NBER WORKING PAPER SERIES

IQ FROM IP: SIMPLIFYING SEARCH IN PORTFOLIO CHOICE

Huaizhi Chen
Lauren Cohen
Umit Gurun
Dong Lou
Christopher Malloy

Working Paper 24801
http://www.nber.org/papers/w24801

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2018

We would like to thank Anup Agrawal, Kenneth Ahern, Gurdip Bakshi, Gennaro Bernile, Bhagwan Chowdry, Eugene Fama, Fu Fangjian, William Goetzmann, Zhiguo He, Petri Jylha, Tobias Moskowitz, Jun Pan, Christo Pirinsky, Andrew Weinstock, Russell Wermers, and seminar participants at the Chicago Booth School of Business, University of Southern California, Singapore Management University, Shanghai Advanced Institute for Finance, the 2018 American Finance Association Meetings in Philadelphia, the 2018 UCLA Anderson Finance Conference, the 2018 QuantCon Conference in New York City, the 2017 University of Miami Behavioral Finance Conference, the 6th Luxembourg Asset Management Summit, and the 2018 Taiwan Finance Association Meetings in Taipei. We are grateful for funding from the National Science Foundation. Please send correspondence to: Lauren Cohen, Harvard Business School, Baker Library 279, Soldiers Field, Boston, MA 02163, phone: 1-617-495-3888, email: lcohen@hbs.edu.

IQ from IP: Simplifying Search in Portfolio Choice
Huaizhi Chen, Lauren Cohen, Umit Gurun, Dong Lou, and Christopher Malloy
NBER Working Paper No. 24801
July 2018
JEL No. G02,G11,G14,G23

ABSTRACT

Using a novel database that tracks web traffic on the SEC's EDGAR servers between 2004 and 2015, we show that mutual fund managers gather information on a very particular subset of firms and insiders, and their surveillance is very persistent over time. This tracking behavior has powerful implications for their portfolio choice, and its information content. An institution that downloaded an insider-trading filing by a given firm last quarter increases its likelihood of downloading an insider-trading filing on the same firm by more than 41.3% this quarter. Moreover, the average tracked stock that an institution buys generates annualized alphas of between 9-18% relative to the purchase of an average non-tracked stock. We find that institutional managers tend to track members of the top management teams of firms (CEOs, CFOs, Presidents, and Board Chairs), and tend to share educational and location-based commonalities with the specific insiders they choose to follow. Collectively, our results suggest that the information in tracked trades is important for fundamental firm value, and is only revealed following the information-rich dual trading by insiders and linked institutions.

Huaizhi Chen
University of Notre Dame
238 Mendoza College of Business
hchen11@nd.edu

Lauren Cohen
Harvard Business School
Baker Library 273
Soldiers Field
Boston, MA 02163
and NBER
lcohen@hbs.edu

Umit Gurun
University of Texas at Dallas
School of Management
800 W Campbell Rd. SM41 75080
Richardson, TX
umit.gurun@utdallas.edu

Dong Lou
Department of Finance
London School of Economics
Houghton Street
London, WC2A 2AE
UK
d.lou@lse.ac.uk

Christopher Malloy
Harvard Business School
Baker Library 277
Soldiers Field
Boston, MA 02163
and NBER
cmalloy@hbs.edu

I. Introduction

There is a fundamental search problem inherent in all portfolio choice. Moreover, in light of the decreasing cost of creating, processing, and transmitting information, the proliferation of information signals has increased greatly in both quantity and dimensionality in recent decades. These forces create a classic signal-noise problem, in which an agent must search ever larger matrices to decipher and create profitable signals. In a Grossman-Stiglitz world, an agent will be happy to collect information up to their private marginal value of expected return from that activity. However, with hundreds of thousands of information signals being produced in any given day, how does an investor reduce the dimensionality of the investment problem sufficiently to know even which subset (or class) of signals have the *potential* to be informative and provide this return in expectation?

In this paper, we propose that this reduction in dimensionality is a critical, and yet understudied, step in the investment process. Using rich, proprietary data provided by the Securities and Exchange Commission (SEC) on every document downloaded from their online site—including the exact timing and the IP address of the agent downloading--we provide new evidence on the search process in delegated portfolio management. In particular, we show that fund managers follow, and download, information on a very particular subset of firms, and that this set of firms stays highly constant over time. Further, their trades on these "tracked" firms (i.e., firms where the fund manager downloads a key filing) are significantly more informative for future operations and future firm performance, relative to their other trades.

The key innovation in our paper is that we are able to explicitly link the monitoring behavior of individual institutional investors (through their download behavior on the SEC's website—which we are able to map to the IP addresses of institutional investment firms, and hence identify them) to specific events on the stocks in their own portfolios. No prior study has been able to examine search behavior at the level of a specific institutional investor. In particular, we focus on how institutional fund managers track the trades of corporate insiders in the stocks they own.

We examine this laboratory for a number of reasons. First, compensation and hiring vs. firing decisions of fund managers – in addition to external human capital valuation such as possible hedge fund transitions -- are often determined by managers' performance relative to their peers. In fact many of the industries' highest profile rankings (e.g., Morningstar, Kiplinger, Barron's, etc.) are relative rankings amongst fund managers competing within a mandate. Thus, career concerns give fund managers a highly incented framework in which to care about the maximization of relative performance. Given this tournament-setup, a natural argument in a fund manager's maximization-function would be to find a signal (or set of signals) on which they have a comparative advantage relative to other managers. This begins to put some structure on the information acquisition problem that managers face.

Turning to insider trades, these are a potentially attractive candidate for relative comparative advantage signals for mutual fund managers. First, insiders are – by definition – a class of agents with privileged access and private information regarding their firms. Second, of all the factors of production – and all the information signals produced on a firm – insider trades are likely amongst the most valuable for unlocking a powerful (and legal) comparative advantage for a given fund manager. For instance, if a firm announces a new product launch, outside of the explicit transmission of material non-public information, it might be difficult (or prohibitively costly in any scalable manner) for an institution to gain a comparative advantage in interpreting this signal relative to other institutions.

However, contrast this with an insider trade within the same firm. The trade itself is public information – a sell, for instance. However, following the publicly disclosed sell, an institutional fund manager who owns the stock and hence has a connection to that firm could feasibly contact someone at the firm and inquire whether the sell was for personal liquidity reasons; for instance, to purchase a vacation home. Once determining that the sell was related to personal liquidity needs – information which the insider is perfectly legally free to tell her connection (i.e., it is not considered material non-public information to speak about vacation home purchases) – the fund manager can more accurately interpret this public signal of the tracked stock and trade accordingly. Importantly, this is an advantage of a connected manager – in that her competitor funds without a connection to the given insider may have a more costly process in gathering the same private information.

To better understand our approach, consider the following example firm and fund manager tracking and trading from our sample. AOK Inc. is a large, publicly traded firm on the NYSE, which in 2018 operated in all 50 states, with over 1,000 stores throughout North America and Asia - shipping to over 80 countries worldwide. The firm was widely held in large position by a number of institutional investors. Moreover, the firm had a number of insiders. In particular, two of these insiders, Mr. Sampson and Dr.

Jenkins (both independent board members), were actively trading over their times at AOK Inc.

In 2013, Mr. Sampson – who lived close to the firm's Headquarters – made an unusually large trade dumping AOK Inc. stock. This trade was reported to the SEC, with the trade showing up on the SEC's site on July 30, 2013. The next day (July 31), mutual fund manager Mr. Thompkins – manager of a large active equity fund located close to Mr. Sampson – checked Mr. Sampson's trade, downloading the file from the SEC's site. This was usual behavior for Mr. Sampson, who had made a habit of keeping particularly close tabs on the trades of Mr. Thompkins. Mr. Thompkins, followed this tracking the very next day (August 1) with an unusually large sale of his position in AOK that he held in his fund for three years. This turned out to be a very savvy move. Following Mr. Sampson's insider sale – echoed by Mr. Thompkins dumping – AOK plummeted in the following months. It dropped 8% in the month following, and burned down nearly 28% in the quarter following the trade. AOK was downgraded by two large banks covering the firm, then on its next earnings announcement missed on profits, revenues, and same store sales, along with guiding downward.

Now consider the trading of Dr. Jenkins. Dr. Jenkins was also an independent director at AOK, a few years following Mr. Sampson. In March of 2015, she made an untimely purchase of AOK on March 14. Quickly thereafter, a fund manager, Mr. Rothman, tracked this trade and made an outsized purchase in the same direction. Now, not only were Mr. Rothman and Dr. Jenkins located closely to each other, they also shared the same alma mater, the Harvard Business School. In the month following Dr. Jenkins trade, and the paired tracked buy by Mr. Rothman, AOK had a large number of positive realization. It marched up over 7% the first month, and climbed over 41% in the quarter following the buys. Moreover, it upped guidance, picked up initiation by a new investment bank, and beat on both earnings and same store sales at its next earnings announcement.

Had one taken the simple strategy of replicating Mr. Rothman's and Mr. Thompkins' tracked trades, (which we term as "tracked buys" and "tracked sells" throughout the paper), one would have made abnormal returns of over 7% in the following month, in each case, and even larger returns accruing in the subsequent months. Moreover, these` returns did not reverse, as - in sharp contrast to any overreaction pattern - they represented fundamental information that was revealed and incorporated into prices (e.g., real quantities of same store sales and profit margins).

In this paper we demonstrate that the above example of AOK Inc. – and its tracking fund managers - represents a much more systematic pattern across the entire universe of mutual fund manager behavior and informed trade profitability. We are able to systematically and predictably identify fund manager's profitable trading on their "tracked" firms and insiders throughout our sample period (2004-2015), using novel data from the Securities and Exchange Commission. Further, our classification scheme is able to richly identify fund manager informed trading even out of seeming identical behavior (i.e., informed buying and selling vs. all other buying and selling). These abnormal returns exist through the present day, and are even slightly *stronger* in point estimate in the most recent period.

In order to do this for the full sample, we first document that mutual fund managers have a very specific set of firms (and insiders) that they track. We then find that their tracking and monitoring activities, again measured using their downloading behavior, have powerful implications for their portfolio choices and future returns. For instance, the fact that an institution downloaded an insider-trading filing by a given firm last quarter increases her likelihood of downloading an insider-trading filing from the same firm by more than 41.3% (t = 30.92) this quarter. For reference, the unconditional probability of an institution downloading at least one insider trading filing in a quarter from any firm in her existing portfolio is 4.8%. In other words, our persistence result – an 8 times increase in probability - is not only statistically significant, but also economically important. We find this is driven by persistence at the individual insider-level tracking. For instance, an increase in the probability of 18.7% (t = 24.72) of downloading say Jamie Dimon's insider trading filing if the manager downloaded the same insider's filing the prior quarter.

