CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

Centre for Econometric Analysis

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Informed Trading, Liquidity Provision, and Stock Selection by Mutual Funds

Zhi Da, Pengjie Gao and Ravi Jagannathan

CEA@Cass Working Paper Series

WP–CEA–02-2009

Comments Welcome

Informed Trading, Liquidity Provision, and Stock Selection by Mutual Funds*

Zhi Da, Pengjie Gao and Ravi Jagannathan†

First Draft: May 10, 2007
This Draft: September 11, 2009

Abstract

We show that a mutual fund's stock selection skill can be decomposed into additional components that include impatient "informed trading" and "liquidity provision." We validate our method by verifying that liquidity providing trades are the primary source of value for the Dimensional Fund Advisors U.S. Micro Cap fund, and index funds lose on liquidity absorbing trades, as documented in the literature. We find that past performance predicts future performance better among funds trading in stocks affected more by information events: past winners earn a risk adjusted excess return (four factor alpha) of 0.35% per month in the future. Most of that superior performance comes from impatient informed trading. We also find that informed trading is more important for growth-oriented funds while liquidity provision is more important for younger funds with income orientation.

*We thank Robert Battalio, Jonathan Berk, Roger Edelen, Craig Holden, Paul Irvine, Frank de Jong, Tim Loughran, Rick Mendenhall, David Musto, Ľuboš Pástor, Christine Parlour, Paul Schultz, Clemens Sialm, Matthew Spiegel, Laura Starks, Sheridan Titman, Charles Trzcinka, Lance Young, an anonymous referee, and seminar participants at Northwestern University, IU/ND/Purdue Finance Symposium, University of Illinois at Urbana-Champaign, University of Michigan, Barclays Global Investors, Tilburg University, the 4th Vienna Symposium on Asset Management, FRA 2007 Annual Meeting, NBER Asset Pricing Meeting, the Western Finance Association 2008 Annual Meeting and AFA 2009 Annual Meeting for comments. This paper has been previously circulated under the title "When does a mutual fund's trade reveal its skill?" Patricia Andersen provides excellent editorial support.

†Correspondence: Zhi Da, Finance Department, Mendoza College of Business, University of Notre Dame; Pengjie Gao, Finance Department, Mendoza College of Business, University of Notre Dame; and Ravi Jagannathan, Finance Department, Kellogg School of Management, Northwestern University and NBER. E-mail: zda@nd.edu, pgao@nd.edu and rjaganna@kellogg.northwestern.edu.

As of 2007, US mutual fund managers collectively had over $12 trillion under their management with more than half of it in stocks. A significant portion of this amount is actively managed, as indicated by a turnover rate in excess of 50% for stock funds.[1] During 1980 to 2006, investors paid over 0.67 percent of portfolio value per year to the active portfolio managers (French, 2008). Naturally, investors would like to understand how active fund managers add sufficient value to justify their higher fees and trading costs relative to passively managed index funds. For that purpose, the common practice is to attribute the performance of a portfolio manager to two sources: security selection and asset allocation (also known as market timing). Knowing what securities a manager held makes attributing the performance of a manager to these two components easier, and the approach has become standard industry practice.

In this paper, we show that the security selection component of the performance can be further decomposed into performance arising from (a) recent liquidity absorbing impatient informed trading, (b) recent liquidity providing trades, (c) positions in securities taken earlier, and (d) an adjustment term for inflows and outflows. We believe that such a decomposition will facilitate individual and institutional investors understand the strengths of an active portfolio manager and the extent to which such strengths will continue to be of value in the future.

Ultimately, an active mutual fund manager's skill comes from superior ability to process valuation-relevant information on a stock that helps correctly identify potential mispricing. How a manager with superior skill trades to add value will depend on how long it takes for the market to realize that the manager is right. Based on how long the informational advantage lasts, a manager's trades can be classified into the following three types.

First, the manager can add value from long-term "value investing" by taking a position in a stock expecting the market to eventually agree with her view in, say, a few years.[2] Second, the manager can add value from medium-term informed trading by transacting in "mispriced" stocks expecting the market to agree with her view within, say, a quarter.[3] Third, the manager can

[1] These numbers are taken from Figures 2.1 and 2.9 the Investment Company Fact Book (2008) published by the Investment Company Institute.

[2] Using fundamental analysis, Mario Gabelli, a money manager, realized that the stock of Hudson General Corp (HGC) was heavily undervalued at around $25 in early 1994 and started to accumulate shares of HGC for his Gabelli Funds (see Figure 1A). The investment paid off after two years, when the stock price reached $40. The market eventually agreed with Mr. Gabelli, after Lufthansa took over HGC at $76 per share. See Greenwald, Kahn, Sonkin and Biema (2001) for details on this case.

[3] The year-to-year same store sales growth reported by Starbucks every month is a widely watched number, and is considered about as important as the company's quarterly earnings announcements for valuation purposes. For

add value from short-term liquidity provision by taking the other side of a trade when liquidity is most needed.[4] Since fund managers often hold an inventory of stocks in order to track their performance benchmarks, they have a natural advantage in making a market in those stocks. Moreover, the superior knowledge about the stocks covered by a manager will help in the market making activities by minimizing potential losses that may arise by trading with those having an information advantage.[5]

In the first case, the exact timing of trades would not be critical. Evaluating the stock selection skill of such a portfolio manager who makes a few concentrated long term bets will be difficult based only on quarterly observations on what the manager holds. Therefore, in this paper our focus is on identifying how much of the value added by a manager comes from the latter two activities.

While in theory knowledge of what the manager holds should help evaluate a portfolio manager's skill better, the fact that mutual fund stock holdings data are available only at infrequent intervals (quarterly in most cases) makes it difficult to assess a manager's abilities when the manager trades actively in between two holdings reporting points in time. As Kacperczyk, Sialm and Zheng (2007) and Elton, Gruber, Krasny and Ozelge (2006) show, "unobservable" actions (trades that cannot be inferred from quarterly holdings) by mutual funds could be important for some funds. Because of that, we are only able to capture the partial effect of a liquidity shock that persists over a calendar quarter-end.

In spite of this limitation, a mutual fund's recent trades inferred from its quarterly holding changes, still contain valuable information about a manager's abilities, especially in the medium-

January to September 2005, Starbucks' reported sales growth rates were in the range of 7% to 9%. Most analysts were of the view that a large part of that growth rate was attributable to the 3% sales price increase that took effect in October 2004, and that this price increase would not help with respect to same-month year-to-year sales growth rates beginning with October 2005. That probably explains the much smaller anticipated growth rate (analyst consensus was 3.6%). However, a careful analysis of sales breakdown would have indicated that the 3% price increase in October 2004 explained little of the sales growth during January-September 2005. So, the October sales growth figure should be more like that for the early months of 2005. While most mutual funds decreased their holdings of Starbucks stock during Q3 2005 in anticipation of an announcement of a drop in same-store sales growth for October, Putnam Voyager Fund actually accumulated more shares (see Figure 1B). On November 3, 2005 Starbucks reported unexpectedly strong sales growth of 7% for October, and its share price jumped. Details on this case can be found in Blumenthal (2007).

[4] It is well known that when index funds trade following index rebalancing, their trades tend to demand liquidity from the market (see Blume and Edelen, 2004). Active fund managers taking the other side of those trades will benefit from liquidity provision.

[5] Sometimes managers may not be directly motivated by the "liquidity provision" objective. For example, consider a mutual fund with a policy of not investing more than a certain percentage of its assets in any one stock. The fund may decrease its holdings of a stock that experiences a recent sharp price increase in order to satisfy its portfolio weighting constraints. Such trades are likely to provide liquidity and will therefore be classified as "liquidity provision" even when liquidity provision was not the motivation behind the trade.

term informed trading, if we can separate the medium-term informed trading from the short-term liquidity provision in order to reduce the noise in the data. This becomes possible when we recognize that informed trading tends to be liquidity-absorbing on average because information loses value over time. Based on that insight, we develop a method for decomposing the stock selection skill of a manager (the characteristic selectivity measure as in Daniel, Grinblatt, Titman and Wermers 1997) into a liquidity absorbing informed trading component, a liquidity providing component, and other components using a mutual fund's holdings.

We examine several empirical properties of the decomposition that lend support for its validity. First, we show that mutual funds are likely to provide liquidity on average only when they reduce their holdings, consistent with our conjecture that it is easier to provide liquidity on stocks that one currently owns. Second, we verify that the decomposition results for the two cases, (a) Dimensional Fund Advisors (DFA) and (b) a group of index funds, are consistent with what one would expect based on the findings reported in the literature.[6] Third, as expected, we find that the informed trading component is more important than the liquidity provision component in explaining cross-sectional variation in the CS measures; and informed trading becomes relatively more important for growth-oriented funds while liquidity provision becomes relatively more important for income-oriented funds. Fourth, we find that funds with higher "return gaps" – defined in Kacperczyk, Sialm and Zheng (2007) to capture the benefit of "unobservable" actions of mutual funds – add value through liquidity provision.

Having demonstrated the effectiveness of our decomposition method, we then apply it to analyze the performance of a large sample of active US equity mutual funds. To analyze the different channels through which a fund manager can add value, one first needs to identify skillful fund managers. Ultimately, an active mutual fund manager's success derives from his or her superior skill in processing valuation-relevant information about a stock, a skill that should allow the identification of potential mispricing. Thus, it is reasonable to expect such skills to be more valuable when stocks the manager can invest in are affected by more value-relevant information events. To the extent that rational managers have the option not to trade such stocks when they know that they do not

[6] Keim (1999) finding that the small-cap equities "9-10 fund" of Dimensional Fund Advisors (DFA) outperformed its benchmark by about 2.2% during the period between 1982 to 1995, illustrates how skillful trade execution can enhance fund performance. Cohen (2002) documents that managers at DFA add value by systematically providing liquidity to those who want to trade small cap stocks for non-information based reasons. We verify that most of the value added by DFA through stock selection indeed comes from the liquidity provision component (CS^{liq}).

have an advantage in analyzing the information affecting a stock, we should expect to find that managers who choose to trade earn higher returns on average. To measure the frequency and intensity of information events, we focus on a market microstructure based measure: the Probability of Informed Trading (PIN) proposed by Easley, Kiefer, O'Hara and Paperman (1996) although we obtain very similar results using several alternative measures of information events. We compute a $trade\_PIN$ variable by value-weighting the PIN of stocks traded by the fund during the quarter using the dollar value of the trade. Intuitively, funds that buy or sell more high-PIN stocks during a quarter should have higher $trade\_PIN$ measures in that quarter.

