variable	Yes	Yes	Yes	Yes
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.131	0.145	0.132	0.145
Num of Observations	3,339	3,339	3,339	3,339

Table 4. Propensity Score Matching Estimation

This table documents descriptive statistics in the year of 2000 before the 2001 regulation change in Panel A and average treatment effects in Panel B for samples matched by logit propensity score. Treatment sample consists of 99 mutual funds with valid matching variables. Treated mutual funds are matched to at most four unique mutual funds in control sample by using a nearest neighbor logit propensity score match with a 0.005 caliper. 366 mutual funds in control sample are matched on variables listed in Panel A. All variables are defined in Table A1. Standard errors are reported in parentheses below the estimated coefficient and robust to clustering at fund level in Panel B. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Descriptive Statistics in Matched Samples (# of Treat: 99; # of Control: 366)

	Treatment		Matched Control			
Variable	Mean	Std Dev	Mean	Std Dev	Diff	*t*
Indep Inv Exp	0.23	0.27	0.22	0.23	0.01	(0.44)
Dep Inv Exp	0.90	0.16	0.87	0.29	0.03	(1.43)
Indep Chair	0.16	0.37	0.12	0.33	0.04	(0.91)
Board Size	5.63	2.25	5.83	2.06	-0.20	(-0.79)
Indep Ownership	0.05	0.04	0.05	0.03	0.00	(-0.50)
Dep Ownership	0.07	0.04	0.07	0.04	0.00	(-0.89)
Director Age	57.10	5.12	57.43	6.58	-0.33	(-0.51)
Compensation	0.01	0.02	0.02	0.02	0.00	(-0.58)
Inst Ownership	0.25	0.41	0.28	0.44	-0.03	(-0.69)
Expense Ratio	1.18	0.66	1.10	0.58	0.08	(1.12)
Fund Age	7.66	8.16	7.43	7.81	0.24	(0.25)
Fund TNA	0.32	1.18	0.51	5.12	-0.19	(-0.60)

Panel B. Average Treatment Effect for Matched Samples

Focal Control Variable	Cum Ret	Ret Gap	Port Ret	Idio Risk
Treat × After	0.055	-0.003	0.006	-0.033
(Unconditional)	(1.21)	(-0.65)	(1.35)	(-0.69)
Treat × After × Indep Inv Exp	0.350*	0.043***	-0.011	-0.135
(Conditional)	(1.68)	(2.67)	(-0.60)	(-0.54)

Table 5. Investment Practice Restriction and Board Independence

This table presents the effect of the SEC amendment in 2001 on investment practice restriction, conditional on investment experience of independent directors, using the regression specifications of (1) and (2). The sample includes all mutual funds listed on SEC's EDGAR database on December 2000 and July 2002, which have information on the CRSP Mutual Fund Database. For brevity, I report difference-in-difference estimators only. All variables are defined in Table A1. Standard errors are reported in parentheses below the estimated coefficient and robust to heteroskedasticity and clustered at fund level. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
Treat × After	-0.911**	0.498	0.784	1.445**
	(-2.17)	(1.04)	(1.36)	(2.06)
Treat	3.464***	1.401**	3.597***	1.476*
	(8.59)	(2.47)	(6.55)	(1.88)
After	-0.389***	-0.457	-0.412***	-0.464
	(-5.38)	(-1.35)	(-4.03)	(-1.32)
Treat × After × Indep Inv Exp			-6.674***	-4.420***
			(-4.30)	(-2.66)
Treat × Indep Inv Exp			-7.715***	-0.299
			(-5.28)	(-0.14)
After × Indep Inv Exp			-0.125	0.025
			(-0.35)	(0.06)
Indep Inv Exp		0.026	7.083***	0.045
		(0.07)	(27.99)	(0.12)
Control Variable	No	Yes	No	Yes
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.013	0.247	0.108	0.247
Num of Observations	10,104	7,845	10,104	7,845

Table 6. Contractual Management Fee and Board Independence

This table presents the effect of the SEC amendment in 2001 on the association of contractual management fee with fund performance, conditional on investment experience of independent directors, using the regression specifications of (1) and (2). The sample includes all mutual funds listed on SEC's EDGAR database on December 2000 and July 2002, which have information on the CRSP Mutual Fund Database. For brevity, I report difference-in-difference estimators only. All variables are defined in Table A1. Standard errors are reported in parentheses below the estimated coefficient and robust to heteroskedasticity and clustered at fund level. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

	(1)	(2)	(3)	(4)
Treat × After	-0.087*	-0.102*	-0.095*	-0.138*
	(-1.88)	(-1.84)	(-1.69)	(-1.94)
Treat	0.096**	0.075	0.102*	0.101*
	(2.17)	(1.47)	(1.92)	(1.69)
After	0.024	0.036	0.028	0.051
	(1.41)	(1.44)	(1.17)	(1.45)
Treat × After × Indep Inv Exp			0.144	0.514*
			(0.69)	(1.83)
Treat × Indep Inv Exp			-0.154	-0.449*
			(-0.75)	(-1.72)
After × Indep Inv Exp			-0.018	-0.070
			(-0.27)	(-0.75)
Indep Inv Exp		0.047	-0.001	0.096
		(0.77)	(-0.02)	(0.91)
Control Variable	No	Yes	No	Yes
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.000	0.037	-0.004	0.035
Num of Observations	1,164	909	1,164	909

Table 7. Investment Experience of Existing/New/Resigned Independent Directors

This table explores the effect of board independence, conditional on investment experience of independent directors who (1) hold the position, (2) get hired, and (3) are resigned in the grace period of the 2001 amendment. I replicate the regression analysis of 2 as in Table 2 and Table 3, adding the three types of investment experience. For brevity, I report difference-in-difference estimators only. All variables are defined in Table A1. Panel A and Panel B report the coefficient estimates from the regression models without and with control variables, respectively. Standard errors are reported in parentheses below the estimated coefficient and robust to heteroskedasticity and clustered at fund level. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Without control variable	Cum Ret	Ret Gap	Port Ret	Idio Risk
Treat × After × Existing Indep Inv Exp	0.503***	0.039***	0.002	-0.158
	(2.69)	(2.66)	(0.14)	(-0.77)
Treat × After × New Indep Inv Exp	0.221	0.029	-0.035	0.331
	(0.80)	(1.37)	(-1.17)	(1.05)
Treat × After × Resigned Indep Inv Exp	0.109	0.025	0.010	-0.031
	(0.72)	(1.40)	(0.55)	(-0.21)
Existing Indep Inv Exp	0.006	0.016***	-0.011***	0.210***
	(1.12)	(10.83)	(-8.32)	(16.56)
New Indep Inv Exp	0.043***	0.012***	-0.006**	0.230***
	(2.80)	(4.00)	(-2.34)	(6.65)
Fired Indep Inv Exp	-0.009	0.013***	-0.006*	0.234***
	(-0.53)	(3.25)	(-1.94)	(5.88)
Control Variable	No	No	No	No
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.014	0.122	0.137	0.091
Num of Observations	8,590	4,249	4,249	8,590

Panel B. With control variable	Cum Ret	Ret Gap	Port Ret	Idio Risk
Treat × After × Existing Indep Inv Exp	0.478**	0.037**	-0.000	-0.070
	(2.29)	(2.16)	(-0.00)	(-0.30)
Treat × After × New Indep Inv Exp	0.713***	0.018	0.027	0.546
	(2.75)	(0.58)	(0.75)	(1.08)
Treat × After × Resigned Indep Inv Exp	-0.031	0.027	-0.006	-0.118
	(-0.15)	(1.41)	(-0.29)	(-0.43)
Existing Indep Inv Exp	0.001	0.006***	-0.005**	0.044**
	(0.10)	(2.93)	(-2.46)	(2.56)
New Indep Inv Exp	0.051***	0.002	0.001	0.056
	(2.73)	(0.46)	(0.32)	(1.54)
Fired Indep Inv Exp	0.005	0.003	0.003	0.091**
	(0.20)	(0.66)	(0.79)	(2.04)
Control Variable	Yes	Yes	Yes	Yes
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.022	0.131	0.145	0.123
Num of Observations	6,808	3,339	3,339	6,808

Table 8. Investment/Finance-But-Investment/Non-Finance Industry Experience

This table explores the effectiveness of board independence, conditional on independent directors with (1) investment experience, (2) working experience in the finance industry except an investment company, and (3) working experience in non-finance industries. I replicate the regression analysis of 2 as in Table 2 and Table 3, adding the three types of working experience. For brevity, I report difference-in-difference estimators only. All variables are defined in Table A1. Panel A and Panel B report the coefficient estimates from the regression models without and with control variables, respectively. Standard errors are reported in parentheses below the estimated coefficient and robust to heteroskedasticity and clustered at fund level. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A. Without control variable				
	Cum Ret	Ret Gap	Port Ret	Idio Risk
Treat × After × Indep Inv Exp	0.268**	0.020*	0.001	-0.075
	(2.52)	(1.96)	(0.11)	(-0.67)
Treat × After × Indep Fin-But-Inv Exp	-0.126	-0.009	-0.008	0.097
	(-1.54)	(-1.12)	(-0.94)	(0.99)
Treat × After × Indep Non-Fin Exp	-0.111*	-0.005	-0.002	0.070
	(-1.70)	(-0.72)	(-0.22)	(1.23)
Indep Inv Exp	0.002	0.004***	-0.002**	0.058***
	(0.30)	(3.00)	(-1.96)	(4.10)
Indep Fin-But-Inv Exp	-0.000	0.002	-0.000	0.025**
	(-0.04)	(1.26)	(-0.01)	(2.12)
Indep Non-Fin Exp	-0.004	0.002	-0.002*	0.019**
	(-0.98)	(1.56)	(-1.80)	(2.52)
Control Variable	No	No	No	No
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.013	0.143	0.145	0.126
Num of Observations	8,590	4,249	4,249	8,590
Panel B. With control variable				
	Cum Ret	Ret Gap	Port Ret	Idio Risk
Treat × After × Indep Inv Exp	0.375**	0.021**	0.005	-0.117
	(2.44)	(2.05)	(0.35)	(-0.65)
Treat × After × Indep Fin-But-Inv Exp	-0.179*	-0.021*	0.002	-0.092
	(-1.91)	(-1.90)	(0.19)	(-0.87)
Treat × After × Indep Non-Fin Exp	-0.049	-0.010	0.006	0.061
	(-0.75)	(-1.45)	(0.83)	(0.94)
Indep Inv Exp	0.004	0.005***	-0.003*	0.060***
	(0.40)	(2.75)	(-1.71)	(3.73)
Indep Fin-But-Inv Exp	-0.004	0.001	-0.000	0.014
	(-0.56)	(0.80)	(-0.17)	(1.04)
Indep Non-Fin Exp	0.000	0.002	-0.001	0.027***
	(0.05)	(1.31)	(-1.01)	(2.77)
Control Variable	Yes	Yes	Yes	Yes
Fund Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R-Square	0.020	0.135	0.145	0.127
Num of Observations	6,808	3,339	3,339	6,808

Table A1. Variable Definition

Variable	Definition	Source
(1) *Board Information*		
Indep Dir Proportion	The proportion of independent directos on a fund board.	EDGAR: Form N-1A
Dir Inv Experience		EDGAR: Form N-1A
- Indep Inv Experience	The proportion of independent directors on a fund board who have professional investment experience in an investment company as an officer, an employee, or a general partner or private investment experience as a private investor.	EDGAR: Form N-1A
- Dep Inv Experience	The proportion of dependent directors on a fund board who have professional investment experience in an investment company as an officer, an employee, or a general partner or private investment experience as a private investor.	EDGAR: Form N-1A
Dir Non-inv Experience		EDGAR: Form N-1A
- Indep Non-inv Experience	The proportion of independent directors on a fund board who have experience in an investment company but in the finance industry	EDGAR: Form N-1A
- Dep Non-inv Experience	The proportion of independent directors on a fund board who have experience in an investment company but in the finance industry	EDGAR: Form N-1A
Dir Non-fin Experience		EDGAR: Form N-1A
- Indep Non-fin Experience	The proportion of independent directors on a fund board who have experience neither in an investment company nor in the finance industry	EDGAR: Form N-1A
- Dep Non-fin Experience	The proportion of independent directors on a fund board who have experience neither in an investment company nor in the finance industry	EDGAR: Form N-1A
Indep Chair	A dummy variable equal to a value one if a chairman is an independent director.	EDGAR: Form N-1A
Board Size (#)	The number of directors assigned to a board.	EDGAR: Form N-1A
Dir Ownership (M$)	The amount of mutual fund shares in millions of dollars invested by directors assigned to a board. I report the average of director's investment in a fund. Form N-1A lists director's investment in the following dollar ranges: none, $1-10,000, $10,001-50,000, $50,001-100,000, over $100,000. I set a director's investment in a fund as the midpoint of the listed ranges except 'none' and 'over $100,000'.	EDGAR: Form N-1A
Director Age (Yr)	Age of a director	EDGAR: Form N-1A
Compensation (M$)	Compensation in million dollars from a fund complex	EDGAR: Form N-1A

(Continued)

Table A1 – *Continued*

Variable	Definition	Source
(2) *Performance Information*		
FF5-Adj Cum Ret	Cumulative risk-adjusted fund returns with respect to the Fama-French 5-factors over four years before 2000 or after July 2002.	CRSP Mutual Fund
FF5-Adj Return Gap	I first compute the difference between quarterly cumulative FF5-adjusted fund return and the quarterly buy-and-hold FF5-adjusted return on a portfolio that invests in the most recently disclosed stock positions. Then, I take an average of the quarterly difference over four years before and after the amendment.	CRSP Mutual Fund
FF5-Adj Port Ret	Similar to the return gap, I compute the quarterly portfolio return first and average it over four years before and after the amendment.	CRSP Mutual Fund
FF5 Idio Risk	The standard deviation of the residual from a regression of monthly fund returns on the Fama-French 5-factor model. This measure is estimated over four years before 2000 or after July 2002.	CRSP Mutual Fund
(3) *Fund Information*		
Restriction (#)	The number of investment restrictions stated in Question 70.	EDGAR: Form N-SAR
- Derivative (#)	The number of investment restrictions stated in Question 70B to 70I.	EDGAR: Form N-SAR
- Leverage (#)	The number of investment restrictions stated in Question 70O, 70Q, and 70R.	EDGAR: Form N-SAR
- Illiquidity (#)	The number of investment restrictions stated in Question 70J.	EDGAR: Form N-SAR
- C-Score (#)	Constraint score proposed by Almazan *et al.* (2004)	EDGAR: Form N-SAR
Expense (%)	The percentage of expense in fund total net assets.	CRSP Mutual Fund
- 12b-1 Fee (%)	The percentage of actual 12b-1 fee in fund total net assets.	CRSP Mutual Fund
- Mgmt Fee (%)	The percentage of actual management fee in fund total net assets.	CRSP Mutual Fund
- K-Mgmt fee (%)	The percentage of contractual management fee in fund total net assets. It is reported in Question 48, 48A to 48K. Question 48 reports the management fee rate when it is a single rate fee. Otherwise, it reports management fees in each bracket, Question 48A to 48K, with certain ranges. In case, I take a range-weighted average of management fees across brackets.	EDGAR: Form N-SAR
Inst Ownership	The proportion of fund assets in institutional share classes to fund total net assets.	CRSP Mutual Fund
Fund Age (Yr)	The number of years since fund inception date.	CRSP Mutual Fund
Fund TNA (B$)	Fund total net assets in billion dollars	CRSP Mutual Fund
Dom Equity Fund	A dummy variable equal to a value one if a fund is a domestic equity fund.	CRSP Mutual Fund
Int'l Equity Fund	A dummy variable equal to a value one if a fund is an international equity fund.	CRSP Mutual Fund
Fix Income Fund	A dummy variable equal to a value one if a fund is a fixed income fund.	CRSP Mutual Fund
Mixed Fund	A dummy variable equal to a value one if a fund is mixed with equity and fixed income.	CRSP Mutual Fund
Index Fund	A dummy variable equal to a value one if a fund is an index fund.	CRSP Mutual Fund

II. How Smart Is Institutional Trading?

A. Introduction

Institutional investors are at the center of academic finance research as the professional asset management business continues to see unprecedented growth in the United States and worldwide. Many researchers study institutional investors' stock picking skills by analyzing their performance using publicly available low frequency data such as quarterly institutional holdings or monthly fund returns.[10] The low frequency data, however, are not able to capture the exact timing of institutional trades. Therefore, it remains challenging to examine the role of institutional trading in price discovery, the corresponding transactions cost, or how institutional investors may trade based on short-lived information. To circumvent limitations on institutional trading data, some studies estimate institutional order flow (IOF, hereafter) using high frequency data on the premise that institutions are more likely to place large orders, e.g. Lee (1992), Lee and Radhakrishna (2000), Battalio and Mendenhall (2005), Malmendier and Shanthikumar (2007), and Hvidkjaer (2008). However, as order splitting becomes common, arising from a concern for trading costs, the reliability of these size-based algorithms has been cast under serious doubt in recent sample periods, as pointed out by Cready, Kumas, and Subasi (2014). Alternatively, several recent studies employ proprietary data sets of institutional trading records for empirical analysis.[11] However, these proprietary data have limited coverage and can contain sample biases that impair external validity. In this article, we present a new method for estimating *aggregate* order flow at the daily level for institutional investors and for subgroups of institutions only.

[10]For example, Carhart (1997), Wermers (2000), Chen, Jegadeesh, and Wermers (2000), Barras, Scaillet, Wermers (2010), and Edelen, Ince, and Kadlec (2016) find weak evidence on institutional investors' stock picking skills using the low frequency data. On the contrary, Grinblatt and Titman (1992), Hendricks, Patel, and Zeckhauser (1993), Kosowski, *et al.* (2006), Avramov and Wermers (2006), Alexander, Cici, and Gibson (2007), Kacperczyk, Sialm, and Zheng (2008), and Elton, Gruber, and Blake (2011) show that institutions have skills in assembling portfolios using the same databases.

[11]For example, Puckett and Yan (2011), Anand, *et al.* (2012), and Henry and Koski (2017), Chakrabarty, Moulton, and Trzcinka (2017) use actual institutional trade data from Abel Noser, which covers about 10% of total institutional trading as estimated by Puckett and Yan (2011). Irvine, Lipson, and Puckett (2007) and Busse, Clifton, and Jegadeesh (2012) use data from Plexus Group. Griffin, Harris, and Topaloglu (2003), Boehmer and Kelley (2009), and Hendershott, Livdan, and Schurhoff (2015) use the Consolidated Equity Audit Trail Data from NYSE.

We estimate daily aggregate institutional order flow by extrapolating the quarterly relation between institutional ownership change in 13F and observable order flow variables constructed using TAQ and Ancerno data. Specifically, and similar to the IOF estimation method proposed by Campbell, Ramadorai, and Schwatz (CRS, 2009), we first evaluate the quarterly association of change in aggregate institutional holdings and aggregate order imbalance from several transaction size bins obtained from both TAQ and Ancerno data. Then, we fit the quarterly relation using the daily TAQ order flow and the Ancerno order flow to retrieve daily aggregate institutional order flow. The underlying assumption of our method is that institutional investors systematically differ from retail investors in submission strategies for orders of different sizes, similar to what is presented by CRS. This assumption is largely left unverifiable by CRS using TAQ data because there is not a trader-type flag in TAQ. Unlike CRS, however, we bring actual institutional trading records from Ancerno to enhance the linkage between institutional trading and trade size. The simple addition of Ancerno data therefore provides a nontrivial contribution to estimation accuracy despite the potential limitations of Ancerno data coverage.

As the full set of institutional order flow is unobservable, we evaluate the performance of our proposed order flow measure in a joint-hypothesis test using return prediction regressions. There is consensus in the literature that institutions are apt to be better informed than retail investors because institutions have direct communication channels with firms, better relationships with sell-side analysts, and better capabilities for processing financial information. Therefore, institutional trading can be a significant contributor to the price formation process. In the same spirit of prior studies, we hypothesize that a less noisy IOF is more informative about the future stock price. In our sample of common stocks listed on NYSE, AMEX, and Nasdaq markets between January 1999 and March 2012, we find that our proposed IOF estimate, termed HH, significantly and positively predicts the returns on the following day with a coefficient of 0.122 and a t-statistic of 11.71 in the Fama-MacBeth (1973) regressions. A one-standard deviation change in HH is associated with a subsequent return of 1.928 basis points (bp). This return predictive ability outperforms the other IOF measures we consider in the study, including the aggregate institutional order flow in Ancerno data, Campbell, Ramadorai, and Schwatzs' estimate, and the estimate based on a size cut-off rule following Lee and Radhakrishna (2000). We obtain the same conclusion in direct horse race tests of multiple IOFs in the same return prediction model as well as for investment analysis. These results suggest that combining TAQ and Ancerno data yields a more accurate IOF estimate in the cross-section.

Because we can identify institution types in 13F and Ancerno, it is possible to apply the same method to only a subgroup of institutions such as: hedge funds, mutual funds, short-term institutional investors, and active institutions. These types of institutions can differ considerably on investment strategies, tools, and payoff structures. Therefore, looking at institutional investors as a whole may discard important heterogeneity across institution types. Taking advantage of estimation flexibility in our method, we choose to estimate and examine institutional order flow from hedge funds and non-hedge funds as an example to shed more light on the stock trading skills of institutional investors.

We investigate the systematic difference between aggregate hedge fund and non-hedge fund trading in three ways: First, we conjecture that hedge funds execute their orders at lower costs than non-hedge funds, hence generating smaller contemporaneous price impact because hedge fund managers can time market liquidity (Cao, *et al.*, 2013) and are likely to provide liquidity (Jame, 2017). Using the same sample in IOF performance evaluation, we find that a one-standard deviation increase in the net hedge fund order flow increases the contemporaneous stock price by 2.016 bp while the contemporaneous price impact reaches 17.158 bp for the non-hedge fund order flow. A comparison of statistical significance leads to the same conclusion, that hedge funds generate limited price pressure in the market.

Second, there is a well-documented performance difference in the literature examined at monthly or quarterly horizons. The superior performance of hedge funds is persistent (Jagannathan, Malakhov, and Novikov, 2010) and their confidential holdings exhibit positive performance for up to twelve months (Agarwal, *et al.*, 2013 and Argon, Hertzel, and Shi, 2013). To understand hedge fundsâĂŹ investment skills better, we test if their trades contain more private informative about short-term future returns than trades from other institutions. Indeed, in Fama-MacBeth regressions, we find that the hedge fund order flow generates a coefficient of 1.603 with a t-statistic of 17.18 while the non-hedge fund order flow has a slightly negative and insignificant coefficient when predicting the next day's returns. Moreover, the pricing effect of hedge fund trades is permanent without subsequent reversals, while the non-hedge fund trades generate significantly negative price impact over longer horizons. This result is robust in several subsample tests based on firm size, bid-ask spread, time periods, and exchange markets. The investment strategies based on hedge fund order flow and non-hedge fund order flow generate average daily return differentials of 0.082 (t-statistic = 10.95) and 0.043 (t-statistic = 5.95), respectively. We also find in an event study that the return predictive ability of hedge fund order flow is related to fundamental information flow around earnings announcements and large price jumps.

Finally, we examine how the two types of institutions respond to well-known return anomalies. Akbas, *et al.* (2015) and Kokkonen and Suominen (2015) show that fund flow into hedge funds mitigates mispricing. However, direct within-month arbitrage trading activities by hedge funds in response to predictable return anomalies are not documented in existing studies. We construct a daily updated mispricing index, similar to the one used by Stambaugh, Yu, Yuan (2012, 2015). In panel regressions of net IOFs on the lagged mispricing index, we find that hedge fund trades are positively associated with the mispricing index from the previous day, implying that they buy undervalued stocks and sell overvalued stocks. On the contrary, the coefficient estimate of lagged mispricing is negative and significant when we apply net non-hedge fund order flow as the dependent variable, suggesting that other institutions, on average, trade in a direction that exacerbates return anomalies.

We make two contributions to the finance literature. First, we introduce a new method of estimating institutional order flow for individual stocks at the daily level. Empirical analysis shows that this new method has robust and stronger performance in terms of return predictability than prior methods proposed by Lee and Radhakrishna (2000) and Campbell, Ramadorai, and Schwartz (2009). Our method can be applied in empirical studies that examine institutional trading behavior at the daily, or longer, frequency. Second, by investigating the contribution of different types of funds to price discovery, we find that hedge funds have better trading skills because their order flow generates smaller contemporaneous price pressure, and is more informative about future returns. We also find that hedge funds actively trade on well-known stock return anomalies daily, while the other institutions, on average, trade against those anomalies. These findings complement the studies using longer-horizon observations such as Akbas, *et al.* (2014) by providing direct evidence at a finer granularity.

The rest of the paper is organized as follows: Section II.B describes our estimation method of aggregate institutional order flow and sample selection. Section II.C tests the performance of our IOF estimate. Section II.D investigates the systematic difference between hedge fund and non-hedge fund trading. Finally, Section III.E presents our conclusions..