We show that these "tracked" insider trades have predictability for future firm operations and returns. In particular, when an institution buys a stock following a tracked insider buy, this portfolio outperforms other fund manager purchases by a risk-adjusted 4-factor alpha of 354 basis points per quarter (t=2.35), or over 14% annualized abnormal return per year. Similarly, when an institution sells a stock following a tracked insider sell, this portfolio underperforms other institutional sales by a risk-adjusted4-factor alpha of 209 basis points per quarter (t=3.74), or over 8% annualized abnormal return per year. These returns are unaffected by known risk determinants or factor model adjustment chosen (i.e. 4-factor and DGTW).

In addition, if the results we find reflect institutional managers exhibiting a true comparative advantage in their tracked stocks, we might expect these managers to know when *not* to "follow" the tracked insiders' behavior. For example, if the institution can decipher that the given trade was for liquidity reasons (as opposed to information-based), the manager would not want to mimic that trade of the tracked insider. This implies that when we observe institutions choosing not to follow the trades of their tracked insiders, these insider trades should have less predictive ability for future returns. We find this pattern in the data. In particular, firms in which institutions buy alongside tracked insider purchases tend to outperform those in which tracked insiders buy but the institutions choose *not* to buy alongside the insiders; we find analogous results on the sell side as well. A final counterfactual we explore is the returns to following all insider buys and sells; we show that the returns we document are not simply earned by following all insider buys (or sells), but rather that managers' tracking behavior appears to allow them to engage in valuable stock selection above and beyond simply mimicking all insider trades.

We show that the outperformance that we document on these tracked trades continues for a number of quarters following the tracked insider (and institution) sell. Importantly, it never reverses, suggesting that the information in tracked trades is important for fundamental firm value, and is only revealed following the information-rich dual trading by insider and linked institution.

Further, since our institutional holdings data is primarily quarterly in nature—as opposed to the SEC download files and insider transaction data which both contain precise timestamps -- we also use more granular data from Ancerno on the daily trades of institutional investors to directly map the timing of institutional trades to the timing of insider transactions. We show that institutions that trade in the direction of recent insider transactions do so relatively quickly, with a large proportion of these trades (almost 50%) occurring within 30 days after the insider trade; far less occurs before the insider trades, or after this 30-day window. Moreover, a sizable percentage of the return occurs in the direct proximity following the tracked trade of the insider, underscoring the importance of the real-time tracking in this linked relationship between the institution and a given insider.[1]

We explore the mechanism at work behind our results in a variety of ways. First we show that institutional managers tend to track members of the top management teams of firms (CEOs, CFOs, Presidents, and Board Chairs) and accountants, and shy away from tracking outside directors and insiders with PhDs. Next we take our institutional holdings data and isolate the fund managers within each institutional investment firm whose holdings correlate most with the insider trades tracked by their fund company, and explore these "cherry picking" fund managers in greater depth. We find that these cherry picking fund managers are more likely to: a) be located in California and Massachusetts, and b) have an educational or location-based link to the insider in question. These findings are consistent with the idea that fund managers choose to track and mimic the trades of the specific insiders who not only possess the most valuable information, but also those with whom they have lower-cost channels for obtaining private information.

One note on our finding of strong outperformance on both tracked insider buys and insider sales couched in the universe of insider trading. The literature on insider trading has found a few systematic empirical facts regarding its profitability. Insider buys

[1] We also explore alternate (conservative) timing conventions to show the robustness to these alternative conventions.

are followed by, on average, systematic positive abnormal returns of roughly 50-100 basis points in the following month (this literature dates back to Lorie and Niederhoffer (1968) and Jaffe (1974)). However, the literature has found statistically zero returns to the average insider sale (Jeng et al. (2003) and Lakonishok and Lee (2001)). Reasons for this include potential liquidity, diversification, and other motives that could the information content of insider selling behavior. The reason this dynamic is critical is that insider sales – with their average statistically zero return – make up over 80% of all insider trades (Seyhun (1998) and Cohen et al. (2011)). Thus, in only finding robust predictability of insider buys, this comprises less than 20% of all insider trading activity. Importantly, given that we find evidence in the paper of profitable tracking and trading of both insider buying and insider selling, it suggests that fund managers appear able to exploit the rich information in the entirety of insider trading (and not solely the less than 20% linked to buys).

The paper proceeds as follows. Section II lays out the background for the setting we examine in the paper. Section III presents our data collection procedures, and summary statistics. Section IV provides our main results on the robust behavior of institutions tracking particular insiders and the return predictability pattern of institutions tracked insider trades. Section V explores the mechanism driving our findings in greater depth; Section VI presents results using alternative timing conventions, as well as some robustness checks; and Section VII concludes.

II. Literature Review

Our work relates to a several strands of the literature, including papers analyzing the investment performance of mutual fund managers, articles exploring the characteristics and profitability of insider trading, and a slew of studies documenting gradual information diffusion and limited attention in the stock market.

The area of the mutual fund literature most closely related to our paper is the collection of work examining whether mutual fund managers possess stock-picking ability. This remains an open question, because while many papers (Jensen (1968), Malkiel (1995), Gruber (1996), and Carhart (1997)) find that active managers fail to outperform passive benchmark portfolios (even before expenses), several others (Grinblatt and Titman (1989, 1993), Grinblatt, Titman, and Wermers (1995), Daniel et al. (1997), and Wermers (1997)) find that active managers do exhibit stock-picking ability. In terms of specific characteristics known to correlate with superior performance, Chevalier and Ellison (1999) use biographical data on managers to show that fund managers from undergraduate institutions with higher average SAT scores earn abnormal returns.[2] Other evidence shows that fund managers tend to overweight nearby companies (Coval and Moskowitz (1999), and earn higher returns on their local holdings (Coval and Moskowitz (2001)). Closest to this paper is Cohen, Frazzini, and Malloy (2006), who find that fund managers place bigger bets on firms they are connected to through their education network, and perform significantly better on these holdings relative to their non-connected holdings. Hong, Kubik, and Stein (2005) also document word-of-mouth effects between same-city mutual

[2] Massa and Simonov (2005) also document a relation between the portfolio choices of individual investors and their past educational backgrounds.

fund managers with respect to their portfolio choices. We add to this list by exploring to what extent mutual funds actively investigate the insider trades on stocks within their own portfolios. Our approach highlights another channel through which fund managers earn abnormal returns.

Our paper is also closely related to a large literature examining the behavior of corporate insiders. Many of these papers study the cross-sectional return forecasting ability of insider trades aggregated at the firm level. Numerous papers (see, for example Lorie and Niederhoffer (1968), Jaffe (1974), Seyhun (1986, 1998), Rozeff and Zaman (1988), Lin and Howe (1990), Bettis, Coles, and Lemmon (2000), Lakonishok and Lee (2001), and Marin and Olivier (2008)) focus on the abnormal returns to firms in relation to various metrics of firm-level insider trading. Seyhun (1998) summarizes this evidence and reports that several different trading rules lead to abnormal returns. In addition, Jeng et al. (2003) show that insider purchases earn abnormal returns of more than 6% per year, while insider sales fail to earn significant abnormal returns.

Several papers take a more granular approach and examine individual insider-level data in order to identify which insiders are truly informed. For example, Cohen, Malloy, and Pomorski (2010) show that the past trading records of insiders can be used to identify which insiders are likely to be trading on information and which are not. In addition, Piotroski and Roulstone (2005) demonstrate that insider trades reflect both contrarian beliefs as well as private information about future cash flows, and Ke, Huddard and Petroni (2003) demonstrate that insiders trade before significant accounting disclosures. Kahle (2000) shows that long-run stock returns associated with seasoned equity offerings (SEOs) are significantly related to measures of insider trading, and Clarke, Dunbar, and

Kahle (2001) provide evidence consistent with insiders exploiting windows of opportunity by trying to issue overvalued stock. Finally, Jagolinzer (2009) presents more evidence of strategic trading by insiders by focusing on a subset of insiders who publicly disclose 10b5-1 plans; he shows that insiders initiate sales plans before negative returns and terminate sales plans before positive returns.

Our paper can also be situated within the large and growing literature on limited attention, and the slow diffusion of information in the stock market. Many of these papers argue that if investors have limited resources and capacity to collect, interpret, and finally trade on value-relevant information, we should expect stock prices to incorporate information only gradually. For instance, because of gradual information diffusion (Hong and Stein, 2007) and/or gradual capital diffusion (slow moving capital (Duffie, 2010)), this information may be impounded into stock prices slowly. Meanwhile, there is a substantial literature studying investors' limited attention to information. Theoretical papers such as Merton (1987), Hong and Stein (1999), and Hirshleifer and Teoh (2003), argue that with investors subject to binding attention and resource constraints, delayed information flows can lead to expected returns that are not explained by traditional asset pricing models. Numerous empirical studies find supporting evidence for these models. For example, Huberman and Regev (2001), Barber and Odean (2006), DellaVigna and Pollet (2006), Hou (2006), Hong, Torous, and Valkanov (2007), Cohen and Frazzini (2008), and Cohen and Lou (2011) find that investors respond quickly to salient, eye-catching information, but tend to ignore information that is less obvious yet nonetheless essential to firm value.

Since our work utilizes the log file from the SEC Edgar database, our paper is also related to a few recent papers that use this data to explore different, but related issues in corporate finance and asset pricing. For example, Loughran and McDonald (2016) provide a first descriptive analysis of this dataset and show that--after sifting out robot requests--the average publicly-traded firm has their annual report requested only 28.4 total times by investors immediately after the 10 K-filing; they conclude that the “lack of annual report requests suggests that investors generally are not doing fundamental research on stocks.” Meanwhile Lee, Ma, and Wang (2016) apply a “co-search” algorithm to the SEC log file in order to identify economically-related peer firms; they show that firms appearing in chronologically adjacent searches by the same individual are fundamentally similar on multiple dimensions. Finally, Drake, Roulstone, and Thornock (2016) show that EDGAR activity is positively related with corporate events (particularly restatements, earnings announcements, and acquisition announcements), poor stock performance, and the strength of a firm’s information environment; EDGAR activity is also related to, but distinct from, other proxies of investor interest such as trading volume, business press articles, and Google searches. Meanwhile Bozanic, Hooppes, Thornock, and Williams (2017) focus on IRS use of the Edgar data; Gibbons, Iliev, and Kalodimos (2018) explore sells-side analysts’ downloading activity; Iliev, Kalodimos, and Lowry (2018) examine mutual fund family downloads of proxy filings and the governance implications of this type of monitoring. While all of these papers explore the SEC logfile data in various ways, none of them explicitly link the users of the data to individual institutional investors (i.e., the 13F filers that feature in our analysis) in order to explore the investment implications of SEC searches in a direct and granular way.