We find that funds trading high-PIN stocks outperform those trading low-PIN stocks by 53 bps per quarter *before fees* (t-value = 2.87) after controlling for stock characteristics such as size, book-to-market ratio and return momentum. However, the after fee alphas are either negative or not significantly different from zero, consistent with the findings in Fama and French (2008). Easley, Hvidkjaer and O'Hara (2002) document that High-PIN stocks earn higher returns on average. They interpret this as being compensation for risk associated with private information, i.e., PIN-risk. That does not drive our results. Both stocks that mutual funds buy and sell have about the same PIN values, but stocks bought by mutual funds tend to outperform those sold by mutual funds. In addition, after controlling for PIN risk directly, we obtain very similar results. Further, we show that our findings are not driven by momentum trading rules described in the literature. Interestingly, a large fraction of the superior stock selection skill of managers trading high-PIN stocks comes from impatient informed trading. In contrast, liquidity provision appears more important for funds trading in low-PIN stocks where there is little adverse selection risk as studied by Glosten and Harris (1998).

The early literature on portfolio performance evaluation finds that most managed portfolios earn close to zero or negative risk-adjusted returns especially after taking fees into account[7]. More recent studies find that some funds do add value and it is possible to identify those funds based on past performance and the stocks they hold.[8] Interestingly Mamaysky, Spiegel and Zhang (2007b) show that superior performance can be identified based on past performance alone by imposing

[7] See Jensen (1968), Elton, Gruber, Das, and Hlavka (1993), Brown and Goetzmann (1995), Gruber (1996) and Carhart (1997).

[8] See Grinblatt and Titman (1989, 1993), Wermers (1997), Daniel, Grinblatt, Titman and Wermers (1997), Chen, Jegadeesh and Wermers (2000), Schultz (2007)

restrictions implied by economic reasoning. However, during the more recent sampling period from 1983-2004, both Kosowski, Timmermann, Wermers and White (2006) and Fama and French (2008) document that historical alpha alone does not reliably pick up funds with positive (after-fee) alphas going forward.

We conjecture that past superior performance is more likely to be an indication of future performance for a manager who attained such performance by trading stocks associated with more information events, i.e., we should expect stronger fund performance persistence among funds that traded in High-PIN stocks recently. Consistent with that conjecture, when we combine our $trade\_PIN$ variable, and the filters proposed in Mamaysky, Spiegel and Zhang (2007b), we are able to construct a portfolio of mutual funds that significantly outperforms the market going forward at monthly frequency even during the more recent sampling period. In particular, the past winners identified using the methods in Mamaysky, Spiegel and Zhang (2007a, 2007b) methods among funds trading high-PIN stocks outperform the market after fees with a statistically significant alpha of 35 bps per month (t-value = 3.33.) Further, most of that outperformance coming from impatient informed trading. In contrast, past winners among all funds in our sample earn only 19 bps per month, not significantly different from zero.

It is well recognized that several mutual fund characteristics are related to superior stock-selection skills. For example, funds that follow "aggressive growth" and "growth" styles (Daniel, Grinblatt, Titman and Wermers, 1997); hold stocks of firms whose headquarters are located geographically closer to the fund's headquareters (Coval and Moskowitz, 2001); have more industry concentration in their holdings (Kacperczyk, Sialm and Zheng, 2004); have less diversification in their holdings (Baks, Busse and Green, 2006); have larger deviations from passive index or larger "active shares" (Cremers and Petajisto, 2006); and less dependency on analyst's recommendation (Kacperczyk and Seru, 2007) tend to perform better. In addition funds that are smaller in size (Chen, Hong, Huang and Kubik, 2004) perform better after controlling for mutual fund family size. We contribute to this literature by showing that the informed trading component of stock selection skill is more important for growth-oriented funds whereas the liquidity provision component is more important for income-oriented funds.

Throughout the paper, we infer the trades of mutual funds by comparing their quarter-end holdings over consecutive quarters. The procedure adds noise to several of our empirical exercises.

First, since we ignore interim trading and other unobservable actions of mutual funds within a quarter, our $trade\_PIN$ measure is computed with noise. Second, inferring mutual fund trades using their quarterly holdings also adds noise when classifying trades as "informed-trading" or "liquidity-provisiding." We therefore examine the extent to which we lose information by relying on quarterly holdings with the help of data on mutual fund trades (for a subset of funds) from the Plexus Group. Since we observe the actual transactions of fund managers in the Plexus database, we can pin down the error that arises from inferring a mutual fund's trades by observing only its quarterly holdings. We find that a fund's $trade\_PIN$ calculated by observing only its quarterly is in fact very accurate, with an average absolute estimation error of 1% of the size of a fund's $trade\_PIN$. However, our classification of trades as "informed trading" or "liquidity provision" is less precise. We find that on average, only 66% of a fund's trades inferred from its quarterly holdings data would be correctly classified as "informed-trading" or "liquidity-provision." This percentage of correct assignment is statistically significantly higher than 50%, confirming that trade classification based on quarterly data contains useful information. Not surprisingly, the accuracy of the assignment is higher among funds that conduct little interim trading. For these funds, 70% of their trades are correctly assigned to be either "informed-trading" or "liquidity-provision." Interestingly, such funds tend to trade stocks with higher PINs on average, implying that our results involving high $trade\_PIN$ funds are less likely to be affected by noise in the classification process.

The remainder of this paper is organized as follows. We develop the decomposition method in section 1. We describe the data sources and the sample construction procedure in section 2. In section 3, we evaluate the validity of the decomposition method by examining Dimensional Funds US Micro Cap Portfolio and a group of index funds, and relate the decomposition procedure to the "return gap" concept developed by Kacperczyk, Sialm and Zheng (2007). In section 4, we then empirically examine the channels through which funds that trade during information events add value, thereby illustrating the use of our decomposition. In section 5, we analyze the noise associated with infering mutual trades using quarterly holding data using actual fund transactions in the Plexus Group data. We conclude in section 6. The appendices contain a numerical example on our decomposition, a brief discussion on the variance decomposition approach, and a short note on various measures of private information events.

1 Decomposing Mutual Fund's Stock Selection Skills

In order to separate the value-added through security selection by a mutual fund manager into different components, we start with the characteristics-based performance measure — characteristic selectivity (CS) — developed by Daniel, Grinblatt, Titman and Wermers (DGTW, 1997) and Wermers (2004). The CS measure of a mutual fund during quarter $t+1$, based on its actual stock holdings at the end of quarter t, can be computed as:

$$\begin{aligned} CS_{t+1} &= \sum_j w_{j,t} \left[R_{j,t+1} - BR_{t+1}(j,t)\right] \\ &= RH_{t+1} - BR_{t+1}. \end{aligned}$$

where $R_{j,t+1}$ is the return on stock j during quarter $t+1$; $w_{j,t}$ is the dollar value weight of stock j held by the mutual fund at the end of quarter t; and $BR_{t+1}(j,t)$ is the benchmark portfolio return during quarter $t+1$ to which stock j is matched at the end of quarter t based on its size, book-to-market equity ratio, and past 12-month return. In addition RH_{t+1} denotes the implied return of the fund holdings during quarter $t+1$ based on the fund holding at the end of quarter t and BR_{t+1} is the return on the benchmark portfolio with matching stock characteristics. Intuitively, the CS measure detects whether managers are able to select stocks that outperform average stocks with similar characteristics.

We can further decompose the CS measure. A numerical illustration of such decomposition is provided in Appendix A. Suppose mutual funds rebalance only at discrete points in time, $t = 1, 2, 3, . ., T$. For convenience, we assume time periods are measured in quarters. Let N_t be a column vector of mutual fund stock holdings (in number of shares, split-adjusted) at the end of quarter t. By comparing N_{t-1} and N_t, we can define three stock portfolios:

1. Hold portfolio, which has stock holdings:

$$N_t^H = \min\left(N_{t-1}, N_t\right)$$

where the operator min() calculates the element-by-element minimum; and N_t^H captures holdings that appear in both quarters.

2. Buy portfolio, which has stock holdings:

$$N_t^B = N_t - N_t^H$$

The Buy portfolio holds stocks bought by the fund during quarter t.

3. Sell portfolio, which has stock holdings:

$$N_t^S = N_{t-1} - N_t^H$$

The Sell portfolio contains stocks sold by the fund during quarter t.

Over time, the mutual fund stock holdings change as follows:

$$N_t = N_{t-1} - N_t^S + N_t^B$$

Figure 2 contains a graphical representation of these three portfolios. Let the smaller pie represent the fund's holdings at the end of quarter $t-1$ and the bigger pie represent the fund's holdings at the end of quarter t. The intersection of the two pies thus represents the Hold portfolio which contains stocks untouched by the fund during quarter t. The portion of the smaller pie excluding the Hold portfolio represents the Sell portfolio which contains stocks sold by the fund during quarter t. The portion of the bigger pie excluding the Hold portfolio represents the Buy portfolio which contains stocks recently bought by the fund during quarter t.

Let P_t be a column vector of corresponding stock prices at the end of quarter t. Let us denote the market value of Hold, Buy and Sell portfolios as H_t, B_t and S_t, respectively. Accordingly, we have:

$$\begin{aligned} H_t &= P_t' N_t^H \\ B_t &= P_t' N_t^B \\ S_t &= P_t' N_t^S \end{aligned}$$

At the end of quarter t, the mutual fund's stock holdings are a combination of the Hold portfolio and the Buy portfolio (or the bigger pie in Figure 2). The fund CS measure for quarter $t+1$ is therefore the value-weighted average of CS measures on the Hold portfolio and Buy portfolio for quarter $t+1$:

$$CS_{t+1} = \frac{H_t}{H_t + B_t} CS_{H,t+1} + \frac{B_t}{H_t + B_t} CS_{B,t+1},$$

where $CS_{H,t+1}$ and $CS_{B,t+1}$ denote CS measure on Hold and Buy portfolios for quarter $t+1$.