B. Data and variable description

B.1. Sample selection

We merge four databases into our sample using eight-digit CUSIP and Monthly TAQ master file (1) daily CRSP data for stock return, share volume turnover ratio, relative bid-ask spread, market capitalization, a primary exchange code, and share type code; (2) TAQ data to calculate total order imbalance in different trade size bins; (3) Thomson Reuters 13F Ownership data for quarterly holding of institutional investors; and (4) Ancerno institutional trade data to calculate stock-day Ancerno order imbalance. The sample includes $10,221,329$ stock-day observations from January 1999 to March 2012. Every observation in our sample meets the following criteria to be eligible: (1) it is a common stock; (2) it is primarily listed on NYSE, AMEX, or Nasdaq; (3) it is available in Monthly TAQ master file; (4) it is reported in 13F and CRSP; (5) the relative close bid-ask spread is positive and less than one half; and (6) the stock price is greater than \$5. We winsorize all types of estimated institutional order flow measures, relative bid-ask spread, and share volume turnover ratio at one and 99 percent levels in the cross-section for each day.

B.2. Estimating institutional order flow

Cut-off rule from Lee and Radhakrishna (2000)

Many studies have identified individual and institutional trades using transaction size. For example, Lee (1992) considers transactions less than $\$10,000$ to be initiated by individual investors. More recently, Malmendier and Shanthikumar (2007) categorize individual and institutional investors by using $\$20,000$ and $\$50,000$ thresholds. The cut-off rule is intuitively appealing because institutional investors are more likely to be large investors who can buy or sell large amount of stocks in a single transaction. Lee and Radhakrisha (LR, 2000) show that, according to TORQ data for 144 NYSE stocks from November 1990 to January 1991, the cut-off rule effectively separates institutional trades from individual trades. Also, Griffin, Harris, and Topaloglu (2003) report that institutional trades consist of 85.99% of block trades (more than $10,000$ shares in a transaction) and 18.14% of small trades (less than 500 shares in a transaction) in a hundred of Nasdaq stocks from May 2000 to February 2001. However, the cut-off rule is subject to a misclassification concern rooted in order splits by institutional investors. Using the TORQ data, LR disprove it by showing that 94% of total orders are executed in a single transaction. When it comes to more recent periods, however, Campbell, Ramadorai, and Schwatz (2009) find that institutional trades significantly involve small- sized trades (less than $\$2,000$ in a transaction) for small firms, using quarterly institutional holding and TAQ data from 1993 to 2000. In addition, Cready, Kumas, and Subasi (2014) document that the majority of institutional trades are categorized as small-sized trades (less than 500 shares in a transaction) in Ancerno institutional data from 2003 to 2010. These empirical findings challenge the accuracy of the cut-off rule.

To construct a benchmark IOF measure following LR, we apply a $\$5,000$ cut-off rule to identify institutional trades in TAQ data from January 1999 to March 2012[12]. First, we determine a trade direction using the Lee and Ready (1991) algorithm[13]. We put a positive (negative) sign on a

[12]LR show that the $5,000 cut-off can capture 82% of institutional trades, the highest percentage among their chosen cut-offs. Using alternative thresholds does not alter our conclusion about the effectiveness of the cut-off rule. These results are available upon request.

[13]Lee and Ready (1991) algorithm classifies a trade as buyer-initiated (seller-initiated) if its execution price is higher (lower) than a mid-point of National Best Bid and Offer (NBBO) prices. If the execution price is the same as the

buyer-initiated (seller-initiated) transaction. Then, we categorize a transaction greater than \$5,000 as an institutional trade. Lastly, we aggregate the number of signed shares of institutional trades at stock-day level and divide by the number of shares outstanding for cross-sectional normalization to produce an estimated institutional order imbalance, LR.

Institutional order flow from Campbell, Ramadorai, and Schwatz (2009)

Campbell, Ramadorai, and Schwatz (2009) develop a non-linear model to estimate institutional order flow at daily frequency, extrapolated from the quarterly relation of institutional ownership change in 13F and order imbalance in different trade sizes from TAQ. They highlight two advantages of the method in comparison to the cut-off rule. One is that it exploits information on net order flow in the full range of transaction sizes, rather than based on a single threshold, to account for the fact that institutional investors can split orders either to hide their trades or to reduce transaction costs. Another advantage is that their method allows size-categorized order imbalance to be positively or negatively associated with an institutional ownership change because institutions can strategically choose to take or provide liquidity, depending on the size of the trades. For example, an institution can submit several small-sized *limit* orders for the purpose of efficient order execution. In this case, order flow in the small-sized transaction would be *negatively* associated with the change of institutional ownership.

The estimation is based on the following equation:

$$\Delta Y_{i,q} = \alpha_q + \rho \Delta Y_{i,q-1} + \phi Y_{i,q-1} + \beta^U U_{i,q} + \beta^{UY} Y_{i,q-1} \times U_{i,q} + \sum_{Z=1}^{19} \beta(Z, Y_{i,q-1}) F_{Z,i,q} + \epsilon_{i,q}, \quad (4)$$

where for a stock i in a quarter q, α is a set of four quarter dummies, Y is aggregate institutional ownership in 13F, U is aggregate unclassified trades scaled by shares outstanding for which the Lee and Ready (1991) algorithm cannot determine the direction, and F_Z is aggregate order imbalance scaled by shares outstanding in a trade-size bin Z. CRS assign trades into nineteen size bins whose lower limit points are \$0, \$2,000, \$3,000, \$5,000, \$7,000, \$9,000, \$10,000, \$20,000, \$30,000,

mid-point, the trade is classified as buyer-initiated (seller-initiated) if the execution price is higher (lower) than the previous trade price.

$50,000, $70,000, $90,000, $100,000, $200,000, $300,000, $500,000, $700,000, $900,000, and $1 million. Since large trades for a small stock are rare, the coefficients on order flow in large trade sizes are likely swung by the small number of unusual trades. To address the potential mis-estimation issue, CRS borrow a yield curve model for $\beta(Z, Y_{i,q-1})$ from Nelson and Siegel (NS, 1987):

$$\beta(Z, Y_{i,q-1}) = b_{01} + b_{02}Y_{i,q-1} + (b_{11} + b_{12}Y_{i,q-1} + b_{21} + b_{22}Y_{i,q-1})[1 - e^{-Z/\tau}]\frac{\tau}{Z} - (b_{21} + b_{22}Y_{i,q-1})e^{-Z/\tau}, \tag{5}$$

where τ is a constant to be estimated. The NS model smooths out the coefficient variation across transaction size by putting less weight on order flow in larger trade-size bins. Also, CRS divide the sample into quintile portfolios based on NYSE breakpoints of market capitalization at the start of each quarter, and estimate the coefficients in Equation (4) for each quintile portfolio using non-linear least squares to fit the model to maximize the adjusted R-squared over different values of τ[14].

Next, taking the estimates in Equation (4), CRS calculate the expected change of institutional ownership, E[$\Delta Y_{i,d}$], as institutional order flow on a day.

$$\Delta Y_{i,d} = \alpha_d + \rho \Delta Y_{i,d-1} + \phi Y_{i,d-1} + \beta^U U_{i,d} + \beta^{UY} Y_{i,d-1} \times U_{i,d} + \sum_{Z=1}^{19} \beta(Z, Y_{i,d-1}) F_{Z,i,d} + \epsilon_{i,d}, \tag{6}$$

where d indexes a day. CRS set to zero unobservable daily variables of aggregate institutional ownership such as $\Delta Y_{i,d-1}$, $Y_{i,d-1}$, and $Y_{i,d} \times U_{i,d}$ as well as a set of daily dummies, α_d. The frequency conversion is possible under an exogeneity assumption that the error terms, $\epsilon_{i,d}$, are not correlated with all of its leads and lags within a quarter.

Institutional order flow from Ancerno institutional trading data

Ancerno (AN) institution data is transaction-level institutional trading data provided by Abel Noser Solutions, formerly Ancerno, Ltd., which offers consulting services for trading as well as

[14]There could be a convergence problem in the estimation model of (4). For example, we have observed the adjusted R-squared asymptotically approaching to its maximum as τ increases. For the infinite τ case, we set the arbitrary maximum value of τ to 100,000.

transaction cost analyses to institutional asset owners, investment managers, and brokers. A material advantage to using AN is that it contains comprehensive records of an institution's trading history since that institution started to employ Abel Noser's consulting services. The data contain a client manager/broker/trader code, a trade date, CUSIP of a traded stock, a trade direction, an execution price, the number of execution shares, commissions, and fees for each trade[15]. However, AN has a potential sample selection bias because it covers a limited portion of institutional trades. Puckett and Yan (2011) estimate that the data fail to account for approximately 90% of total institutional trades. We also find that the time-series correlation of total 13F institutional ownership changes with the aggregate Ancerno institutional order flow is 0.249, on average, at quarterly frequency for our sample period from January 1999 to March 2012. Those statistics cast doubt on whether AN trades sufficiently represent the entire set of institutional trades.

As an alternative IOF measure, we calculate the imbalance in Ancerno trades. Specifically, we calculate the difference between buyer execution shares and seller execution shares in Ancerno for a stock on the same day scaled by number of shares outstanding, termed AN. Note that Ancerno does not document records of AN client institutions that do not engage in trading of a stock on a particular day. We treat the non-trading stock-day observations as missing[16].

Our proposed institutional order flow

We estimate institutional order flow in the same spirit as CRS. Exploiting the quarterly association of an institutional ownership change with TAQ order flow as well as AN institutional order flow, we fit their relation to retrieve the institutional ownership change on a daily basis. The utilization of AN institutional order flow has a non-trivial impact on our estimation. The actual institutional order flow allows us to reduce potential measurement errors in CRS estimation stemming from a noisy TAQ order flow that groups together individual and institutional trades.

We use non-linear least square fit over different values of τ to evaluate the coefficients in the

[15]For more detailed information, see Puckett and Yan (2011), Anand, *et al.* (2012), Busse, Clifton, and Jegadeesh (2012), Cready, Kumas, and Subasi (2014), Chakrabarty, Moulton, and Trzcinka (2017), Henry and Koski (2017), or Jame (2017).

[16]We also do the same test, setting the value of the non-trading observations to zero. The test results are qualitatively the same as those reported in this paper.

following model:

$$
\begin{aligned}
\Delta Y_{i,q} &= \alpha_q + \rho \Delta Y_{i,q-1} + \phi Y_{i,q-1} + \beta^{U} U_{i,q} + \beta^{UY} Y_{i,q-1} \times U_{i,q} \\
&+ \sum_{Z=1}^{19} \beta^{F}(Z, Y_{i,q-1}) F_{Z,i,q} + \sum_{Z=1}^{19} \beta^{D}(Z, Y_{i,q-1}) D_{Z,i,q} + \epsilon_{i,q},
\end{aligned} \tag{7}
$$

where $D_{Z,i,q}$ is aggregate AN institutional order imbalance calculated as the difference of buy-side execution shares and sell-side execution shares in a transaction size bin Z for a stock i at a quarter q and other variables are the same as (4). $\beta^{F}(Z, Y_{i,q-1})$ and $\beta^{D}(Z, Y_{i,q-1})$ have the same functional form as $\beta(Z, Y_{i,q-1})$ of (5). We also use the same nineteen size bins as CRS,and estimate Equation (7) in each size quintile portfolio based on NYSE breakpoints separately.

Then, we recover the expected institutional ownership change, E[$\Delta Y_{i,d}$], on a daily basis, by fitting the estimates of (7) into the following model:

$$
\begin{aligned}
\Delta Y_{i,d} &= \alpha_d + \rho \Delta Y_{i,d-1} + \phi Y_{i,d-1} + \beta^{U} U_{i,d} + \beta^{UY} Y_{i,d-1} \times U_{i,d} \\
&+ \sum_{Z=1}^{19} \beta^{F}(Z, Y_{i,d-1}) F_{Z,i,d} + \sum_{Z=1}^{19} \beta^{D}(Z, Y_{i,d-1}) D_{Z,i,d} + \epsilon_{i,d}.
\end{aligned} \tag{8}
$$

We also set to zero unobservable daily variables of aggregate institutional ownership such as $\Delta Y_{i,d-1}$, $Y_{i,d-1}$, and $Y_{i,d} \times U_{i,d}$ as well as a set of daily dummies, α_d. The estimated daily change of institutional ownership is termed HH for the whole group of institutional investors.

Hedge fund and non-hedge fund order flow

We estimate two institutional order flows for hedge funds and non-hedge funds to explore the heterogenous trading behavior among institutions. Our estimation is, by construction, flexible enough to measure the expected institutional order flow for different types of institutional investors: One can evaluate institutional order flow for any type of institution by putting quarterly ownership of a target institution group in the estimation model of Equation (7). In this paper, we choose hedge funds and non-hedge funds as an example.

We replicate the estimation of Equation (7) to get coefficients regarding hedge funds and non-hedge funds, replacing the aggregate institutional ownership with aggregate hedge fund ownership and non-hedge fund ownership, respectively[17]. Then, we fit the estimates of Equation (7) into the retrieval procedure in Equation (8) to measure the ownership change of hedge funds and non-hedge funds at daily frequency as hedge fund order flow (HH^{HF}) and non-hedge fund order flow (HH^{NHF}), respectively.

B.3. Control variables

We calculate total order imbalance (TOI), share volume turnover ratio ($TURN$) and relative bid-ask spread ($SPRD$) for a stock on a particular day as control variables. The detailed definitions are the following:

- TOI: The number of buyer-initiated shares minus the number of seller-initiated shares in TAQ based on the Lee and Ready (1991) algorithm, scaled by the number of shares outstanding in CRSP.
- $TURN$: Daily trading volume over the number of shares outstanding.
- $SPRD$: The difference between the bid and ask prices scaled by the average bid and ask prices.

In addition, we use the risk-adjusted mid quote stock return (RET) with respect to the Fama-French (2015) five factors as our measure of stock return.

B.4. Summary statistics

Panel A of Table 19 presents the time-series average of cross-sectional statistics for institutional order flow (IOF) measures and control variables. Our sample covers, on average, $3,073.2$ stocks per trading day for $3,326$ trading days from January 1999 to March 2012.

[Place Table 19 about here]

[17]We thank Yuehua Tang for providing an identifier for hedge funds in Thomson Reuters 13F Ownership data used in Agarwal, *et al.* (2013).

First, and for purposes of comparison, four versions of IOF measures are reported. Our proposed IOF (HH) has a mean of 0.025, which is five times its median of 0.005. HH has a standard deviation of 0.158 and is positively skewed with a minimum of -0.535 and a maximum of 0.671. AN is well-balanced as the mean and median are close to zero at 0.004 and 0.002, respectively. The average number of observations per day for AN is $2,043.1$ due to limited coverage of AN data. CRS has a mean of 0.024 similar to HH, but a smaller standard deviation of 0.074. LR has a mean of 0.014, smaller than HH.

Statistics for hedge and non-hedge IOFs are reported next. The average numbers of observations per trading days for HH^{HF} and HH^{NHF} have $2,893.5$ and $3,071.6$, respectively. HH^{HF} has a mean of 0.006 and a standard deviation of 0.020, both smaller than those of HH^{NHF} (mean = 0.019; standard deviation = 0.121). The result is consistent with the fact that the average ownership of hedge funds (9.30%) is about one-third of non-hedge funds (30.00%) in our sample period.

Lastly, descriptive statistics for control variables are presented. Total order imbalance (TOI) has very similar statistics to LR: a mean order imbalance of 0.015 and a standard deviation of 0.179. The mean bid-ask spread ($SPRD$) is 0.009 and the mean turnover ratio ($TURN$) is 0.649. The mean and median of risk-adjusted mid-quote return are 0.000 and 0.027, respectively.

Panel B of Table 19 presents the time-series average of cross-sectional correlation between estimated IOFs and control variables. Although we use both TAQ and AN order imbalances to estimate HH using the same dependence structure, the loadings of HH on TAQ and AN data seem to differ considerably. The correlation between HH and AN is 0.875, almost seven times the correlation between HH and TOI at 0.127, suggesting that HH puts more weight on AN imbalance than TAQ imbalance in estimation. This average daily correlation of 0.875 is also much higher than the quarterly correlation between aggregate AN institutional order flow and 13F ownership change at 0.249. The quarterly correlation between aggregate TAQ order imbalance and 13F ownership change is 0.099. HH's correlations with CRS and LR are 0.352 and 0.142, respectively. Therefore, we find significant impact of including the AN data to our estimation of institutional order flow. The resulting HH is sufficiently different from the existing IOF estimates. In addition, non-hedge fund IOF (HH^{NHF}) behaves similarly to HH with a high correlation of 0.957 between the two, because non-hedge funds make up the majority of the investment industry and also because 63.8 percent of their trades are correlated with hedge funds' trades.

We also find AN institutional order flow is weakly correlated with CRS, LR, and TOI. The correlation of CRS with LR is 0.306, similar to its correlation with TOI as LR and TOI are almost identical with a correlation coefficient of 0.946. All estimated institutional order flows except AN are positively associated with stock liquidity. For example, HH is negatively (positively) associated with relative bid-ask spread (share volume turnover ratio) with a correlation of −0.052 (0.148). AN is statistically independent on stock liquidity. Additionally, IOF measures make a positive impact on a contemporaneous price. In particular, CRS has the weakest influence on a current stock price with a correlation of 0.069.

Interestingly, hedge funds are more sensitive to liquidity and their trades make weaker contemporaneous price impact than non-hedge funds as the correlation tests suggest. Specifically, the correlation of hedge funds' order flow is significantly more negative (positive) with relative bid-ask spread (share volume turnover ratio) than the correlation of non-hedge funds' order flow. The difference is −0.011 (0.128) with a t-statistic of −22.27 (106.77). The contemporaneous price impact of hedge fund order flow is significantly weaker than that of non-hedge fund order flow. The correlations of HH^{HF} and HH^{NHF} with contemporaneous return are 0.088 and 0.119, respectively. The difference in the correlations is −0.030 with a t-statistic of −45.83. We will dive deeper into the impact of hedge fund order flow on contemporaneous prices in Section II.D.1.

C. Performance of institutional order flows

We compare our proposed institutional order flow HH with the other IOF estimates in this section. As the total institutional order flow is not directly observable on a daily basis, we evaluate the effectiveness of different IOF measures by testing a joint hypothesis that a more accurate IOF is more informative about future stock prices in the cross-section. Institutional trading can contain price information for the following reasons: First, many institutions, as large shareholders, have direct communication with firms. Therefore, they can ask firms to disclose more information formonitoring purposes (Boone and White, 2015). Second, they maintain good relationships with sell-side analysts - a relationship that allows institutions to receive additional tips on a particular stock from the analysts (Irvine, Lipson, and Puckett, 2007). Third, they are capable of correctly interpreting financial information, such as earnings. Cohen, Gompers, and Vuolteenaho (2002) find evidence that institutions trade in the profitable direction following earnings news. Additionally, Piotroski and Roulstone (2004) and Campbell, Ramadorai, and Schwartz (2009) show that institutions can anticipate future earnings surprises. Market microstructure studies largely confirm the role of institutional trading in price discovery. For example, Hendershott, Livan, and Schurhoff (2015) find that institutional trading enhances price efficiency by spreading information prior to a news announcement. Boehmer and Kelly (2009) also find that stocks with high institutional ownership have more efficient prices. Based on the premise that institutional trading conveys price information, we examine the predictive power of different measures of IOF in the following subsection. We also compare the performance of investment strategies based on different IOFs in the subsequent subsection.

C.1. Cross-sectional return prediction

We begin our return prediction analysis with Fama-MacBeth (1973) regressions in our sample consisting of all common stocks listed on NYSE, AMEX, and Nasdaq between January 1999 and March 2012. Table 10 reports the estimated coefficients of the following model:

$$
\begin{aligned}
R_{i,t} &= \alpha_t + \sum_{k=1}^{5} \beta_{t,k}^{I} IOF_{i,t-k} + \sum_{k=1}^{5} \beta_{t,k}^{T} TOI_{i,t-k} \\
&+ \gamma_t^{B} SPRD_{i,t-1} + \gamma_t^{T} TURN_{i,t-1} + \sum_{k=1}^{5} \gamma_{t,k}^{R} R_{i,t-k} + \epsilon_{i,t},
\end{aligned}
$$

where for stock i on day t, R is risk-adjusted mid-quote stock return with respect to Fama-French (2015) five factors; IOF is an institutional order flow including HH, AN, CRS, and LR. We include lagged TOI in the model to examine the additional pricing effect from institutional order imbalance beyond the total order imbalance. Also, lagged relative bid-ask spread ($SPRD$) and lagged share volume turnover ratio ($TURN$) control for liquidity effects in the model, and lagged returns are included to control for return reversals. To account for serial correlations, we use Newey-West (1987) standard errors with eight lags to calculate the t-statistics.

[Place Table 10 about here]

The first column shows that the coefficient of HH is 0.122 at the first lag with a t-statistic of 11.71. In terms of economic significance, a one-standard deviation increase of HH is associated with an increase of 1.928 basis point (bp) in the next day's stock return in the cross-section. However, this large price impact gradually reverses in subsequent days as the coefficients at longer lags are all negative and significant. After five days, the positive initial price impact completely dies out. In the second column, the number of observations reduce from $10,015,095$ to $5,033,554$ because the AN data do not fully cover all the stock-day observations in our sample. Similar to HH, AN also has a positive coefficient (0.097, t-statistic = 6.39) at the first lag. The economic significance of AN is slightly weaker than HH as a standard-deviation increase in AN is associated with an increase of 1.329 bp in stock return on the next day. The coefficients of AN also turn negative at longer lags, though only marginally significant, indicating that the price impact from AN lasts for more than a week. In the third column, CRS has significantly negative coefficients in all lags, indicating this IOF measure captures price pressure only in the cross-section[18]. The last column shows that the first lagged term of LR has significantly negative price impact, but it flips to a positive sign in the second lag. The other lags are statistically indifferent from zero.

All the control variables other than total order imbalance have the expected signs of their coefficients. Total order flow has mixed signs in the first lagged term across columns with longer lags always having negative and significant coefficients. Spread and turnover ratio both have positive and significant coefficients, consistent with Amihud and Mendelson (1986) and Gervais,

[18]In unreported tests, we find that CRS positively predicts future returns on the following day in regressions without the control variables. We have also achieved similar results in CRS (2009) using the same vector auto-regression specification.

Kaniel, and Mingelgrin (2001), respectively. We also observe significant return reversals in all regressions.

The results in Table 10 suggest that HH is a powerful predictor for stock return. Focusing on the price impact on the next day only, HH outperforms the other IOF estimates in both statistical and economic significance. The strong predictive power of HH indicates that AN institutional trades significantly complements the TAQ data in extracting valuable price information. However, the pricing effect from HH turns negative at longer horizons while AN does not in a restricted sample. To compare the performance of HH with other IOF measures in a fair test, we conduct additional regression analysis by using HH and another IOF estimate in the same regression and report results in Table 11. For brevity, we present the estimated coefficients of institutional order flow and total net order flow only, while the regressions always include the full set of control variables.

[Place Table 11 about here]

The first column presents the horse race result between HH and AN. HH has a coefficient of 0.073 with a t-statistic of 3.24 at the first lag and the subsequent reversals also become smaller in magnitude compared to the result in Table 10. AN, however, fails to retain coefficients significantly different from zero at any lag in the horse race. It seems the predictive ability of AN is subsumed by HH in the same cross-section of stocks. In the second and third columns, the coefficients of HH remain positive and significant at the first lag while CRS and LR yield only negative and significant coefficients, similar to those in Table 10.

In summary, the predictive power of our proposed IOF estimate, HH, outperforms the other IOF estimates in the cross-section. The estimated coefficient of HH is always positive and significant in the first lag, regardless of inclusion of competing institutional order flow. Note that the pricing effect of HH is not persistent but completely reverses in the five subsequent trading days. IOF behavior implies institutions are likely to rely on short-lived information for their trades as a whole. AN institutional order flow is also positively associated with stock return on the following day and its price impact is persistent at least for five trading days. When it comes to the horse race with HH

in Table 11, however, AN loses its predictive ability for stock returns. CRS and LR only generate a negative price impact on future stock returns in the regression, controlling for stock illiquidity and return reversals. If institutional investors are better informed than an average investor, their order flow should contain price information. Consistent with this view, we find that HH generates persistent and robust return predictive ability for returns, while the other IOF measures do not.

C.2. Investment analysis

In this subsection, we perform univariate sorts to examine the profitability of order flow strategies. We sort all stocks into decile portfolios based on one of the IOF measures for each day and calculate the average equal-weighted portfolio returns on the following day. Table 12 reports the average decile portfolio returns as well as the return differentials and Fama-French (2015) alphas between the top and bottom order imbalance decile portfolios. All t-statistics in Table 12 are calculated based on Newey-West (1987) standard errors with eight lags.

[Place Table 12 about here]

We find that the investment strategy of buying stocks in the top HH decile and selling stocks in the bottom decile generates an average return of 5.2 bp on the following day with a t-statistic of 6.81. The alpha with respect to the Fama-French (2015) five factors (FF5) is at the same level with an even larger t-statistic. The long-short strategy based on AN also produces a significantly positive average return of 5.9 bp on the next day with a t-statistic of 6.96[19]. Strategies based on CRS and LR, however, do not generate positive abnormal returns. In the case of CRS, the return differential is even significantly negative. The univariate portfolio sort results in this subsection are consistent with Fama- MacBeth regression results in the previous subsection.

[19]The performance difference between the HH-based and AN-based strategies is -0.007 with a t-statistic of -1.10, statistically indifferent from each other. Note that there is the difference in sample size: $10,221,329$ for HH, CRS, and LR versus $6,790,481$ for AN.