III. Data and the Setting

We combine data from a variety of sources in order to execute the empirical tests in this paper. We use CRSP to obtain stock related information, the Thompson Reuters Insiders database to obtain insider transactions, and Thompson Reuters Ownership data to obtain stock holdings of fund families. We also construct three unique and novel datasets using various sources. In this section, we describe in detail how we construct these three datasets, which entail: (1) matching IP addresses to 13-F organizations, (2) matching 13-F organization names to daily trades reported in the Ancerno daily mutual fund holdings database (described further below), and (3) collecting biographical background information both for corporate insiders and mutual fund managers.

A. Matching IP Addresses to 13-F Organizations

To match IP address to 13-F organizations, we follow a four-step procedure. First, we obtain the data on SEC filings and the IP addresses of their viewers from the SEC at the log file website (https://www.sec.gov/data/edgar-log-file-data-set). The log files are available from 2003 onwards, and are posted by the SEC on a quarterly basis with a 6-months delay. We use mainly the data logs from 2004 to 2015 due to the limitation of the IP geolocation data.

The IP addresses in the dataset are partially anonymized using a static cipher. The data describe the access of filings by different IP addresses. That is, each row of the data corresponds to a certain IP address (24.145.236.jcf) viewing a specific filing coded by an accession number (0000891020-04-000160) at a specific time and date (12:00 am on April 31, 2004).

In order to match the data to organization-level information, we first de-anonymized the ciphers. This is done by using another set of server logs from a private but well-trafficked website. Assuming that the intersection of IP addresses for existing IP are similar between two servers, we are able to deduce which cipher, say aaa, corresponds to what number between 0 and 255 using the frequency of potential matches.

For instance, if 1.1.1.aaa visited the SEC server and 1.1.1.111 visited the private website on January 1st, 2004, then the cipher aaa to 111 has 1 additional match. The number of matches between a cipher and its most frequently matched last IP octet is distinct for the vast majority of the ciphers. Out of the 256 ciphers, over 230 have a most frequently matched octet that does not intersect with any other potential pairing. The last digit octet is the most frequent for only 1 cipher. The rest of the pairings are matched using a process of elimination. For example, if 001 is the most frequent octet for both aaa and aab, but 001 has many more matches with aaa than aab, then aaa is matched to 001. In this case, aab is then linked to its next most frequent octet until all 256 pairs are matched.

Once the IP data is deciphered, we connect the specific filing IP address to a set of organizations using a dataset of organization IP addresses from MaxMind. The IP of organizations data is released on a periodic basis. The IP address linked to each viewing from Edgar is matched with the last available organization data for that IP address at the time of the viewing.

In the second step, we hand-match names of the 13-F organizations to the list of potential organizations from MaxMind by Research Assistants. We start off with two thousand 13-F organizations with the largest average AUM between Q1 2004 and Q4 2015. Since IPs are non-static and the MaxMind data also changes from period to period, certain institutions appear more frequently and longer than others. In Table 1, we illustrate the cipher table we used in order to de-anonymize the IP addresses of the 13-F organizations in our sample.

In the third step, we identify the link between each 13-F filings institution and their IP addresses, and focus on the documents accessed in the EDGAR system. We use the WRDS accession filing database to link each IP viewing to a specific filing. This specific accessing filing contains a mapping of each EDGAR document to a COMPUSTAT firm. After this step, we are able to observe which institution tracked which filing in the EDGAR, i.e. Fidelity at xxx.xxx.xxx.xxx viewed a specific Form 3 for Apple on a particular date.

In the final step, we scrape the insider trading filings from the SEC website in order to obtain the datacodes recorded in each form. This datacode in each insider form allow us to obtain the accession numbers necessary to match to the Thompson Insider database. For example, an insider trade Form 3 (0000891020-04-000160) represents Tim Cook's unloading of shares. After following these steps, we are able to observe which of the identified IP addresses of 13-F organizations accessed which particular insider trading forms on the Edgar server from Q1 2004 to Q4 2015.

B. Matching 13-F Organization Names to Daily Trades Reported in the Ancerno Database

In this section we describe how we match the institution identifier from the Ancerno database (managercode) to the 13-F institution identifier (mgrno), which we use to aggregate the tracking of insiders by individual IP addresses. The primary data source consists of detailed institutional stock transactions from Ancerno (formerly Abel Noser), a leading consulting firm that works with institutional investors to monitor and optimize their equity trading costs. Ancerno's clients include major pension plan sponsors such as California Public Employees' Retirement System (CalPERS) and United Airlines, mutual fund families (i.e., money managers), such as Fidelity Investments and Putman Investments, and a small number of brokerage firms. This dataset is also used in other studies, such as Goldstein, Irvine, Kandel, and Wiener (2009), Puckett and Yan (2011), Hu, Mclean, Pontiff, and Wang (2013), and Cohen, Lou, and Malloy (2016).

Our sample period for the Ancerno data is Q1 2004 to Q3 2011. In this period, we start off with 11,649 *managercode-year-quarter* observations in the Ancerno database. This reports the name of the institutional manager for each of its client portfolios. Using this information, we are able to manually match the trading by institutions from the Ancerno database to the available institutions gathered from the Edgars SEC data. As a result of this procedure, we are able to identify 91 of the 779 mgrnos from the 13-F data in Ancerno.

C. Collecting biographical background information both for corporate insiders and mutual fund managers.

To collect biographical information on corporate insiders, we use the BoardEx database, which collects information about individuals who have been on the board of or assumed an executive manager role at a publicly traded firm or a major private firm. The set of personal information includes academic qualifications, current and past job positions, and memberships in professional and other groups. To collect the biographical backgrounds of mutual fund managers, we use the Morningstar database, which contains fund-level performance measures as well as manager profiles. We manually parse through the education information of fund managers after searching for each fund name in the database, and pay special attention to make sure that we capture the fund manager who is in charge of the fund at the time of the trade. In total, we have collected biographical information on 53,744 corporate insiders and 225 mutual fund managers associated with our sample of insiders and institutional portfolios.

Table 2 reports the summary statistics of the forms downloaded by matched financial institutions, and the frequency of these forms in the Edgar database. The first three columns report the top accessed forms, their respective number of downloads by all matched institutional IP addresses, and the relative frequency of these downloads. The next three columns remove Mass Download IP addresses- IP addresses that access more than 3000 filings per day- and report the same three statistics. The last three columns are the top most filed forms, their respective number of filings, and the relative frequency of these forms in the Edgar Database.

There were over 400 million form file requests over the sample period. The two most requested filing types were corporate 8-Ks and insider trading filings (Form 4s). 8-Ks are required to be filed by firms to notify shareholders of material events transpiring at the firm. After removing mass downloads, Form 4 downloads still represent over 7% of the total downloads. Form 4s also represent over 36% of all forms filed in the Edgar Database.

Panel B of Table 2 reports the correlation between the holding firm checked by institutions through 10-K forms, 8-K forms, and Insider Trading Forms (4, 5, 6, and their amendments). This panel shows that there is a positive correlation between an institution's tendency to check the insider filings along with the 10-K and 8-K forms.

Table 3 lists the top 30 active institutions (in terms of total value of the active portfolio last observed) that we are able to link to an IP address. Virtually all of the large fundamental stock picking investment firms are represented in this list. Note that we only include active stock pickers / non-diversified institutions in our tests; for example, we exclude all institutions with more than 500 stock positions across all their funds, which removes index providers like Vanguard.

IV. Empirical Results

A. Persistence in Tracking Behavior

The main thesis of our paper is that investors, considering their resource constraints, should optimally choose to focus their information gathering efforts on a subset of the firms and a subset of the signals where they have a comparative advantage in terms of collecting and interpreting the information. To illustrate, if investor A has a comparative edge in interpreting information from the healthcare industry (due to, for example, her prior work experience), we expect the investor to focus her research activity, and consequently her portfolio holdings, in this industry.

Moreover, since comparative advantages in information processing are accumulated (developed) through years of experience and interactions with other economic agents, and are thus unlikely to change rapidly over time, we expect persistent patterns in investors' information-gathering activity. We start our empirical analysis by examining the following question: conditional on investor A searching for regulatory filings by company X in one period, do we see searches by the same investor on the same firm in the *next* period?

Table 4 reports the persistence in institutions' search behavior for insider trading filings (Forms 4, 5, and 6) on the EDGAR server. We conduct a panel OLS regression where the dependent variable is a dummy that equals one if an institution downloads at least one insider-trading filing by a given firm in quarter t. The main independent variable of interest is a similar dummy defined in quarter t-1. As can be seen from Column 1 of Panel A, there is substantial persistence in institutions' search behavior. The fact that an institution downloaded an insider-trading filing by a given firm in quarter t-1 increases her likelihood of downloading an insider-trading filing from the same firm by more than 41% (t = 30.92) the next period. For reference, the unconditional probability of an institution downloading at least one insider trading filing in a quarter from any firm in her existing portfolio is 4.8%. In other words, our persistence result is not only statistically significant, but also economically important.

In Columns 2 and 3, we further include a host of control variables, as well as portfolio fixed effects and year-stock fixed effects. Our results remain economically large. For example, in Column 3 (with the full set of controls and fixed effects), the coefficient on lagged search dummy is 0.255 (t = 20.75)) — i.e., an institution that downloaded insider filings of a given firm in the prior quarter has a 25.5% higher chance of downloading insider filings by the same firm again in the following quarter.

In the next three columns of Table 4, we narrow in on the specific insiders. In other words, we track not only institutions' searching for insider filings by Apple, but also the specific filings by Tim Cook. The results are consistent with those shown in the first three columns. As can be seen from Column 4, an institution that downloaded an insider trading filing by a given executive in a quarter has an 18.7% (t = 24.72) higher likelihood of downloading an insider trading filing by the same executive in the following quarter. Again, including portfolio and year-times-insider fixed effects has little impact on our results. For example, in Column 6, where we include the full set of controls and fixed effects in our regression, the coefficient on lagged search behavior drops only slightly to 13.3% (t = 21.48).

B. Contemporaneous Trading of Fund Managers and Corporate Insiders

After establishing that institutions' search behavior on EDGAR is highly persistent (that is, each institution tends to follow the same group of firms and insiders over time), we next turn to institutions' trading decisions. In particular, we examine whether institutions trade in the same direction as the insiders that they follow. To this end, we classify trading in each stock by each institution as either a buy or a sell, based on changes in the number of shares.[3]

In Table 5 we report the results of panel regressions of the direction of contemporaneous trading by institutions on the direction of trading by the insiders checked by these institutions. The dependent variable, Direction, is set equal to -1 if the position is a sell, and 1 if the position is a buy. Direction_Checked_Insider is set equal to -1 if the insiders that were checked by the institution sold in net, and 1 if the insiders bought in net. Similarly, the variable Buy_Checked_Insider is equal to 1 if the insiders checked bought in net, and Sell_Checked_Insider is equal to 1 if the insiders checked sold in net. Meanwhile Direction_All_Insider is set to -1 if *all* the insiders from the firm sold in net during the quarter, and 1 if all the insiders from the firm bought in net. And again, Buy_All_Insider is set equal to 1 if all the insiders from the firm bought in net during the quarter; and Sell_All_Insider is set to 1 if all the insiders from the firm sold in net during the quarter. The panel is weighted by the inverse of the number of positions in each portfolio (so that a portfolio with many positions does not dominate the regression). Only non-diversified institutions, which we classify as institutions with less than 500 unique stock positions across all their funds, are included in the regressions.