We then decompose the CS measure into three components:

$$\begin{aligned} CS_{t+1} &= CS^{O}_{t+1} + CS^{T}_{t+1} + CS^{adj}_{t+1} \\ CS^{O}_{t+1} &= \frac{H_t}{H_t + S_t} CS_{H,t+1} + \frac{S_t}{H_t + S_t} CS_{S,t+1} \\ CS^{T}_{t+1} &= \frac{B_t}{H_t + B_t} CS_{B,t+1} - \frac{S_t}{H_t + S_t} CS_{S,t+1} \\ CS^{adj}_{t+1} &= \frac{H_t}{H_t + B_t} \frac{S_t - B_t}{H_t + S_t} CS_{H,t+1} \end{aligned} \tag{1}$$

The first component, the *old* component (CS^{O}_{t+1}), can be interpreted as the CS measure of the fund as if the fund did not balance its portfolio at all during quarter t. If nothing happens to the fund during quarter t, its stock holdings would remain unchanged ($N_t = N_{t-1}$), and thus would be composed of stocks in the Hold portfolio and Sell portfolio (or the smaller pie in Figure 2). Consequently, the CS measure for quarter $t+1$ would be the value-weighted average of CS measures on the Hold portfolio and Sell portfolios. Intuitively, this captures the value-added to the fund during quarter $t+1$ from fund investments prior to quarter t, and likely corresponds to the benefit from long-term investment.

The second component, the *trade* component (CS^{T}_{t+1}), measures the characteristics-adjusted returns on the most recent mutual fund stock trades during quarter t. Finally, the *adjustment* component (CS^{adj}_{t+1}) represents a small adjustment term whenever $S_t \neq B_t$, which could happen whenever there is inflow or outflow to the fund.

The trade component (CS^{T}_{t+1}) measures value-added from both medium-term informed trading and short-term liquidity provision. Since mutual fund holdings are typically reported at quarterly frequency at most, in order to make a reasonable attempt to separate them, we rely on a key difference between these two types of trades. Informed trading, unlike the liquidity provision trade, is likely to demand liquidity since the value of information erodes quickly over time, so timely execution becomes important. Given this intuition, we can further decompose the trade component CS^T_{t+1} into two components by comparing the sign of quarterly mutual fund holding change and the sign of market order imbalance for each stock traded by the fund (the stocks in the buy or sell portfolios) during quarter t. The stock-level market order imbalance is defined as the total number of buyer-initiated trades minus the total number of seller-initiated trades in the quarter for the individual stock. Following the standard practice in the literature, we implement the trade classification using the algorithm in Lee and Ready (1991). We then classify stock trades where the two signs are identical into one group, denoted by superscript "+"; and where the two signs are different into another group, denoted by superscript "−". As a result, the characteristics-adjusted returns on trades from these groups sum up to CS^T_{t+1}:

$$\begin{aligned} CS^T_{t+1} &= CS^{inf}_{t+1} + CS^{liq}_{t+1} \\ CS^{inf}_{t+1} &= \frac{B^+_t}{H_t + B_t} CS^+_{B,t+1} - \frac{S^+_t}{H_t + S_t} CS^+_{S,t+1} \\ CS^{liq}_{t+1} &= \frac{B^-_t}{H_t + B_t} CS^-_{B,t+1} - \frac{S^-_t}{H_t + S_t} CS^-_{S,t+1} \end{aligned} \tag{2}$$

Given that the aggregate market order imbalance is a good measure of the direction of liquidity needs of a stock, CS^{inf}_{t+1} measures the characteristics-adjusted return on mutual fund trades that on average absorb market liquidity (see Chordia and Subrahmanyam, 2004). Such trades are likely driven by information and are therefore classified as "informed trading." CS^{liq}_{t+1}, on the other hand, measures the characteristics-adjusted return on mutual fund trades that on average supply market liquidity, and hence are classified as "liquidity provision". In the extreme case where the fund manager trades only one stock and when the time interval is one minute rather than one quarter, CS^{liq}_{t+1} will closely resemble the realized spread of Huang and Stoll (1996), which measures the reward to market makers' liquidity provision activities. To summarize, we decompose the fund CS measure as:

$$\begin{aligned} CS_{t+1} &= CS^O_{t+1} + CS^{adj}_{t+1} + CS^T_{t+1}, \\ CS^T_{t+1} &= CS^{inf}_{t+1} + CS^{liq}_{t+1}. \end{aligned} \tag{3}$$

2 Data and Sample Construction

We employ data from several sources. The mutual fund holding data come from the CDA/Spectrum $S12$ mutual fund holding database, which collects the holding information from the N30-D filings to the Security and Exchange Commission (SEC). A detailed description of the database can be found in Wermers (1999). We exclude index funds and lifecycle funds as the latters are hybrid funds.[9] Following the standard practice in the mutual fund literature, we also omit international funds, sector funds, bond funds, and domestic hybrid funds based on the self-reported fund style in the CDA/Spectrum database. Thus, we only keep funds that are self-reported as Aggressive Growth (AGG), Growth or Growth and Income (GNI). To ensure that the funds we examine are reasonably active, we only include fund / quarter observations if the fund trades at least 10 stocks and turns over at least 10% of its holdings during that quarter. Finally, we only include fund / quarter observations for which the fund holdings *at the end of previous quarter* are also available so holding changes can be computed over consecutive quarters. We obtain the information on the after-fee performance of the fund and other fund characteristics from the Center for Research in Security Prices (CRSP) survivor-bias-free mutual fund database.

The CDA/Spectrum mutual fund holding data are matched to CRSP Mutual Fund data using the MFLINKS database produced by the Wharton Research Data Service (WRDS). An appealing feature of MFLINKS database is that it allows us to map different share classes of the same fund, that are recorded as distinct funds in the CRSP Mutual Fund database, to the corresponding mutual fund holdings data in the CDA/Spectrum database. For multiple share classes in CRSP that correspond to the same fund in the CDA/Spectrum database, we aggregate those share classes into one large portfolio.

The stock data come from CRSP. We include all common stocks (CRSP share codes 10 and 11) traded on the NYSE, AMEX, and NASDAQ. The accounting information comes from COMPUSTAT database. To link COMPUSTAT and CRSP, we use CRSP-LINK produced by CRSP. The tick-by-tick stock transaction data come from ISSM (1983 to 1992) and TAQ (1993 to 2004) databases.

Overall, there are $4,654$ distinct funds in our sample during the period from 1983 to 2004. On

[9] Specifically, we exclude a fund if its name contains any of the following: "INDEX", "INDE", "INDX", "S&P", "DOW JONES", "MSCI" or "ISHARE".

average, there are about 701 distinct funds every quarter. The number of funds per quarter increases from about 134 in 1983 to about $1,700$ towards the end of the sample as shown in Table 1. About 61% of the funds in our sample are self-reported as "Growth" funds, about 26% are reported as "Growth and Income (GNI)" and the remaining 13% are reported as "Aggressive Growth (AGG)".

We collect two groups of fund-level characteristics every quarter. First, we obtain common fund characteristics from CRSP mutual fund database. These characteristics include: *age* (the age of the fund in months since inception, in terms of percentile rank in the cross-section);[10] *turnover* (the turnover rate of the fund); *expense* (the expense ratio of the fund); TNA (the total net assets under management by the fund in millions US$); and pct_flow (the net fund flows in percentage defined as $\frac{TNA(t)-TNA(t-1)*(1+Ret(t-1,t))}{TNA(t-1)}$). Second, we aggregate stock characteristics at fund level by value-weighing them for stocks held by the fund using the quarter-end dollar values of the holdings. These characteristics include: $fund\_holding$ (average percentage of total number of shares outstanding of stocks held by the fund); $fund\_size$ (average market capitalization of stocks held by the fund, in billion dollars); $fund\_bm$ (average book-to-market ratio of stocks held by the fund), $fund\_mom$ (average past one-year return on stocks held by the fund) and $fund\_amihud$ (average Amihud (2002) illiquidity measure, in terms of percentile rank in the cross-section, of stocks held by the fund).[11]

3 Validating the Decomposition Approach

There are several potential empirical issues associated with the implementation of our decomposition procedure. First, because we use quarter-end mutual fund stock holdings for the decomposition of stock holdings, we will miss high-frequency turnovers by mutual funds; see Kacperczyk, Sialm and Zheng (2007) and Elton, Gruber, Krasny and Ozelge (2006).[12] To the extent that short-term liquidity provision occurs within a calendar quarter, by using quarter-end holdings only, we may

[10] We use percentile age ranks to remove a time-series (increasing) trend in the age variable.

[11] Amihud illiquidity measure is defined as the average ratio between absolute daily return and daily dollar volume. We use percentile Amihud ranking for two reasons. First, there is a time-series (downward) trend in the Amihud measure due to an increase in trading volume; second, the Amihud measure may be extreme and subject to outliers. Using percentile ranking alleviates these issues.

[12] Campbell, Ramadorai and Schwartz (2007) attempts to infer institutional transactions within a given quarter by selecting trade sizes that best match quarterly holdings changes. Relying on a unique regulation governing mutual fund trade disclosure in Canada, Christoffersen, Keim and Musto (2006) investigate essentially all trades of 210 Canadian mutual funds between 2001 and 2003.

underestimate the benefit from liquidity provision.

Second, the division of informed trading and liquidity provision is imprecise. On the one hand, not all informed trading is liquidity demanding, especially when the trader is very patient and trades in small quantities over a relatively long period of time. While trading relatively large quantities quickly to take advantage of the time value of information, however, it is extremely hard not to absorb liquidity. As a result, liquidity-demanding trades are still likely information-driven on average. On the other hand, not all liquidity-demanding trades are information driven. For example, distress stocks sales by mutual funds (see Da and Gao, 2008) and assets fire-sales due to extreme flows (see Coval and Stafford, 2007) are likely to absorb liquidity but have nothing to do with mispricing trading motives. As distressed stocks are typically of small market capitalization, the impact of transactions will be alleviated, as each component of the CS measure is computed using the value-weighted average. When we leave out value-adding informed trading that is not liquidity demanding and including value-destorying distressed trading that is liquidity demanding, we are underestimating the benefit of informed-trading, and overestimating the benefit of liquidity provision. Finally, our classification of informed trading and liquidity provision depends on quarterly data, which could also be noisy. Collectively, these noises may prevent us from finding any significant results.

In spite of these challenges, we find that the data contains valuable information about the contribution of impatient informed trading and short-term liquidity provision to a mutual fund's stock selection ability. In what follows we first empirically validate our decomposition approach.

3.1 Type of Trades and Average Order Imbalances

In our decomposition, *order imbalance* plays a critical role – we assume that a mutual fund that, on average, trades with the order imbalance demands liquidity and a fund that trades against the order imbalance provides liquidity. According to the conventional wisdom, funds will be relatively impatient when opening or closing out a position, since such actions will more often be based on short-lived superior information. In contrast, funds will be relatively patient when they increase of decrease their holdings thereby providing liquidity to the market for the securities they hold in inventory.