D. *Trading behavior of hedge/non-hedge funds*

One advantage of our IOF estimation method is its flexibility to estimate order flow for different types of institutions. As an example, we study potential heterogenous trading behavior between hedge funds and other institutional investors in this section. Specifically, we examine the following three aspects in stock trading:

1. Execution cost:

   We conjecture that hedge funds execute their orders at low costs. Aragon and Strahan (2012) find that a hedge fund investor can act as a liquidity provider per LehmanâĂŹs bankruptcy in 2008. Cao, *et al.* (2013) also find evidence that hedge funds can time market liquidity using hedge fund returns. The direct comparison of execution costs between hedge funds and other institutions in a large cross-section is still absent in finance literature due to unobservable order flow. To test this, we examine the contemporaneous price impact of hedge fund and non-hedge fund order flow estimates, constructed using our method.

2. Informed trading:

   We expect that institutional order flow from a hedge fund is more informative about future stock price. Prior studies using observations at low frequencies also support this view. For example, Jagannathan, Malakhov, and Novikov (2010) find persistency of superior performance in hedge funds after mitigating a self-reporting concern in hedge fund databases. Agarwal, *et al.* (2013) and Argon, Hertzel, and Shi (2013) report positive and significant abnormal returns from the 13F confidential holdings of hedge funds. More recently, Cao, *et al.* (2017) provide evidence that stocks bought by hedge funds subsequently improve in price efficiency. However, analysis at a daily frequency using hedge fund order flow in a large cross-section is scarce. To fill the gap in the literature, we investigate the return predictability of hedge fund order flow and non-hedge fund order flow in: 1) cross-sectional regressions replicating Table 10; 2) an investment analysis replicating Table 12; and 3) an event study around information shocks associated with earnings announcements and price jumps.

3. Arbitrage trading:

We conjecture that hedge funds trade on mispriced stocks. Akbas, *et al.* (2015) and Kokkonen and Suominen (2015) find a positive relation between monthly flow to hedge funds and return spread between under-valued and over-valued stock portfolios, suggesting that hedge fund flow attenuates market mispricing. Taking advantage of the finer granularity of our estimated hedge fund IOF, we conduct a stock-level analysis on arbitrage trading by hedge funds and other institutions at a daily frequency.

D.1. Contemporaneous price impact

In this subsection, we compare the contemporaneous price impact of hedge fund order flow (HH^{HF}) versus non-hedge fund order flow (HH^{NHF}) in the same sample of common stocks on NYSE, AMEX, and Nasdaq between January 1999 an March 2012. Table 13 presents the estimated coefficients of the following model:

$$\begin{aligned} R_{i,t} &= \alpha_t + \beta_t^{HF} HH_{i,t}^{HF} + \beta_t^{NHF} HH_{i,t}^{NHF} \\ &+ \beta_t^T TOI_{i,t} + \gamma_t^B SPRD_{i,t-1} + \gamma_t^T TURN_{i,t-1} + \sum_{k=1}^{5} \gamma_{t,k}^R R_{i,t-k} + \epsilon_{i,t}, \end{aligned}$$

where HH^{HF} and HH^{NHF} are hedge fund order flow and non-hedge fund order flow, respectively, and the other variables are the same as defined in Table 19.

In univariate regressions, we find that both hedge fund and non-hedge fund IOFs have large and significant contemporaneous price pressure. Specifically, the coefficient of HH^{HF} is 11.142 with a t-statistic of 50.44 in Column (1). HH^{NHF} has a coefficient of 2.636 with a t-statistics of 60.89 in Column (2). Although the coefficient estimate of HH^{HF} is greater than that of HH^{NHF}, the economic significance of HH^{HF} is lower because of the smaller standard deviation in HH^{HF}. A one standard-deviation increase in HH^{HF} is associated with an increase of 22.3 bp in the contemporaneous return, while the economic significance of HH^{NHF} reaches 31.9. In Column (3), we include both HH^{HF} and HH^{NHF} in the same regression model. The multivariate regression results show that the coefficient of HH^{HF} decreases to 2.043 with a much smaller t-statistic of 5.53, but the coefficient of HH^{NHF} remains at the same magnitude with a t-statistic of 38.23. The price pressure further decreases in Column (4) when we include TAQ order imbalance, illiquidity, and lagged returns in the regression. Nonetheless, we still find that the contemporaneous price effect of HH^{HF} is weaker than HH^{NHF} in both statistical and economic significance. The coefficient of HH^{HF} (HH^{NHF}) is 1.008 (1.418) with t-statistics of 2.91 (29.41), implying corresponding economic significance at 2.016 (17.158) bp on the same day. The results support our conjecture that hedge funds execute their orders at lower costs than other institutions.

D.2. Cross-sectional return prediction

In this subsection, we turn into return predictability for hedge fund and non-hedge fund trades. We replicate the regression analysis in Table 10 and Table 11. For brevity, Table 14 report the coefficient estimates of order flow only, while the regressions always include the full set of control variables.

[Place Table 14 about here]

In Column (1), we examine the information content of hedge fund order flow first. The estimated coefficient is 1.508 at the first lag with a t-statistic of 18.49. Similar to the total institutional order flow, HH^{HF} also experiences significant reversal in its pricing effect on subsequent days as the coefficient estimates become negative and significant on days $t-3$ to $t-5$. However, even after five days, the cumulative price impact is still positive and significant, suggesting that hedge fund trades can capture permanent information flow. In Column (2), we find that non-hedge fund order flow also has predictive power for stock returns on the subsequent day with a coefficient of 0.139 and a t-statistic of 10.25. However, the pricing effect of non-hedge fund trades is only transitory, similar to the HH price impact in Table 10. This is not surprising given the high correlation between HH and HH^{NHF}. When we use both HH^{HF} and HH^{NHF} in the same regression in Column (3), we find that the discrepancies are enhanced. Specifically, HH^{HF} obtains similar predictive ability at the first lag without significant reversals later, while HH^{NHF} loses its initial positive price impact with almost the same negative pricing effects at longer horizons. Total order imbalance behaves in a consistent manner across all columns. It generates positive price impact on the next day, which fully reverses quickly, consistent with results presented in Table 10. These results indicate that hedge fund trades have permanent price impact while non-hedge fund trades have transitory price pressure only, supporting our conjecture that hedge funds have an information advantage over other institutions.

Next, we investigate the predictive power of HH^{HF} and HH^{NHF} in subsamples of stocks as a robustness test. In Table 26, we begin by conditioning based on firm size in the first three columns. We sort all stocks into tertile portfolios based on daily market capitalization and report the estimated coefficients of institutional net order flows from our base regression model for each portfolio. The control variables are included in the regressions, but omitted in reporting for brevity. The results can be summarized as follows: First, focusing on the pricing effect on the subsequent day, HH^{HF} has positive and significant coefficient estimates in all three firm groups, while HH^{NHF} only has significant pricing effects for the middle-sized firm group. Second, HH^{HF}'s pricing effects on the subsequent day adversely depends on the firm size for both statistical and economic significance. Specifically, HH^{HF} has a coefficient of 4.260 (t-statistic = 16.60) for small stocks and 0.719 (t-statistic = 5.53) for large stocks. In terms of economic significance, a one-standard deviation increase in HH^{HF} leads to an increase of 7.25 bp in the next dayâĂŹs stock return for small stocks and 1.38 bp for large stocks. Third, the pricing effect of HH^{HF} is permanent in all three groups, with a small reversal in the small firm group, no reversal in the median group, and a continuation of positive effects in the large firm group over subsequent days.

[Place Table 26 about here]

In the next three columns, we examine subsamples of stocks based on liquidity measured by the relative bid-ask spread. We find the following results: First, HH^{HF} has positive and significant coefficient estimates at the first lag in all three spread groups, while HH^{NHF} only generates insignificant or negative price impact at the first lag in all stock groups. Second, hedge fundâĂŹs pricing effects are stronger for stocks with wide spread. For example, the coefficient estimate of HH^{HF} is 3.110 (t-statistic = 13.79) for wide spread stocks and 0.834 (t-statistic = 6.59) for narrow spread stocks. A one-standard deviation increase in HH^{HF} associated with an increase in the subsequent return by 6.50 bp for stocks with wide spreads and by only 1.37 bp for those with narrow spreads. The difference between the spread groups is consistent with the notion that stocks with wide spreads are less transparent because of higher arbitrage costs.

We break the sample into an early period from 1999 to 2004, and a late period from 2005 to 2012 to investigate the time-series variation in return predictability. We report estimated coefficients in the subsequent two columns for these subperiods. HH^{HF} positively and significantly predicts the next dayâĂŹs return in both subperiods. The coefficient of HH^{HF} at the first lag slightly reduces from 1.742 (t-statistic = 10.64) in the early period to 1.488 (t-statistic = 14.44) in the late period. The coefficients suggest that a one-standard deviation increase of HH^{HF} leads to a return increase of 3.54 bp in the early period and 2.91 bp in the late period. However, HH^{NHF} has an insignificant coefficient estimate in both periods. Additionally, at the second lag, HH^{HF} shifts from a positive and significant pricing effect in the early period to a mild reversal in the late period. The difference between two subperiods is consistent with the idea that liquidity improvement is accompanied by the rise of high frequency trading that, in turn increases price efficiency over time[20].

Finally, we separate the sample according to the primary stock exchange in the last two columns. Because the Nasdaq market is traditionally a dealers' market and has a different market structure from NYSE and AMEX, we examine Nasdaq listed and non-Nasdaq listed stocks separately. The results show that HH^{HF} has positive and significant predictive ability in both Nasdaq and non-Nasdaq subsamples, while the coefficient of HH^{NHF} is negative and significant in the non-Nasdaq subsample and insignificant in the Nasdaq subsample. Specifically, HH^{HF} has a coefficient of 1.718 with a t-statistic of 9.24 for NYSE and AMEX stocks, and a coefficient of 1.956 with a t-statistic of 14.91 for Nasdaq stocks.

Given the strong return predictive ability from HH^{HF}, we analyze investment strategies based

[20]See Hendershott, Jones, and Menkveid (2011) and Broggard, Hendershott, and Riordan (2014).

on lagged HH^{HF} and HH^{NHF}. Replicating Table 12, we sort all stocks into decile portfolios based on IOF measures for hedge funds and non-hedge funds every day and calculate the average equal-weighted portfolio returns on the next day. Table 16 reports the average decile portfolio returns as well as the return differentials and Fama-French (2015) alphas between the top and bottom order imbalance decile portfolios.

[Place Table 16 about here]

We find that both strategies based on hedge fund and non-hedge fund trades generate positive and significant abnormal returns on the following day. The HH^{HF} strategy produces a positive alpha of 8 bp with a t-statistic of 11.04. Similarly, the HH^{NHF} strategy also generates a positive and significant alpha of 0.042 (t-statistic = 6.00), much lower than the HH^{HF} strategy.

The results in this subsection are summarized into three findings: First, hedge fund trades can predict a stock price on the subsequent day and make a permanent price impact. Second, the return predictability of hedge funds outperforms other institutions. Lastly, the superior power of return prediction is robust in diverse subsamples. Such findings also suggest that hedge funds may capture fundamental information and exploit their information advantage in trading. For this reason, we study corporate events in the next subsection to understand if the predictive power is related to fundamental information.

D.3. Corporate event studies

Is the predictive ability of IOFs related to fundamental information flow? We answer this question using an event study of important information shocks. To do that, we construct a sample of quarterly earnings announcements-arguably the most important scheduled corporate announcements-and permanent price jumps not related to earnings news that capture other material information. Quarterly earnings announcements are collected from the I/B/E/S database. A permanent price jump is defined as a two standard-deviation shock that does not fully reverse in the ten days to follow. After merging the event data with our sample, we have $376,771$ stock-day observations for the abovementioned corporate events.

Table 27 reports the estimated coefficients of the following model using ordinary least squares regressions:

$$
\begin{aligned}
CAR_{i,t-1,t+1} \quad = \quad & \alpha_t + \sum_{k=2}^{6} \beta_{k,t}^{H} HH_{i,t-k}^{HF} + \sum_{k=2}^{6} \beta_{k,t}^{N} HH_{i,t-k}^{NHF} \\
+ \quad & \sum_{k=2}^{6} \gamma_{k,t}^{TOI} TOI_{i,t-k} + \gamma_t^{T} TURN_{i,t-2} + \gamma_t^{S} SPRD_{i,t-2} + \sum_{k=2}^{6} \gamma_{k,t}^{R} RET_{i,t-k} + \epsilon_{i,t},
\end{aligned}
$$

where for each event i on day t, CAR is the cumulative abnormal return from days $t-1$ to $t+1$, and all the explanatory variables are the same as defined in Table 10 with event subscription i instead of firm subscription. We cluster the standard errors around firm and date in calculating the t-statistics. We report results for the main variables in Table 27 and exclude the control variables for brevity.

[Place Table 27 about here]

We first investigate the two events separately in the first two columns and then combine them in the last column. Regardless of the event samples, we find consistent results in Table 27. The estimated coefficients of hedge fund order flow are always positive and significant on day $t-2$, prior to our event window. However, non-hedge fund order flow obtains only negative and significant coefficients on day $t-2$. In the case of extreme price movements and the combination of both events, the coefficients of HH^{HF} are significantly positive even on event days $t-4$. The results suggest that the return predictability from hedge fund order flow is indeed related to fundamental information flow pertaining to significant corporate events.

D.4. Arbitrage trading

In this subsection, we study how institutions exploit well-known arbitrage opportunities documented in the literature. We construct a mispricing index, similar to Stambaough, Yu, Yuan (2012, 2015), that is updated daily. We identify mispricing based on the following nine return anomalies:[21]

[21]We document how we calculate the nine return anomalies in the Appendix.

1. Net stock issues as in Ritter (1991) and Loughram and Ritter (1995)
2. Composite equity issues as in Daniel and Titman (2006)
3. Total accruals as in Sloan (1996)
4. Net operating assets as in Hirshleifer, et al. (2004)
5. Momentum as in Jegadeesh and Titman (1993)
6. Gross profitability as in Novy-Marx (2013)
7. Asset growth as in Cooper, Gulen, and Schill (2008)
8. Return on assets as in Fama and French (2006) and Chen, Novy-Marx, and Zhang (2011)
9. Investment-to-asset ratio as in Titman, Wei, and Xie (2004)

Using daily CRSP-Compustat merged data from January 1999 to March 2012, we update accounting-based anomalies on the reporting date. To construct the mispricing index, we first rank all stocks into decile portfolios from over-valued stocks to under-valued stocks based on each anomaly factor and for each day. Then, we assign the portfolio rank to every stock in that portfolio as an index for that particular anomaly. Therefore, stocks with higher mispricing index values are more likely to be under-valued. We take the average of the ranks of the nine anomaly signals as the mispricing index for a stock on a particular day.[22] Consistent with the anomaly literature, we examine only common stocks with prices above $5 on NYSE, AMEX, and Nasdaq markets and exclude all stocks in financial and utility industries. After merging the anomaly index into our main sample, we have $1,800,557$ stock-day observations.

We examine the arbitrage trading of hedge funds and non-hedge funds using a pooled regressions. If institutional investorsâĂŹ trades exploit arbitrage opportunities implied by these anomalies, their order flow should be positively associated with the mispricing index. Table 18 report the least squares dummy variables (LSDV) estimators which remove firm fixed effects in the

[22]We only consider stock-day observations for which at least five anomaly factors can be calculated.

following model:

$$
\begin{aligned}
OF_{i,t} &= \beta MISP_{i,t-1} + \gamma^{F}\text{FirstWeek}_{i,t-1} + \gamma^{L}\text{LastWeek}_{i,t-1} \\
&+ \gamma^{FX}\text{FirstWeek}_{i,t-1} \times MISP_{i,t-1} + \gamma^{LX}\text{LastWeek}_{i,t-1} \times MISP_{i,t-1} + \epsilon_{i,t},
\end{aligned}
$$

where for each firm i on day t, OF is the estimated institutional order flow as indicated on the top of each column, FirstWeek (LastWeek) is a dummy variable for the first (last) week in a month, and $MISP$ is the mispricing index we compose. Firm fixed effect is included using the dummy variable approach.

[Place Table 18 about here]

In the univariate regression of HH^{HF} on lagged mispricing index, we find that the coefficient estimate of the mispricing index is 0.833 with a t-statistic of 4.34. When we include the week dummies and their interactions with the mispricing index, we find the association between HH^{HF} and $MISP$ remains positive and significant. Moreover, unconditionally, hedge funds seem more likely to purchase stocks in the first week of a month and sell stocks in the last week of a month given the significant coefficient estimates of 0.049 and -0.110 on the two week dummies, respectively. We also find the hedge fund purchase at the beginning of a month is not likely to be motivated by mispricing as the interaction of $MISP$ and the first-week dummy has a negative and significant coefficient estimate, fully offsetting the positive coefficient on $MISP$.

Turning to non-hedge funds, on the other hand, we find that $MISP$ has a coefficient of -4.315 with a t-statistic of -3.74. This negative relation is also robust in the multivariate regression with week dummies. Interestingly, although the impact of $MISP$ is completely the opposite for HH^{HF} and HH^{NHF}, the seasonality of order flow is similar for both types of funds. HH^{NHF} also becomes significantly more positive in the first week and significantly more negative in the last week of a month.

The results in this table suggest that hedge fund managers trade on asset pricing anomalies on a daily basis. However, the non-hedge fund institutions trade in the opposite way, potentially exacerbating anomalies despite well-documented patterns of abnormal returns from such trading strategies.

E. Conclusion

In this article, we propose a new method to estimate net institutional order flow (IOF) that extrapolates the quarterly relation between institutional ownership change and quarterly order flow information from TAQ and Ancerno with a higher frequency–daily in our case. We test the performance of our estimated IOF estimate, termed HH, by comparing its return predictive ability to three other IOF measures for common stocks listed on NYSE, AMEX, and Nasdaq markets from January 1999 to March 2012. The alternative IOF methods include: the size cut-off rule as seen in Lee and Radhakrisha (LR, 2000); Campbell, Ramadorai, and Schwartz's (CRS, 2009) method, which is similar to ours', but uses only the TAQ order imbalance; and the aggregate imbalance in Ancerno data (AN). We find that HH outperforms the other IOF measures in predicting future stock returns in the cross-section. The superior performance of HH comes from the fact that we utilize the trade information made available in both TAQ and Ancerno. The CRS method relies exclusively on the heterogeneity of institutional trading patterns across trade-size that can be discerned using TAQ data. Though TAQ records all transactions on public exchanges, the measurement error in trader identity is enormous. Ancerno, however, provides cleanly identified institutional trade records, while its coverage is rather limited. Combining these two data sources, we construct a more informative measure of institutional order flow than presented previous studies.

Our estimation method can also be applied to net order flow estimation for particular types of institutions. As an example, we investigate trading skills of hedge fund and non-hedge fund investors, as such direct empirical comparison at a relatively high frequency in a large cross-section is still absent from the literature. Our results indicate that hedge funds have superior trading skills as opposed to non-hedge fund institutions. Specifically, hedge fund trades create a weaker contemporaneous price impact than non-hedge fund trades. Hedge fund order flow is more informative about future stock prices with a positive and permanent impact, while non-hedge fund trades exhibit only transitory price pressure. Moreover, we find that hedge fund trades are positively associated with a mispricing index constructed using well-known stock return anomalies on the previous day, indicating that hedge funds purchase under-valued stocks and sell over-valued stocks. In sharp contrast, the non-hedge fund order flow is negatively associated with the same mispricing index. The results suggest that hedge funds exploit arbitrage opportunities from stock return anomalies while non-hedge funds trade in the opposite direction to exacerbate the anomalies. Overall, our findings indicate that the superior trading skills of institutional investors are largely driven by hedge funds.

The proposed estimation method for institutional order flow can be applied in other empirical studies that require institutional order flow estimates at a daily frequency. We apply the method to hedge funds and non-hedge funds only in this study, but the method can also be applied to other types of institutional trading, such as long-term and short-term investors, or active and passive investors. Our analysis is based on stock-level institutional order imbalance, and it would be interesting to combine it with fund-level analysis to gain a more complete picture of the wider-reaching effects of institutional trading. We leave these questions to future studies.

Appendix. Return anomalies

1. Net stock issues (Ritter (1991) and Loughram and Ritter (1995)): We calculate net stock issues as the growth rate of split-adjusted shares outstanding in the past 252 trading days between days $t - 252$ and t.

2. Composite equity issues (Daniel and Titman (2006)): We obtain composite equity issues by subtracting log change of shares outstanding from a cumulative log share adjustment factor in the past five-year window. The adjustment factor is equal to the log number of shares if one reinvestment all cash distributions back into the stock.

3. Total accurals (Sloan (1996)): We measure total accruals as non-cash working capital less than the depreciation expense, scaled by total assets in the previous fiscal year.

4. Net operating assets (Hirshleifer, et al. (2004)): We measure net operating assets by subtracting all operating assets from all operating liability and scaling it by the total assets in the previous fiscal year.

5. Momentum (Jegadeesh and Titman (1993)): Momentum factor is calculated as the cumulative return between days $t - 252$ and $t - 42$.

6. Gross profitability (Novy-Marx (2013)): We use gross profitability scaled by total assets in the same fiscal year.

7. Asset growth (Cooper, Gulen, and Schill (2008)): Asset growth is defined as the growth rate of total assets from the previous fiscal year to the current fiscal year.

8. Return on assets (Fama and French (2006) and Chen, Novy-Marx, and Zhang (2011)): We calculate return on assets as quaterly earnings scaled by book equity value in the last quarter.

9. Investment-to-asset (Titman, Wei, and Xie (2004)): We measure investment-to-asset by adding gross property, plant, and equipment to change in inventories. We scale it by book equity value in the last fiscal year.

REFERENCES

[1] Agarwal, Vikas, W. E. I. Jiang, Yuehua Tang, and Baozhong Yang, 2013, Uncovering hedge fund skill from the portfolio holdings they hide, The Journal of Finance 68, 739-783.

[2] Akbas, Ferhat, Will J. Armstrong, Sorin Sorescu, and Avanidhar Subrahmanyam, 2015, Smart money, dumb money, and capital market anomalies, Journal of Financial Economics 118, 355-382.

[3] Alexander, G. J., G. Cici, and S. Gibson, 2007, Does motivation matter when assessing trade performance? An analysis of mutual funds, Review of Financial Studies 20, 125-150.

[4] Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid ask spread, Journal of Financial Economics 17, 223-249.

[5] Anand, A., P. Irvine, A. Puckett, and K. Venkataraman, 2012, Performance of institutional trading desks: An analysis of persistence in trading costs, Review of Financial Studies 25, 557-598.

[6] Aragon, George O., Michael Hertzel, and Zhen Shi, 2013, Why do hedge funds avoid disclosure? Evidence from confidential 13f filings, Journal of Financial and Quantitative Analysis 48, 1499-1518.

[7] Aragon, G. O., and P. E. Strahan, 2012, Hedge funds as liquidity providers: Evidence from the lehman bankruptcy, Journal of Financial Economics 103, 570-587.

[8] Avramov, D., and R. Wermers, 2006, Investing in mutual funds when returns are predictable, Journal of Financial Economics 81, 339-377.

[9] Barras, L., O. Scaillet, and R. Wermers, 2010, False discoveries in mutual fund performance: Measuring luck in estimated alphas, Journal of Finance 65, 179-216.

[10] Battalio, R. H., and R. R. Mendenhall, 2005, Earnings expectations, investor trade size, and anomalous returns around earnings announcements, Journal of Financial Economics 77, 289-319.

[11] Boehmer, Ekkehart, and Eric K. Kelley, 2009, Institutional investors and the informational efficiency of prices, Review of Financial Studies 22, 3563-3594.

[12] Boone, Audra L., and Joshua T. White, 2015, The effect of institutional ownership on firm transparency and information production, Journal of Financial Economics 117, 508-533.

[13] Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-frequency trading and price discovery, Review of Financial Studies 27, 2267-2306.

[14] Busse, Jeffrey A., T. Clifton Green, and Narasimhan Jegadeesh, 2012, Buy-side trades and sell-side recommendations: Interactions and information content, Journal of Financial Markets 15, 207-232.

[15] Campbell, John, Tarun Ramadorai, and Allie Schwatz, 2009, Caught on tape, Journal of Financial Economics 92, 26.

[16] Cao, Charles, Yong Chen, Bing Liang, and Andrew W. Lo, 2013, Can hedge funds time market liquidity?, Journal of Financial Economics 109, 493-516.

[17] Cao, Charles, Bing Liang, Andrew W. Lo, and Lubomir Petrasek, 2017, Hedge fund holdings and stock market efficiency, The Review of Asset Pricing Studies rax015-rax015.

[18] Carhart, M. M., 1997, On persistence in mutual fund performance, Journal of Finance 52, 57-82.

[19] Chakrabarty, Bidisha, Pamela C. Moulton, and Charles Trzcinka, 2017, The performance of short-term institutional trades, Journal of Financial and Quantitative Analysis 52, 1403-1428.

[20] Chen, H. L., N. Jegadeesh, and R. Wermers, 2000, The value of active mutual fund management: An examination of the stockholdings and trades of fund managers, Journal of Financial and Quantitative Analysis 35, 343-368.

[21] Chen, Long, Robert Novy-Marx, and Lu Zhang, 2011, An alternative three-factor model, Working paper.

[22] Cohen, R. B., P. A. Gompers, and T. Vuolteenaho, 2002, Who underreacts to cash-flow news? Evidence from trading between individuals and institutions, Journal of Financial Economics 66, 409-462.