[3] Our results are largely unchanged if we instead define trading using changes in portfolio weights.

As shown in the first row of Table 5, institutions' trading behavior is highly correlated with the insider trading of the insiders they actively track. The coefficient on Direction_Checked_Insider is large and positive (=0.1671, t-sta$t$=3.97), and implies that the probability that an institution trades a given stock it holds in a given direction increases by 41% if one of the firms in its portfolio has net tracked insider trading in that direction in that quarter. From Columns 2 and 3, we see nearly symmetrically large and significant trading responses of fund managers to both tracked insider buys and tracked insider sales from. Moreover, in Columns 4-6 we control for both: i.) overall insider trading of the firm (i.e., not just the insiders being tracked by the given manager at the firm in question), and ii.) lagged trading of the given manager in the given firm the prior quarter (i.e., was the fund manager buying or selling the given stock last quarter, in net). While lagged trading does predict future trading, insider trading of the entire firm does not significantly predict fund manager behavior. Importantly, tracked insider trades remain large and significant predictors of fund manager behavior even controlling for these. These results highlight the particularly important information that appears to be embedded in tracked insider trades for mutual fund manager trading behavior.

C. Portfolio Returns to Active Insider Tracking

Given that institutions tend to trade in the same direction as the subset of managers they follow, a natural question to ask is whether institutions earn abnormal returns from these trades. If institutions correctly choose which firms/managers' insider trades to follow based on their comparative advantages to process/interpret information, we should expect these trades to generate positive abnormal returns.

To test this idea, we form equal-weighted calendar-time portfolios using various sorting criteria, and compute the return differentials across these portfolios.[4] The timing of these tests is as follows: a) in months *t*-3 to *t*-1 we observe the trading behavior of institutions at the quarterly frequency (note that later in the paper we explore daily trading records for a subset of our sample in pinpoint the timing even more directly), as well as the exact insider trading date and corresponding download date by these investors, and b) in months *t+*1 to months *t+*3 we compute the quarterly returns to these funds managers' trades.

In the daily trading tests that we present later in the paper, we demonstrate that the vast majority of fund manager trades that occur around insider trades occur *after* the reported insider trade date, so we do not believe our results are driven by investors trading before the insider trades, and before they download the SEC file. However, to further allay this concern we also employ an alternate timing convention in the Appendix where we force the insider trade and download dates to occur in the quarter *prior* to any subsequent trades by the fund manager, and then explore returns in the quarter following those fund manager trades. Given that the insider trading literature has documented that the returns to following insider trades primarily accrue in the 30 days immediately after the reported trade dates (a fact we replicate in our sample), the idea of forcing our portfolio tests to wait at least 3 full months after the insider trade dates to measure any return benefits of insider following seems overly conservative. However, as shown in the Appendix, even using that conservative timing convention we still find abnormal returns for fund managers who trade in the direction of their tracked insider trades.

[4] We discuss the value-weight return results below.

In Table 6 Panel B we report the raw, DGTW-, and 4-factor-adjusted returns of portfolios relating to insiders who bought stocks followed through the Edgar system. We begin by computing the total portfolio return of an average institution in our sample. As shown in the first row of Table 6, during our sample period, the average institution earns a marginally significant and economically small DGTW-adjusted quarterly return of 18bps ($t = 1.73$).

Next we construct a variety of counterfactuals and benchmark portfolios to try to isolate the information content of fund managers' purchases following insider trades that they actively follow. For instance, we divide each institution's entire portfolio into several sub-components: the first sub-portfolio---shown in row 2, and which serves as one possible benchmark--includes all holdings where the institution does not check/download a Form 4 filing after an insider trade, or (i.e., the union of) all holdings where the fund managers tracks an insiders who chooses to sell stock. Row 3 then computes a long-short portfolio, which is held for one quarter, where the long portfolio consists of the future quarterly return to stocks where a fund manager downloads an insider purchase filing, and the short portfolio consists of the portfolio in row 2. This long-short portfolio, while non-trivial in magnitude (around 1 percent per quarter), is nonetheless insignificantly different from zero. However, in row 4 if we instead compute a long-short portfolio where the long portfolio consists of all stocks where a fund manager checks the insiders' buy transaction and *also purchases* shares in that stock during the same quarter, while keeping the same short portfolio represented in row 2, the long-short portfolio earns a statistically and economically significant 3.44 percent in the following quarter (t-sta$t$=2.32). Then in row 5 we change the benchmark and instead construct a long-short portfolio where we compare the returns to a portfolio where the fund manager checks an insider purchases and also buys the stock (the long portfolio) to the returns to all other stocks that fund managers also buy in that same quarter (the short portfolio). Note that here the benchmark (i.e., the short portfolio) consists of all other stock purchases by active alpha-seeking fund managers, and yet Row 5 shows that this long-short portfolio still earns a large 3.41 percent in the following quarter (t-sta$t$=2.32), indicating the large information content of these tracked insider trades.

In rows 6 and 7 of Table 6 Panel B we then construct an alternate benchmark that consists of all stocks where a fund manager downloads the filing after an insider purchase, but chooses *not* to buy the stock in that quarter. As noted earlier, if our results truly reflect institutional managers exhibiting a comparative advantage in their tracked stocks, we might expect these managers to know when *not* to "follow" the tracked insiders' behavior. For example, if the institution can decipher that the given trade was for personal reasons (as opposed to information-based), the manager would not want to follow that trade of the tracked insider. Indeed we find that these insider trades have less predictive ability for future returns; for instance, in row 6 when we use this counterfactual as our short portfolio and use the same long portfolio as in rows 3-5, the long-short spread is large (but statistically insignificant); in row 7 if we further restrict the long portfolio to only situations where the fund manager makes a large increase in the stock position from a zero or insignificant initial position, the long-short spread increases to over 4% per quarter and becomes statistically significant or at least marginally significant.

Our final counterfactual/benchmark portfolio involves looking at the returns to all stocks that are not checked by a fund manager but where the insider purchases shares. Interesting the portfolio of tracked insider buys with a corresponding fund manager purchase again outperforms this benchmark of all other insider purchases by up to 3.68%

per quarter (t=2.48). Collectively, our findings point to large and significant return predictability from fund managers' actively tracking and following specific insider trades. This result holds for a variety of natural benchmarks: relative to all other (untracked) insider trades, relative to all other fund manager purchases, and relative to insider trades that they check but do not mimic.

Table 6 Panel C presents the analogous results around insider sales, as opposed to insider purchases. The results are remarkably consistent with the purchase results shown in Panel B, albeit at somewhat smaller magnitudes for the L/S portfolios in rows 5 through 8—more in the 6.5 to 8.5 percent annualized range, as opposed to 9-18 percent annualized range for the purchase results. Note that these magnitudes are still quite large for 4-factor annualized alphas. For instance, consider row 5 where the long side requires both the institution and its tracked insiders to sell the security in question, and where the benchmark portfolio includes all holdings which are sold by the institution but not by tracked insiders; in this specification the portfolio underperforms the benchmark by 175bps (t-stat = 3.10) to 209 bps (t-sta$t$=3.74) in the following quarter. In row 6, we again define the benchmark as including all holdings that are not sold by the institution but are sold by tracked insiders. Similar to the result from Panel B, the return differential for the tracked stocks that are sold relative to the tracked stocks that institutions choose *not* to sell is economically large and either significant or marginally significant depending on the precise specification. Finally, if the counterfactual benchmark portfolio is instead defined as the set of insider sells which are not downloaded by the institution, the L/S spread is again large and significant (ranging from 1.74 to 2.02% per quarter in abnormal returns).

Collectively, the results in Table 6 indicate that institutions are able to identify the most informative trades by closely tracking the trades of specific corporate insiders.

V. Mechanism

In this section we explore the mechanism behind our key results in greater depth. We examine the characteristics of the insiders who are tracked, the characteristics of the fund manager doing the tracking, and the characteristics of the matched insider-institution pairs, in order to investigate the drivers of the return predictability we document above.

A. Characteristics of Insiders Who Are Tracked By Institutions

First we examine the profiles of the specific insiders who are being searched in our sample. The first column of Table 7 contains the names of the various background characteristic associated with the firm insiders. The second column reports the percentage of each type of profile across all BoardEx reported insiders. The third column reports the percentage of each type for profiles that are actively checked by institutional investors, and computes differences between the percentage of checked insiders relative to all insider profiles.

Table 7 indicates that institutional managers tend to track members of the top management teams of firms (CEOs, CFOs, Presidents, and Board Chairs) and accountants, and shy away from tracking outside directors and insiders with PhDs. For instance, CEO-related insider trade reports account for 28.7% of the total downloads by institutions, even though they make up only 9.8% of all insider transactions, a difference of 18.8% (t-stat=41.99). We find similarly large differences for Presidents, Board Chairs, and CFOs, suggesting that institutional managers perceive these filings to be more value-relevant than ordinary director filings (which they download at a significantly lower rate than their overall incidence in the population of filings).

B. Characteristics of Fund Managers Who Track and Mimic Insider Trades

Next we take our institutional holdings data and attempt to isolate the fund managers within each institutional investment firm whose holdings correlate most with the insider trades tracked/downloaded by their institution. We label these highly correlated fund managers as "cherry picking" fund managers. Specifically, for each portfolio *j* in the 13-F family *k*, we run the following regression:

$$Sell_{j,i,t+1} = \alpha_j + \beta_j \cdot Insidersell_{i,t} + \epsilon_{j,i,t}$$

- $Insidersell_{i,t}$ is an indicator for whether the insiders at firm *i* observed by the 13-F family *k* had in net sold shares at *t*.
- If β_j is >90% significance level, we call the fund *j* a Cherry Picker.