In order to examine this hypothesis, for funds in our sample we first examine their holding changes over two consecutive quarters and categorize them into four groups: (1) open (holdings increase from zero to positive); (2) close (holdings decrease from positive to zero); (3) increase (holdings increase but not from zero) and (4) decrease (holdings decrease but not to zero). For each group, we then compute the average dollar holding change as a percentage of the total dollar holding change of the fund (computed using prices at quarter-end); and the average order imbalance measure. The average order imbalance measure is defined as the difference between total number of buyer-initiated shares brought and total number of seller-initiated shares sold divided by total number of shares traded during the quarter, the resulting number is then cross-sectionally demeaned. The associated t-value is computed using the time-series average with Newey-West adjustment for autocorrelations.

The results are provided in Panel A of Table 2. The average order imbalance measure for each trade type tells us whether the trade is on average demanding liquidity. We find that when fund managers open new positions and close out standing positions, they are likely to absorb market liquidity. That is consistent with our expectation that those trades are likely motivated by considerable mispricing perceived by fund managers who are willing to pay for the price of immediacy. We also find that when fund managers adjust their holdings, they are likely to provide liquidity on average. Again, that is, consistent with our conjecture that it is easier to provide liquidity on stocks a fund currently owns.

3.2 Dimensional Fund Advisors (DFA)

Dimensional Fund Advisor (DFA) is an asset management firm founded in 1981. Allegedly, the firm does not pick stocks via fundamental analysis. Instead, the firm helps its clients get exposure to certain segments of the asset markets via passive indexing or enhanced indexing. Anecdotal evidence suggests that a subset of the funds managed by DFA create value by systematically providing liquidity to those who want to trade small stocks for non-information related reasons.[13] If it is the case, using our decomposition procedure, one would expect to find a positive liquidity provision component in DFA's CS measure and an informed trading component close to zero. Of course, since we examine one specific fund over a limited time span, the statistical significance could be weak.

[13] See the case studies by Keim (1999) and Cohen (2002).

We examine the quarterly stock holdings of DFA's US Micro Cap Portfolio during the period from 1983 to 2004 and decompose its CS measure. The results are provided in Table 2 Panel B (upper panel). The overall CS measure for the fund is 36.1 bps per quarter but not statistically significant (t-value = 1.72), indicating that the fund does not seem to select stocks that outperform those with similar characteristics. As expected, the largest component of the overall CS measure is due to liquidity provision (20.5 bps per quarter) which is significant at 10% level (t-value = 1.84). In contrast, the informed trading component is very close to zero and statistically insignificant, which is consistent with what firm's investment policy claims. The liquidity provision component is larger than the informed trading component – and almost significant at conventional levels using a one tailed test.

3.3 Index Funds

Since the majority of index funds are formed to track the market index or other broad indices with the objective of minimizing tracking errors, we do not expect them to have a large CS measure. Index funds are most likely to trade during index rebalancing and demand liquidity in those trades (see Blume and Edelen, 2004). These trades would be incorrectly classified as "Informed Trading" within our decomposition framework, and the Informed Trading component, if different from zero, is likely to be negative. It is therefore less appropriate to apply the decomposition to index funds. For that reason we will focus only on actively managed funds for the remaining parts of the paper. Nevertheless, examining index funds provides another opportunity to test the validity of our decomposition approach.

We identify the index funds by their fund names recorded in CDA/Spectrum $S12$ mutual fund holding database. During the period from 1983 to 2004, there are about 11 domestic index fund portfolios identified each quarter on average from the holding database, starting from 1 fund each quarter in 1983 to about 25 funds each quarter after 2000. Using their stock holdings, we apply our decomposition to each fund and the results are then equally-weighted across funds during every quarter. The results are again presented in Table 2 Panel B (lower panel). The overall CS measure for index funds as a group is almost exactly zero. The index fund group has a positive although not significant CS^O component of about 25 bps per quarter on average (t-value = 0.93). In addition, the index funds on average make some profit (although not significant) from providing liquidity, as evident from a positive CS^{liq} component of about 6 bps per quarter (t-value=0.36). Interestingly, the positive CS^{O} and CS^{liq} are offset by a negative Informed Trading component (CS^{inf}= -35 bps) which is statistically significant (and also significantly less than the informed trading component using a one tailed test,) indicating a sizable price for liquidity paid by the index funds for trades that arise due to index rebalancing, new money flowing in, and redemptions.

3.4 Fund Styles

Panel C of Table 2 suggests that, overall, active fund managers seem to have some stock selection skill that requires trading with the order imbalance in the market. The average character selectivity measure is 23.5 basis points per quarter (t-value = 1.91), indicating the stocks selected by fund managers outperform stocks with similar characteristics. Of the 23.5 basis points, 13.9 basis points come from the passive buy-and-hold strategy and 14.2 basis points come from stocks recently traded by the funds. The adjustment component is small in absolute term (-1.8 basis points) but significant, potentially driven by fund flow to managers with skills as empirically documented by Chevalier and Ellison (1997), and Sirri and Tufano (1998), among others, and theoretically analyzed by Berk and Green (2004).[14] Finally, although both the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) are positive, neither is significant.

Across fund styles, we expect growth-oriented funds and funds to have more shorter-term informed trading opportunities. Income-oriented funds are more likely to augment their returns through liquidity provision. In addition, we should also expect relatively more variation in informed trading component within growth funds and relatively more variation in liquidity provision component within income-oriented funds. The results in Panel C of Table 2 confirm our conjectures. On average, growth-oriented funds have a larger informed trading component and income-oriented funds a larger liquidity provision component.

We examine the relative importance of different components of CS measures using the variance decomposition procedure detailed in Appendix B. In a nutshell, the variance decomposition delineates how much the cross-sectional variation in the total CS measure can be attributed to the cross-sectional variation in each of its four components. The results are reported in Panel C

[14] When managers have skill (CS^{P} is likely to be positive), fund inflow is more likely ($B > S$); when managers have no skill (CS^{P} is likely to be negative), fund outflow is more likely ($S > B$). Both effects lead to a negative CS^{adj} as in equation (1).

of Table 2 for the full sample of all active US equity funds, and across three style subsamples. As we expected, the informed trading component, CS^{inf} varies relatively more across funds among growth-oriented funds while the liquidity provision component varies relatively more CS^{liq} becomes relatively more across funds in the income-oriented category.

3.5 Return Gap

Since we observe a fund's holdings only at the end of each quarter, there is little we can say regarding what a fund did during the quarter except whether such actions added value. To examine how much additional value a fund adds relative to the return on a hypothetical buy-and-hold strategy based on the quarterly holdings reported by the mutual fund, we use the "return gap" measure in Kacperczyk, Sialm and Zheng (2007). Return gap is defined as the difference between the return available to investors and the return the fund would have generated if its holdings changed only at the end of each quarter net of fees:

$$RG_t = RF_t + EXP_t - RH_t$$

where RG_t is the return gap during quarter t, RF_t is the fund return available to investors during quarter t, RH_t is the implied return of the fund holdings during quarter t and the holding of the fund is based on the fund holding at the end of quarter $t-1$, EXP_t is the fund's expenses.

The return gap captures the net benefits and costs associated with the unobserved actions of the mutual fund managers. Kacperczyk, Sialm, and Zheng (2007) show that funds with high return gaps tend to have higher alphas in the future. They conjecture that the higher alpha could be in part due to the liquidity provision and market making activities that funds are engaging in during a quarter which are not captured by end-of-quarter holdings. Such activities can lead to "negative" price impact thus contribute to the fund's alpha. If that is true, and to the extent liquidity provision and market making activities take place uniformly over the quarter, we should be able to find a positive correlation between the return gap and the liquidity provision component of the characteristic selectivity measure (CS).

Both the return gap (RG) and the CS measure are likely to contain a component reflecting the fund manager's skill which persists over time. However, RG_t and CS_t are likely to be negatively correlated contemporaneously due to idiosyncratic shocks to RH_t which do not affect RF_t. Consider an example in which a fund sold a stock before a sharp idiosyncratic price run-up within a quarter t. This price run-up will lead to a large RH_t since the stock is held by the fund in the beginning of the quarter. It will not affect the characteristics-benchmark portfolio return (BR_t) since the price run-up is idiosyncratic in nature. It will not contribute to RF_t either since the fund sold the stock before the run-up occurs. As a result, RG_t will decrease while CS_t increases.

To examine the contemporaneous relation between the return gap and CS measures, we first sort funds based on the current quarter return gaps and report the average CS components at the portfolio level in the same quarter. The results are presented in Panel D1 of Table 2. The magnitudes of the average monthly return gaps (RG) across quintile portfolios are comparable to those found in Kacperczyk, Sialm and Zheng (2007). We document that return gap is strongly negatively correlated with the the Old CS component (CS^O) and the Informed Trading component (CS^{inf}). Interestingly, despite the negative contemporaneous correlation between the return gap and the CS measures, return gap is strongly positively correlated with the liquidity provision component (CS^{liq}). Such positive correlation is more likely driven by managers' skill which is present in both RG and CS^{liq}, and is consistent with the notion that funds specializing in liquidity provision are more likely to trade a stock within a quarter.

We then examine the relation between the current return gap and future CS measures. Following Kacperczyk, Sialm and Zheng (2007), we sort funds based on their past one-year return gaps and report the average CS components at the portfolio level during the next quarter. The results are presented in Panel D2 of Table 2. We first confirm the findings in Kacperczyk, Sialm and Zheng (2007) that fund's return gaps strongly predict future fund performance. Funds with large return gaps earn 50 bps more than funds with low return gaps in terms of four-factor alphas during the next quarter (t-value = 4.84). Among the four components of the CS measure, only the liquidity provision component (CS^{liq}) seems to be significantly related to past return gaps. The liquidity provision component of funds with large return gaps is 11 bps more than that of funds with low return gaps during the next quarter (t-value = 1.81). Overall, the positive correlation between the return gap and the liquidity provision component lends further support that our decomposition approach helps to isolate different kinds of managers' skills.

In summary, the empirical challenges for our decomposition procedure mainly arise from the lack of high frequency mutual fund transaction data. Nevertheless, we show that the decomposition, in particular, classifying trades into informed trading and liquidity provision, seems to make sense in several settings. In these settings, we demonstrate that the decomposition results are consistent with what one would expect. In the next section, we will take the decomposition procedure further to analyze a larger sample of mutual funds to better understand how skilled fund managers add value.