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

[24] Cready, William, Abdullah Kumas, and Musa Subasi, 2014, Are trade size-based inferences about traders reliable? Evidence from institutional earnings-related trading, Journal of Accounting Research 52, 877-909.

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

[26] Edelen, Roger M., Ozgur S. Ince, and Gregory B. Kadlec, 2016, Institutional investors and stock return anomalies, Journal of Financial Economics 119, 472-488.

[27] Elton, E. J., M. J. Gruber, and C. A. Blake, 2011, Holdings data, security returns, and the selection of superior mutual funds, Journal of Financial and Quantitative Analysis 46, 341-367.

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

[29] Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, Journal of Financial Economics 116, 1-22.

[30] Fama, Eugene F., and James D. Macbeth, 1973, Risk, return, and equilibirum: Empirical tests, Journal of Political Economy 81, 607-636.

[31] Gervais, S., R. Kaniel, and D. H. Mingelgrin, 2001, The high-volume return premium, Journal of Finance 56, 877-919.

[32] Griffin, J. M., J. H. Harris, and S. Topaloglu, 2003, The dynamics of institutional and individual trading, Journal of Finance 58, 2285-2320.

[33] Grinblatt, Mark, and Sheridan Titman, 1992, The persistence of mutual fund performance, The Journal of Finance 47, 1977-1984.

[34] Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2011, Does algorithmic trading improve liquidity?, Journal of Finance 66, 1-33.

[35] Hendershott, T., D. Livdan, and N. Schurhoff, 2015, Are institutions informed about news?, Journal of Financial Economics 117, 249-287.

[36] Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 1993, Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988, The Journal of Finance 48, 93-130.

[37] Henry, Tyler R., and Jennifer L. Koski, 2017, Ex-dividend profitability and institutional trading skill, The Journal of Finance 72, 461-494.

[38] Hirshleifer, D., K. W. Hou, S. H. Teoh, and Y. L. Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, Journal of Accounting & Economics 38, 297-331.

[39] Hvidkjaer, S., 2008, Small trades and the cross-section of stock returns, Review of Financial Studies 21, 1123-1151.

[40] Irvine, P., M. Lipson, and A. Puckett, 2007, Tipping, Review of Financial Studies 20, 741-768.

[41] Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov, 2010, Do hot hands exist among hedge fund managers? An empirical evaluation, The Journal of Finance 65, 217-255.

[42] Jame, Russell, 2017, Liquidity provision and the cross-section of hedge fund returns, Management Science.

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

[44] Kacperczyk, M., C. Sialm, and L. Zheng, 2008, Unobserved actions of mutual funds, Review of Financial Studies 21, 2379-2416.

[45] Kokkonen, Joni, and Matti Suominen, 2015, Hedge funds and stock market efficiency, Management Science 61, 2890-2904.

[46] Kosowski, R., A. Timmermann, R. Wermers, and H. White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, Journal of Finance 61, 2551-2595.

[47] Lee, C. M. C., 1992, Earnings news and small trades - an intraday analysis, Journal of Accounting & Economics 15, 265-302.

[48] Lee, Charles M. C., and Balkrishna Radhakrishna, 2000, Inferring investor behavior - evidence from torq data, Journal of Financial Market 3, 29.

[49] Lee, Charles M. C., and Mark J. Ready, 1991, Inferring trade direction from intraday data, Journal of Finance 46, 733-746.

[50] Loughran, T., and J. R. Ritter, 1995, The new issues puzzle, Journal of Finance 50, 23-51.

[51] Malmendier, U., and D. Shanthikumar, 2007, Are small investors naive about incentives?, Journal of Financial Economics 85, 457-489.

[52] Nelson, C. R., and A. F. Siegel, 1987, Parsimonious modeling of yield curves, Journal of Business 60, 473-489.

[53] Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation cosistent covariance-matrix, Econometrica 55, 703-708.

[54] Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, Journal of Financial Economics 108, 1-28.

[55] Piotroski, Joseph D., and Darren T. Roulstone, 2004, The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firmâĂŘspecific information into stock prices, The Accounting Review 79, 1119-1151.

[56] Puckett, Andy, and Xuemin Yan, 2011, The interim trading skills of institutional investors, Journal of Finance 66, 601-633.

[57] Ritter, Jay R., 1991, The long-run performance of initial public offerings, The Journal of Finance 46, 3-27.

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

[59] Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, Journal of Financial Economics 104, 288-302.

[60] Stambaugh, Robert F., Jianfeng Yu, and Y. U. Yuan, 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle, Journal of Finance 70, 1903-1948.

[61] Titman, S., K. C. J. Wei, and F. X. Xie, 2004, Capital investments and stock returns, Journal of Financial and Quantitative Analysis 39, 677-700.

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

Table 9. Summary statistics

This table shows the time-series averages of cross-sectional statistics for the sample during January 1999 to March 2012. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data. For each institutional order flow estimate, we calculate the daily order imbalance as the aggregate buyer-initiated volume minus the aggregate seller-initiated volume scaled by shares outstanding. HH is our proposed institutional order flow described in Section II.B.2. AN is calculated using Ancerno data. CRS is the estimate proposed by Campbell, Ramadoral, and Schwatz (2009). LR is the estimate of Lee and Radhakrishna (2000) using \$ 5,000 cut-off rule. HH^{HF} and HH^{NHF} are our proposed hedge fund and non-hedge fund order flow, respectively, described in Section II.B.2. TOI is daily total order flow in TAQ based on Lee and Ready (1993) algorithm. $SPRD$ is daily relative spreads measured as twice the distance between daily close offer and bid prices scaled by the quote midpoint. $TURN$ is daily turnover ratio defined as trading volume over the number of shares outstanding. RET is daily risk-adjusted mid-quote stock return with respect to Fama-French (2015) five factors. Panel A reports discriptive statistics. Panal B presents correlation coefficients between institutional order flows and other variables.

Panel A. Descriptive Statistics

	Number of Dates	Avg Num. of Stocks	Mean	Stdev	Min	Med	Max
HH	3326	3073.2	0.025	0.158	-0.535	0.005	0.671
AN	3326	2042.1	0.004	0.137	-0.574	0.002	0.547
CRS	3326	3073.2	0.024	0.074	-0.163	0.006	0.357
LR	3326	3073.2	0.014	0.151	-0.558	0.002	0.646
HH^{HF}	3326	2893.5	0.006	0.020	-0.053	0.002	0.096
HH^{NHF}	3326	3071.6	0.019	0.121	-0.411	0.004	0.521
TOI	3326	3073.2	0.015	0.179	-0.694	0.003	0.779
$SPRD$	3326	3073.2	0.009	0.013	0.000	0.005	0.126
$TURN$	3326	3073.2	0.649	0.788	0.003	0.414	4.877
RET	3326	3068.9	0.000	0.027	-0.122	-0.001	0.153

Panel B. Correlation

	HH	AN	CRS	LR	HH^{HF}	HH^{NHF}	TOI
HH	1.000						
AN	0.875	1.000					
CRS	0.352	0.003	1.000				
LR	0.142	0.055	0.306	1.000			
HH^{HF}	0.712	0.512	0.459	0.134	1.000		
HH^{NHF}	0.957	0.886	0.397	0.164	0.638	1.000	
TOI	0.127	0.061	0.303	0.946	0.144	0.173	1.000
$SPRD$	-0.052	0.002	-0.107	-0.054	-0.064	-0.053	-0.056
$TURN$	0.148	-0.005	0.339	0.149	0.282	0.154	0.153
RET	0.096	0.103	0.069	0.241	0.088	0.119	0.305

Table 10. Daily return prediction in the cross-section

This table presents Fama-MacBeth (1973) regression results for the following equation.

$$R_{i,t} = \alpha_t + \sum_{k=1}^{5} \beta_{t,k}^{I} IOF_{i,t-k} + \sum_{k=1}^{5} \beta_{t,k}^{T} TOI_{i,t-k} + \gamma_t^{B} SPRD_{i,t-1} + \gamma_t^{T} TURN_{i,t-1} + \sum_{k=1}^{5} \gamma_{t,k}^{R} R_{i,t-k} + \epsilon_{i,t},$$

where for stock i on day t, RET is risk-adjusted mid-quote stock return with respect to Fama-French (2015) five factors, IOF is an institutional order flow estimate indicated on the top of each column. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP and Thomson Reuters's 13F data from January 1999 to March 2012. All variables are the same as defined in Table 19. All coefficient estimates are multiplied by 100. Corresponding t-statistics based on New-West (1987) standard errors are reported in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent level, respectively.

	HH	AN	CRS	LR
Intercept	-0.013**	-0.038***	-0.008	-0.015***
	(-2.57)	(-6.62)	(-1.61)	(-2.88)
IOF_{t-1}	0.122***	0.097***	-0.155***	-1.349***
	(11.71)	(6.39)	(-6.36)	(-14.31)
IOF_{t-2}	-0.020**	-0.006	-0.091***	0.125***
	(-2.29)	(-0.42)	(-5.04)	(2.83)
IOF_{t-3}	-0.032***	-0.024*	-0.074***	0.027
	(-3.98)	(-1.71)	(-4.10)	(0.59)
IOF_{t-4}	-0.048***	-0.031**	-0.098***	-0.000
	(-5.19)	(-2.39)	(-5.30)	(-0.00)
IOF_{t-5}	-0.039***	-0.018	-0.105***	0.027
	(-4.57)	(-1.37)	(-5.67)	(0.62)
TOI_{t-1}	0.123***	-0.039***	0.152***	1.337***
	(11.91)	(-3.39)	(13.99)	(15.22)
TOI_{t-2}	-0.102***	-0.070***	-0.092***	-0.210***
	(-12.47)	(-7.39)	(-10.89)	(-5.11)
TOI_{t-3}	-0.049***	-0.024**	-0.040***	-0.074*
	(-5.80)	(-2.35)	(-4.71)	(-1.77)
TOI_{t-4}	-0.026***	-0.013	-0.017**	-0.033
	(-3.13)	(-1.27)	(-2.00)	(-0.79)
TOI_{t-5}	-0.026***	-0.003	-0.017**	-0.045
	(-2.94)	(-0.28)	(-1.97)	(-1.13)
$SPRD_{t-1}$	2.364***	7.080***	2.264***	2.409***
	(10.38)	(9.29)	(9.97)	(10.56)
$TURN_{t-1}$	0.069***	0.056***	0.083***	0.066***
	(10.11)	(9.20)	(11.46)	(10.38)
RET_{t-1}	-1.215***	-0.727***	-1.205***	-1.545***
	(-7.51)	(-4.35)	(-7.56)	(-10.11)
RET_{t-2}	-1.081***	-1.042***	-1.060***	-1.041***
	(-9.21)	(-7.02)	(-9.12)	(-8.83)
RET_{t-3}	-0.653***	-0.423***	-0.663***	-0.648***
	(-6.78)	(-3.36)	(-6.96)	(-6.72)
RET_{t-4}	-0.404***	-0.409***	-0.430***	-0.435***
	(-3.92)	(-3.15)	(-4.20)	(-4.23)
RET_{t-5}	-0.205**	-0.257**	-0.235**	-0.226**
	(-2.00)	(-2.01)	(-2.32)	(-2.21)
Adjusted R^2	0.025	0.039	0.026	0.027
Observation	10,015,095	5,033,554	10,015,095	10,015,095

Table 11. Horse races in daily return predictive ability

This table presents Fama-MacBeth (1973) regression results with competing institutional order flow estimates. We replicate the regression analysis in Table 10 using HH and another institutional order flow in the same regression. For brevity, we only report the coefficient estimates of order flow, while the regressions always include the full set of control variables. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Nasdaq from January 1999 to March 2012. All coefficient estimates are multiplied by 100. Corresponding t-statistics based on New-West (1987) standard errors are reported in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent level, respectively.

	AN	CRS	LR
Intercept	-0.035***	-0.009*	-0.013**
	(-5.93)	(-1.69)	(-2.58)
HH_{t-1}	0.073***	0.188***	0.156***
	(3.24)	(18.23)	(15.59)
HH_{t-2}	-0.018	-0.012	-0.023***
	(-1.02)	(-1.41)	(-2.60)
HH_{t-3}	-0.022	-0.027***	-0.034***
	(-1.22)	(-3.24)	(-4.26)
HH_{t-4}	-0.040**	-0.038***	-0.046***
	(-2.03)	(-3.95)	(-4.96)
HH_{t-5}	-0.026	-0.027***	-0.039***
	(-1.49)	(-3.03)	(-4.66)
IOF_{t-1}	0.008	-0.289***	-1.407***
	(0.27)	(-10.98)	(-14.81)
IOF_{t-2}	0.014	-0.087***	0.124***
	(0.58)	(-4.73)	(2.83)
IOF_{t-3}	0.006	-0.060***	0.037
	(0.22)	(-3.17)	(0.83)
IOF_{t-4}	0.018	-0.073***	0.011
	(0.68)	(-3.88)	(0.24)
IOF_{t-5}	0.018	-0.083***	0.042
	(0.75)	(-4.32)	(0.98)
TOI_{t-1}	-0.034***	0.156***	1.376***
	(-2.81)	(14.33)	(15.59)
TOI_{t-2}	-0.072***	-0.093***	-0.209***
	(-7.25)	(-11.11)	(-5.12)
TOI_{t-3}	-0.028***	-0.041***	-0.080*
	(-2.68)	(-4.81)	(-1.91)
TOI_{t-4}	-0.012	-0.017**	-0.038
	(-1.13)	(-1.98)	(-0.91)
TOI_{t-5}	-0.005	-0.017*	-0.054
	(-0.43)	(-1.91)	(-1.35)
Adjusted R^2	0.041	0.027	0.028
Observation	5,033,554	10,015,095	10,015,095

Table 12. Investment strategies based on institutional order flow

This table reports the performance of investment strategies based on estimated institutional order flows (IOFs) as indicated on the top of each column. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data from January 1999 to March 2012. We sort all stocks into decile portfolios based on IOFs every day. The average portfolio returns on the next day are reported. Also reported are the return differentials and alphas with respect to the Fama-French (2015) five factors between the top and bottom decile portfolios. All variables are the same as defined in Table 19. Corresponding t-statistics based on New-West (1987) standard errors are reported in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent level, respectively.

	HH	AN	CRS	LR
Low (%)	0.103	0.106	0.167	0.191
2	0.077	0.026	0.147	0.126
3	0.115	0.006	0.151	0.105
4	0.110	0.019	0.084	0.096
5	0.026	0.047	0.009	0.072
6	0.062	0.072	0.052	0.038
7	0.102	0.099	0.079	0.032
8	0.113	0.114	0.095	0.062
9	0.125	0.143	0.098	0.098
High	0.155	0.165	0.105	0.168
HML (%)	0.052***	0.059***	-0.062***	-0.023
	(6.81)	(6.96)	(-6.86)	(-1.59)
FF5 Alpha (%)	0.049***	0.056***	-0.063***	-0.007
	(6.92)	(6.99)	(-8.13)	(-0.67)

Table 13. Contemporaneous price impact of institutional trading

This table presents Fama-MacBeth (1973) regression results of risk-adjusted mid-quote stock return on contemporaneous hedge fund order flow (HH^{HF}), non-hedge fund order flow (HH^{NHF}) and various control variables. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data from January 1999 to March 2012. All variables are the same as defined in Table 19. All coefficient estimates are multiplied by 100. Corresponding t-statistics based on New-West (1987) standard errors are reported in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent level, respectively.

	(1)	(2)	(3)	(4)
Intercept	-0.034***	-0.016***	-0.027***	-0.038***
	(-6.89)	(-3.40)	(-5.44)	(-6.53)
HH_t^{HF}	11.142***		2.043***	1.008***
	(50.44)		(5.53)	(2.91)
HH_t^{NHF}		2.636***	2.363***	1.418***
		(60.89)	(38.23)	(29.41)
TOI_t				4.403***
				(90.02)
$SPRD_{t-1}$				5.252***
				(17.99)
$TURN_{t-1}$				-0.149***
				(-14.55)
RET_{t-1}				-2.953***
				(-19.74)
RET_{t-2}				-1.894***
				(-16.52)
RET_{t-3}				-1.230***
				(-13.16)
RET_{t-4}				-0.825***
				(-8.31)
RET_{t-5}				-0.543***
				(-5.49)
Adjusted R^2	0.011	0.016	0.018	0.127
Observation	9,610,012	10,202,019	9,417,101	9,417,101

Table 14. Private information in hedge fund and non-hedge fund

This table replicates Fama-MacBeth (1973) regression in Table 10 using our proposed hedge fund (HH^{HF}) and non-hedge fund (HH^{NHF}) order flow estimates. For brevity, we only report the coefficient estimates of order flow variables, while the regressions always include the full set of control variables. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data from January 1999 to March 2012. All variables are the same as defined in Table 19. All coefficient estimates are multiplied by 100. Corresponding t-statistics based on New-West (1987) standard errors are reported in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent level, respectively.

	(1)	(2)	(3)
Intercept	-0.016***	-0.013**	-0.016***
	(-3.00)	(-2.54)	(-3.06)
HH^{HF}_{t-1}	1.508***		1.603***
	(18.49)		(17.18)
HH^{HF}_{t-2}	-0.070		0.065
	(-1.03)		(0.79)
HH^{HF}_{t-3}	-0.175***		-0.026
	(-2.74)		(-0.32)
HH^{HF}_{t-4}	-0.322***		-0.096
	(-4.65)		(-1.24)
HH^{HF}_{t-5}	-0.233***		-0.020
	(-3.38)		(-0.24)
HH^{NHF}_{t-1}		0.137***	-0.015
		(10.25)	(-0.98)
HH^{NHF}_{t-2}		-0.024**	-0.028**
		(-2.07)	(-2.03)
HH^{NHF}_{t-3}		-0.035***	-0.032**
		(-3.34)	(-2.37)
HH^{NHF}_{t-4}		-0.063***	-0.055***
		(-5.36)	(-4.28)
HH^{NHF}_{t-5}		-0.055***	-0.049***
		(-5.04)	(-3.71)
TOI_{t-1}	0.083***	0.118***	0.085***
	(8.09)	(11.48)	(8.30)
TOI_{t-2}	-0.101***	-0.099***	-0.099***
	(-12.58)	(-12.10)	(-12.22)
TOI_{t-3}	-0.045***	-0.046***	-0.044***
	(-5.27)	(-5.46)	(-5.12)
TOI_{t-4}	-0.029***	-0.025***	-0.026***
	(-3.45)	(-2.95)	(-3.07)
TOI_{t-5}	-0.023***	-0.024***	-0.019**
	(-2.58)	(-2.76)	(-2.17)
Adjusted R^2	0.027	0.025	0.027
Observation	9,414,134	10,009,538	9,408,616

Table 15. Return prediction in subsamples

This table examines the predictive ability of estimated institutional order flows in subsamples based on firm characteristics. We replicate the regression analysis in Table 14. For brevity, we only report the coefficient estimates of our proposed hedge fund order flow (HH^{HF}) and non-hedge fund order flow (HH^{NHF}), while the regressions always include the full set of control variables. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data from January 1999 to March 2012. In first (second) three columns, we separate our sample into tertile subsamples based on market capitalization (relative bid-ask spread). In subsequent two columns, we divide the sample into into an early (1999-2004) and a late (2005-2012) subperiods. In the last two columns, we compare stocks listed on NYSE and AMEX versus Nasdaq. All variables are the same as defined in Table 19. All coefficient estimates are multiplied by 100. Corresponding t-statistics based on New-West (1987) standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent lavel, respectively.

	Size			Spread			Time		ExMarket	
	Small	Middle	Large	Wide	Medium	Narrow	Early	Late	NYMX	NASDAQ
Intercept	-0.033***	-0.037***	-0.019***	-0.024***	-0.054***	-0.038***	-0.032***	-0.003	-0.044***	-0.003
	(-4.30)	(-5.11)	(-3.08)	(-3.11)	(-8.64)	(-6.44)	(-3.58)	(-0.48)	(-6.29)	(-0.41)
HH^{HF}_{t-1}	4.260***	1.066***	0.719***	3.110***	1.773***	0.834***	1.742***	1.488***	1.718***	1.956***
	(16.60)	(6.67)	(5.53)	(13.79)	(12.30)	(6.59)	(10.64)	(14.44)	(9.24)	(14.91)
HH^{HF}_{t-2}	-0.560**	0.013	0.447***	-0.178	-0.008	0.286**	0.340**	-0.161*	0.197	-0.029
	(-2.36)	(0.09)	(3.51)	(-0.84)	(-0.07)	(2.45)	(2.40)	(-1.78)	(1.39)	(-0.25)
HH^{HF}_{t-3}	-0.269	-0.122	0.320**	-0.011	0.100	0.055	0.112	-0.139	0.138	-0.089
	(-1.22)	(-0.93)	(2.53)	(-0.05)	(0.81)	(0.48)	(0.87)	(-1.36)	(0.88)	(-0.78)
HH^{HF}_{t-4}	0.053	-0.033	0.128	0.037	-0.028	-0.246**	-0.142	-0.058	0.021	0.018
	(0.22)	(-0.25)	(1.02)	(0.19)	(-0.22)	(-2.30)	(-1.11)	(-0.62)	(0.11)	(0.16)
HH^{HF}_{t-5}	-0.308	-0.133	0.388***	-0.247	0.110	0.154	0.045	-0.073	0.205	-0.057
	(-1.30)	(-1.01)	(3.33)	(-1.30)	(0.90)	(1.29)	(0.32)	(-0.74)	(1.26)	(-0.50)
HH^{NHF}_{t-1}	-0.198***	0.068**	-0.023	0.005	-0.050**	-0.020	-0.045	0.010	-0.062***	0.007
	(-3.52)	(2.38)	(-1.32)	(0.11)	(-2.11)	(-1.04)	(-1.62)	(0.61)	(-3.08)	(0.31)
HH^{NHF}_{t-2}	-0.023	-0.034	-0.045**	-0.015	-0.011	-0.045**	-0.007	-0.046***	-0.021	-0.025
	(-0.43)	(-1.31)	(-2.52)	(-0.37)	(-0.50)	(-2.42)	(-0.28)	(-3.13)	(-1.08)	(-1.15)
HH^{NHF}_{t-3}	-0.063	-0.035	-0.039**	-0.110***	-0.049**	-0.018	-0.032	-0.032**	-0.028	-0.036*
	(-1.21)	(-1.39)	(-2.43)	(-2.76)	(-2.13)	(-1.00)	(-1.40)	(-2.02)	(-1.57)	(-1.78)
HH^{NHF}_{t-4}	-0.127**	-0.130***	-0.030*	-0.081**	-0.073***	-0.022	-0.054**	-0.056***	-0.047**	-0.071***
	(-2.42)	(-5.42)	(-1.78)	(-2.02)	(-3.26)	(-1.33)	(-2.33)	(-4.13)	(-2.17)	(-3.46)
HH^{NHF}_{t-5}	-0.019	-0.085***	-0.036**	-0.078**	-0.044**	-0.051***	-0.067***	-0.035**	-0.015	-0.081***
	(-0.37)	(-3.61)	(-2.20)	(-2.06)	(-2.11)	(-2.75)	(-2.81)	(-2.46)	(-0.86)	(-3.97)
Adjusted R^2	0.031	0.035	0.062	0.032	0.036	0.049	0.024	0.030	0.043	0.028
Observation	2,810,254	3,289,716	3,308,646	3,277,090	3,225,654	2,905,872	4,481,597	4,927,019	4,226,645	5,181,971

Table 16. Investment analysis of hedge/non-hedge fund order flow

This table documents the performance of investment strategies based on our proposed hedge fund order flow (HH^{HF}) and non-hedge fund order flow (HH^{NHF}) described in Section II.B.2. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data from January 1999 to March 2012. We sort all the stocks into decile portfolios based on IOFs every day. The average portfolio returns on the next day are reported. Also reported are the return differentials and alphas with respect to the Fama-French (2015) five factors between the top and bottom decile portfolios. All variables are the same as defined in Table 19. Corresponding t-statistics based on New-West (1987) standard errors are reported in parentheses. ***, **, and * indicate statistical significance at 1, 5, and 10 percent level, respectively.

	HH^{HF}	HH^{NHF}
Low (%)	0.111	0.107
2	0.083	0.084
3	0.098	0.129
4	0.045	0.114
5	0.028	0.016
6	0.061	0.053
7	0.077	0.092
8	0.099	0.113
9	0.128	0.130
High	0.192	0.150
HML (%)	0.082***	0.043***
	(10.95)	(5.95)
FF5 Alpha (%)	0.080***	0.042***
	(11.04)	(6.00)

Table 17. Predicting CAR around corporate events

This table presents ordinary least squares regression results for the following equation,

$$
\begin{aligned}
CAR_{i,t-1,t+1} &= \alpha_t + \sum_{k=2}^{6} \beta_{k,t}^{H} HH_{i,t-k}^{HF} + \sum_{k=2}^{6} \beta_{k,t}^{N} HH_{i,t-k}^{NHF} \\
&+ \sum_{k=2}^{6} \gamma_{k,t}^{TAQ} TOI_{i,t-k} + \gamma_t^{T} TURN_{i,t-2} + \gamma_t^{S} SPRD_{i,t-2} + \sum_{k=2}^{6} \gamma_{k,t}^{R} RET_{i,t-k} + \epsilon_{i,t},
\end{aligned}
$$

where for each corporate event i announced on day t, CAR is cumulative abnormal return from day $t-1$ to $t+1$, HH^{HF} and HH^{NHF} are our proposed hedge fund and non-hedge fund order flow, and all the explanatory variables are the same as defined in Table 19 with event subscription i instead of firm subscription. For brevity, we only report the coefficient estimates of the hedge/non-hedge fund order flow with the full set of control variables in the regressions. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data from January 1999 to March 2012. The corporate events we study are quarterly earnings announcements in the first column, extreme price movement exceeding two standard deviations of daily returns and not followed by return reversal for at least ten days in the second column, and combination of the two corporate events in the third column. All coefficient estimates are multiplied by 100. Corresponding t-statistics based on firm and date clustered standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively.