We then examine the geographic and educational characteristics of these fund managers in Table 8. Panel A of Table 8 reports the distribution of top locations (by state) of all matched mutual fund managers, and specifically the fund managers we dub to be cherry pickers. Panel A shows that cherry pickers are more likely to reside in Massachusetts, which has a high concentration of fundamental stock pickers in the Boston area (e.g., at Fidelity, Wellington, etc.), and also California (where Capital Group and many other institutional stock selection houses are located). Panel B then records the distribution of education backgrounds of matched mutual fund managers relative to the cherry pickers. The universities with the highest absolute differences between manager and cherry picker distributions are reported. Interestingly the single most under-represented school relative to its overall distribution is the University of Chicago, the home of "efficient markets," where students are repeatedly taught that fundamental stock pickers generally do not add alpha to an otherwise diversified equities portfolio.

C. Exploring the Source of Fund Manager - Insider Links

In Table 9 we try to pinpoint the mechanism more cleanly by focusing on the manager-insider links that we observe in the data. Specifically, we try to better understand why certain fund managers choose to follow certain corporate insiders. To do so, we explore the connection links between a cherry picking mutual fund manager and an insider, based on commonalities in location and educational backgrounds. Again we define cherry pickers as fund managers whose trades each quarter correlate with net checked insider trading at the 90% significance level. The Match Indicator is a dummy variable for whether the insider at a public firm that a cherry picker searched for had commonalities in educational background and location with that fund manager. The sample includes all potential insiders linked by each fund manager's portfolio holdings. Table 9 shows that commonalities in educational background and location are strong predictors of the likelihood of a tracking match, even controlling for past matching, as well as time fixed effects. Collectively, the findings in Tables 8 and 9 are consistent with the idea that fund managers choose to track and mimic the trades of the specific insiders who not only possess the most valuable information, but also those with whom they have the lowest-cost channels for obtaining private information.

VI. Alternate Timing Conventions, Daily Trading Records, and Robustness

We also explore a variety of alternate timing conventions, as well as novel data on the daily trading behavior of institutions in order to verify the robustness of our findings, and solidify the interpretation of our results.

First, in Appendix Table 1, we re-run the return predictability analysis from Table 6, but impose a longer lag structure in the return tests. In this Table A1 we require the insider trade and download dates to occur in the quarter *prior* to any subsequent trades by the fund manager, and then explore returns in the quarter following those fund manager trades. Note that this builds in a minimum of 90 days (and on average, far longer) from the insider trade date to the beginning of the measured future return period. As noted earlier, given that the insider trading literature has documented that the returns to following insider trades primarily accrue in the 30 days immediately after the reported insider trade dates, the idea of forcing our portfolio tests to wait a minimum of 3 full months after the insider trade dates to measure any return benefits of insider following seems conservative. Nonetheless, Appendix Table A1 indicates that using this alternate timing convention still reveals meaningful abnormal returns for fund managers who trade in the direction of their tracked insider trades, albeit at a smaller magnitude and with only marginal significance.

To further motivate the timing convention used in Table 6, we collect an alternate source of institutional holdings at the *daily* level from the Ancerno database described earlier. This data contains daily stock-level holdings and trades at the institution level for a subset of our overall sample (the description of our matching of this institutional data to the precise fund-manager level is contained in Section III). Using this data, we explore the oft-cited anecdotal claim that institutions generally follow insiders, but rarely trade ahead of them. To test this conjecture, Figure 1 tabulates the proportion of gross purchases of stock by institutions over a number of trading days after checking that insiders had purchased a stock in net. Figure 2 presents the analogous results for sales of stocks, as opposed to purchases. Both Figures 1 and 2 paint a similar picture: most institutional trading that follows in the direction of insider trading happens quickly, within 30 days, and far less happens right before or right after that 30-day window post-insider trade. These results suggest that our return predictability results in Table 6 are highly unlikely to be an artifact of institutions trading before the insider trade filing that they subsequently download.

Moreover, we use the Ancerno sample to explore the return results of tracked trades in a setting where we can measure the tracking and precise trade timing of fund managers. In particular, we see the exact date on which the trade is executed by the manager following the tracking of the insider trade. Although this is a much smaller sample – both in the cross-section and time-series – and so the power is lower, we still see large and significant predictive power of these tracked trades for future returns. In particular, if a fund buys a stock after the checked insider had bought the stock, these stocks outperform a stock that wasn't bought after checking by 1.88% (t=2.15) over the next 20 trading days. Moreover, if a fund sells a stock after the checked insider sold the stock, these stocks underperform a stock that wasn't sold after checking by 3.22% (t=4.70) over the next 20 trading days

In Figure 3 we also investigate the long-run returns to tracked insider trades by fund managers. Specifically, we show in this figure that the outperformance that we document on these tracked trades continues for a number of months following the tracked insider and institution trading in the same direction. Importantly, the cumulative returns never reverse, suggesting that the information in tracked trades is important for fundamental firm value, and is only revealed following the information-rich dual trading by insider and linked institution.

VII. Conclusion

With the proliferation of information signals in both quantity and dimensionality in recent decades, investors face an increasingly complex portfolio choice problem. Most investors simply do not have enough resources and time to comb through all the information available to them. With hundreds of thousands of information signals being produced in any given day, how does an investor reduce the dimensionality of the investment problem sufficiently to know which signals to track and collect?

In this paper, we show one way that investors attempt to reduce the dimensionality of this increasingly complex investment search problem. Namely, using web traffic on the SEC's EDGAR servers between 2004 and 2015, we find that mutual fund managers track a very particular subset of firms and insiders. In addition, fund manager tracking activity not only remains persistent over time, but also has powerful implications for their portfolio choice and subsequent performance. For example, when an institution buys a stock following a tracked insider buy, this portfolio outperforms other fund manager purchases by a highly significant 14% alpha. Equivalently, when an institution sells a stock following a tracked insider sell, this portfolio underperforms other institutional sales by an annualized alpha of over 8% (t=3.74). Moreover, managers seem able to know precisely *which* of their tracked insiders' trades (e.g., Jamie Dimon's trades) to follow, and which not to, as the trades they track and choose to act upon significantly outperform those that they track and choose not to trade along with. Lastly, the abnormal returns that following these tracked trades continue to accrue for the quarters following the trades and do not reverse, suggesting that the information contained in the trades is important for fundamental firm value, and is revealed and incorporated into firm value only following the information-rich tracked trades.

Mutual fund managers choose to track very specific profiles of insiders. On average, they significantly tilt their tracking and tracked-trading to: CEOs, CFOs, and Board Chairs, while at the same time shying away from tracking non-chair directors. Moreover, while a number of factors likely determine the precise choice of connection, two that we find are significant predictors are close proximity and school alumni network connection between the fund managers and firm insiders.

Stepping back, as the costs of producing, disseminating, collecting, and processing information signals continue to fall, signal proliferation will only accelerate. This will make dimensionality reduction, if anything, a growing problem facing investors for the foreseeable future. We believe that our study - using novel, rich, micro-data on fund manager tracking and trading behavior – is a first-step to micro-founding and understanding successful attempts to do precisely this. Future research should push ahead even further to establish alternate ways that investors can solve this problem, and continue to engage as important information collectors and price-setters in modern capital markets.

References:

Barber, B., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies 21, 785–818.

Bettis, Carr, Jeffrey Coles, and Michael Lemmon, 2000, Corporate policies restricting trading by insiders, Journal of Financial Economics 57, 191-220.

Bozanic, Z., Hoopes, J. L., Thornock, J. R. and Williams, B. M. (2017), IRS Attention. Journal of Accounting Research, 55: 79-114.

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

Chevalier, Judith, and Glenn Ellison, 1999, Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance, Journal of Finance 54, 875-899.

Clarke, Jonathan, Craig Dunbar, and Kathleen M. Kahle, 2001, Long-run performance and insider trading in completed and canceled seasoned equity offerings, Journal of Financial and Quantitative Analysis 36, 415-430.

Cohen, L. and D. Lou, 2012, Complicated firms, Journal of Financial Economics, 83, 383-400.

Cohen, L, D. Lou, and C. Malloy, 2016, Cloaked trading, Journal of Investment Consulting 17, 207–248.

Cohen, L., Frazzini, A., 2008. Economic links and predictable returns. Journal of Finance 63, 1977–2011.

Cohen, L., Frazzini, A., and Malloy, C., 2008, The small world of investing: Board connections and mutual fund returns, Journal of Political Economy 116, 951-979.

Cohen, Lauren, Christopher Malloy, and Lukasz Pomorski. "Decoding inside information." The Journal of Finance 67.3 (2012): 1009-1043.

Coval, Joshua, and Tobias Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, Journal of Finance 54, 2045-2074.

Coval, Joshua, and Tobias Moskowitz, 2001, The geography of investment: Informed trading and asset prices, Journal of Political Economy 109, 811- 841.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, Journal of Finance 52, 394-415.

DellaVigna, S., Pollet, J., 2006. Investor inattention, firm reaction, and Friday earnings announcements. Journal of Finance 64, 709–749.

Drake, Michael S, Darren T. Roulstone, and Jacob R. Thornock, The Determinants and Consequences of Information Acquisition via EDGAR, 2016, Contemporary Accounting Research, Forthcoming

Duffie, D., 2010. Asset price dynamics with slow-moving capital. Journal of Finance 65, 1237–1267.

Gibbons, Brian and Iliev, Peter and Kalodimos, Jonathan, Analyst Information Acquisition via EDGAR. Penn State University Working Paper.

Grinblatt, Mark, and Sheridan Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, Journal of Business 62, 394-415.

Grinblatt, Mark, and Sheridan Titman, 1993, Performance measurement without benchmarks: An examination of mutual fund returns, Journal of Business 66, 47-68.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, American Economic Review 85, 1088-1105.

Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, Journal of Finance 51, 783-810.

Hirshleifer, D., S. Lim, and S. Teoh, 2009, Driven to distraction: Extraneous events and underreaction to earnings news, Journal of Finance 64, 2289–2325.

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

Hong, H., Tourus,W., Valkanov, R., 2007. Do industries lead the stock market? Journal of Financial Economics 83, 367–396.

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

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

Huberman, G., Regev, T., 2001. Contagious speculation and a cure for cancer: a nonevent that made stock prices soar. Journal of Finance 56, 387–396.

Iliev, Peter and Kalodimos, Jonathan and Lowry, Michelle, Investors Attention to Corporate Governance. Penn State University Working Paper.

Jaffe, Jeffrey, 1974, Special information and insider trading, Journal of Business 47, 410-428

Jagolinzer, Alan, 2009, SEC Rule 10b5-1 and insiders' strategic trade, Management Science 55, 224-239

Jeng, Leslie, Andrew Metrick, and Richard Zeckhauser, 2003, Estimating the returns to insider trading: a performance-evaluation perspective, The Review of Economics and Statistic 85, 453-471.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945- 1964, Journal of Finance 23, 389-416.

Kahle, Kathleen, 2000, Insider trading and the long-run performance of new security issues, Journal of Corporate Finance 6, 25-53.