4 Information Events and Mutual Fund Stock Selection

Before we can examine the different channels through which skilled fund managers add value, we have to first identify who those skillful fund managers are. Ultimately, an active mutual fund manager will be successful if he or she has superior skill in processing valuation-relevant information on a stock, which helps in identifying potential mispricing. Having superior information processing abilities is not enough. There has to be opportunities as well. More opportunities arise when stocks are affected by information events. Because rational managers can choose not to trade such stocks if they do not have an advantage in analyzing information, we should expect those managers who choose to trade them to earn superior returns on average. The assumption is that those who take the opposite side are "noise traders" who trade for a variety of reasons that we do not fully understand.

To identify the occurrence of information events, we first make use of the Probability of Informed Trading measure (PIN), of Easley, Kiefer, O'Hara and Paperman (1996), and Easley, Kiefer and O'Hara (1997). In their model, there are two types of traders: informed traders and uninformed traders. In the absence of information events, only uninformed traders trade (for unspecified "liquidity" reasons,), and the order is equally likely to be a buy or a sell, resulting in an order imbalance measure close to zero on average, and a low PIN measure. On the other hand, when there are significant information events and informed traders also trade, there will be large amounts of buy orders *or* sell orders (depending on the nature of the information), resulting in a large order imbalance and a high PIN measure.[15] Empirically, PIN decreases with trading volume, size

[15] A more detailed description of the PIN measure and its estimation procedure is contained in Appendix C.

and analyst coverage, but increases with bid-ask spread, and insider and institutional ownership, consistent with it being a reasonable measure of private information events.

To estimate PIN, we use tick-by-tick transaction data for each quarter from 1983 to 2004, employing the entire three-month data to ensure precision of the estimates. Overall, we have on average 4110 stocks with PIN measures in a quarter. Due to data availability from ISSM, NASDAQ stocks enter the sample in 1987 and account for a large portion of the sample afterwards. The mean of PIN measures in our sample is 25.8% with an associated standard deviation of 12.1%. Consistent with Easley, Hvidkjaer and O'Hara (2002), we find that high-PIN stocks are likely to be smaller and less liquid stocks.

In each quarter and for each fund, we then compute a $trade\_PIN$ variable by value-weighting the PIN of stocks traded by the fund during the quarter using the dollar value of the trade. Specifically, we compute $trade\_PIN$ for the j-th mutual fund at the end of quarter t in our sample as:

$$trade\_PIN_{j,t} = \frac{\sum_{i=1}^{N} PIN_{i,t} \times d_{i,j}}{\sum_{i=1}^{N} d_{i,j}} \tag{4}$$

where $PIN_{i,t}$ is the estimated PIN measure of the i-th stock traded by mutual fund j during quarter t, and $d_{i,j}$ is the absolute dollar value (using the stock price at the end of the quarter) of the holding change during quarters t as reported by the mutual fund j. Intuitively, funds that buy or sell more high-PIN stocks should have higher $trade\_PIN$ measures.

To evaluate the mutual fund performance, we use both factor-adjusted returns and holding-based characteristics-adjusted returns performance. Our first factor-based performance measure is based on the Carhart (1997) four-factor model, which augments the Fama-French three factor model (1993) with a momentum factor (UMD). We sum across the four-factor-adjusted alphas within a quarter to obtain the mutual fund's quarterly four-factor-adjusted return.

It is important to note that the momentum factor is not a pervasive risk factor but a style factor capturing what a fund manager does in order to create value (see Pastor and Stambaugh, 2002a and 2002b). Whereas momentum trading involves price impact that has to be taken into account (see Korajczyk and Sadka, 2005), the UMD factor does not. Therefore, a fund that has a positive three-factor alpha but a negative four-factor alpha may still be preferred by an investor. In view of that, we also report the three-factor alphas. In order to examine whether a mutual funds three-factor alpha may be from activities that expose the fund to economy-wide pervasive liquidity risk, we also examined performance by augmenting the four-factor model with the aggregate liquidity factor of Pastor and Stambaugh (2003). Since the five- and four- factor alphas are very similar, we do not report the five-factor alphas.

Our second performance measure is the DGTW "Characteristic Selectivity (CS)" measure which is computed using the quarterly fund stock holdings. The factor-adjusted mutual fund returns differ from the characteristics-adjusted mutual fund returns in several important aspects. First, while the four-factor adjusted-returns are after fees and expenses, the characteristic-adjusted returns are before fees and expenses. This difference explains why the characteristic-adjusted returns are typically higher than the factor-adjusted returns. Second, unlike the factor-adjusted fund returns, the characteristics-adjusted fund returns ignore possible cash, stock holdings below reporting threshold and other non-stock holdings of the mutual fund. However, such holdings are usually small, accounting for less than 5% of the fund holdings on average in our sample. In addition, factor-adjustment and characteristics-adjustment generate similar return spreads between the top and bottom fund deciles, which indicates that non-stock holdings by mutual funds are unlikely to introduce any systematic biases to our results.

4.1 Information Events and Informed Trading

At the end of each quarter over 1983 - 2004, we sort all mutual funds in the sample into deciles according to their $trade_PIN$s and examine the factor-adjusted and the DGTW characteristics-adjusted mutual fund portfolio returns (CS measures) in the next four quarters after portfolio formation within each decile. The results are summarized in Table 3.

The central message in Table 3 is that funds trading more high-PIN stocks outperform the funds trading low-PIN stocks. Using the four-factor model for risk adjustment, we find that funds in the top $trade\_PIN$ decile outperform funds in the bottom $trade\_PIN$ decile by 48 basis points in the next quarter with a t-value of 3.15. The return spread is 46 basis points in the second quarter with a t-value of 2.98. Return spreads are 35 basis points in the third quarter (t-value = 2.24) and 35 basis points in the fourth quarter (t-value = 2.26). Thus, within a one-year horizon after portfolio formation, funds within the highest $trade\_PIN$ decile outperform the funds within the lowest trade-PIN decile by roughly 1.6 percentage points. Using three-factor alphas which arguably are better measures of fund managers' skill, we find that funds trading high-PIN stocks outperform those trading low-PIN stocks by 56 basis points during the first quarter and by 2.1 percentage points during the first one year.

In general, we see a positive relation between future risk-adjusted fund returns and the $trade\_PIN$ variable: the lowest five trade-PIN decile portfolios usually have large negative and statistically significant factor-adjusted return during four quarters after portfolio formation; in contrast, the highest five trade-PIN decile portfolios have small negative and in most cases statistically insignificant factor-adjusted returns.

The results are similar for the characteristics-based risk adjustment – with the caveat that characteristics-based risk adjustment is done to pre-fee returns not after fee returns. On average, the top trade-PIN decile portfolio of funds outperform the bottom decile portfolio by 53 and 40 basis points per quarter in the first and second quarter after portfolio formation. These differences are significant at 1% significance level. The return spreads between the top and bottom deciles attenuate to 19 to 9 basis points in the third and fourth quarter. Notice that the CS measure exhibits less persistence than the three- and four-factor alphas. That may be due to the fact that the CS measure does not capture value added trough within-quarter trading activities. To the extent that the manager's skill in creating value through trading within quarters is likely to persist, we should expect more persistence in the three- and four-factor alphas when compared to the persistence in the CS measure.

4.2 Why Do Funds Trading High-PIN Stocks Have Higher Alphas?

PIN Risk?

Easley, Hvidkjaer, and O'Hara (2002) document that high-PIN stocks earn higher returns in order to compensate the agents for the risk of trading with informed traders. Since high $trade\_PIN$ funds may also hold high-PIN stocks, the high returns they earn might simply be due to higher risk that is not captured by the four-factor model or the DGTW benchmark characteristics risk adjustment.

To address this concern, we first directly control for PIN risk in the risk adjustment. In the case of the factor-risk adjustment, we augment the benchmark four-factor model with a PIN risk factor - PIN_t. Similar to Easley, Hvidkjaer, and O'Hara (2004), we construct the PIN risk factor as the high-PIN decile portfolio return minus the low-PIN decile portfolio return. The resulting five-factor-adjusted mutual fund return thus controls for any systematic PIN risk. In the case of characteristics-based risk adjustment, we construct characteristics benchmark portfolios by matching along size, book-to-market, past return, and PIN characteristics simultaneously. At the end of each quarter, we sort all stocks into 81 portfolios using a 3 by 3 by 3 by 3 sequential sort based their sizes, book-to-market ratios, past 12 month returns and PIN measures (in that order). We then compute a new characteristics-adjusted fund return or characteristic selectivity measure (CS^*) using the 81 benchmark portfolio returns.

Both the five-factor-adjusted fund returns and the new characteristics-adjusted returns during the next quarter in the $trade\_PIN$ sorted fund deciles are presented in the first two columns of Panel A, Table 4. In general, the risk-adjusted fund returns increase with the $trade\_PIN$ measure. Funds trading high-PIN stocks perform better than those trading low-PIN stocks, even after directly controlling for PIN risk. The spread between the returns on the high-$trade\_PIN$ funds and the low-$trade\_PIN$ funds, after directly controlling for PIN risk, narrows slightly but remains positive and statistically significant. The five-factor alpha spread is 45 basis points per quarter with a t-value of 2.85, while the new characteristics-adjusted return spread is 43 basis points per quarter with a t-value of 2.11.

In a further check, we show that our results are robust to three alternative measures of the number of information events. We describe these three alternative measures in Appendix C. The first measure we consider is the asymmetric information component ($adjPIN$) of the modified PIN measure proposed by Duarte and Young (2007), which removes the illiquidity component of the original PIN measure. Duarte and Young (2007) show that the pricing of PIN risk is driven by the illiquidity component, while $adjPIN$ is not priced in the cross-section. The second measure is the information asymmetry component of the bid-ask spread (*theta*) as proposed in Madhavan, Richardson, and Roomans (1997). In addition to causing large order imbalance, informed-trading will also force the market maker to increase the bid-ask spread which can be captured by a higher *theta* measure. Finally, assuming that significant information events usually lead to abnormal trading in a stock, we use a measure of abnormal turnover (*aturn*) calculated following Chordia, Huh and Subrahmanyam (2006).

To measure the average number of information events associated with stocks traded by mutual funds during a quarter, we compute $trade\_adjPIN$, $trade\_theta$, and $trade\_aturn$ in the same fashion as $trade\_PIN$ by replacing PIN with $adjPIN$, $theta$, and $aturn$ in (4) accordingly. The results appear in Table 4, Panel A. For brevity, we do not report the three-factor alphas which in fact generate even larger spreads between the extreme fund deciles.