	Earnings	Price Jump	Combination
Intercept	0.232***	0.543***	0.420***
	(5.66)	(11.81)	(12.24)
HH_{t-2}^{HF}	6.014***	3.758**	4.481***
	(3.39)	(2.24)	(3.50)
HH_{t-3}^{HF}	-1.116	2.771*	1.675
	(-0.61)	(1.89)	(1.46)
HH_{t-4}^{HF}	2.645	4.847***	4.364***
	(1.46)	(2.62)	(3.12)
HH_{t-5}^{HF}	0.640	0.878	1.014
	(0.39)	(0.56)	(0.84)
HH_{t-6}^{HF}	-1.992	1.203	0.241
	(-1.09)	(0.79)	(0.20)
HH_{t-2}^{NHF}	-0.950***	-0.910***	-0.829***
	(-3.48)	(-3.41)	(-4.16)
HH_{t-3}^{NHF}	0.057	-0.786***	-0.504**
	(0.20)	(-2.94)	(-2.50)
HH_{t-4}^{NHF}	-0.414	-0.753***	-0.675***
	(-1.39)	(-2.60)	(-3.22)
HH_{t-5}^{NHF}	-0.574**	-0.547**	-0.540***
	(-2.09)	(-2.09)	(-2.80)
HH_{t-6}^{NHF}	0.062	-0.735***	-0.446**
	(0.22)	(-2.88)	(-2.33)
Adjusted R^2	0.011	0.005	0.006
Observation	128,454	230,149	345,339

Table 18. Institutional trading on asset pricing anomalies

This table presents panel regression results with firm-fixed effect using the least squares dummy variable (LSDV) method for the following equation,

$$\begin{aligned} OF_{i,t} &= \beta MISP_{i,t-1} + \gamma^{F}\text{FirstWeek}_{i,t-1} + \gamma^{L}\text{LastWeek}_{i,t-1} \\ &+ \gamma^{FX}\text{FirstWeek}_{i,t-1} \times MISP_{i,t-1} + \gamma^{LX}\text{LastWeek}_{i,t-1} \times MISP_{i,t-1} + \epsilon_{i,t}, \end{aligned}$$

where for each firm i on day t, OF is an institutional order flow estimate as indicated on the top of each column, FirstWeek (LastWeek) is a dummy variable for the first (last) week in a month, and $MISP$ is a mispricing index composed from nine categorical variables for well-known stock return anomalies including net stock issues of Loughran and Ritter (1995); composite equity issues of Daniel and Titman (2006); momentum of Jegadeesh and Titman (1993); gross profitability of Novy-Marx (2013); return on assets of Chen, Novy-Marx, and Zhang (2010); total accurals of Sloan (1996); net operating assets of Hirshleifer, Hou, Teoh, and Zhang (2004); asset growth of Cooper, Gulen, and Schill (2008); investment-to-assets of Titman, Wei, and Xie (2004). HH^{HF} and HH^{NHF} are our proposed hedge fund and non-hedge fund order flow, respectively, described in Section II.B.2. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars that have information in TAQ, CRSP, and Thomson Reuters's 13F data from January 1999 to March 2012. All coefficient estimates are multiplied by 100. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent level, respectively.

	HH^{HF}		HH^{NHF}	
$MISP$	0.833***	0.988***	-4.315***	-3.434***
	(4.34)	(4.53)	(-3.74)	(-2.62)
First-Week		0.049**		0.359***
		(2.18)		(2.66)
Last-Week		-0.110***		-0.298**
		(-5.57)		(-2.52)
First-Week $\times MISP$		-1.213***		-5.104**
		(-3.09)		(-2.16)
Last-Week $\times MISP$		0.123		-0.442
		(0.36)		(-0.21)
Firm Fixed Effect	Yes	Yes	Yes	Yes
Adjusted R^2	0.051	0.051	0.026	0.026
Observation	1,730,798	1,730,798	1,790,153	1,790,153

III. Informed Trading and Price Discovery in 90 Years

A. Introduction

Informed trading is at the center of price discovery in models such as Kyle (1985), Glosten and Milgrom (1985), and Easley and O'Hara (1987). Empirical analysis of the effect of informed trading, however, is difficult because traders do not disclose their motivation in reality. Econometricians typically measure informed trading by order imbalance calculated using tick data as in Chordia and Subrahmanyam (2004), Bernile, Hu, and Tang (2016), and Baruch, Panayides, and Venkataraman (2017).[23] This approach relies on availability of tick data, hence inapplicable in the U.S. equity market before Institute for the Study of Security Markets (ISSM) started in 1983. Most other markets have either much shorter histories of tick data or no such data at all. Alternatively, researchers have examined activities of certain groups of investors that are likely to have advanced information.[24] Nonetheless, these methods become feasible only recently when the corresponding data become available. It remains challenging to empirically examine informed trading in the long history of financial markets and beyond the U.S. equity market.

In this study, we propose a measure of informed trading based on an uncontroversial idea that informed trading positively affects security prices. This idea can be illustrated using the Kyle (1985) model. Kyle's key result can be summarized by the price updating rule of a risk neutral market maker as $\Delta P = \lambda x$, where x is the order imbalance the market maker receives and λ is the price sensitivity termed as "market depth" by Kyle. Determined by the relative amount of expected informed trading, λ essentially measures illiquidity of an asset. The pricing relation can be rewritten as $x = \Delta P/\lambda$. The price change ΔP is observable but illiquidity λ is a latent variable. Aided by the literature on low frequency proxies for λ starting from Roll (1984), this transformation

[23]Several recent studies such as Brennan, Huh, and Subrahmanyam (2018) and Back, Crotty, and Li (2018) use order flow to estimate posterior probabilities of informed trading from structural models instead of examining raw order imbalance.

[24]For example, informed trading is identified by activities of corporate insiders (Finnerty, 1976), institutional investors (Boehmer and Kelley, 2009; Edmans, 2009), short sellers (Christophe, Ferri, and Angel, 2004; Boehmer, Jones, and Zhang, 2008; Engelberg, Reed, and Ringgenberg, 2012), and options traders (Easley, O'Hara, and Srinivas, 1998).

leads to an implied order imbalance (IOI) that can be used to identify informed trading. The IOI also contains the imbalance from noise trading. As Duarte and Young (2009) point out, the order imbalance dynamic can result from liquidity shocks too. Both Odders-White and Ready (2008) and Back, Crotty, and Li (2018) extend the original Kyle model to incorporate trading sessions without information shocks. Following the same logic, we remove the impact of liquidity shocks on IOI to arrive at our measure of informed trading. Unlike those studies, however, we adopt a reduced-form approach given our focus on the pricing effect in the cross-section. Specifically, we regress individual stock's IOI on illiquidity and returns in the cross-section every week. We define the residual from this regression as our implied informed trading (IIT). Intuitively, IIT approximates the weekly order imbalance orthogonalized to stock returns and illiquidity in the cross-section. If it captures private information, IIT should obtain positive return predictive ability.

Although our measure of informed trading is motivated by the Kyle (1985) model, it is important to note that the asserted price impact from order imbalance is a generic result in market microstructure theories. Indeed, in sequential trade models such as Glosten and Milgrom (1985) and Easley and O'Hara (1987), the asset price set by the market maker is updated after order arrival to reflect the probability of new information in the order flow. Motivated by inventory cost rather than asymmetric information, the dynamic inventory models such as Ho and Stoll (1983) and Spiegel and Subrahmanyam (1995) also investigate the price changes following transactions as a result of the price pressure faced by market makers. Therefore, it is possible to apply our measure of informed trading to various types of markets. One limitation of IIT is that in some rare cases, informed traders obtain monopoly over advanced information in distanced future, and choose to trade patiently using limit orders only. If the other market participants fail to learn about the information from increased limit orders, this type of informed trading would not lead to a positive price impact, therefore undermining effectiveness of IIT.[25] In empirical tests, this imperfection makes it more challenging for IIT to obtain statistically significant predictive power and our findings can be interpreted as a lower bound of pricing effects from informed trading.

[25]For informed traders' choice of limit or market orders, see e.g. Bloomfield, O'Hara, and Saar (2005), Kaniel and Liu (2006), Collin-Dufresne and Fos (2015), and Baruch, Panayides, and Venkataraman (2017).

This measure of informed trading distinguishes itself from existing measures by requiring only past returns and illiquidity in calculation. While illiquidity in the form of Kyle's λ has to be estimated using high-frequency data, Chordia, Huh, and Subrahmanyam (2009) show that many illiquidity measures calculated using low-frequency price and volume data have high cross-sectional correlations to and similar pricing effects as the Kyle λ. Therefore, it appears reasonable to use alternative illiquidity measures to calculate IIT for studying effects of informed trading in the cross-section. This feature is particularly desirable in empirical studies of large samples because now it enables us to measure informed trading as long as stock price and volume data are available, for example, in the pre-ISSM period in the U.S. or in international markets. Other measures of informed trading are often unavailable in such applications due to data limits.

In our empirical analysis, we calculate IIT using six well-known illiquidity measures that can be calculated each day without using intraday data in the full history of Center for Research in Security Prices (CRSP) between 1927 and 2016. We show that regardless of the illiquidity measure used, IIT always significantly contributes to subsequent price discovery. We conduct all analyses at the weekly frequency in the paper. Our internet appendix to the paper shows that the results are qualitatively the same using daily observations. In univariate portfolio sorts in the full CRSP sample, the top IIT decile portfolio outperforms the bottom IIT decile portfolio by 29.7 to 98.9 basis points (bp) in the following week with strong statistical significance. The return differential retains similar magnitude and statistical significance after we adjust for Fama-French (1993) risk factors as well as a momentum factor. The pricing effect of IIT is persistently positive and significant during subperiods of the 90-year sample, and becomes larger during NBER recession periods and financial crises. Double sorting analyses confirm that the predictability of IIT does not result from return reversals or liquidity effects. The predictive ability of IIT is robust in Fama-MacBeth (1973) regressions controlling for illiquidity and past returns at the same time. We also find that the positive pricing effect of IIT is permanent as there is no subsequent reversal. Rather, for some IIT measures, the positive predictive ability extends beyond one week and even exists up to four weeks ahead, suggesting slow price adjustment consistent with the literature on investor underreaction to information shocks such as Foster, Olsen, and Shevlin (1984) or Bernard and Thomas (1989).

To better understand the nature of information captured by IIT, we conduct three more tests. The first is a horse race between IIT and existing measures of informed trading including aggregate order flow, institutional trades, short selling activities, and options trading volume. In Fama-MacBeth (1973) regressions, the predictive ability of IIT is not fully absorbed by these existing informed trading measures although the magnitude decreases. The weakened effect of IIT is not only a result of having more control variables but also because of more restricted samples in these regressions. The result indicates that our measure of informed trading captures additional private information absent in the other measures while the information content also overlaps among different measures.

Our second test is about cross-sectional variation of the effect from informed trading. If IIT captures activity of informed traders, it should be more effective for stocks with higher degrees of information asymmetry because private information can be more valuable for such stocks. Consistent with this conjecture, we find that the predictability of IIT is significantly stronger for more opaque stocks characterized by low market capitalization, little analyst coverage, or low institutional ownership.

The third test concerns fundamental information flow associated with significant corporate events, when IIT is expected to be most effective with concentrated informed trading. To identify value-related events, we construct a comprehensive sample consisting of earnings announcements, analyst recommendation updates, scheduled 13D filings, and permanent price jumps between 1993 and 2016. Our event study results show that all of the IIT measures obtain positive and significant predictive ability of abnormal returns around these corporate event days. Moreover, when we add an event week dummy and its interactions with lagged IIT to the Fama-MacBeth (1973) regressions of returns in the same period, we find that the interactions also have positive and significant coefficients in addition to IIT's significant predictive power. As the predictive ability of IIT strengthens before corporate events, it suggests that a large portion of the private information captured by IIT is indeed related to fundamental value of stocks as identified by ex post events.

We complete our analysis by investigating the role of informed trading in international equity markets. Such empirical analysis is challenging using traditional measures of informed trading. Using data from Datastream, we are able to construct three IIT measures in the other G10 countries out of the six measures that we use in the U.S. sample. Because of data limitation on the bid, ask, high, and low prices in those markets, we cannot calculate IIT using the relative bid-ask spread, high-low spread, or close-high-low spread for the full international samples. We find that the three IIT measures based on Amihud illiquidity, market microstructure invariance, and turnover generate positive and significant abnormal returns in weekly portfolio sorts in all the markets like in the U.S. market. This result shows that our measure of informed trading is effective in capturing private information even in international markets.

Our study contributes to the finance literature by providing a simple and effective way of measuring trading activity of informed traders. This measure can be calculated using only stock returns and liquidity proxies. Therefore, the proposed IIT can be employed in empirical studies where traditional measures of informed trading are unavailable. We show that regardless of the illiquidity proxy used, IIT is effective in extracting private information in various samples across time periods and markets. Moreover, the information content of IIT complements that in existing measures by providing additional return predictability. This method can be useful for both asset pricing and corporate finance research concerning informed trading in a wide variety of applications.

We also contribute to the literature of informed traders' role in price discovery. While the effect of informed trading is well understood since Kyle (1985) and Glosten and Milgrom (1985), empirical evidence exists largely in restricted samples only. Using ISSM and New York Stock Exchange's Trade and Quote (TAQ) data, researchers can examine informed trading using stock order flow only after 1983. Other methods involving institutional trading, short selling, and options trading lead to further reduced sample periods and smaller cross-sections. Taking advantage of the flexibility of IIT, we extend our analysis of informed trading well beyond the commonly analyzed sample. We provide evidence that informed trading is a universal phenomenon and its contribution to subsequent price discovery is significant. Therefore, our findings generalize the conclusion of prior literature such as Chordia and Subrahmanyam (2004). Our result also directly speaks to the concern of informed trading using limit orders. Although this phenomenon invalids many measures of informed trading *risk* as Collin-Dufresne and Fos (2015) show, in our extensive samples, it appears that an informed trader is more likely to use market orders on average, leading to return continuation with the same sign of private information.

Related to our study, Campbell, Grossman, and Wang (1993) and Llorente, Michaely, Sarr, and Wang (2002) examine the interaction of return and turnover to study time-series relations between trading volume and subsequent returns at the market level and for individual stocks, respectively. Tetlock (2010) uses the same interaction to study the impact of news on return autocorrelations in the cross-section. Unlike those studies, we do not examine the interaction of return and turnover directly. Instead, we use its component orthogonal to return and liquidity to measure informed trading. Moreover, turnover is only one of the thirteen illiquidity measures we use to construct IIT in the main paper and internet appendix. We show that this choice of illiquidity measure is not critical to obtain our results. Therefore, our results should not be interpreted as a conditional effect of return reversal or momentum related to trading volumes.

The rest of the paper is organized as follows. Section III.B describes the theoretical background of our measure of informed trading. Section III.C describes the data used in the study and sample selection. Section III.D conducts price discovery tests using our proposed measure of informed trading. Section III.E concludes.

B. Implied Informed Trading

We are interested in empirical evidence of how private information is impounded into asset prices. To arrive at a flexible measure of private information in the wide cross-section of stocks, we use the seminal model of Kyle (1985) to illustrate the idea. In Kyle's model, there are three types of traders: a single insider who exclusively has information on liquidation value of a tradable asset, uninformed noise traders who transact randomly, and a risk-neutral market maker. The insider and noise traders simultaneously decide the direction and quantity of trades first. The insider sets the quantity to maximize her profit in the direction consistent with the private information. Upon observing net order flow from the insider and noise traders, the market maker sets a price equivalent to the expected liquidation value of the asset conditioning on the net order flow, and clears market orders at the price. In equilibrium, the market maker updates the asset price as follows:

$$P = p_0 + \lambda(\tilde{x} + \tilde{u}), \tag{9}$$

where P is the transaction price set by the market maker in the current round of trading, p_0 is the price from the previous round of trading, λ is the inverse of market depth determined by the expectation of intensity of informed trading, $\tilde{x}$ is the signed quantity of orders from the insider and $\tilde{u}$ is the signed quantity of orders from noise traders. That is, price innovation is the product of inverse market depth and net order flow observed by the market maker.

Rearranging Equation (9) with respect to net order flow gives us

$$Net\ Order\ Flow = (\tilde{x} + \tilde{u}) = (P - p_0) \times \frac{1}{\lambda}. \tag{10}$$

Because ex post price changes are directly observable, Equation (10) allows us to estimate order flow as a fraction of the price change and an illiquidity proxy on the premise that such illiquidity measure is correlated with λ. Indeed, Chordia, Huh, and Subrahmanyam (2009) show that illiquidity measures using daily prices and trading volumes have high cross-sectional correlations with λ estimates using high frequency data. Therefore, it is reasonable to use Equation (10) to calculate implied order imbalance (IOI) using existing low frequency illiquidity measures in the cross-section.

However, net order flow alone is insufficient to identify informed trading because noise trading also contributes to the imbalance. That is why in each round of trading, the market maker achieves only partial price adjustment facing contaminated order imbalance, which leads to further price response when the liquidation value is revealed later. To refine information content in net order flow, we construct a measure of informed trading that explicitly removes liquidity effect from IOI. Specifically, we regress IOI on contemporaneous illiquidity and stock return in the cross-section. The residual from this regression is our measure for implied informed trading (IIT). Intuitively, IIT is the component in IOI orthogonal to illiquidity and stock returns. Given our interest in price discovery due to informed order flow, this orthogonalization is necessary to exclude pricing effects from illiquidity such as in Amihud (2002) and return reversal as in Lo and Mackinlay (1990). Note that although IIT can potentially capture information more accurately than net order imbalance, it requires ex post price updates in the large cross-section to calculate. Therefore, while this measure benefits econometricians studying price discovery, it is unlikely to benefit market makers in price setting.

Theoretically, this measure of informed trading can be applied to any trading interval with constant illiquidity. Therefore, the performance of this method could be more suitable for relatively short trading periods because liquidity shocks occur frequently in longer horizons (Bali, Peng, Shen, and Tang (2014) and Chordia, Hu, Subrahmanyam, and Tong (2019). Because we want to avoid using tick data for maximum flexibility, choosing short horizons may run into another problem of measurement error due to few observations available in a measurement interval. To balance the two considerations, we apply our method to weekly stock market observations. Our internet appendix shows that the results are qualitatively the same when we examine daily observations or use illiquidity measures calculated using longer horizons.

We use six well-known low frequency illiquidity measures in the literature to calculate IIT, including the turnover based Amihud's (2002) illiquidity measure as in Brennan, Huh, and Subrahmanyam (2013), market microstructure invariance as in Kyle and Obizhaeva (2016) and Fong, Holden, and Tobek (2018), inverse share volume turnover ratio, relative bid-ask spread, high-low spread as in Corwin and Schultz (2012), and close-high-low spread as in Abdi and Ronaldo (2017).[26] The calculation details of these illiquidity measures can be found in the appendix to the

[26]We have also experimented with Roll's (1984) effective spread, Gibbs estimate from Hasbrouck (2004, 2009), effective tick size and effective spread from Holden (2009), Lesmond, Odgen and Trzcinka's (1999) effective spread, the frequency of zero returns from Lesmond, Odgen and Trzcinka (1999), and Fong, Holden, and Trzcinka's (2017) effective spread. The results using these alternative illiquidity measures are consistent with the main findings. However,

paper. For Amihud's illiquidity (AMI), share volume turnover ($TURN$), relative percentage spread ($BASPRD$), and high-low spread ($CSSPRD$), we first calculate their values for every stock-day observation and then take the average over a week starting every Thursday. The market microstructure invariance implied illiquidity ($INVAR$) and close-high-low spread ($ARSPRD$) are estimated using daily observations in the same week. We require at least three daily observations in a week to calculate weekly illiquidity to reduce measurement errors.

After obtaining weekly illiquidity, we then calculate corresponding IOI by dividing contemporaneous stock returns and illiquidity in the same week. Note that for turnover ratio, IOI is the product of stock returns and turnover. We use stock returns rather than price changes in our calculation for cross-sectional normalization. Because stock returns can also result from systematic risk exposure, we use the risk-adjusted return with respect to the Fama-French (1993) factors and Carhart's (1997) momentum factor. In the internet appendix to the paper, we show that our main findings are robust to alternative stock return specifications. Specifically, we experiment with raw return, quote mid-range return, mid quote return, open-to-close transaction return, and open-to-close mid quote return as well as alternative weeks starting from different days of a week. We also find similar results using risk-adjusted returns with respect to Fama-French (2015) five factors and Pastor and Stambaugh's (2003) liquidity factor in the sample from 1965 and 2016.

Finally, we run cross-sectional regressions every week:

$$IOI_{i,w} = \alpha_w + \beta_w ILLIQ_{i,w} + \gamma_w RET_{i,w} + \epsilon_{i,w}, \tag{11}$$

where for stock i in week w, $ILLIQ$ is an illiquidity measure corresponding to the IOI measure, RET is weekly risk-adjusted stock return using Fama-French-Carhart four factors. The residuals from the regressions are then defined as implied informed trading (IIT). With six illiquidity measures, we arrive at the following six IITs:

1. $AIIT$: IIT based on the turnover-based Amihud illiquidity (AMI).

these illiquidity estimates generate more missing values and the resulting smaller samples are not directly comparable to the measures reported in the paper. Therefore, we report those results in the internet appendix for interested readers.

2. $IIIT$: IIT based on market microstructure invariance implied illiquidity ($INVAR$).
3. $TIIT$: IIT based on share volume turnover ratio ($TURN$).
4. $BAIIT$: IIT based on percentage bid-ask spread ($BASPRD$).
5. $CSIIT$: IIT based on high-low spread ($CSSPRD$).
6. $ARIIT$: IIT based on close-high-low spread ($ARSPRD$).

Because informed trading leads to only partial price adjustment in Kyle's (1985) model and many other market microstructure theories, a good measure of informed trading should be correlated with subsequent price movement when the private information it captures becomes public. Therefore, our hypothesis is that IIT positively predicts future stock returns in the cross-section. Alternatively, if the concentration of informed traders is sufficiently high, there might be full price adjustment to informed order flow and there is no subsequent pricing effect. If informed traders overreact to their private signals or mainly use limit orders, the return autocorrelation can even turn negative, as well as correlation between our measure of informed trading and future returns. Therefore, testing return predictive ability of IIT can shed light on the underlying market dynamics behind the data. Such empirical facts are important to market microstructure studies as well as any application of heterogeneously informed investors in financial markets.

C. Data

C.1. Sample Selection

The advantage of our implied informed trading measure is that it can be calculated without using tick data and greatly enhances the flexibility of empirical analysis on informed trading. We construct our sample of all common stocks listed on NYSE, AMEX, and Nasdaq covered by CRSP between 1927 and 2016 with prices above \$5. After the close bid and ask data become available in 1984, we also exclude stock-days with percentage bid-ask spreads below zero or greater than a half. If stock return is missing on one of the days in a week, we exclude the stock-week observation from our sample. Trading volume of stocks primarily listed on Nasdaq exchange markets is adjusted following Gao and Ritter (2010). Specifically, trading volumes of Nasdaq stocks are divided by 2 prior to Februrary 1, 2001, by 1.8 between February 1 and December 31, 2001, and by 1.6 for 2002 and 2003. Due to limited availability of close bid and ask prices, we can only compute $BASPRD$ and $BAIIT$ from 1984 onwards.

C.2. Summary Statistics

After we construct the illiquidity and implied order imbalance variables, we winsorize all the variables at the 1 and 99 percentiles in the cross-section every week. Then, we estimate implied informed trading (IIT) to minimize the potential effect of outliers. The estimated IIT is also cross-sectionally winsorized at the 1 and 99 percentiles every week. Panel A of Table 19 presents the time-series averages of cross-sectional statistics of IITs.

[Place Table 19 about here]

The sample includes $10,841,939$ stock-week observations on $4,682$ weeks between April 1927 and December 2016. On average, there are $2,316$ stocks per week. The number of observations varies slightly for five of the illiquidity measures and corresponding IITs that can be calculated every week during the sample period. $BASPRD$ and $BAIIT$ can only be calculated from 1984. Therefore, the variables are available in only 1,706 weeks although the average number of stocks per week exceeds 3,000 at this later period. Despite the different illiquidity measures employed, the IITs share many common characteristics. Both the mean and median are always statistically indifferent from zero with the standard deviation at a much larger magnitude. The minimum and maximum of the IITs are around three to four times the standard deviation. Note that the IITs we calculate do not have the uniform unit because the illiquidity proxies used have different units. Therefore, we use these IITs to capture signed intensity of informed trading in the cross-section and focus on its predictive power of future returns.