Ke, Bin, Steven Huddard, and Kathy Petroni, 2003, What insiders know about future earnings and how they use it: Evidence from insider trades, Journal of Accounting and Economics 35, 315-346.

Lakonishok, Josef, and Inmoo Lee, 2001, Are insiders' trades more informative?, Review of Financial Studies 14, 79-111.

Lee, Charles M. C., Paul Ma, and Charles Wang, "Search-Based Peer Firms: Aggregating Investor Perception through Internet Co-Searches,"Journal of Financial Economics 2015, Vol. 116, Issue 2, Pages 410-431.

Lin, Ji-Chai, and John Howe, 1990, Insider trading in the OTC market, Journal of Finance 45, 1273-1284.

Lorie, James, and Victor Niederhoffer, 1968, Predictive and statistical properties of insider trading, Journal of Law and Economics 11, 35-53.

Loughran, Tim and McDonald, Bill, The Use of EDGAR Filings by Investors (December 3, 2016), Journal of Behavioral Finance forthcoming.

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971- 1991, Journal of Finance 50, 549-572.

Massimo Massa, Andrei Simonov; Is College a Focal Point of Investor Life?, Review of Finance, Volume 15, Issue 4, 2011, Pages 757–797.

Marin, Jose, and Jacques Olivier, 2008, The dog that did not bark: Insider trading and crashes, Journal of Finance 63, 2429-2476.

Merton, R., 1987. A simple model of capital market equilibrium with incomplete information. Journal of Finance 42, 483–510.

Piotroski, Joseph, and Darren Roulstone, 2005, Do insider trades reflect contrarian beliefs and superior knowledge about cash flow realizations?, Journal of Accounting and Economics 39, 55-81

Rozeff, Michael, and Mir Zaman, 1988, Market efficiency and insider trading: New evidence, Journal of Business 61, 25-44

Seyhun, H. Nejat, 1986, Insiders' profits, costs of trading, and market efficiency, Journal of Financial Economics 16, 189-212.

Seyhun, H. Nejat, 1988, The information content of aggregate insider trading, Journal of Business 61, 1-24.

Wermers, Russ, 1997, Momentum investment strategies of mutual funds, performance persistence, and survivorship bias, University of Colorado Working Paper.

Table 1. Cipher Table (Pre-Publication)

This table reports the mapping of IP addresses' hidden octet (code) to actual octet. The procedure we follow to identify these one-to-one mapping is as follows: We match the first 9 digits (the unciphered portion) of each IP addresses on the Edgar Server to the first 9 digits of IP addresses on a separate web trafficked server. Each hidden octet (code) is matched to each actual octet a number of times. We choose the most frequent matching actual octet as the deciphering octet. We mask the octets below, planning to make the entire table of ciphers available to all researchers post peer-review.

Code	Octet	Code	Octet	Code	Octet	Code	Octet	Code	Octet	Code	Octet	Code	Octet	Code	Octet
ghf	0	---	32	---	64	---	96	---	128	---	160	---	192	---	224
jbj	1	---	33	---	65	---	97	---	129	---	161	---	193	---	225
jdd	2	---	34	---	66	---	98	---	130	---	162	---	194	---	226
ggf	3	---	35	---	67	---	99	---	131	---	163	---	195	---	227
---	4	---	36	---	68	---	100	---	132	---	164	---	196	---	228
---	5	---	37	---	69	---	101	---	133	---	165	---	197	---	229
---	6	---	38	---	70	---	102	---	134	---	166	---	198	---	230
---	7	---	39	---	71	---	103	---	135	---	167	---	199	---	231
---	8	---	40	---	72	---	104	---	136	---	168	---	200	---	232
---	9	---	41	---	73	---	105	---	137	---	169	---	201	---	233
---	10	---	42	---	74	---	106	---	138	---	170	---	202	---	234
---	11	---	43	---	75	---	107	---	139	---	171	---	203	---	235
---	12	---	44	---	76	---	108	---	140	---	172	---	204	---	236
---	13	---	45	---	77	---	109	---	141	---	173	---	205	---	237
---	14	---	46	---	78	---	110	---	142	---	174	---	206	---	238
---	15	---	47	---	79	---	111	---	143	---	175	---	207	---	239
---	16	---	48	---	80	---	112	---	144	---	176	---	208	---	240
---	17	---	49	---	81	---	113	---	145	---	177	---	209	---	241
---	18	---	50	---	82	---	114	---	146	---	178	---	210	---	242
---	19	---	51	---	83	---	115	---	147	---	179	---	211	---	243
---	20	---	52	---	84	---	116	---	148	---	180	---	212	---	244
---	21	---	53	---	85	---	117	---	149	---	181	---	213	---	245
---	22	---	54	---	86	---	118	---	150	---	182	---	214	---	246
---	23	---	55	---	87	---	119	---	151	---	183	---	215	---	247
---	24	---	56	---	88	---	120	---	152	---	184	---	216	---	248
---	25	---	57	---	89	---	121	---	153	---	185	---	217	---	249
---	26	---	58	---	90	---	122	---	154	---	186	---	218	---	250
---	27	---	59	---	91	---	123	---	155	---	187	---	219	---	251
---	28	---	60	---	92	---	124	---	156	---	188	---	220	---	252
---	29	---	61	---	93	---	125	---	157	---	189	---	221	---	253
---	30	---	62	---	94	---	126	---	158	---	190	---	222	---	254
---	31	---	63	---	95	---	127	---	159	---	191	---	223	---	255

Table 2. Summary Statistics

In Panel A of this table, we report the summary of the forms downloaded by matched financial institutions, and the frequency of forms in the Edgar database. The first three columns report the top accessed forms, their respective number of downloads by all matched institutional IP addresses, and the relative frequency of these downloads. The next three columns remove Mass Download IP addresses- IP addresses that access more than 3000 filings per day- and report the same three statistics. The last three columns are the top most filed forms, their respective number of filings, and the relative frequency of these forms in the Edgar Database. In Panel B, we report the correlation between holding firm checked by institutions through 10-K forms, 8-K forms, and Insider Trading Forms (4, 5, 6, and their amendments).

Panel A. Summary statistics on the download frequency by form type

All Edgar Downloads			After Removing Mass Downloads			Edgar Forms		
Form	#Downloads	Frequency	Form	#Downloads	Frequency	Form	# of Forms	Frequency
4	291,592,283	55.6%	8-K	19,871,337	17.8%	4	6,006,779	36.6%
8-K	68,471,817	13.0%	10-K	19,523,529	17.5%	8-K	1,413,387	8.6%
10-Q	29,551,127	5.6%	10-Q	19,468,598	17.4%	SC 13G/A	591,842	3.6%
10-K	24,935,559	4.8%	4	8,043,938	7.2%	3	552,843	3.4%
13F-HR	13,889,856	2.6%	6-K	2,563,774	2.3%	10-Q	536,219	3.3%
4/A	9,268,341	1.8%	DEF 14A	2,120,153	1.9%	497	375,547	2.3%
8-K/A	7,874,539	1.5%	424B2	2,073,561	1.9%	SC 13G	347,617	2.1%
SC 13G	6,357,204	1.2%	424B5	1,943,211	1.7%	6-K	336,734	2.1%
SC 13G/A	5,706,116	1.1%	S-1/A	1,891,509	1.7%	424B3	258,821	1.6%
6-K	3,984,856	0.8%	424B3	1,862,154	1.7%	SC 13D/A	206,492	1.3%
DEF 14A	2,847,715	0.5%	13F-HR	1,832,043	1.6%	13F-HR	202,615	1.2%

Panel B. 8K-10K Correlation.

	Correlation		
	10K	8K	Insiders
10K	1.00	0.28	0.11
8K	0.28	1.00	0.17
Insiders	0.11	0.17	1.00

Table 3. Top 30 Linked Institutions

This table reports the top 30 active institutions (in terms of total value of the active portfolio last observed) that we are able to link to an IP address.

	Mgrno	Institution Name	Total Value of Active Portfolio
1	12740	CAPITAL RESEARCH & MGMT	$644,502,498,763
2	27800	FIDELITY MGMT & RESEARCH (US)	$623,465,559,994
3	55390	MELLON BANK NA	$479,115,002,653
4	71110	T. ROWE PRICE ASSOCIATES, INC.	$435,541,810,456
5	62890	BANK OF AMERICA CORPORATION	$377,795,325,262
6	91910	WELLINGTON MGMT CO, L.L.P.	$357,154,732,394
7	11836	CAPITAL WORLD INVESTORS	$352,914,042,331
8	25610	AXA FINANCIAL, INC.	$352,084,782,849
9	90457	VANGUARD GROUP, INC.	$351,551,062,880
10	58950	MSDW & COMPANY	$316,603,792,017
11	58835	JPMORGAN CHASE & COMPANY	$316,062,544,205
12	65260	NORTHERN TRUST GLOBAL INVTS	$305,692,675,586
13	11835	CAPITAL RESEARCH GBL INVESTORS	$299,354,230,407
14	72400	PUTNAM INVESTMENT MGMT, LLC	$285,307,501,875
15	41260	GOLDMAN SACHS & COMPANY	$245,836,983,966
16	10586	AMVESCAP PLC LONDON	$237,493,511,896
17	7800	DEUTSCHE BK AKTIENGESELLSCHAFT	$226,994,949,895
18	48170	JANUS CAPITAL MANAGEMENT LLC	$225,037,238,864
19	50160	LEGG MASON INC	$215,319,086,243
20	84900	CITIGROUP INVESTMENTS INC.	$200,525,553,607
21	39300	FRANKLIN RESOURCES, INC.	$196,253,825,506
22	11371	NORGES BK INVT MGMT (NBIM)	$188,412,117,446
23	54600	MASSACHUSETTS FINL SERVICES CO	$179,313,992,099
24	10039	GEODE CAPITAL MGMT, L.L.C.	$169,824,996,700
25	8350	BERKSHIRE HATHAWAY	$153,892,447,818
26	18265	COLLEGE RETIRE EQUITIES	$145,381,192,291
27	45639	COLUMBIA MGMT INV ADVISERS LLC	$145,135,260,505
28	37700	WACHOVIA CORPORATION	$141,242,468,650
29	56780	MERRILL LYNCH CAPITAL MARKETS	$139,563,896,834
30	23270	DODGE & COX, INC.	$135,031,475,226

Table 4. Persistence of Insider Checking by Institutions

In Panel B of this table, we report the persistence of institutions monitoring of insider activities. We regress a dummy (Checked Firm at t+1) that indicates whether the institutions checked the insider filing on its lagged value and a set of fixed effects, such as portfolio, year x stock and year x stock x insider fixed effects. We require the institutions to be matched in consecutive quarters. T-statistics, clustered quarterly, is reported in parenthesis. Panel B reports the unconditional checking probabilities calculated by quarter.

Panel A.