We obtain very similar results for these alternative measures of the amount of information. The next-quarter mutual fund risk-adjusted returns (using benchmark risk adjustment models) in general increase with these alternative measures. In addition, funds trading stocks associated with more information events outperform funds trading stocks associated with fewer information events by about 48 basis points per quarter, similar to the results using the original PIN measure. These risk-adjusted return spreads are highly significant in the cases, and independent of whether we use factor adjustment or characteristics adjustment to account for risk. The fact that we obtain very similar results using $adjPIN$ instead of PIN provides further support that PIN risk is not driving our results.

Finally, we directly examine the average PIN of stocks bought and sold by the funds separately in Panel B of Table 4. In each quarter and for each fund, we first compute the value-weighted average PIN of stocks in the "buy" portfolio (stocks recently bought by the fund) and the "sell" portfolio (stocks recently sold by the fund). These PINs are then averaged across funds in the same $trade\_PIN$ decile and across time. Among funds trading high-PIN stocks, the stocks they recently bought and sold have very similar average PINs (0.2294 *vs.* 0.2301 and their difference is not statistically significant). However, we confirm in Table 5 that the trade component of the CS measure is positive and significant for funds trading high-PIN stocks. Recall that the trade component measures the value-added from the most recent mutual fund stock trades (both buy and sell). A positive trade component thus suggests that stocks recently bought by those funds outperform those recently sold by them although these stocks are of similar PINs. This finding again suggests that the alpha generated by funds trading high-PIN stocks comes from superior stock selection skill rather than the PIN risk.

Momentum Trading?

Grinblatt, Titman, and Wermers (1995) document that mutual funds often use momentum as a stock selection criterion, so momentum effects can significantly influence mutual fund performance (see also Carhart, 1997). Panel B of Table 4 shows that funds trading high-PIN stocks hold more recent winners than funds trading low-PIN stocks, resulting in a higher $fund\_mom$ on average. A natural question arises: could the difference in the CS measures between funds trading high- and low-PIN stocks be driven by the momentum effect? We believe that the answer is *no* for several reasons.

First, factor-adjusted and characteristics-adjusted fund returns are computed throughout after adjusting for momentum effects. Second, when we later regress the risk-adjusted fund returns on several fund characteristics in a cross-sectional regression, we find $fund\_mom$ to be insignificant, while $trade\_PIN$ is still highly significant (see Table 9), confirming that the higher return associated with funds trading high-PIN is not driven by the momentum effect. Finally, we directly examine the average past return characteristics of stocks bought and sold by the funds separately in Panel B of Table 4. In each quarter and for each fund, we first compute the value-weighted average past one-year return of stocks in the "buy" portfolio (stocks recently bought by the fund) and the "sell" portfolio (stocks recently sold by the fund). These past returns are then averaged across funds in the same $trade\_PIN$ decile and across time. Although high-$trade\_PIN$ funds seem to buy more recent winners than low-$trade\_PIN$ funds (the average past one-year return in the "buy" portfolio is 34.3% for high-$trade\_PIN$ funds *vs.* 20.9% for low-$trade\_PIN$ funds), high-$trade\_PIN$ funds also sell more extreme recent winners at the same time (the average past one-year return in the "sell" portfolio is 46.6% for high-$trade\_PIN$ funds); they thus are not momentum traders in the traditional sense. In addition, funds in $trade\_PIN$ deciles 7 to 9 seem to buy or hold even more winners than funds in the top $trade\_PIN$ decile. If the momentum effect drives the high CS measure, we would expect funds in $trade\_PIN$ deciles 7 to 9 to have higher risk-adjusted returns on average. This is clearly not the case. In what follows, we therefore proceed to examine the relative importance of informed trading and liquidity provision components to total performance of these funds.

Informed Trading or Liquidity Provision?

Table 5 presents results of the decomposition applied to the decile portfolios of funds sorted on $trade\_PIN$. This reveals interesting differences in value-added between funds trading high-PIN stocks and funds trading low-PIN stocks. For high-$trade\_PIN$-funds, most of the before fee characteristics adjusted returns (character selectivity (CS) measure) comes from active trading during the previous quarter ($CS^T = 31.2$ basis points with a t-value of 2.83.) While positions put in earlier contribute almost an equal amount it is not statistically significant, since $CS^O = 26.5$, but $t = 1.43$ The funds trading medium-$trade\_PIN$ stocks (deciles 5 and 6) also earn significantly positive before fees characteristics adjusted returns, and their value addition comes mostly from CS^O, i.e., positions they put in earlier. Funds trading in low$-trade_PIN$ stocks lose on informed trading with the result their before fees characteristics adjusted return is slightly negative though not significantly different from zero. We can confirm that the stocks bought by mutual funds (the buy portfolio) and the stocks sold by mutual funds (the sell portfolio) have very similar average PIN measures. The trade component, which should be less subject to PIN risk, is positive and significant for funds trading in high PIN stocks (deciles 8, 9 and 10). Although both the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) are positive for high-$trade\_PIN$-funds, only the informed trading component is significant (20.4 basis points with a t-value of 2.25), and it is twice the size of the liquidity provision component (10.4 basis points). This is consistent with our conjecture. When skillful managers absorb liquidity by trading high-PIN stocks, they are likely to have valuation-relevant information, and thus make money on informed trading. For them, there is less of an added cost of demanding immediacy in the market than there is a benefit from superior information, as Grossman and Stiglitz (1980) would predict. In terms of liquidity provision, not all of them can perform well consistently. As a result, although the liquidity provision component is positive on average, it is much smaller and not significant, perhaps because of the possibility of trading against informed traders and the noise associated with identifying liquidity provision using quarterly mutual fund holdings data .

The low-$trade\_PIN$-funds, despite near zero stock selection skill on average, seem to possess some skill in liquidity provision. The liquidity provision component (16.2 basis points) is significant (t-value = 2.57). This is because when fund managers trade low-PIN stocks, they are likely to trade with uninformed traders. When they trade against market order imbalance, they are likely to make money by providing the needed liquidity. Although the reward for liquidity provision on these stocks is lower than that on the high-PIN stocks, the risk of adverse selection is also lower, making liquidity provision more easily detected. The positive liquidity provision component is partly offset by a negative informed trading component, resulting in a close-to-zero CS measure. As can be seen from the last column the difference between the informed trading component and liquidity provision component is negative and statistically significant for funds trading low $trade\_PIN$ stocks.

To sum this up, the decomposition exercise reveals interesting patterns in how mutual fund's trades can add value. While informed trading is more likely to add value at times when the stock traded are associated with information events, liquidity provision is more likely to add value (or be easily detected in a statistical sense) when the stocks traded are associated with few information events.

5 Identifying Funds With Positive Future Alphas

5.1 Sorting on Historical Alpha

A key challenge in mutual fund studies is to distinguish skill from luck. If the superior performance of a mutual fund is due to its manager's skill, to the extent that such skill persists in the near future, we would expect past winner funds to continue outperforming the past loser funds. If the superior performance of a mutual fund is due to luck, we would not expect persistence in fund performance. In this section, we examine the fund performance persistence in our sample.

At the end of each month from 1983 to 2004, we sort all mutual funds in our sample into deciles according to their four-factor alphas (Fama-French three factors augmented by Carhart's momentum factor) estimated using the previous 5 years of monthly mutual fund return data. Each fund decile portfolio is then held for one month and rebalanced next month. We then compute the three-factor and four-factor alphas on these rebalanced fund decile portfolios. The results are presented in Table 6. While past winners continue to outperform past losers by about 34 bps per month (t-value = 2.93), this outperformance is mainly driven by persistence in poor performance by the past losers. Funds associated with the lowest alphas in the past continue to underperform the market by 32 bps per month (t-value = 3.48). On the other hand, funds associated with the highest alphas in the past do not significantly outperform the market in the future. Their average four-factor alpha in the next month is only 2 bps (t-value =0.21). The results are similar with three-factor alphas.

Mamaysky, Spiegel and Zhang (2007b) point out that historical alphas contain large measurement errors, and they propose three filters which substantially help predict future alphas of funds in their sample. We repeat the sorting exercise after including those three filtersSpecifically, for any fund to be included in any decile the following must be true: (1) the absolute value of alpha must be less than 2% per month, (2) the CAPM beta must be in between 0 and 2, and (3) in the previous month the forecasted alpha the difference between the realized return and the market return must have the same sign. Consistent with the findings in Mamaysky, Spiegel and Zhang (2007b), fund performance becomes more persistent after applying these filters as shown in Table 6. Past winners now outperform past losers by almost 78 bps per month (t-value = 4.12). In addition, funds associated with the highest alphas in the past outperform the market by a larger amount (15 bps per month). However, the outperformance is not statistically significant (t-value =1.19). In other words, by sorting funds into deciles based on their past alphas in the more recent sampling period (1983-2004), we are not able to identify fund portfolios that, on average, outperform the market going forward into the future after fees and expenses. This is consistent with the findings reported in Kosowski, Timmermann, Wermers and White (2006) and Fama and French (2008).[16] A potential explanation is given by Berk and Green (2004): erosion in the after fee performance of superior mutual funds due to inflows coupled with the diseconomies of scale, and the entry of new funds that on average have little superior abilities makes it difficult to identify funds that provide superior returns to investors going forward. The failure of historical alpha alone in predicting positive future alphas suggests the need for bringing in additional information. We take this up next.

5.2 Information Events and Performance Persistence

The managers' skill ultimately come from their superior ability in processing information. As we argued earlier, that skill is likely to have more value in stocks that are affected by information events. To the extent that a manager's skills are likely to persist for some time, past superior performance is more likely to be an indication of future performance for a manager who attained that performance by trading stocks associated with more information events. In other words, we would expect stronger fund performance persistence among funds that traded in High-PIN stocks

[16] If we examine a longer sampling period from 1970-2002 as in Mamaysky, Spiegel and Zhang (2007b), the top 10% of funds (after applying the filters) associated with the highest past alpha indeed outperform the market in the future by 3% per year (t-value = 2.68). Also consistent with the findings in Mamaysky, Spiegel and Zhang (2007b), computing alphas using Kalman Filters (Mamaysky, Spiegel and Zhang, 2007a) instead of rolling-window OLS regressions does not improve the fund performance persistence test.

recently and we will be able to identify funds that outperform the market in the future only in this group.