Panel B of Table 19 presents the time-series averages of cross-sectional correlations between IITs. The correlation analysis uncovers two results. First, most of the pairwise correlations are positive with two exceptions between $TIIT$ and $CSIIT$ and between $TIIT$ and $ARIIT$. The general positive correlations indicate that the measures are likely to capture a common component of private information, if any. Second, the magnitude of correlations depends on the type of the underlying illiquidity measure. Specifically, volume-related IITs including $AIIT$, $IIIT$, and $TIIT$ have correlations ranging from 0.190 to 0.659 between them. Similarly, spread-related IITs including $BAIIT$, $CSIIT$, and $ARIIT$ have in-between correlations from 0.142 to 0.335. The correlation between volume-related and spread-related IITs, however, is usually lower. This could be due to different natures of volume-related and spread-related illiquidity measures. Therefore, testing the same hypothesis using IITs based on different illiquidity measures sets a high hurdle for statistical significance.

D. Results

In this section, we test the performance of proposed implied informed trading (IIT) measures by investigating its predictive power for future stock return.

D.1. Portfolio Sorts

We begin our analysis with univariate portfolio sorts. At the end of every Wednesday, we sort all stocks in our sample into decile portfolios based on one of the IIT measures and calculate the average equal-weighted portfolio returns in the next week. Table 20 reports the average decile portfolio returns as well as the return differentials and Fama-French-Carhart alphas between the top and bottom decile portfolios in 90 years between 1927 and 2016. All t-statistics in Table 20 are calculated using Newey-West (1987) standard errors with nine lags.

[Place Table 20 about here]

We find that the trading strategies based on IITs are highly profitable regardless of the illiquidity measure employed. For example, the decile portfolios sorted on $AIIT$ generate monotonically increasing returns and the strategy that buys stocks in the highest decile and sells stocks in the lowest decile generates an average return of 0.538% in the following week with a t-statistic of 22.26. The alpha with respect to the Fama-French-Carhart factors is at the same level with an even larger t-statistic. Strategies based on the other IITs also generate significant abnormal returns with alphas of 0.395% per week for $IIIT$, 0.549% for $TIIT$, 0.500% for $BAIIT$, 0.297% for $CSIIT$, and 0.989% for $ARIIT$.[27] While the portfolio sorts generate significant abnormal return, one should interpret such abnormal returns with caution because transaction costs are not taken into consideration in this analysis. Given the fact that the strategies require weekly rebalancing, potential trading costs can be relatively high, especially in the early sample periods. Therefore, we focus more on the qualitative results from the investment analysis.

Next, we examine time-series pattern in performance of IIT strategies. In Table 21, we first divide the full sample into three subperiods: Early (1927-1950), Middle (1951-1983), and Late (1984-2016). Then we replicate the univariate portfolio sorts in Table 20 in each subsample. For brevity, Table 21 only reports the return differentials between the top and bottom decile portfolios. This analysis yields two findings. First, all IIT strategies achieve significant abnormal returns in all of the subperiods. Second, except $ARIIT$, the IIT strategy is the most profitable in the Early period before 1950. The average abnormal return then decreases during the Middle period between 1951 and 1983, and increases again after 1983. $BAIIT$ is available only in the Late subsample though. This time-series pattern indicates that the superior performance of IIT strategies is persistent in the extensive sample period of 90 years. Concerning impact of macroeconomic conditions on informed trading, we also examine subsamples of NBER Recession and Non-recession periods as well as Crisis (Great Depression (1929-1932), Dot-com Bubble (1999-2001), and Subprime Mortgage Bubble (2007-2009)) and Non-crisis periods. Table 21 shows that all IIT strategies generate larger abnormal returns during NBER recessions and Crisis periods and again, the performance is robust in all subsamples. It is consistent with the notion that private information becomes more valuable

[27]In the internet appendix, we report mid-range portfolio returns following Abdi and Ranaldo (2017) to address concerns about potential bid-ask bounce. The portfolio performance is the qualitatively same as Table 20. Stock mid-range return is measured by the change of mid-range price, defined as the mean of the daily high and low prices, from Thursday to following Wednesday. This alternative return measure can be calculated during our full sample period.

when market uncertainty surges.

[Place Table 21 about here]

Our last portfolio analysis is double sorting on illiquidity (past returns) and IIT to control for any liquidity premium or return reversal effect. We first sort stocks into quintile portfolios based on a control variable. Then, we further sort stocks into quintile portfolios based on IIT within each control variable's quintile portfolio. For reporting brevity, Table 22 reports average returns of the same IIT quintile portfolio across all five quintile control portfolios as well as the return differentials between the top and bottom quintiles and alphas. The control variable is the corresponding illiquidity in Panel A and past returns in Panel B. Regardless of the IIT specification and control variable in the bivariate sorts, we find that the return differentials between the high and low IIT quintile portfolios are always positive and significant, ranging from 15.5 basis points (bp) to 78.4 bp per week with t-statistics between 8.01 and 49.56. The portfolio returns also show clear monotonic increase in most columns along the IIT quintile portfolios. The FF4 alphas are commensurate with the return differentials and have even larger t-statistics. The results suggest that the predictive ability of IIT is different from well-known illiquidity premium and stock return reversal in short horizons.

[Place Table 22 about here]

D.2. Multivariate Regression Analysis

In this subsection, we use Fama-MacBeth (1973) regressions at the stock level to examine return predictability. These regressions can also control for more variables at the same time. Table 23 presents the coefficient estimates of the following model,

$$RET_{i,w} = \alpha_w + \sum_{k=1}^{4} \beta_{k,w}^{IIT} IIT_{i,w-k} + \beta_w^{ILLIQ} ILLIQ_{i,w-1} + \sum_{k=1}^{4} \beta_{k,w}^{RET} RET_{i,w-k} + \epsilon_{i,w}, \qquad (12)$$

where for stock i in week w, RET is the four factor risk-adjusted stock return and $ILLIQ$ is the corresponding illiquidity measure to IIT. To account for serial correlations, we use Newey-West (1987) standard errors with nine lags to calculate the t-statistics.

[Place Table 23 about here]

Consistent with results from portfolio sorts, Table 23 suggests that, regardless of how the IIT measure is calculated, it always has a positive and persistent price impact at weekly horizons in our sample period over 90 years. In the first column, the coefficient of $AIIT$ is 0.037 with a t-statistic of 17.30. In terms of economic significance, a one-standard deviation shock in $AIIT$ would lead to a change of 23.12 bp in the following week's stock return. In this long sample, we find that price information takes longer to be fully incorporated because $AIIT$'s coefficients are 0.016 at the second lag (t-stat = 10.04), 0.011 at the third lag (t-stat = 7.29), and 0.006 at the fourth lag (t-stat = 4.67). We find similar results using $IIIT$, $TIIT$, and $ARIIT$. $BAIIT$ and $CSIIT$ have positive and significant coefficients of 0.070 (t-stat = 10.88) and 0.022 (t-stat = 13.86) at the first lag, respectively. Their coefficients turn insignificant at longer horizons. For robustness checks, we run the same multivariate regressions at daily frequency but only report those results in the internet appendix. All the results are qualitatively the same as in Table 23.

The control variables behave in a consistent manner across all columns. The turnover-based Amihud's illiquidity, microstructure invariance, bid-ask spread, and high-low spread are negatively associated with future stock return, while turnover ratio is positively associated with future returns. The negative pricing effect of liquidity in the short term is consistent with delayed response to liquidity shocks as in Bali et al. (2014) and Chordia et al. (2019). Abdi-Ranaldo spread is insignificant pricing effect on the following week. Past returns are always negatively associated with future returns.

D.3. Controlling for Other Measures of Informed Trading

Next, we investigate if IIT contains unique information flow not captured by existing measures of informed trading. Our alternative measures come from four categories including high frequency order flow, institutional trading, short selling, and options market.

Net order flow has been widely used as an indicator of informed trading in studies such as Chordia and Subrahmanyam (2004). Using data in ISSM and TAQ from 1983 to 2016, we follow Lee and Ready (1991) and Easley, de Prado, and O'Hara (2016) to calculate aggregate order flow termed OI and BVC, respectively. The Lee and Ready (1991) algorithm signs a trade as buyer-initiated (seller-initiated) if the price is above (below) the midpoint of prevailing bid and ask prices. If the trade price is the same as the quote midpoint, the tick rule will be applied. Specifically, a trade will be classified as buyer-initiated (seller-initiated) if the price is higher (lower) than the last different trade price. When calculating OI, we apply a 5-second lag in matching quotes and trades in ISSM and TAQ before 1998 to account for reporting delay. The Easley et al.'s (2016) trade signing algorithm essentially applies the tick rule on either time or volume bulks of trades. We use 50 volume buckets for every stock-day when calculating BVC.

Given various advantages of institutional investors over retail counterparts, there is ample empirical evidence that institutional trades are informative in e.g., Boehmer and Kelly (2009). We use three measures of institutional trading. AN is the difference between buyer execution shares and seller execution shares in Ancerno institutional trade data between 1983 and 2012. LR is the institutional order flow measure of Lee and Radhakrishna (2000) using a \$5,000 cut-off rule. CRS is the institutional order flow measure of Campbell, Ramadorai, and Schwartz (2009). This method first estimates the relation between quarterly total institutional trade imbalance in 13F Filings and publicly observable order imbalance in different trade size bins in the TAQ data. The quarterly relation is then extrapolated to daily observations to estimate daily institutional trade imbalance. We estimate LR and CRS using data in ISSM and TAQ from 1983 to 2016.

Finally, short sellers and options traders have been shown to possess advanced information as their trading volumes predict stock returns (see, e.g., Boehmer, Jones, and Zhang (2008) and

Johnson and So (2012)). To measure short sellers' activity, we calculate days to cover ratio ($DTCR$), defined as the number of shares borrowed scaled by daily trading volume, from Data Explorers from 2007 to 2013.[28] We also calculate an options volume to stock volume ratio (OS) following Johnson and So (2012) to measure informed trading activity in the options market. The need of options volume data in OptionMetrics restricts the sample period to 1996 to 2016.

We first calculate the time-series averages of cross-sectional correlations between IIT and alternative informed trading measures. Note that due to sample restrictions, the results are achieved in different samples. Table 24 shows that most of these correlations are positive, suggesting that different measures of informed trading can capture a common component. Because options trading volume negatively predicts returns, the negative correlations between OS and IITs are consistent with overlapping information. The magnitude of the correlations, however, is weak in general. The average of all correlations reported in Table 24 is 0.036 and the maximum is 0.094. IITs have higher correlations with order flow based measures including OI, BVC, AN, LR, and CRS. This is not surprising because IIT is also rooted in implied order imbalance. Short selling and options market activity, on the other hand, have much lower correlations with IIT.

[Place Table 24 about here]

Next, we estimate the following equation using Fama-MacBeth (1973) regressions:

$$RET_{i,w} = \alpha_w + \sum_{k=1}^{4} \beta_{k,w}^{IIT} IIT_{i,w-k} + \sum_{k=1}^{4} \beta_{k,w}^{INF} INF_{i,w-k} + \beta_w^{ILLIQ} ILLIQ_{i,w-1} + \sum_{k=1}^{4} \beta_{k,w}^{RET} RET_{i,w-k} + \epsilon_{i,w}, \tag{13}$$

where INF is an alternative measure of informed trading.

Table 25 reports only the coefficient estimates of IITs for brevity while the full set of control variables is always used in the regressions.[29] In Panel A, we first add four lags of net order

[28]We have experimented with alternative measures for short selling including short interest ratio, stock lending fees, fee spreads, loanable shares, utilization ratio, and the supply-demand shocks as in Cohen et al. (2007). Those results are qualitatively the same to our main findings. We report the results using $DTCR$ out of conservativeness because $DTCR$ has the strongest predictive ability among all short selling measures in our sample and reduces the predictive power of IIT the most.

[29]The full regression results can be found in our internet appendix.

imbalance based on the Lee and Ready (1991) algorithm into the regressions. We find that these additional control variables of Lee and Ready order imbalance do not affect the predictive ability of IIT. At the first lag, IIT always has a positive and significant coefficient with t-statistics ranging from 9.45 to 23.11. For most IITs, the positive return predictability continues to multiple weeks ahead similar to Table 23. When we use order imbalance based on Easley et al.'s (2016) algorithm, we find similar results on IIT in Panel B, except that its predictive ability is even strengthened for $BAIIT$ and $CSIIT$.

[Place Table 25 about here]

Next, we control for institutional trading measures in Panels C to E. Despite different methods and sample periods used, the results in these three panels clearly show that the predictive ability of IIT is almost intact controlling for institutional trades. Interestingly, the predictive power of IIT and other order flow based measures of informed trading largely complements each other although the correlations between these variables are relatively high.

Finally, we turn to short selling in Panel F and options trading in Panel G as additional controls. Note that the cross-section as well as the sample period decreases significantly in these two panels compared to the early panels due to limited data coverage. Nonetheless, we find that the positive return predictive ability of IIT largely survives these control variables. Positive and significant coefficient estimates of IITs exist in most columns although the magnitude and statistical significance reduce compared to the main results in Table 23. Only $IIIT$ has positive but insignificant coefficients in Panel G controlling for OS. However, even this drawback in performance of IIT comes from sampling rather than additional control variables. In unreported tests, we find almost the same results on IIT in the restricted samples without controlling for $DTCR$ or OS. The results are not surprising given low correlations between IIT and $DTCR$ and OS in Table 24. In summary, the results in this subsection indicate that the information content in IIT is at least partially distinguished from that captured by the other measures of informed trading. Therefore, IIT as a novel information measure can help uncover informed traders' activity and enhance return predictability.

D.4. Subsample Tests on Information Asymmetry

Having documented robust predictive power of IIT, we move on to investigate its cross-sectional variation conditioning on the degree of information asymmetry. We use three proxies for information environment, including market capitalization, sell-side analyst coverage, and institutional ownership. Sell-side analyst coverage is calculated as the number of analysts that have updated earnings estimates at least once within a year in the I/B/E/S database. Institutional ownership is the proportion of outstanding shares held by institutions in the Thomson Reuters 13F database. Although market capitalization is available during our full sample period, sell-side analyst coverage and institutional ownership samples start from only 1980, leading to $7,240,218$ stock-week observations. For each information proxy, we divide the sample into two subsamples based on the median value every week. Then, we compare the predictive ability of IIT between the two subsamples by replicating Table 20 in subsamples. Table 26 reports the average return differentials between the top and bottom IIT quintile portfolios. Such portfolio results also allow us to compare the profitability in subsamples directly at the bottom of each panel.

[Place Table 26 about here]

In Panel A of Table 26, we first divide our sample by market capitalization. Large stocks attract greater attention from investors and financial intermediaries, hence less prone to asymmetric information among investors. Consistent with this conjecture, we find that the abnormal return to IIT strategies is larger in the subsample with low market capitalization. The difference of abnormal returns between small and large stock subsamples is always positive and significant, ranging from 0.175% to 0.729% (t-statistics between with 7.92 and 22.83). Note that although the IIT strategies are more profitable for small firms, they also generate significant abnormal returns for large firms except $BAIIT$.

Similarly, we divide the sample by analyst coverage and institutional ownership in Panels B and

C, respectively. Because stocks with more analysts following and higher institutional ownership are under greater scrutiny, these stocks can be more transparent. Therefore, we expect IIT strategies to be more profitable for stocks with low analyst coverage and low institutional ownership. The results in Panels B and C support this conjecture with significant performance difference between subsamples. Moreover, the IIT strategies always generate positive and significant abnormal returns in all of the subsamples in these two panels.

In summary, the results in this subsection show that the IIT strategies have robust performance in subsamples and the performance is stronger for stocks that likely have higher information asymmetry. These findings reinforce our interpretation of the return predictability of IIT as private information gradually getting impounded into asset prices.

D.5. Corporate Event Studies

To further strengthen the role of IIT as an information measure, we investigate return predictability around corporate events when the value of private information is greater. We construct a comprehensive sample of salient corporate events including quarterly earnings announcements, analyst recommendation changes, schedule 13-D filings, and permanent price jumps. Earnings announcements are arguably the most important scheduled events when public firms release fundamental information. Analyst recommendation updates and schedule 13-D filings reflect important valuation opinions from sell-side and buy-side respectively, and often have significant market impact. We collect quarterly earnings announcement and analyst recommendation change dates from the I/B/E/S database. We extract 13-D filings from the WRDS SEC Analytics Suite. Finally, there can be value related news not covered by these data. We rely on ex post price jumps to identify the remaining events. Specifically, we examine price movement exceeding two standard deviation shocks that does not fully reverse within ten days. Our event sample only begins in 1994 due to limited coverage of 13D filings. After merging the event data with CRSP data, we have $2,007,161$ observations in this event sample.

We first test the predictive power of IIT for announcement returns. If informed traders are rational, they should be able to predict the market reaction to the news. Therefore, pre-event informed trading should be consistent with the abnormal announcement return around the event. We estimate the following equation using ordinary least squares (OLS) regressions,

$$AR_{i,w} = \alpha_w + \sum_{k=1}^{4} \beta_{k,w}^{IIT} IIT_{i,w-k} + ILLIQ_{i,w-1} + \sum_{k=1}^{4} \beta_{k,w}^{AR} AR_{i,w-k} + \epsilon_{i,w}, \tag{14}$$

where for each event i in week w, AR is the risk-adjusted return calculated using event weeks $w-26$ to $w+26$, and all the explanatory variables are the same as defined in Table 23 with event subscription i instead of firm subscription. Time fixed effects are included. We cluster the standard errors around firm and year in calculating the t-statistics.

Table 27 shows the regression results. All IIT measures have positive and significant coefficient estimates for event weeks $w-1$ and $w-2$. For example, $AIIT$ has coefficients of 0.007 and 0.004 (t-statistic = 19.51 and 10.78) in weeks $w-1$ and $w-2$, respectively. The other IITs obtain coefficients with similar statistical significance. At longer horizons, there is no reversal effect. The results suggest that IIT is able to predict abnormal returns before fundamental information arrives.

[Place Table 27 about here]

Next, we examine whether the price sensitivity to IIT also changes around corporate events. To do so, we run the following Fama-MacBeth (1973) regressions in the CRSP sample between 1994 and 2016,

$$RET_{i,w} = \alpha_w + \beta_w^{EVENT} Event_{i,w} + \sum_{k=1}^{4} \beta_{k,w}^{INTER} Event \times IIT_{i,w-k} \tag{15}$$

$$+ \sum_{k=1}^{4} \beta_{k,w}^{IIT} IIT_{i,t-k} + ILLIQ_{i,w-1} + \sum_{k=1}^{4} \beta_{k,w}^{RET} RET_{i,w-k} + \epsilon_{i,w}, \tag{16}$$

where $Event$ is a dummy variable equal to one if we identify a corporate event in the week and zero otherwise, and all the explanatory variables are the same as defined in Table 23. We report the coefficient estimates of the event dummy, IITs, and their interactions in Table 28. For brevity, the control variables are not reported.[30]

[Place Table 28 about here]

Table 28 shows that consistent with our main findings, IIT has significant and positive predictive ability. The event dummy always has positive and significant coefficients, suggesting more positive than negative corporate events in our sample. More importantly, we find that the interaction terms of the event dummy and the first lag of IIT are always positive and significant across all columns in Table 28. For example, the interaction with $AIIT_{w-1}$ has a coefficient of 0.004 (t-statistic = 3.94), half the coefficient of $AIIT_{w-1}$ itself (0.008). Similar results can be found in the other columns, suggesting that the predictive power of IIT increases by about 50% approaching a significant corporate event. The strengthened predictive ability of IIT further extends to several weeks ago. At the same time lag, the interaction is more likely to be significant than the corresponding IIT, suggesting that informed traders become more aggressive if their advanced information is about an upcoming corporate event. Taken the results in this subsection together, we find that the return predictability from IIT is closely related to fundamental information flow to financial markets.

D.6. International Study in G10 Countries

Our last test of IIT takes advantage of its flexibility by replicating the portfolio analysis in international equity markets. Long-history of tick data is generally unavailable for most markets outside the US, making it difficult to examine the effect of informed trading in international markets. Our method can be a potential solution. We collect daily stock data of non-US G10 countries from Datastream. Specifically, using Datastream's constituent lists, we extract the date, equity identifier, daily equity price, daily trading volume, and shares outstanding for both alive and dead equities from stock exchanges in those countries. Because the bid-ask spread and daily high and low prices are unavailable for many countries over a long period, we exclude $BAIIT$, $CSIIT$, and $ARIIT$ in this analysis.

[30]We report the full results in the internet appendix.

Following Griffin, Kelly, and Nardari (2010), we filter the Datastream sample in the following way. We choose stocks that can be traded in local markets in each country via local currency. We consider a stock to be dead when it is not traded for over sixty trading days. To circumvent data errors, we exclude stock-days with daily returns greater than 200 percent, and observations showing that the current or next day's return exceeds 100 percent but the cumulative return of the two subsequent returns is less than 20 percent. We include only country-days with at least ten active firms. The sample goes backward from December 2016 to the first date of coverage by Datastream in each country. The longest history belongs to Canada dating back to January, 1973 and the shortest history is Switzerland starting from April, 1990. In terms of size of the cross-section, Japan has the largest number of stocks $(2,607)$ per week and Belgium has the smallest cross-section with only 149 stocks per week. We use the same method to calculate illiquidity but use weekly transaction returns instead of weekly risk-adjusted returns to calculate IIT. Potentially, the IITs become noisier in this international sample.

Given relatively small number of stocks in some markets, we sort all stocks in a country into quintile portfolios based on each IIT every week. For brevity, we only report the return differentials between the top and bottom quintile portfolios in the following week for different IIT strategy and country combinations in Table 29. We find that despite different IIT measures and markets examined, all IIT strategies generate positive and significant abnormal returns in all the countries except for $IIIT$ in Italy and $TIIT$ in the United Kingdom. Among the significant abnormal returns, the highest is achieved by $AIIT$ in Canada with an average weekly return of 1.863% and a t-statistic of 27.07 and the lowest is $IIIT$ in the United Kingdom (12.2 bp and t-statistic = 2.51).

[Place Table 29 about here]

We also confirm that the positive performance of IIT strategies in G10 countries is not due to illiquidity or past stock returns by replicating the double sorting analysis in Table 22. Table 30 reports the return differentials between the high and low quintile portfolios controlling for the IIT-corresponding illiquidity in Panel A and past returns in Panel B. We find that the return differential is positive and significant at one or five percent significance level in all non-US G10 countries except in the UK, after controlling for illiquidity and past stock return.[31] The return differentials range from 11.1 bp (t-stat = 2.15) for $TIIT$ in Netherland to 258.6 bp (t-stat = 30.51) for $AIIT$ in Canada.

[Place Table 30 about here]

The results from international markets indicate that despite significant cross-border differences in market microstructure and regulation, IIT has consistent and pervasive predictive ability for future returns, suggesting that the information measure we propose can be applied in various scenarios of empirical research.

E. Conclusion

In this study, we propose a simple method to estimate informed traders' activity motivated by the Kyle (1985) model without using tick data. Viewing contemporaneous price impact as the result of order imbalance and illiquidity, we calculate implied order imbalance (IOI) as the fraction of stock return and well-known illiquidity measures estimated using daily aftermarket data. We arrive at our measure of implied informed trading (IIT) by taking the residual term from the regression of IOI on contemporaneous liquidity and stock returns, which removes the impact of liquidity trading on net order imbalance. Given the flexibility and convenience of IIT calculation, we are able to conduct empirical tests on informed trading in the CRSP universe from April 1927 to December 2016.

We test the performance of implied informed trading by examining its return predictability in this extensive sample. We find that regardless of the illiquidity measure used to construct IIT, the proposed IIT always shows positive and significant predictive power for stock returns in subsequent weeks. In univariate portfolio sorts, the top IIT decile portfolio return significantly exceeds the bottom IIT decile portfolio return by 27.4 to 101.9 bp in the following week. The

[31] We find that an unusually high proportion (26.56%) of stock-day observations in the UK have zero return but non-zero trading volume in Datastream.

return predictability of IITs is persistent during subperiods of the 90-year sample period and becomes stronger during financial crises and NBER recessions. The pricing effect of IIT survives well-known measures of informed trading including short sale interest, institutional order flow, aggregate order flow, and option trading volume in multivariate regressions. Additionally, we find that the IIT strategy's profitability is higher for stocks with higher degrees of information asymmetry. We also document that IIT is informative about abnormal announcement returns surrounding significant corporate events, suggesting that the return predictability is associated with fundamental information flow. Finally, we construct an international sample for non-US G10 countries using data from Datastream. We find that the long-short IIT strategies consistently generate positive and significant abnormal returns in all international equity markets. These results suggest that IIT measures are effective in capturing private information in a wide variety of markets. The fact that we use different illiquidity proxies to find the same results reinforces our interpretation of the discovered return predictability as evidence of return continuation due to partial adjustment upon informed trading, consistent with classical market microstructure models such as Kyle (1985).

A notable advantage of our approach is to utilize only aftermarket data but not high-frequency tick data. The proposed IIT measures can be easily calculated for large samples, hence suitable for empirical studies involving informed trading in different markets. Although we mainly examine price discovery at the weekly frequency in our study, it is possible to apply the same method in intraday studies. It is also possible to extend the analysis to other asset classes such as corporate bond and derivatives markets under additional assumptions, where order flow data are more difficult to acquire. We leave these questions to future studies.