	Checked Firm at t+1			Checked Insider at t+1		
Checked Firm at t	**0.413**	**0.259**	**0.255**			
	(30.92)	**(20.90)**	**(20.75)**			
Checked Insider at t				**0.187**	**0.142**	**0.133**
				(24.72)	**(23.52)**	**(21.48)**
Portfolio Fixed Effect	No	Yes	Yes	No	Yes	Yes
Year x Stock Fixed Effect	No	No	Yes			
Year x Stock x Insider Fixed Effect				No	No	Yes
Adj. R^2	0.19	0.27	0.28	0.04	0.08	0.10
No.	1,338,919	1,338,919	1,338,919	11,190,087	11,190,087	11,190,087

Panel B.

	Checked Firm at t+1	Checked Insider at t+1
Unconditional Checking	4.80%	1.13%

Table 5. Contemporaneous Trading

This table reports the panel regression of the direction of contemporaneous trading by institutions and the direction of trading by the insiders checked by these institutions. The dependent variable, Direction, is -1 if the position is a sell and 1 is a buy. Insider_Direction is -1 if the insiders checked by the institution sold in net and 1 if the insiders bought in net. Insider_Buy is 1 if the insiders checked bought in net, and Insider_Sell is 1 if the insiders checked sold in net. All_Insider_Direction is -1 if the insiders from the firm sold in net during the quarter and 1 if the insiders from the firm bought in net. All_Insider_Buy is 1 if all the insiders from the firm bought in net during the quarter. All_Insider_Sell is 1 if all the insiders from the firm sold in net during the quarter. The panel is weighted by the inverse of the number of positions in each portfolio (so a portfolio with many positions won't dominate the regression). Only non-diversified institutions, that is institutions with less than 500 stock positions, are included in the regression. The t-stats are clustered quarterly.

	Direction (-1 for sell, 1 for buy)					
Insider_Direction	**0.1671**			**0.1413**		
	(3.97)			**(3.27)**		
Insider_Buy		**0.1401**			**0.1230**	
		(2.38)			**(2.02)**	
Insider_Sell			**-0.1753**			**-0.1548**
			(-3.50)			**(-3.77)**
All_Insider_Direction				0.0095		
				(1.76)		
All_Insider_Buy					0.0123	
					(1.82)	
All_Insider_Sell						-0.0010
						(-1.37)
Lag_Direction				**0.1335**	**0.1329**	**0.125**
				(8.41)	**(8.74)**	**(7.93)**
Adj. R^2	0.029%	0.0044%	0.025%	2.070%	2.095%	2.047%
No.	516,288	516,288	516,288	516,288	516,288	516,288

Table 6. Portfolio Returns

This table displays various calendar time portfolios strategies related to the insiders followed through the Edgars Filing System. In terms of timing, we require that institutions both view the insider trading record and trade the underlying stock in quarter t, and analyze the returns to the trades in quarter t+1. Panel A records the timing of the portfolio formation and the subsequent returns. Panel A records the Raw, DGTW, and 4 Factor adjusted returns of portfolios relating to insiders who bought stocks followed through the Edgar system. Panel B records the quarterly Raw, DGTW, and 4 Factor adjusted returns of portfolios relating to insiders who sold stocks followed through the Edgar system. The sample period is from Q1 2004 to Q4 2015.

Panel A. Tracking the insider purchases

Equal Weighted Portfolios Averaged by TNA	Raw Return	DGTW	4 Factors	L % of Assets	S % of Assets
1) All Portfolios	**2.99%**	0.18%	0.28%	100%	
	(2.08)	(1.73)	(1.81)		
2) Unchecked Portfolios Or Checked Insider Selling	**2.98%**	0.18%	0.27%	98.9%	
	(2.08)	(1.66)	(1.81)		
3) Checked Insider Buying vs 2)	1.33%	1.43%	1.33%	1.1%	98.9%
	(1.09)	(1.23)	(1.14)		
4) Checked Insider Buying and Bought vs 2)	**3.36%**	**3.44%**	**3.54%**	0.6%	98.9%
	(2.18)	**(2.32)**	**(2.33)**		
5) Checked Insider Buying and Bought vs. Rest Bought	**3.36%**	**3.41%**	**3.54%**	0.6%	50.9%
	(2.19)	**(2.32)**	**(2.35)**		
6) Checked Insider Buying and Bought vs. Checked and Not Bought	3.31%	3.04%	3.87%	0.6%	0.5%
	(1.50)	(1.46)	(1.74)		
7) Checked Insider Buying and Bought vs. Checked and Not Bought	4.13%	3.87%	**4.77%**	0.2%	0.2%
(Small to Zero Initial Positions)	(1.78)	(1.78)	**(2.04)**		
8) Checked Insider Buying and Bought vs. Not Checked Insider	**3.48%**	**3.68%**	**3.60%**	0.6%	2.8%
	(2.39)	**(2.48)**	**(2.41)**		

Panel B. Tracking the insider sales

Equal Weighted Portfolios Averaged by TNA	Raw Return	DGTW	4 Factors	L % of Assets	S % of Assets
1) All Portfolios (Only if holding exists)	**2.98%**	0.19%	0.28%	100%	
	(2.09)	(1.91)	(1.89)		
2) Unchecked Portfolios Or Checked Insider Buying	**2.98%**	0.19%	0.29%	93.2%	
	(2.09)	(1.93)	(1.92)		
3) Checked Insider Selling vs 2)	**-1.05%**	**-0.93%**	**-1.19%**	6.8%	93.2%
	(-2.61)	**(-2.46)**	**(-3.20)**		
4) Checked Insider Selling and Sold vs 2)	**-2.01%**	**-1.75%**	**-2.08%**	4.4%	93.2%
	(-3.65)	**(-3.23)**	**(-3.77)**		
5) Checked Insider Selling and Sold vs. Rest Sold	**-2.03%**	**-1.75%**	**-2.09%**	4.4%	46.9%
	(-3.65)	**(-3.19)**	**(-3.74)**		
6) Checked Insider Selling and Sold vs. Checked and not Sold	**-1.75%**	-1.26%	**-1.75%**	4.4%	2.4%
	(-2.39)	(-1.75)	**(-2.29)**		
7) Checked Insider Selling and Sold vs. Checked and not Sold	**-1.98%**	-1.44%	**-2.09%**	0.9%	0.8%
(Small Initial Positions)	**(-2.33)**	(-1.73)	**(-2.39)**		
8) Checked Insider Selling and Sold vs. Not Checked Insider	**-1.96%**	**-1.74%**	**-2.02%**	4.4%	44.3%
Selling and Sold	**(-3.62)**	**(-3.18)**	**(-3.65)**		

Table 7. Characteristics of Checked Insider Profiles

This table reports the characteristics of profiles of the insiders being searched in our sample. The first column contains the names of background characteristic associated with the firm insiders. The second column reports percentage of each type of profile over all BoardEx reported insiders. The third column reports the percentage of each type for profiles being checked by institutional investors. The differences between the percentage of the checked and all profiles are calculated in the third column.

	All Profile	Checked Profile	Difference
Accountants	11.25%	15.55%	**4.30%**
			(11.34)
MBAs	34.53%	34.57%	0.04%
			(0.12)
PhDs	7.19%	6.46%	**-0.73%**
			(-4.29)
MDs	2.13%	1.92%	**-0.21%**
			(-2.75)
Directors	70.23%	54.58%	**-15.65%**
			(-34.88)
CEOs	9.83%	28.71%	**18.88%**
			(41.99)
Presidents	19.91%	33.69%	**13.78%**
			(35.02)
Chairman	9.49%	27.28%	**17.80%**
			(35.21)
CFO	9.95%	16.18%	**6.24%**
			(24.31)
Harvard	9.92%	11.65%	**1.73%**
			(7.73)
Princeton	1.53%	1.91%	**0.38%**
			(2.96)
Yale	1.80%	2.01%	0.21%
			(1.71)

Table 8. Characteristics of mutual fund managers whose trades correlate with those of insiders

This table reports the distribution of fund managers within the identified 13-F institutions. Cherry Pickers are fund managers whose trades each quarter correlate with net insider trading at 90% significance level. Panel A contains the distribution of top locations of all matched mutual fund managers and specifically the cherry pickers. Panel B records the distribution of education backgrounds of matched mutual fund managers and specifically the cherry pickers. The universities with the highest absolute differences between manager and cherry picker distributions are reported.

Panel A. Distribution of top locations of all matched mutual fund managers and the cherry pickers

State	Distribution of Managers	Distribution of Cherry Pickers	Diff
AZ	1.0%	2.0%	1.0%
CA	4.4%	8.0%	3.6%
CT	0.3%	1.1%	0.8%
FL	0.9%	1.5%	0.6%
IL	1.5%	1.5%	0.0%
KS	1.5%	4.3%	2.8%
MA	20.3%	29.1%	8.8%
MD	6.8%	9.7%	2.9%
MO	1.1%	2.2%	1.1%
NC	1.8%	2.8%	1.0%
NJ	1.6%	1.3%	-0.3%
NY	12.8%	8.3%	-4.5%
OH	4.2%	1.8%	-2.3%
PA	5.2%	2.7%	-2.5%
TX	1.6%	1.5%	-0.2%
WA	1.2%	1.2%	-0.1%
WI	1.0%	1.5%	0.5%

Panel B. Distribution of education backgrounds of matched mutual fund managers and the cherry pickers

School	Distribution of Managers	Distribution of Cherry Pickers	Diff
University of Wisconsin	2.1%	3.9%	1.8%
Massachusetts Institute of Technology	1.6%	3.3%	1.7%
Wellesley College	0.7%	2.2%	1.5%
Colgate University	0.7%	2.0%	1.2%
Rensselaer Polytechnic Institute	0.4%	1.5%	1.2%
Emory University	0.7%	1.8%	1.1%
Harvard University	6.0%	7.0%	1.0%
Harvard Business School	1.7%	2.6%	0.9%
London University	0.4%	1.3%	0.9%
Georgetown University	1.1%	0.9%	-0.2%
Northeastern University	1.2%	0.8%	-0.4%
New York University	3.8%	3.4%	-0.4%
Princeton University	2.2%	1.7%	-0.5%
University of Michigan	1.9%	1.0%	-0.9%
University of Chicago	5.3%	4.0%	-1.3%

Table 9. Checked Profiles

This table explores potential connection links between a Cherry Picking mutual fund manager and commonalities in location and education backgrounds. Cherry Pickers are fund managers whose trades each quarter correlate with net insider trading at 90% significance level. The Match Indicator is a dummy variable for whether the Cherry Picker had searched for the insider at a public firm on commonalities in education and location. The sample includes all potential insiders linked by each fund manager's portfolio holdings. The t-stats are clustered at the quarterly level.