To evaluate this conjecture, at the end of each quarter from 1983 through 2004, we first sort all mutual funds in the sample into two groups according to their $trade\_PIN$ during the quarter. Within each group, we then further sort funds into quintiles according to their four-factor alphas estimated using the previous 5 years of monthly returns as in Mamaysky, Spiegel and Zhang (2007b). Overall, we have 10 fund portfolios, which is comparable to the earlier exercise using fund deciles. These ten fund portfolios are then held for one month and rebalanced next month. We then compute the monthly four-factor alphas on these rebalanced fund portfolios. For each $trade\_PIN$ group, we report the results for the fund groups with high, medium and low past alphas in Panel A of Table 7. Among funds that trade low-PIN stocks recently, the performance persistence is very weak. Funds with high past alphas do not significantly outperform those with low past alphas during the next month. In contrast, among funds that trade high-PIN stocks, funds that earn high past alphas have higher alphas in the next month than funds that earn low past alphas. The spread between four-factor alphas is 36 bps per month (t-value = 5.10). Such spread is unlikely driven by the PIN risk since funds with high and low past alphas in the same $trade\_PIN$ group have very similar $trade_PIN$s.[17] Without the filters in Mamaysky, Spiegel and Zhang (2007b), we are not able to identify a fund portfolio that significantly outperforms the market going forward at monthly frequency using $trade\_PIN$ variable alone. The past winner within the high-$trade\_PIN$ group does outperform the market after fee (four-factor alpha is 13 bps per month) but the alpha is not significant (t-value = 1.65).

Once we combine the use of $trade\_PIN$ with the filters proposed in Mamaysky, Spiegel and Zhang (2007b), the results improve significantly. First, we document much stronger performance persistence. Among funds that trade high-PIN stocks, funds that earn high past alphas have much higher alphas in the next month than funds that earn low past alphas. The spread between four-factor alphas is 69 bps per month (t-value = 4.61). More importantly, we are now able to identify a fund portfolio that significantly outperforms the market going forward at monthly frequency. The past winner within the high-$trade\_PIN$ group now outperforms the market after fee with

[17] In the tests presented in this section, we verify that controlling for the PIN risk directly using either a fifth PIN risk factor or using PIN characteristic-adjustement produces very similar results.

a statistically significant alpha of 35 bps per month (t-value = 3.33). In fact, among the high-$trade\_PIN$ funds, the performance persistence is as strong on the past winners as on the past losers. We obtain similar results using the Fama-French three-factor alphas as in Panel B of Table 7.

Berk and Green (2004) argue that fund managers with skills command higher fees. As a result, zero after-fee alpha does not necessarily imply that the manager has no skill. In another words, mutual funds' pre-fee performance and their fee growth rates could serve as additional measures of manager skills. When we examine factor-alphas before fees and expenses in Panel B, we find that funds with high past alphas and trading high-PIN stocks earn a higher four-factor alpha of 45 bps per month (t-value = 4.24) on average. Among funds that trade high-PIN stocks, those with high past alphas experience a much higher total-dollar-fee growth rate of 8.49% during the first quarter. The fee growth rate for the funds with low past alpha during the same quarter, by contrast, is only 1.42%.

Sources of Future Performance

Table 8 gives the decomposition of the pre-fee CS measures of these fund portfolios. Past winner funds in the high trade-PIN group have a total CS measure of 107 bps in the following quarter. About half of that (48 bps) indeed comes from informed trading, and the rest mostly come from positions taken from earlier quarters. The positive and significant CS^{inf} and CS^{O} explain why these past winners are able to outperform the other funds in the future. They are indeed very good at informed tradings. Interestingly, the past loser funds in the low trade-PIN class have a significantly positive liquidity provision component (CS^{liq}) and a significantly negative informed trading component (CS^{inf}) – only that the two are about equal (in absolute term) and cancel each other. A potential explanation is that these funds do not possess skill and are trading stocks associated with little information. When they demand liquidity, they pay a price, resulting in a negative informed trading component. When they supply liquidity, they benefit from the price concession, resulting in a positive liquidity provision component. Since the stocks they trade are associated with little adverse selection risk, it is easier to detect both components in a statistical sense.

5.3 Fund Characteristics and Performance Components

To examine the relation between fund characteristics and the CS measures, we use a Fama-MacBeth (1973) cross-sectional regression approach. Specifically, we regress the next-quarter CS measure and its components on several fund-level characteristics for each quarter from 1983 to 2004. All right side variables are measured as deviations from their corresponding cross-sectional means, standardized to have unit variance, and winsorized at the 1st and 99th percentiles to alleviate the effect of outliers. In addition, the regression intercept can be interpreted as the average effect of having a Growth-and-Income ("GNI") fund style. Finally, the regression coefficients are averaged across time, and the associated t-values are computed using the Newey-West correction with lag terms of 8 quarters to account for autocorrelations in the error terms. The regression results are reported in Table 9.

When we regress the total CS measure on fund characteristics, we find $trade\_PIN$ to be significant even in the presence of many other fund-level characteristics, indicating that the difference in stock selection skill between funds trading high-PIN stocks and those trading low-PIN stocks is not entirely driven by other correlated fund characteristics. We also find that $flow$ has a positive and significant coefficient, consistent with the notion that money is chasing skilled fund managers rationally. In addition, the significance of $dummy\_AGG$ means that funds with an "Aggressive Growth (AGG)" investment style are better in selecting stocks, confirming earlier findings by Daniel, Grinblatt, Titman and Wermers (1997).

We also focus on the two particular components of the total CS measures: the informed trading component (CS^{inf}), and the liquidity provision component (CS^{liq}). Interestingly, fund-characteristics associated with informed trading, and with liquidity provision are quite different. When we regress CS^{inf} on fund characteristics, we find $trade\_PIN$ to be even more significant, indicating that the positive relation between stock selection skill and high $trade\_PIN$ is likely driven by informed trading. In addition, $flow$ and $dummy\_AGG$ remain significant, indicating that informed trading is more prevalent in funds with an "Aggressive Growth (AGG)" investment style and money is chasing fund managers who are good at informed trading. In contrast, regressing CS^{liq} on fund characteristics causes different patterns to emerge. First, $trade\_PIN$ is now negatively related to CS^{liq} (although not significantly). Second, intercept and age are significant, indicating that younger funds and funds with "Growth and Income (GNI)" investment styles are likely to be more highly rewarded, via liquidity provision.

6 Measuring the Noise in Mutual Fund Trades Inferred From Quarterly Data

Throughout the paper, we infer the trades of mutual funds by comparing quarter-end holdings over consecutive quarters. This procedure adds noise for the following reasons. First, we miss interim trading by the funds within a quarter leading to measurement errors in the calcuation of $trade\_PIN$. Second, inferring mutual fund trades using their quarterly holdings also adds noise to the classification of trades as belonging to the "informed-trading" category or the "liquidity-provision" category. In this section we examine the potential impact of these noises on our inference using a proprietary institutional trading dataset provided by the Plexus Group, a consulting firm for institutional investors that monitors the cost of institutional trading.

Plexus Group customers consist of over 200 financial institutions that collectively transacted over $4.5 trillion in equity trading prior to its acquisition by ITG, Inc. By early 2003, Plexus Group had analyzed 25% of exchange-traded volume worldwide. The Plexus Group dataset we use covers 1991Q2-1993Q1and 1995Q4-1998Q2. Prior research that uses the Plexus Group data include Keim and Madhavan (1995) and Conrad, Johnson, and Wahal (2003).

The dataset records the details (date, size, buy/sell indicator, type of order) of every institutional order that was executed for all the institutions that Plexus Group monitors. Therefore, for every institutional money managers in Plexus Group data, we we are able to compute their $trade\_PIN$, and accurately classify their trades as belonging to the "informed-trading" or "liquidity-provision" category. However, note that the Plexus Group Data covers only a subset of the mutual funds we examined earlier, and that too only for a relatively short time period. Further, since we do not observe the actual identities of these managers, we are not able to relate their trades to their actual performance and other manager and fund characteristics.

Based on the actual transactions of institutional managers recorded in Plexus Group data, we are able to quantitatively evaluate how the noise (measurement error) in our procedure for inferring mutual fund trades using their quarterly holdings affect our conclusions. Other things being the same, we should expect our procedure to introduce more noise when dealing with fund managers who conduct more within-quarter transactions. Therefore, each quarter, we sort managers into quintiles based on the intensity of interim trading where interim trading is defined as the ratio of the total value of all within-quarter round-trip trades and the total value of all net trades in that quarter. Examining the net trades of a manager during a quarter is equivalent to inferring her trades by comparing her quarter-end holdings over consecutive quarters. We then compute $trade\_PIN$ in two ways: one based on the actual trades, and the other based the net trades which miss out interim trading. The difference between the two estimates measures the noise introduced by using the quarterly holdings data.

We find that using quarterly data to infer mutual fund trades has little impact on the calculation of $trade\_PIN$. The results are reported in Table 10. First, the estimation error has a mean close to zero for all fund quintiles with different intensities of interim trading. Second, the magnitude of the estimation error is also small. The average absolute estimation error in $trade\_PIN$ is only 1% of the average $trade\_PIN$. Even for funds associated with the largest amount of interim trading, the average absolute estimation error is less than 3% of the average $trade\_PIN$.

Classifying trades into either “informed-trading” or “liquidity-provision” group using quarterly data is more noisy. When we examine the actual fund transactions in the Plexus Group data, we report in Table 10 that on average, 66% of these trades are correctly assigned to be either “informed-trading” or “liquidity-provision” even with the quarterly data. The percentage is higher when measured using total value of the trades. The percentage of correct assignment is very significantly higher than 50% for all fund quintiles, suggesting that trade classification using quarterly data is still sufficiently informative. Not surprisingly, the accuracy of the assignment is better among funds that conduct little interim trading. For these funds, 70% of their trades are correctly assigned to be either “informed-trading” or “liquidity-provision.” Interestingly, these funds trade stocks with higher PINs on average, implying that our results involving high $trade\_PIN$ funds are less likely to be affected by noise introduced by ignoring interim trades due to the use of quarterly holdings data.

Overall, using the actual transactions of institutional money managers, we find our procedure for inferring mutual fund trades using quarterly holding data to have little impact on the calculation of a fund’s $trade\_PIN$. Although the classification of trades into either “informed-trading” or "liquidity-provision" categories is affected to a larger extent by ignoring interim trades by a fund, the classification scheme remains sufficiently informative. To sum up, our general conclusions regarding informed trading being the source of alpha for funds trading high $trade\_PIN$ stocks should remain valid.