Appendix. Low frequency illiquidity measures

In this appendix, we explain the low frequency illiquidity measures used to construct our implied informed trading measures.

Turnover-based Amihud Illiquidity

Amihud (2002) proposes an illiquidity proxy by measuring the daily expected price response to one dollar of trading volume. Brennan, Huh, and Subrahmanyam (2013) note that Amihud's illiquidity inherently contains firm size effect because the denominator, trading activity, is calculated as the product of stock price and trading volume while the numerator, stock return, is a standardized variable in the cross-section. To mitigate the size effect, Brennan, Huh, and Subrahmanyam (2013) propose to use turnover ratio instead of dollar trading volume in calculating the Amihud illiquidity. We follow this approach and define our turnover-based Amihud illiquidity as:

$$AMI_{i,w} = \text{Average}_{i,w}\left(\frac{|R_{i,d}|}{\text{Share Volume Turnover Ratio}_{i,d}}\right), \tag{17}$$

where for stock i in week w, d is a trading day, R is the daily stock return, and share volume turnover ratio is a ratio of daily trading volume to total number of shares outstanding. We set AMI to missing if there are less than three observations for a stock-week.

Implied Illiquidity by Microstructure Invariance of Kyle and Obizhaeva (2016)

Kyle and Obizhaeva (2016) introduce an illiquidity measure from the perspective of risk transfer. The theory leads to two joint hypotheses of invariance of dollar risk transfers and invariance of transaction costs. Such microstructure invariances imply an illiquidity measure ($INVAR$) as following:

$$INVAR_{i,w} = \left(\frac{\sigma^2_{i,w}}{\text{Dollar Trading Volume}_{i,w}}\right)^{1/3}, \tag{18}$$

where for stock i in week w, σ is the standard deviation of daily stock returns in the week, and dollar trading volume is the sum of daily trading volume over a week. Empirically, we require at least three daily observations to calculate $INVAR$.

Share Volume Turnover Ratio and Relative Bid-Ask Spread

Share volume turnover ratio and bid-ask spread have been widely used as liquidity measures. When the turnover ratio is low, it becomes hard to encounter potential trading partners. Therefore, a transaction is rarely immediately executed and the stock liquidity becomes low (Karpoff, 1986). We take inverse share volume turnover ratio as an illiquidity measure. Bid-ask spread directly measures the transaction costs from immediate execution and should increase when the market is more illiquid (Amihud and Mendelson, 1986) We use percentage bid-ask spread, the closing bid-ask spread scaled by the midpoint of the bid and ask prices. We calculate share volume turnover ratio and percentage bid-ask spread for every stock everyday. Then, we take the average over a week if there are at least three observations for a stock in a week.

High-Low Spread of Corwin and Schultz (2012)

Corwin and Schultz (2012) develop a bid-ask spread estimator using daily high and low transaction prices. These extreme prices are useful to estimate the effective spread because the high (low) price is almost always at ask (bid) price. Therefore, the difference between high and low prices is a result of both stock return variance and bid-ask spread. Assuming that stock return variance increases proportionally with time interval but spread does not, Corwin and Schultz derive a spread estimator ($CSSPRD$) as follows:

$$CSSPRD_{i,d} = \left(\frac{2e^{\alpha_d} - 1}{1 + e^{\alpha_d}}\right), \tag{19}$$

where for stock i on day d, $\alpha_d = \frac{\sqrt{2\beta_d} - \sqrt{\beta_d}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma_d}{3-2\sqrt{2}}}$ with β_d being the sum of squared one-day high-low ratios over two consecutive days d and $d + 1$, and γ_d being a squared two-day high-low ratio. We estimate the high-low spread for every stock everyday using the SAS code provided on Shane Corwin's website (https://www3.nd.edu/ scorwin/). Then, we take the average $CSSPRD$ if there are at least three observations for a stock in the week. Otherwise, the spread is treated as missing.

Close-High-Low Spread of Abdi and Ranaldo (2017)

Abdi and Ranaldo (2017) develop a method to estimate the effective bid-ask spread using daily close, high, and low prices. Motivated by Roll's (1984) model, they show that the squared distance between the close log price and the mid-range of high and low log prices is a function of the efficient-price variance and the squared effective spread. Rearranging the squared distance, they arrive at a simple measure for effective spread as following:

$$ARSPRD_{i,w} = \text{Average}_{i,w}\left(\sqrt{\max\left[4(c_d - \eta_d)(c_d - \eta_{d+1}), 0\right]}\right), \tag{20}$$

where for stock i in week w, d is a trading day, c is the close log-price, and η is the average of high and low log-prices. We set $ARSPRD$ to missing if there are less than three observations for a stock-week.

REFERENCES

[1] Abdi, Farshid, and Angelo Ranaldo, 2017, A simple estimation of bid-ask spreads from daily close, high, and low prices, The Review of Financial Studies 30, 4437-4480.

[2] Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, Journal of Financial Markets 5, 31-56.

[3] Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid ask spread, Journal of Financial Economics 17, 223-249.

[4] Back, Kerry, Kevin Crotty, and Tao Li, 2018, Identifying information asymmetry in securities markets, The Review of Financial Studies 31, 2277-2325.

[5] Bali, Turan G., Lin Peng, Yannan Shen, and Yi Tang, 2014, Liquidity shocks and stock market reactions, The Review of Financial Studies 27, 1434-1485.

[6] Baruch, Shmuel, Marios Panayides, and Kumar Venkataraman, 2017, Informed trading and price discovery before corporate events, Journal of Financial Economics 125, 561-588.

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

[8] Bernile, Gennaro, Jianfeng Hu, and Yuehua Tang, 2016, Can information be locked up? Informed trading ahead of macro-news announcements, Journal of Financial Economics 121, 496-520.

[9] Bloomfield, R., M. O'Hara, and G. Saar, 2005, The "make or take" decision in an electronic market: Evidence on the evolution of liquidity, Journal of Financial Economics 75, 165-199.

[10] Boehmer, E., C. M. Jones, and X. Y. Zhang, 2008, Which shorts are informed ?, Journal of Finance 63, 491-527.

[11] Boehmer, Ekkehart, and Eric K. Kelley, 2009, Institutional investors and the informational efficiency of prices, Review of Financial Studies 22, 3563-3594.

[12] Brennan, Michael, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2013, An analysis of the amihud illiquidity premium, The Review of Asset Pricing Studies 3, 133-176.

[13] Brennan, Michael J., Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2018, High-frequency measures of informed trading and corporate announcements, The Review of Financial Studies 31, 2326-2376.

[14] Campbell, John, Tarun Ramadorai, and Allie Schwatz, 2009, Caught on tape, Journal of Financial Economics 92, 26.

[15] Campbell, John Y.; Grossman, Sanford J.; Wang, Jiang, 1993, Trading volume and serial correlation in stock returns, The Quarterly Journal of Economics 108, 35.

[16] Carhart, M. M., 1997, On persistence in mutual fund performance, Journal of Finance 52, 57-82.

[17] Chordia, Tarun, Jianfeng Hu, Avanidhar Subrahmanyam, and Qing Tong, 2019, Order flow volatility and equity costs of capital, Management Science 65, 1520-1551.

[18] Chordia, T., S. W. Huh, and A. Subrahmanyam, 2009, Theory-based illiquidity and asset pricing, Review of Financial Studies 22, 3629-3668.

[19] Chordia, Tarun, and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, Journal of Financial Economics 72, 485-518.

[20] Christophe, S. E., M. G. Ferri, and J. J. Angel, 2004, Short-selling prior to earnings announcements, Journal of Finance 59, 1845-1875.

[21] Cohen, L., K. B. Diether, and C. J. Malloy, 2007, Supply and demand shifts in the shorting market, Journal of Finance 62, 2061-2096.

[22] Collin-Dufresne, Pierre, and Vyacheslav Fos, 2015, Do prices reveal the presence of informed trading?, The Journal of Finance 70, 1555-1582.

[23] Corwin, Shane A., and Paul Schultz, 2012, A simple way to estimate bid-ask spreads from daily high and low prices, Journal of Finance 67, 719-759.

[24] Duarte, J., and L. Young, 2009, Why is pin priced?, Journal of Financial Economics 91, 119-138.

[25] Easley, David, Marcos Lopez de Prado, and Maureen O'Hara, 2016, Discerning information from trade data, Journal of Financial Economics 120, 269-285.

[26] Easley, David, and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, Journal of Financial Economics 19, 69-90.

[27] Easley, D., M. O'Hara, and P. S. Srinivas, 1998, Option volume and stock prices: Evidence on where informed traders trade, Journal of Finance 53, 431-465.

[28] Edmans, A., 2009, Blockholder trading, market efficiency, and managerial myopia, Journal of Finance 64, 2481-2513.

[29] Engelberg, J. E., A. V. Reed, and M. C. Ringgenberg, 2012, How are shorts informed? Short sellers, news, and information processing, Journal of Financial Economics 105, 260-278.

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

[31] Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, Journal of Financial Economics 116, 1-22.

[32] Fama, Eugene F., and James D. Macbeth, 1973, Risk, return, and equilibirum: Empirical tests, Journal of Political Economy 81, 607-636.

[33] Finnerty, Joseph E., 1976, Insiders and market efficiency, The Journal of Finance 31, 1141-1148.

[34] Fong, Kingsley Y. L, Craig W. Holden, and Ondrej Tobek, 2018, Are volatility over volume liquidity proxies useful for global or us research?, Kelley School of Business Research Paper No. 17-49.

[35] Fong, Kingsley Y L, Craig W Holden, and Charles A Trzcinka, 2017, What are the best liquidity proxies for global research?*, Review of Finance 21, 1355-1401.

[36] Foster, George, Chris Olsen, and Terry Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, The Accounting Review 59, 574-603.

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

[38] Glosten, Lawrence R., and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 14, 71-100.

[39] Griffin, John M., Patrick J. Kelly, and Federico Nardari, 2010, Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets, Review of Financial Studies 23, 3225-3277.

[40] Hasbrouck, J., 2004, Liquidity in the futures pits: Inferring market dynamics from incomplete data, Journal of Financial and Quantitative Analysis 39, 305-326.

[41] Hasbrouck, J., 2009, Trading costs and returns for us equities: Estimating effective costs from daily data, Journal of Finance 64, 1445-1477.

[42] Ho, Thomas S. Y., and Hans R. Stoll, 1983, The dynamics of dealer markets under competition, Journal of Finance 38, 1053-1074.

[43] Holden, Craig W., 2009, New low-frequency spread measures, Journal of Financial Markets 12, 778-813.

[44] Johnson, Travis L., and Eric C. So, 2012, The option to stock volume ratio and future returns, Journal of Financial Economics 106, 262-286.

[45] Kaniel, Ron, and Hong Liu, 2006, So what orders do informed traders use?, The Journal of Business 79, 1867-1913.

[46] Karpoff, Jonathan M., 1986, A theory of trading volume, Journal of Finance 41, 1069-1087.

[47] Kyle, Albert S., 1985, Continuous auctions and insider trading, Econometrica 53, 1315-1335.

[48] Kyle, Albert S., and Anna A. Obizhaeva, 2016, Market microstructure invariance: Empirical hypotheses, Econometrica 84, 1345-1404.

[49] Lee, Charles M. C., and Balkrishna Radhakrishna, 2000, Inferring investor behavior - evidence from torq data, Journal of Financial Market 3, 29.

[50] Lee, Charles M. C., and Mark J. Ready, 1991, Inferring trade direction from intraday data, Journal of Finance 46, 733-746.

[51] Lesmond, David A., Joseph P. Ogden, and Charles A. Trzcinka, 1999, A new estimate of transaction costs, Review of Financial Studies 12, 1113-1141.

[52] Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002, Dynamic volume-return relation of individual stocks, The Review of Financial Studies 15, 1005-1047.

[53] Lo, Andrew W., and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, The Review of Financial Studies 3, 175-205.

[54] Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation cosistent covariance-matrix, Econometrica 55, 703-708.

[55] Odders-White, E. R., and M. J. Ready, 2008, The probability and magnitude of information events, Journal of Financial Economics 87, 227-248.

[56] Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, Journal of Political Economy 111, 642-685.

[57] Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, Journal of Finance 39, 1127-1139.

[58] Spiegel, Matthew, and Avanidhar Subrahmanyam, 1995, On intraday risk premia, Journal of Finance 50, 319-339.

[59] Tetlock, Paul C., 2010, Does public financial news resolve asymmetric information?, The Review of Financial Studies 23, 3520-3557.

Table 19. Summary statistics

Panel A reports the time-series averages of cross-sectional statistics of weekly implied informed trading (IIT) measures between April 1927 and December 2016. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with close prices above five dollars. For each illiquidity measure, we first calculate implied order imbalance (IOI) as weekly Fama-French-Carhart risk-adjusted returns (RET) times illiquidity. Then IIT is defined as the regression residual of IOI on contemporaneous risk-adjusted returns and illiquidity in the cross-section. $AIIT$ is calculated using turnover-based Amihud illiquidity (AMI) from Brennan, Huh, and Subrahmanyam (2013). $IIIT$ is calculated using market microstructure invariance ($INVAR$) from Kyle and Obizhaeva (2016). $TIIT$ is calculated using share volume turnover ratio ($TURN$). $BAIIT$ is calculated using percentage bid-ask spread scaled by the midpoint ($BASPRD$) at market close from 1984 onwards. $CSIIT$ is calculated using high-low spreads ($CSSPRD$) as in Corwin and Schultz (2012). $ARIIT$ is calculated using close-high-low spread ($ARSPRD$) as in Abdi and Ronaldo (2017). Weekly AMI, $TURN$, $BASPRD$, and $CSSPRD$ are averages of the daily values in a week and weekly $INVAR$ and $ARSPRD$ are estimates using the five days' prices and volumes. Panel B shows the averages of cross-sectional correlations between the main variables.

Panel A. Descriptive statistics

	Number of Weeks	Avg. Num. of Stocks	Mean	Standard Deviation	Min	Med	Max
$AIIT$	4,679	2,116	-0.006	6.248	-23.141	0.011	24.225
$IIIT$	4,679	2,122	0.000	0.559	-1.961	0.020	1.879
$TIIT$	4,679	2,161	-0.001	0.159	-0.603	-0.003	0.680
$BAIIT$	1,712	3,203	0.017	16.268	-57.757	0.399	56.887
$CSIIT$	4,679	2,292	-0.005	2.989	-11.359	0.029	11.945
$ARIIT$	4,679	2,315	-0.001	0.361	-1.427	-0.013	1.949
AMI	4,679	2,138	0.045	0.103	0.001	0.018	0.978

Panel B. Correlations

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
$AIIT$	1.000					
$IIIT$	0.473	1.000				
$TIIT$	0.659	0.190	1.000			
$BAIIT$	0.513	0.714	0.339	1.000		
$CSIIT$	0.104	0.402	-0.076	0.335	1.000	
$ARIIT$	0.033	0.111	-0.112	0.142	0.153	1.000

Table 20. Weekly investment analysis

This table reports the performance of investment strategies based on implied informed trading (IIT) measures as indicated on the top of each column. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars from April 1927 to December 2016. We sort all the stocks into decile portfolios based on IIT measures every week. The average portfolio returns in the next week are reported. Also reported are the return differentials and alphas with respect to the Fama-French-Carhart four factors between the top and bottom decile portfolios. All variables are the same as defined in Table 19. Corresponding t-statistics based on Newey-West (1987) standard errors are reported in parentheses. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Low (%)	0.125	0.143	0.126	0.120	0.208	-0.083
2	0.136	0.164	0.133	0.144	0.190	-0.002
3	0.209	0.204	0.195	0.209	0.215	0.122
4	0.235	0.238	0.237	0.236	0.250	0.194
5	0.252	0.267	0.263	0.255	0.240	0.212
6	0.266	0.286	0.273	0.250	0.268	0.251
7	0.292	0.312	0.286	0.307	0.289	0.309
8	0.354	0.379	0.326	0.387	0.351	0.411
9	0.509	0.501	0.448	0.568	0.454	0.648
High	0.662	0.538	0.675	0.619	0.505	0.905
HML (%)	0.538^a	0.395^a	0.549^a	0.500^a	0.297^a	0.989^a
	(22.26)	(17.86)	(18.66)	(18.92)	(15.85)	(42.56)
FF4 α (%)	0.534^a	0.392^a	0.551^a	0.507^a	0.291^a	0.988^a
	(33.79)	(24.84)	(31.86)	(23.93)	(21.55)	(65.03)

Table 21. Subperiod tests

This table reports the performance of investment strategies based on implied informed trading (IIT) in different sample periods. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars. We divide the sample into subperiods in three ways: (1) Early (1927-1950), Middle (1951-1983), and Late (1984-2016); (2) Recession and Non-recession periods as defined by NBER; (3) Crisis and Non-crisis periods considering Great Depression (1929-1932), Dot-com Bubble (1999-2001), and Subprime Mortgage Crisis (2007-2009). We sort all the stocks into decile portfolios based on IIT measures every week. Reported are the return differentials between the top and bottom decile portfolios. All variables are the same as defined in Table 19. Corresponding t-statistics based on Newey-West (1987) standard errors are reported in parentheses. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

Subperiods	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Early	0.916^a	0.787^a	0.950^a		0.514^a	1.017^a
	(14.53)	(13.99)	(12.31)		(11.58)	(16.64)
Middle	0.286^a	0.234^a	0.219^a		0.119^a	1.046^a
	(15.80)	(12.60)	(10.31)		(5.38)	(38.15)
Late	0.518^a	0.275^a	0.590^a	0.500^a	0.319^a	0.911^a
	(20.89)	(11.99)	(20.04)	(19.64)	(13.45)	(31.80)
Recession	0.779^a	0.667^a	0.754^a	0.638^a	0.356^a	1.258^a
	(10.47)	(10.14)	(9.35)	(6.07)	(7.62)	(17.39)
Non-Recession	0.481^a	0.331^a	0.500^a	0.486^a	0.283^a	0.925^a
	(24.74)	(18.00)	(19.58)	(19.35)	(15.34)	(49.18)
Crisis	1.071^a	0.833^a	1.150^a	0.639^a	0.508^a	1.385^a
	(10.91)	(9.01)	(11.87)	(10.74)	(7.83)	(13.36)
Non-Crisis	0.463^a	0.334^a	0.464^a	0.462^a	0.267^a	0.933^a
	(24.82)	(18.90)	(19.01)	(17.20)	(15.46)	(51.44)

Table 22. Double-sorting investment analysis

This table presents double-sorting portfolio returns of implied informed trading (IIT) measures after controlling for illiquidity and past returns. At the end of each week from 1927 and 2016, we first sort stocks into quintile portfolios based on the corresponding illiquidity in Panel A and risk-adjusted stock returns in Panel B. Then, within each illiquidity or return quintile, we further sort stocks into quintile portfolios based on IIT measures. We report average returns of the same IIT quintile across all quintile portfolios of the conditioning variable in the next week. Also reported are the average of return differentials and alphas with respect to the Fama-French-Carhart four factors between the top and bottom IIT quintile portfolios. All variables are the same as defined in Table 19. Corresponding t-statistics based on Newey-West (1987) standard errors are reported in parentheses. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

Panel A. Dependent sort on IITs controlling for IIT-corresponding illiquidity

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Low (%)	0.111	0.167	0.060	0.142	0.213	-0.032
2	0.233	0.219	0.237	0.210	0.229	0.151
3	0.248	0.257	0.264	0.247	0.241	0.242
4	0.308	0.321	0.297	0.345	0.307	0.370
High	0.605	0.533	0.624	0.603	0.498	0.751
HML (%)	0.494^a	0.365^a	0.564^a	0.462^a	0.285^a	0.784^a
	(32.21)	(27.02)	(31.36)	(24.27)	(21.61)	(49.56)
FF4 α (%)	0.488^a	0.364^a	0.556^a	0.464^a	0.282^a	0.782^a
	(44.74)	(35.16)	(47.97)	(33.08)	(28.62)	(65.30)

Panel B. Dependent sort on residual IITs controlling FF4-adjusted return

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Low (%)	0.174	0.102	0.227	0.107	0.171	0.049
2	0.257	0.241	0.304	0.249	0.270	0.186
3	0.288	0.305	0.288	0.290	0.287	0.262
4	0.303	0.328	0.281	0.355	0.313	0.365
High	0.469	0.513	0.382	0.525	0.444	0.623
HML (%)	0.295^a	0.411^a	0.155^a	0.418^a	0.273^a	0.574^a
	(22.41)	(28.92)	(8.01)	(25.17)	(19.67)	(45.91)
FF4 α (%)	0.294^a	0.409^a	0.170^a	0.422^a	0.269^a	0.573^a
	(30.65)	(40.13)	(11.87)	(33.33)	(27.68)	(56.39)

Table 23. Return prediction using Fama-MacBeth (1973) regressions

This table presents time-series averages of the coefficient estimates from cross-sectional regressions of the following equation,

$$RET_{i,w} = \alpha_w + \sum_{k=1}^{4} \beta_{k,w}^{IIT} IIT_{i,w-k} + \beta_w^{ILLIQ} ILLIQ_{i,w-1} + \sum_{k=1}^{4} \beta_{k,w}^{RET} RET_{i,w-k} + \epsilon_{i,w},$$

where for stock i in week w, RET is risk-adjusted stock return using Fama-French-Carhart four factors, IIT is the implied informed trading measure as indicated on the top of each column, and $ILLIQ$ is an illiquidity measure corresponding to the implied informed trading measure. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars in daily CRSP from April 1927 to December 2016. All variables are the same as defined in Table 19. All coefficient estimates are multiplied by 100. In parentheses, we report t-statistics of the average coefficient based on Newey-West (1987) standard errors. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.059^a	-0.039^a	-0.181^a	-0.076^a	-0.073^a	-0.118^a
	(-7.59)	(-4.43)	(-22.81)	(-5.38)	(-8.65)	(-13.35)
IIT_{w-1}	0.037^a	0.486^a	0.952^a	0.070^a	0.022^a	0.740^a
	(17.30)	(15.87)	(9.18)	(10.88)	(13.86)	(23.20)
IIT_{w-2}	0.016^a	0.160^a	0.633^a	-0.003	0.003^b	0.122^a
	(10.04)	(8.89)	(7.38)	(-0.99)	(2.51)	(7.32)
IIT_{w-3}	0.011^a	0.059^a	0.697^a	-0.003	-0.000	0.048^a
	(7.29)	(4.31)	(9.39)	(-1.38)	(-0.40)	(3.00)
IIT_{w-4}	0.006^a	0.033^b	0.309^a	0.001	-0.001	0.043^a
	(4.67)	(2.28)	(4.08)	(0.57)	(-0.57)	(2.58)
$ILLIQ_{w-1}$	-1.135^a	-0.428^a	0.074^a	-1.477^a	-2.758^a	0.007
	(-15.13)	(-7.21)	(12.69)	(-2.75)	(-6.23)	(0.92)
RET_{w-1}	-7.870^a	-8.075^a	-8.296^a	-5.445^a	-7.726^a	-7.594^a
	(-37.48)	(-37.08)	(-35.32)	(-19.09)	(-33.27)	(-33.44)
RET_{w-2}	-2.924^a	-2.998^a	-2.995^a	-1.565^a	-2.801^a	-2.716^a
	(-21.24)	(-21.57)	(-21.13)	(-8.92)	(-20.56)	(-20.55)
RET_{w-3}	-1.300^a	-1.319^a	-1.217^a	-0.534^a	-1.144^a	-1.066^a
	(-11.67)	(-11.67)	(-10.41)	(-3.71)	(-10.22)	(-9.61)
RET_{w-4}	-0.436^a	-0.428^a	-0.367^a	-0.002	-0.267^a	-0.249^b
	(-4.20)	(-4.14)	(-3.47)	(-0.02)	(-2.62)	(-2.43)
Adjusted R^2	0.027	0.028	0.032	0.018	0.025	0.027
Num of Stocks per Week	2,096	2,094	2,170	3,167	2,182	2,327

Table 24. Correlations of informed trading measures

This table shows the time-series averages of cross-sectional correlations between measures of informed trading. The sample includes all common stocks listed on NYSE, AMEX, and Nasdaq with prices above five dollars from 1983 to 2016. The implied informed trading (IIT) measures are defined in Table 19. OI is net order flow based on the Lee and Ready (1991) algorithm. BVC is net order flow based on bulk volume classification following Easley, de Prado, and O'Hara (2016). AN is the total purchase volume minus total sale volume of institutional trades recorded by Ancerno between 1983 and 2012. LR is an institutional order flow measure based on the $5,000 cut-off rule as in Lee and Radhakrishna (2000). CRS is an estimate of institutional order flow as in Campbell, Ramadorai, and Schwatz (2009). OI, BVC, LR, and CRS are estimated using tick data from the NYSE TAQ database between 1983 and 2016. $DTCR$ is days to cover ratio defined as the average number of shares on loan relative to the average daily trading volume using security lending data from Markit between 2007 and 2013. OS is the options volume to stock volume ratio from Option Metrics as in Johnson and So (2012) between 1996 and 2016.