	Match Indicator		
Education Link	**0.17%**	**0.16%**	**0.14%**
	(3.56)	**(3.87)**	**(3.39)**
Location Link	**0.43%**	**0.44%**	**0.39%**
	(4.45)	**(4.45)**	**(3.98)**
Lag Match Indicator			**14.1%**
			(8.93)
Quarterly Fixed Effect	No	Yes	Yes
R^2	0.02%	0.42%	2.15%
N	3,966,193	3,966,193	3,966,193

Figure 1. This figure illustrates the proportion of gross purchases of stock by institutions over a number of trading days after checking that insiders had purchased a stock in net.

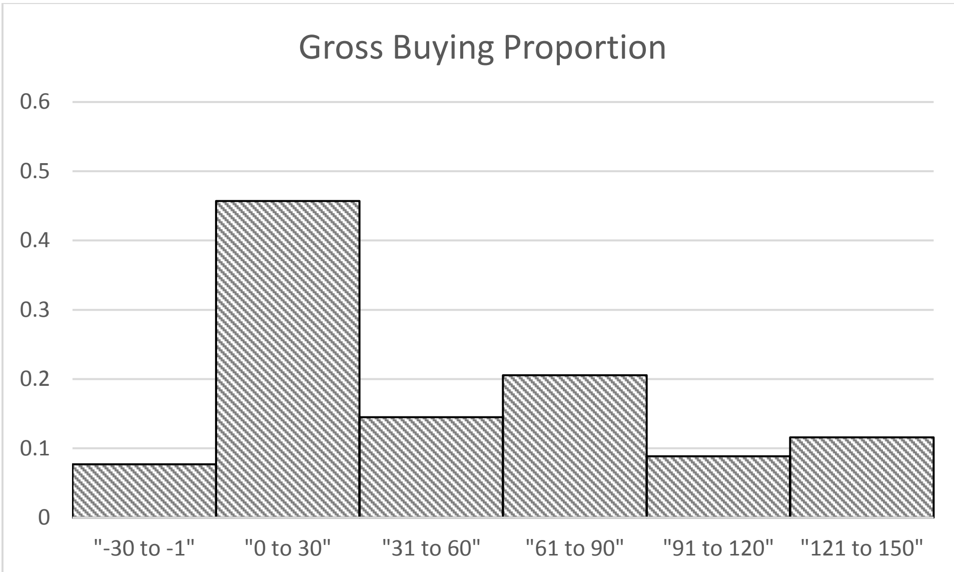


Figure 2. This figure illustrates the proportion of gross selling of a stock by institutions over a number of trading days after checking insiders had sold a stock in net.

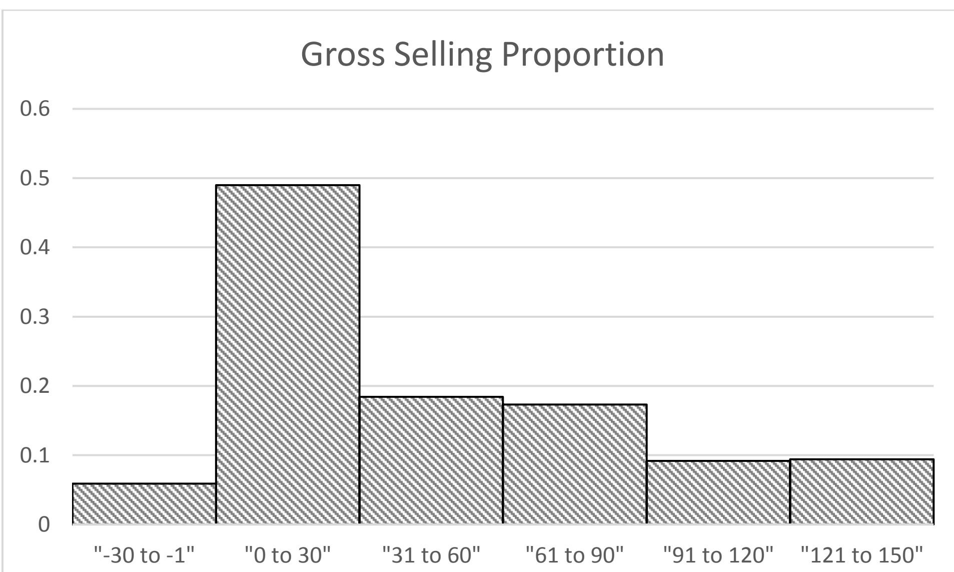
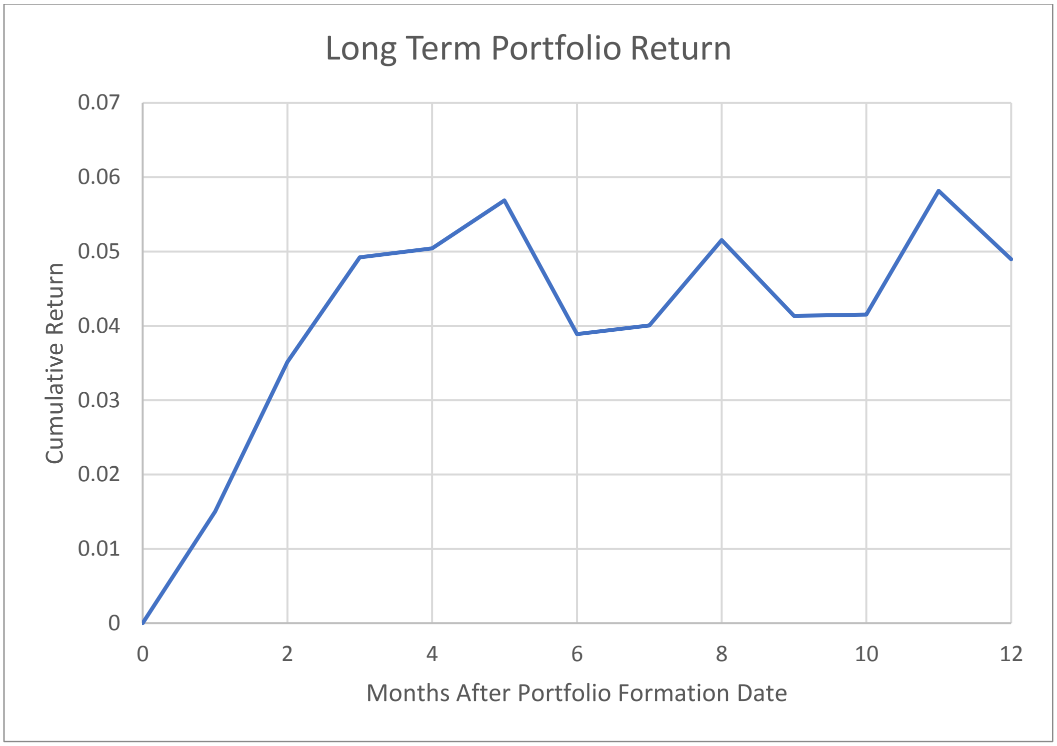


Figure 3. This figure illustrates the long horizon cumulative returns of a combined strategy that trades the buyside portfolio and goes the opposite of the sellside portfolio (row 5).



Appendix Table A1.

This table displays various calendar time portfolios strategies related to the insiders followed through the Edgars Filing System. Panel A records the Raw, DGTW, and 4 Factor adjusted returns of portfolios relating to insiders who bought stocks followed through the Edgar system. Panel B records the quarterly Raw, DGTW, and 4 Factor adjusted returns of portfolios relating to insiders who sold stocks followed through the Edgar system. The sample period is from Q1 2004 to Q4 2015.

Panel A.

Value Weighted Portfolios Averaged by TNA	Raw Return	DGTW	4 Factors	L % of Assets	S % of Assets
1) All Portfolios	**2.61%**	0.15%	0.17%	100%	
	(2.27)	(1.18)	(1.70)		
2) Unchecked Portfolios Or Checked Insider Selling	**2.61%**	0.14%	0.17%	98.9%	
	(2.28)	(1.14)	(1.72)		
3) Checked Insider Buying vs 2)	0.23%	0.54%	-0.37%	1.1%	98.9%
	(0.17)	(0.44)	(-0.29)		
4) Checked Insider Buying and Bought vs 2)	**3.72%**	**3.65%**	3.04%	0.6%	98.9%
	(2.11)	**(2.33)**	(1.79)		
5) Checked Insider Buying and Bought vs. Rest Bought	**3.85%**	**3.76%**	3.22%	0.6%	50.9%
	(2.21)	**(2.40)**	(1.90)		
6) Checked Insider Buying and Bought vs. Checked and Not Bought	3.23%	2.86%	3.20%	0.6%	0.5%
	(1.27)	(1.25)	(1.23)		
7) Checked Insider Buying and Bought vs. Checked and Not Bought	**5.26%**	**5.16%**	**5.60%**	0.2%	0.2%
(Small to Zero Initial Positions)	**(2.09)**	**(2.25)**	**(2.23)**		
8) Checked Insider Buying and Bought vs. Not Checked Insider	**4.60%**	**4.57%**	**4.10%**	0.6%	2.8%
	(2.76)	**(2.87)**	**(2.45)**		

Panel B.

Value Weighted Portfolios Averaged by TNA	Raw Return	DGTW	4 Factors	L % of Assets	S % of Assets
1) All Portfolios (Only if holding exists)	**2.62%**	0.13%	**0.19%**	100%	
	(2.28)	(1.16)	**(2.05)**		
2) Unchecked Portfolios Or Checked Insider Buying	**2.63%**	0.13%	**0.19%**	93.2%	
	(2.28)	(1.18)	**(2.12)**		
3) Checked Insider Selling vs 2)	-0.17%	-0.10%	-0.50%	6.8%	93.2%
	(-0.31)	(-0.23)	(-0.94)		
4) Checked Insider Selling and Sold vs 2)	2.66%	0.16%	**-0.74**	4.4%	93.2%
	(-1.04)	(-0.66)	(-1.41)		
5) Checked Insider Selling and Sold vs. Rest Sold	-0.63%	-0.35%	-0.81%	4.4%	46.9%
	(-1.20)	(-0.77)	(-1.52)		
6) Checked Insider Selling and Sold vs. Checked and not Sold	-0.66%	0.06%	**-0.68%**	4.4%	2.4%
	(-0.87)	(0.09)	(-0.88)		
7) Checked Insider Selling and Sold vs. Checked and not Sold	**-1.98%**	-1.06%	**-2.19%**	0.9%	0.8%
(Small Initial Positions)	**(-1.88)**	(-1.05)	**(-2.04)**		
8) Checked Insider Selling and Sold vs. Not Checked Insider	**-0.65%**	**-0.40%**	**-0.84%**	4.4%	44.3%
Selling and Sold	**(-1.24)**	**(-0.87)**	**(-1.57)**		