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
OI	0.089	0.060	0.075	0.025	0.034	-0.001
BVC	0.093	0.055	0.067	0.052	0.022	0.008
AN	0.076	0.091	0.040	0.074	0.019	0.023
LR	0.094	0.069	0.080	0.031	0.035	0.000
CRS	0.066	0.083	0.018	0.041	0.036	0.009
$DTCR$	0.025	0.018	0.017	0.011	-0.009	-0.006
OS	-0.011	-0.009	-0.008	-0.001	0.004	0.000

Table 25. Fama-MacBeth (1973) regressions controlling for other informed trading measures

This table presents time-series averages of coefficient estimates from cross-sectional regressions of the following equation,

$$RET_{i,w} = \alpha_w + \sum_{k=1}^{4} \beta_{k,w}^{IIT} IIT_{i,w-k} + \sum_{k=1}^{4} \beta_{k,w}^{INF} INF_{i,w-k} + \beta_w^{ILLIQ} ILLIQ_{i,w-1} + \sum_{k=1}^{4} \beta_{k,w}^{RET} RET_{i,w-k} + \epsilon_{i,w},$$

where for stock i in week w, RET is risk-adjusted stock return using Fama-French-Carhart four factors, IIT is the implied informed trading measure as indicated on the top of each column, INF is an alternative informed trading measure, and $ILLIQ$ is an illiquidity measure corresponding to the implied informed trading measure. For brevity, we only report the coefficient estimates of IIT with the full set of control variables in the regressions. The alternative informed trading measure is net order flow using the Lee and Ready (1991) algorithm in Panel A, net order flow using the bulk volume classification method of Easley, de Prado, and O'Hara (2016) in Panel B, institutional trade imbalance in Ancerno database in Panel C, institutional order flow based on a \$5,000 cut-off rule as in Lee and Radhakrishna (2000) in Panel D, institutional order flow as in Campbell, Ramadorai, and Schwatz (2009) in Panel E, short sale interest defined as the average number of shares on loan relative to the average daily trading volume in Panel F, and the options volume to stock volume ratio as in Johnson and So (2012) in Panel G. All IIT variables are the same as defined in Table 19. All coefficient estimates are multiplied by 100. In parentheses, we report t-statistics of the average coefficient based on Newey-West (1987) standard errors. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

Panel A. INF is Lee and Ready (1991) order flow (Sample period: 1983-2016)

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.065^a	-0.052^a	-0.142^a	-0.077^a	-0.081^a	-0.086^a
	(-6.24)	(-4.33)	(-10.76)	(-5.42)	(-5.39)	(-6.05)
IIT_{w-1}	0.013^a	0.065^a	0.336^a	0.051^a	0.025^a	1.249^a
	(15.37)	(10.75)	(9.45)	(10.27)	(13.74)	(23.11)
IIT_{w-2}	0.004^a	0.004	0.155^a	0.001	0.001	0.204^a
	(5.79)	(0.82)	(5.83)	(0.34)	(0.51)	(6.19)
IIT_{w-3}	0.003^a	0.003	0.163^a	-0.003	-0.000	0.078^b
	(5.74)	(0.69)	(6.61)	(-1.38)	(-0.30)	(2.42)
IIT_{w-4}	0.002^a	0.003	0.107^a	0.003	-0.001	0.040
	(3.55)	(0.73)	(4.20)	(1.07)	(-1.04)	(1.11)
INF Control	Yes	Yes	Yes	Yes	Yes	Yes
$ILLIQ$ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.019	0.020	0.022	0.021	0.019	0.020
Num of Stocks per Week	3,478	3,491	3,516	3,502	3,390	3,516

(Continued)

Table 25 – *Continued*

Panel B. INF is Easley, de Prado, and O'Hara (2016) order flow (Sample period: 1983-2016)

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.063^a	-0.060^a	-0.114^a	-0.084^a	-0.084^a	-0.087^a
	(-5.41)	(-4.47)	(-7.67)	(-6.22)	(-4.97)	(-5.30)
IIT_{w-1}	0.011^a	0.055^a	0.386^a	0.029^a	0.020^a	1.280^a
	(13.38)	(8.71)	(11.45)	(7.99)	(10.12)	(21.75)
IIT_{w-2}	0.004^a	0.008	0.205^a	0.006^a	0.003^b	0.192^a
	(5.18)	(1.39)	(6.48)	(4.00)	(2.35)	(5.27)
IIT_{w-3}	0.002^a	0.000	0.134^a	0.001	0.000	0.080^b
	(3.42)	(0.09)	(5.30)	(0.42)	(0.10)	(2.16)
IIT_{w-4}	0.002^a	-0.001	0.149^a	0.000	-0.002	0.026
	(3.44)	(-0.21)	(4.79)	(0.34)	(-1.29)	(0.64)
INF Control	Yes	Yes	Yes	Yes	Yes	Yes
$ILLIQ$ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.022	0.023	0.026	0.021	0.022	0.024
Num of Stocks per Week	3,314	3,324	3,389	3,941	3,320	3,389

Panel C. INF is trade imbalance in Ancerno database (Sample period: 1999-2012)

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.067^a	-0.060^a	-0.108^a	-0.077^a	-0.096^a	-0.115^a
	(-3.68)	(-2.72)	(-4.31)	(-4.06)	(-3.15)	(-3.83)
IIT_{w-1}	0.010^a	0.061^a	0.339^a	0.007^a	0.014^a	1.694^a
	(10.69)	(8.71)	(15.24)	(6.95)	(5.77)	(18.86)
IIT_{w-2}	0.003^a	0.025^a	0.085^a	0.003^a	0.005^b	0.254^a
	(3.96)	(3.84)	(3.92)	(3.41)	(2.36)	(3.92)
IIT_{w-3}	0.002^a	0.012^b	0.062^a	0.001	0.005^b	0.137^b
	(3.69)	(2.05)	(3.50)	(0.90)	(2.49)	(2.09)
IIT_{w-4}	0.001^c	0.005	0.032^c	0.001	-0.001	0.009
	(1.65)	(0.97)	(1.81)	(1.07)	(-0.56)	(0.13)
INF Control	Yes	Yes	Yes	Yes	Yes	Yes
$ILLIQ$ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.020	0.021	0.024	0.021	0.020	0.021
Num of Stocks per Week	3,776	3,783	3,832	3,762	3,699	3,832

(Continued)

Table 25 – *Continued*

Panel D. INF is Lee and Radhakrishna (2000) institutional order flow
(Sample period: 1983-2016)

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.064[a]	-0.050[a]	-0.142[a]	-0.075[a]	-0.080[a]	-0.086[a]
	(-6.21)	(-4.24)	(-10.81)	(-5.35)	(-5.36)	(-6.03)
IIT_{w-1}	0.013[a]	0.066[a]	0.339[a]	0.051[a]	0.025[a]	1.251[a]
	(15.45)	(10.75)	(9.55)	(10.28)	(13.77)	(23.16)
IIT_{w-2}	0.004[a]	0.004	0.155[a]	0.001	0.001	0.206[a]
	(5.80)	(0.86)	(5.82)	(0.31)	(0.60)	(6.23)
IIT_{w-3}	0.003[a]	0.003	0.163[a]	-0.004	-0.000	0.081[b]
	(5.76)	(0.74)	(6.59)	(-1.40)	(-0.23)	(2.50)
IIT_{w-4}	0.002[a]	0.003	0.105[a]	0.003	-0.001	0.039
	(3.52)	(0.75)	(4.15)	(1.05)	(-1.03)	(1.11)
INF Control	Yes	Yes	Yes	Yes	Yes	Yes
$ILLIQ$ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.019	0.019	0.022	0.021	0.019	0.019
Num of Stocks per Week	3,478	3,492	3,516	3,502	3,390	3,516

Panel E. INF is Campbell, Ramadorai, and Schwatz (2009) institutional order flow
(Sample period: 1983-2016)

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.079[a]	-0.067[a]	-0.138[a]	-0.103[a]	-0.096[a]	-0.101[a]
	(-7.29)	(-5.42)	(-10.19)	(-7.30)	(-6.18)	(-6.76)
IIT_{w-1}	0.013[a]	0.063[a]	0.373[a]	0.051[a]	0.024[a]	1.258[a]
	(16.08)	(10.62)	(10.67)	(10.27)	(13.15)	(23.47)
IIT_{w-2}	0.003[a]	0.002	0.148[a]	-0.001	0.001	0.205[a]
	(4.93)	(0.50)	(5.56)	(-0.38)	(0.70)	(5.82)
IIT_{w-3}	0.003[a]	0.005	0.164[a]	-0.003	-0.000	0.072[b]
	(5.42)	(1.16)	(6.71)	(-1.34)	(-0.13)	(2.22)
IIT_{w-4}	0.002[a]	0.005	0.113[a]	0.003	-0.001	0.022
	(3.72)	(1.18)	(4.52)	(1.08)	(-1.04)	(0.58)
INF Control	Yes	Yes	Yes	Yes	Yes	Yes
$ILLIQ$ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.019	0.020	0.022	0.021	0.019	0.020
Num of Stocks per Week	3,329	3,341	3,363	3,338	3,241	3,363

(Continued)

Table 25 – *Continued*

Panel F. INF is days to cover ratio (Sample period: 2007-2013)

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.083[a]	-0.114[a]	-0.068[b]	-0.084[a]	-0.154[a]	-0.167[a]
	(-4.19)	(-4.76)	(-2.38)	(-4.13)	(-4.39)	(-4.90)
IIT_{w-1}	0.003[a]	0.020[b]	0.176[a]	0.002[a]	0.006	1.829[a]
	(3.69)	(2.37)	(5.90)	(3.53)	(1.54)	(11.49)
IIT_{w-2}	0.002[a]	0.024[a]	0.042	0.001[a]	0.008[c]	0.264[c]
	(2.63)	(3.47)	(1.45)	(2.84)	(1.80)	(1.77)
IIT_{w-3}	0.002[a]	0.017[b]	0.061[a]	0.001[c]	0.002	-0.052
	(2.72)	(2.42)	(2.81)	(1.72)	(0.40)	(-0.31)
IIT_{w-4}	0.001	0.004	0.012	-0.000	-0.005	-0.195
	(0.79)	(0.54)	(0.41)	(-0.40)	(-1.61)	(-1.19)
INF Control	Yes	Yes	Yes	Yes	Yes	Yes
$ILLIQ$ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.025	0.026	0.029	0.025	0.026	0.029
Num of Stocks per Week	2,431	2,431	2,431	2,431	2,362	2,431

Panel G. INF is option volume to stock volume ratio
(Sample period: 1996-2016)

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.065[a]	-0.094[a]	-0.078[a]	-0.121[a]	-0.083[b]	-0.080[b]
	(-2.70)	(-3.00)	(-3.12)	(-5.35)	(-2.52)	(-2.56)
IIT_{w-1}	0.004[a]	0.005	0.122[a]	0.003[b]	0.012[a]	1.681[a]
	(5.50)	(0.88)	(5.29)	(1.98)	(5.20)	(19.95)
IIT_{w-2}	0.001	0.008	0.068[a]	0.003[a]	0.004[b]	0.079
	(1.59)	(1.57)	(3.31)	(2.72)	(2.03)	(1.02)
IIT_{w-3}	0.001[b]	0.003	0.077[a]	0.000	0.001	0.084
	(2.04)	(0.59)	(3.98)	(0.42)	(0.46)	(1.20)
IIT_{w-4}	0.001	0.007	0.016	0.001	0.001	0.084
	(1.41)	(1.54)	(0.85)	(0.98)	(0.39)	(1.08)
INF Control	Yes	Yes	Yes	Yes	Yes	Yes
$ILLIQ$ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.025	0.027	0.029	0.026	0.028	0.029
Num of Stocks per Week	2,124	2,124	2,124	2,097	2,033	2,124

Table 26. Investment analysis in subsamples based on stock characteristics

This table examines the performance of investment strategies based on implied informed trading (IIT) in subsamples of stocks. At the end of each week from 1927 to 2016, we first divide the sample into two groups using the cross-sectional median of a conditioning variable. Then, within each subsample, we further sort stocks into quintile portfolios based on IIT measures. We report the average of return differentials between the top and bottom IIT quintile portfolios in the next week as well as the performance difference between the subsamples. All IIT variables are the same as defined in Table 19. The conditioning variables are proxies for the level of information asymmetry, including market capitalization in Panel A, sell-side analyst coverage in Panel B, and institutional ownership in Panel C. Corresponding t-statistics based on Newey-West (1987) standard errors are reported in parentheses. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

Panel A. Market capitalization

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Low (%)	0.637^a	0.472^a	0.568^a	0.748^a	0.326^a	0.992^a
	(24.79)	(22.11)	(19.31)	(24.94)	(14.31)	(40.79)
High	0.249^a	0.120^a	0.254^a	0.02	0.151^a	0.570^a
	(16.81)	(9.23)	(14.68)	(1.19)	(12.92)	(34.76)
Difference (%)	0.388^a	0.352^a	0.314^a	0.729^a	0.175^a	0.421^a
	(16.10)	(15.76)	(12.54)	(22.83)	(7.92)	(20.53)

Panel B. Sell-side analyst coverage

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Low (%)	0.689^a	0.609^a	0.531^a	0.765^a	0.351^a	0.820^a
	(23.38)	(20.85)	(14.58)	(25.35)	(11.73)	(31.98)
High	0.162^a	0.053^a	0.195^a	0.124^a	0.194^a	0.673^a
	(9.14)	(3.26)	(10.40)	(6.27)	(10.28)	(25.29)
Difference (%)	0.527^a	0.556^a	0.336^a	0.642^a	0.157^a	0.148^a
	(17.79)	(19.65)	(10.68)	(22.10)	(6.16)	(8.14)

Panel C. Institutional ownership

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Low (%)	0.695^a	0.620^a	0.552^a	0.816^a	0.353^a	0.866^a
	(21.61)	(20.35)	(13.67)	(26.51)	(11.05)	(32.68)
High	0.196^a	0.067^a	0.216^a	0.148^a	0.191^a	0.713^a
	(13.48)	(4.40)	(11.14)	(8.23)	(11.34)	(29.87)
Difference (%)	0.500^a	0.553^a	0.335^a	0.669^a	0.162^a	0.154^a
	(17.20)	(19.51)	(10.91)	(21.44)	(6.52)	(8.56)

Table 27. Return predictability around corporate events

This table presents ordinary least squares regression results of the following equation,

$$AR_{i,w} = \alpha + \alpha_w + \sum_{k=1}^{4} \beta_k^{IIT} IIT_{i,w-k} + \beta^{ILLIQ} ILLIQ_{i,w-1} + \sum_{k=1}^{4} \beta_k^{AR} AR_{i,w-k} + \epsilon_{i,w},$$

where for each corporate event i announced in week w, AR is the risk-adjusted return calculated using event weeks $w-26$ to $w+26$, and IIT is the implied informed trading measure as indicated on the top of each column, and all the explanatory variables are the same as defined in Table 19. Year fixed effects are included. The corporate events include earnings announcements, analyst recommendation updates, schedule 13D filings, and extreme price movement exceeding two standard deviations and not followed by return reversal for at least ten days. The sample includes all common stocks with prices above five dollars and percentage bid-ask spreads between zero and a half between 1994 and 2016. All coefficient estimates are multiplied by 100. Corresponding t-statistics based on firm and year clustered standard errors are reported in parentheses. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	0.124^a	-0.097^a	0.131^a	-0.193^a	-0.089^a	-0.070^a
	(11.41)	(-7.91)	(12.06)	(-12.32)	(-6.13)	(-5.22)
IIT_{w-1}	0.007^a	0.046^a	0.345^a	0.003^a	0.032^a	2.448^a
	(19.51)	(13.97)	(17.88)	(18.75)	(14.77)	(34.85)
IIT_{w-2}	0.004^a	0.021^a	0.179^a	0.001^a	0.007^a	0.224^a
	(10.78)	(6.37)	(9.24)	(10.82)	(3.37)	(3.23)
IIT_{w-3}	0.001^c	-0.001	0.135^a	-0.000	0.001	0.127^c
	(1.75)	(-0.34)	(6.73)	(-0.04)	(0.65)	(1.79)
IIT_{w-4}	0.001^a	0.002	0.069^a	0.000	-0.002	-0.036
	(3.48)	(0.75)	(3.63)	(0.75)	(-0.69)	(-0.51)
$ILLIQ_{w-1}$	-0.036	3.813^a	0.001	13.903^a	14.262^a	0.182^a
	(-0.56)	(27.67)	(0.98)	(23.97)	(21.78)	(22.17)
RET_{w-1}	-4.920^a	-4.982^a	-4.692^a	-4.934^a	-4.997^a	-4.862^a
	(-34.52)	(-34.59)	(-32.63)	(-34.31)	(-33.35)	(-34.11)
RET_{w-2}	-3.979^a	-3.854^a	-3.864^a	-3.795^a	-3.809^a	-3.847^a
	(-27.74)	(-26.59)	(-26.60)	(-26.22)	(-25.29)	(-26.86)
RET_{w-3}	-2.431^a	-2.310^a	-2.318^a	-2.256^a	-2.321^a	-2.303^a
	(-16.45)	(-15.48)	(-15.53)	(-15.18)	(-15.04)	(-15.65)
RET_{w-4}	-1.813^a	-1.745^a	-1.801^a	-1.697^a	-1.714^a	-1.753^a
	(-12.37)	(-11.79)	(-12.30)	(-11.57)	(-11.11)	(-12.05)
Time-Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.004	0.004	0.004	0.004	0.004	0.005
Observation	1,759,342	1,761,532	1,775,824	1,751,863	1,667,196	1,775,805

Table 28. Return predictability with corporate event dummies

This table presents Fama-MacBeth (1973) regression results with corporate event dummies from 1994 to 2016. We replicate the regression analysis in Panel A of Table 23, adding a corporate event dummy and its interactions with implied informed trading (IIT) measures as indicated on the top of each column. For brevity, we do not report the results for the control variables. The event dummy take value of one if any of the following events occurs in a given week, earnings announcements, recommendation updates, schedule 13D filings, and extreme price movement exceeding two standard deviations and not followed by return reversal for at least ten days. All variables are the same as defined in Table 19. All coefficient estimates are multiplied by 100. In parentheses, we report t-statistics of the average coefficient based on Newey-West (1987) standard errors. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

	$AIIT$	$IIIT$	$TIIT$	$BAIIT$	$CSIIT$	$ARIIT$
Intercept	-0.223^a	-0.259^a	-0.227^a	-0.270^a	-0.276^a	-0.274^a
	(-15.19)	(-12.86)	(-11.56)	(-14.92)	(-10.67)	(-11.45)
$Event_w$	0.322^a	0.349^a	0.320^a	0.352^a	0.345^a	0.333^a
	(9.76)	(9.61)	(10.09)	(9.81)	(9.98)	(9.87)
$Event_w \times IIT_{w-1}$	0.004^a	0.022^a	0.159^a	0.006^a	0.029^a	1.385^a
	(3.94)	(3.39)	(5.75)	(3.12)	(12.76)	(15.49)
$Event_w \times IIT_{w-2}$	0.001	0.013^b	0.082^a	0.005^a	0.003	0.007
	(1.25)	(2.38)	(3.51)	(2.81)	(1.13)	(0.07)
$Event_w \times IIT_{w-3}$	0.002^b	0.015^a	0.074^a	0.004^b	0.005^b	0.206^b
	(2.36)	(2.91)	(3.54)	(2.58)	(2.22)	(2.35)
$Event_w \times IIT_{w-4}$	0.001^b	0.011^b	0.041^c	0.003^b	0.001	-0.067
	(1.99)	(2.16)	(1.83)	(2.01)	(0.60)	(-0.63)
IIT_{w-1}	0.008^a	0.046^a	0.277^a	0.015^a	0.006^a	0.836^a
	(9.96)	(6.86)	(10.66)	(5.22)	(2.96)	(14.56)
IIT_{w-2}	0.002^a	0.006	0.066^a	0.000	0.003	0.186^a
	(4.11)	(1.07)	(3.42)	(0.23)	(1.58)	(3.92)
IIT_{w-3}	0.001^a	-0.004	0.050^a	-0.002^c	-0.001	0.072
	(2.64)	(-0.76)	(2.81)	(-1.72)	(-0.88)	(1.46)
IIT_{w-4}	0.001	-0.005	0.019	-0.002	-0.002	0.014
	(1.16)	(-0.95)	(1.21)	(-1.28)	(-1.15)	(0.23)
ILLIQ Control	Yes	Yes	Yes	Yes	Yes	Yes
RET Control	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.023	0.023	0.029	0.023	0.023	0.024
Num of Stocks per Week	4,085	4,097	4,173	4,119	4,026	4,173

Table 29. Investment analysis in G10 countries

This table reports the performance of investment strategies based on implied informed trading (IIT) measures in non-US G10 countries using Datastream data. At the end of every week, we sort all stocks into quintile portfolios based on IIT measures. We report the return differentials between the top and bottom quintile portfolios in the next week for each market. All IIT variables are the same as defined in Table 19. Corresponding t-statistics based on Newey-West (1987) standard errors are reported in parentheses. We also report the beginning and ending months, number of weeks, and average number of stocks per week for each country. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

Country	Belgium	Canada	France	Germany	Italy	Japan	Netherland	Switzerland	Sweden	United Kingdom
$AIIT$	1.281^a	1.863^a	0.495^a	0.603^a	0.111^b	0.497^a	0.445^a	0.229^a	0.992^a	0.205^a
	(13.09)	(27.07)	(11.32)	(9.95)	(2.52)	(13.92)	(7.91)	(5.92)	(19.65)	(4.90)
$IIIT$	1.231^a	1.547^a	0.323^a	0.688^a	0.063	0.187^a	0.317^a	0.136^a	0.730^a	0.122^b
	(13.02)	(26.89)	(6.97)	(10.06)	(1.33)	(6.19)	(5.53)	(2.99)	(13.76)	(2.51)
$TIIT$	0.613^a	1.337^a	0.449^a	0.203^a	0.166^a	0.481^a	0.401^a	0.276^a	0.809^a	-0.007
	(8.34)	(24.57)	(12.07)	(4.21)	(3.92)	(14.59)	(7.09)	(7.31)	(15.12)	(-0.17)
Start Date	1986.1	1973.1	1988.6	1988.6	1986.7	1984.1	1986.1	1989.1	1982.1	1988.1
End Date	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12
Ave Firm per Week	149	1,977	708	811	247	2,607	169	198	378	1,352

Table 30. Double-sorting investment analysis in G10 countries

This table reports the double-sorting portfolio performance of investment strategies based on implied informed trading (IIT) measures after controlling for corresponding illiquidity and past returns in non-US G10 countries using Datastream data. At the end of every week, we first sort all stocks into quintile portfolios based on the corresponding illiquidity of an implied informed trading (IIT) measure in Panel A and stock returns in Panel B. Then, within each illiquidity or return quintile, we further sort stocks into quintile portfolios based on IIT. We report the average return differentials between the top and bottom IIT quintile portfolios across all quintile portfolios of the conditioning variable in the next week. All variables are the same as defined in Table 19. Corresponding t-statistics based on Newey-West (1987) standard errors are reported in parentheses. We also report the beginning and ending months, number of weeks, and average number of stocks per week for each country. Superscripts ^{a}, ^{b}, and ^{c} indicate statistical significance at the 1, 5, and 10 percent level, respectively.

Panel A. Dependent sort on IITs controlling for corresponding illiquidity

Country	Belgium	Canada	France	Germany	Italy	Japan	Netherland	Switzerland	Sweden	United Kingdom
$AIIT$	1.229^a	2.586^a	0.666^a	1.044^a	0.139^a	0.584^a	0.298^a	0.246^a	1.224^a	0.174^a
	(9.89)	(30.51)	(17.40)	(16.74)	(3.34)	(17.25)	(4.76)	(6.11)	(20.27)	(5.23)
$IIIT$	1.470^a	1.860^a	0.510^a	1.178^a	0.141^a	0.290^a	0.212^a	0.166^a	0.871^a	0.347^a
	(10.58)	(28.18)	(11.55)	(15.66)	(3.13)	(9.83)	(3.33)	(3.81)	(14.56)	(7.39)
$TIIT$	0.664^a	2.281^a	0.617^a	0.478^a	0.223^a	0.640^a	0.111^b	0.190^a	0.992^a	-0.027
	(7.56)	(30.58)	(18.30)	(10.65)	(5.23)	(19.66)	(2.15)	(4.92)	(16.98)	(-0.70)

Panel B. Dependent sort on IITs controlling for stock return

Country	Belgium	Canada	France	Germany	Italy	Japan	Netherland	Switzerland	Sweden	United Kingdom
$AIIT$	1.868^a	1.963^a	0.524^a	1.208^a	0.082^b	0.387^a	0.280^a	0.231^a	0.972^a	0.028
	(16.20)	(31.97)	(14.78)	(17.72)	(2.19)	(13.62)	(4.88)	(6.06)	(18.09)	(0.97)
$IIIT$	2.230^a	2.726^a	0.463^a	1.757^a	0.038	0.363^a	0.148^b	0.190^a	0.986^a	-0.052
	(16.83)	(33.03)	(11.97)	(17.67)	(1.06)	(14.01)	(2.55)	(4.75)	(17.09)	(-1.54)
$TIIT$	1.302^a	1.623^a	0.474^a	0.422^a	0.078^b	0.361^a	0.110^b	0.174^a	0.834^a	-0.045
	(13.24)	(33.56)	(14.36)	(9.04)	(1.97)	(13.83)	(2.35)	(4.84)	(15.17)	(-1.30)
Start Date	1986.1	1973.1	1988.6	1988.6	1986.7	1984.1	1986.1	1989.1	1982.1	1988.1
End Date	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12	2016.12
Ave Firm per Week	149	1,977	708	811	247	2,607	169	198	378	1,352