7 Conclusion

The traditional approach to portfolio performance evaluation is to decompose the skill of a portfolio manager into two components: security selection and market timing. In this paper we suggest a further decomposition of the former based on whether the portfolio manager's trades demand liquidity ("informed trading") or provide liquidity ("liquidity provision"). We develop a method for such a decomposition based on the composition of the portfolio holdings of a mutual fund. We validate and illustrate the use of our decomposition method by empirically examining the stock selection ability of managed mutual funds. We find that informed trading is more important for growth-oriented funds while liquidity provision is more important for younger funds with income orientation.

In practice, our decomposition approach will be more valuable to plan sponsors and large institutional investors who have access to holdings on a daily or weekly basis. Our procedure can help evaluate a manager's contribution on both the informed trading and liquidity provision dimensions with greater precision, thereby contributing to better allocation of resources across different money managers.

We confirm the folk wisdom that a mutual fund manager with superior stock selection ability is more likely to benefit from trading in stocks affected by information-events. In particular, using the Probability of Informed Trading (PIN) measure of Easley, Kiefer, O'Hara, and Paperman (1996) as a quantitative indicator of the amount of information affecting a stock, we show how to construct a portfolio of mutual funds that have a positive alpha. A managed portfolio of past winners *among funds trading high-PIN stocks*, identified using the methods in Mamaysky, Spiegel and Zhang (2007a, 2007b), have a statistically significant after fees alpha of 35 bps per month (t-value = 3.33.) Most of that alpha comes from impatient informed trading. In contrast, the alpha of the corresponding portfolio of past winners *among all funds* is not statistically or economicaly significantly different from zero.

Appendix A: A Numerical Example for the Decomposition of Mutual Fund Stock Selection Skill

Assume there are six stocks (A, B, C, D, E, and F). A mutual fund's holdings in these stocks at the end of quarter $t-1$ (N_{t-1}) and t (N_t), stock prices at the end of quarter t (P_t), and the characteristics-adjusted stock returns during quarter $t+1$ $[R_{j,t+1} - BR_{t+1}(j,t)]$ can be summarized in the following table:

Stock	N_{t-1}	N_t	P_t	$R_{j,t+1} - BR_{t+1}(j,t)$
A	2	1	10	−3%
B	2	0	15	−2%
C	2	2	20	−1%
D	2	2	25	1%
E	2	3	30	2%
F	0	2	35	3%

The Hold, Buy and Sell are then defined by the holdings N_t^H, N_t^B and N_t^S:

Stock	$N_t^H = \min(N_{t-1}, N_t)$	$N_t^B = N_t - N_t^H$	$N_t^S = N_{t-1} - N_t^H$
A	1	0	1
B	0	0	2
C	2	0	0
D	2	0	0
E	2	1	0
F	0	2	0
Value	$H_t = 160$	$B_t = 100$	$S_t = 40$

The portfolio values H_t, B_t, and S_t are determined using the prices at the end of quarter t (P_t). Notice that $B_t > S_t$, and the difference is likely financed by fund inflows, or a reduction in cash position or the sale of other non-stock assets held by the fund. The Hold, Buy, and Sell can be treated as three separate funds whose CS measures can be computed using equation (**??**) and holdings as:

	Hold	Buy	Sell
CS	$CS_{H,t+1}$= 0.63%	$CS_{B,t+1}$= 2.70%	$CS_{S,t+1}$= −2.25%

Given this information, equation (1) then decomposes the total CS measure into three components:

CS_{t+1}	CS^{O}_{t+1}	CS^{T}_{t+1}	CS^{adj}_{t+1}
1.42%	0.05%	1.49%	−0.12%

If we further assume that the fund trades B and F in the same direction as the aggregate order imbalance and trades A and E against the direction of aggregate order imbalance, equation (2) further decomposes the trade component (CS^{T}_{t+1}) into an informed trading component (CS^{inf}_{t+1}) and a liquidity provision component (CS^{liq}_{t+1}):

CS^{T}_{t+1}	CS^{inf}_{t+1}	CS^{liq}_{t+1}
1.49%	1.11%	0.38%

Appendix B: Variance Decomposition of the "Characteristic Selectivity" (CS) Measure

Empirically, we decompose the total "Characteristic Selectivity" (CS) measure (DGTW, 1997) into four components:[18]

$$CS = CS^{O} + CS^{adj} + CS^{inf} + CS^{liq}.$$

Consequently, we have

$$var(CS) = cov(CS, CS^{O}) + cov(CS, CS^{adj}) + cov(CS, CS^{inf}) + cov(CS, CS^{liq}),$$

where $var(\cdot)$ and $cov(\cdot)$ are the cross-sectional variance and covariance, respectively. Dividing both sides of the above equation by $var(CS)$, we then have

$$1 = \beta_{P} + \beta_{adj} + \beta_{inf} + \beta_{liq}.$$

[18] For simiplicity of notation, we omit the time subscript t and fund superscript i.

The term $\beta_{(\cdot)}$ then measures the contribution of component $(\cdot)$ to the cross-sectional variations of CS. The sum of the contribution from the four components is equal to one by construction. β can be measured by regression. For instance, β_P is estimated by regressing CS^O on CS cross-sectionally. Empirically, we have a panel data of cross-sectionally demeaned CS, CS^O, CS^{adj}, CS^{inf} and CS^{liq}. To estimate β, we run a Weighted Least Squares (WLS) regression. In practice, this means deflating the data for each fund-quarter by the number of funds in the corresponding cross-section.

Appendix C: Measures of Private Information Events — A Brief Description

Easley and O'Hara, along with their coauthors, in a series of papers develop this measure to capture the probability of information-based trading. Let α denote the probability that an information event occurs; δ denote low value of underlying asset, conditioning on the occurrence of informational event; μ is the rate of informed trade arrivals; ϵ_b is the arrival rate of uninformed buy orders; ϵ_s is the arrival rate of uninformed sell orders. Easley, Hvdkjaer and O'Hara (2002) propose the following MLE estimation to estimate the parameter vector $\Theta \equiv \{\alpha, \mu, \epsilon_b, \epsilon_s, \delta\}$

$$\begin{aligned} L(\Theta|B,S) &= (1-\alpha)\, e^{-\epsilon_b} \frac{\epsilon_b^B}{B!} e^{-\epsilon_s} \frac{\epsilon_s^S}{S!} \\ &\quad + \alpha\delta e^{-\epsilon_b} \frac{\epsilon_b^B}{B!} e^{-(\mu+\epsilon_s)} \frac{(\mu+\epsilon_s)^S}{S!} \\ &\quad + \alpha(1-\delta)\, e^{-(\mu+\epsilon_s)} \frac{(\mu+\epsilon_b)^B}{B!} e^{-\epsilon_s} \frac{\epsilon_s^S}{S!} \end{aligned} \tag{5}$$

where B and S represent total buy trades and sell trades for the day respectively. Given the above specifications, the probability of information-based trade, PIN, is

$$PIN = \frac{\alpha\mu}{\alpha\mu + \epsilon_b + \epsilon_s}. \tag{6}$$

With some independence assumptions across trading days, the likelihood function (5) becomes

$$L\left(\Theta|\left(B_i, S_i\right)_{i=1}^{i=N}\right) = \prod_{i=1}^{N} L(\Theta|B_i, S_i). \tag{7}$$

The problem with estimation of PIN measure is that later years (since 2001), the number of buy and sell orders becomes extremely large, particularly for some NASDAQ stocks. One way to solve this problem, as in Vega (2006), is to impose the constraint that the arrival rates of informed and uninformed orders are the same,

$$\epsilon_b = \epsilon_s = \epsilon, \tag{8}$$

hence we estimate a modified version of (5),

$$L(\Theta|B,S) = (1-\alpha)\, e^{-2\epsilon} \frac{\epsilon^{B+S}}{B!S!} + \alpha\delta e^{-(\mu+2\epsilon)} \frac{\epsilon^{B}(\mu+\epsilon)^{S}}{B!S!} + \alpha(1-\delta)\, e^{-(\mu+2\epsilon)} \frac{\epsilon^{S}(\mu+\epsilon)^{B}}{B!S!} \tag{9}$$

and consequently, the probability of informed trading, PIN, is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon}. \tag{10}$$

It is interesting to note that the probability that an information event occurs (α) and the rate of informed trade arrivals (μ) enter PIN as a product term ($\alpha\mu$). Although α and μ may be individually estimated rather imprecisely, since estimation errors in these two parameters are usually strongly negatively correlated, the resulting PIN estimate is quite precise. In addition, the variation in α and μ are offsetting, making PIN a much stable measure bounded between 0 and 1.

Duarte and Young (2007) extend (5) to take into account large buy and sell volatilities, and pervasive positive correlation between buy and sell orders. Their model allows the possibility of order flow shocks and different distributions of the number of the buyer-initiated informed trades and the number of the seller-initiated informed trades. With such an extension, one may estimate an adjusted version of the probability of informed trading (AdjPIN) as

$$AdjPIN = \frac{\alpha \times [(1-\delta)\times\mu_b + \delta\times\mu_s]}{\alpha \times [(1-\delta)\times\mu_b + \delta\times\mu_s] + (\Delta_b+\Delta_s)\times\left[\alpha\times\theta' + (1-\alpha)\times\theta\right] + \epsilon_b + \epsilon_s} \tag{11}$$

where the additional parameter θ denotes the probability of symmetric order flow shocks conditional on no arrival of private information event, and θ' denotes the probability of symmetric order flow shocks conditional on the arrival of private information. Δ_b and Δ_s denote the additional arrival rate of buy orders and sell orders conditional on the arrival of the symmetric order flow shocks. Duarte and Young (2007) simplify (11) by restricting $\theta = \theta'$. To reduce the sheer volume of calculations, and to estimate a relatively parsimonious model with fewer parameters, we further impose the constraints that $\mu_b = \mu_s = \mu$ and $\Delta_b = \Delta_s = \Delta$. According to Duarte and Young (2007), the adjusted-PIN estimated with these constraints generate similar results to their full-fledged model.

Thus, the adjusted-PIN measure we estimate is specified as:

$$AdjPIN = \frac{\alpha \times \mu}{\alpha \times \mu + 2 \times \Delta \times \theta + 2 \times \epsilon}$$

In addition to causing large order imbalance, informed-trading will also force the market maker to increase the bid-ask spread. In the structural model of intra day trading costs proposed by Madhavan et. al. (1997), the price change can be captured by:

$$p_t - p_{t-1} = (\phi + \theta)x_t - (\phi + \rho\theta)x_{t-1} + u_t$$

Here x_t is the sign of the order flow (1: trade at ask, -1: trade at bid, 0: trade between bid and ask), ϕ is the market maker's cost of supplying liquidity, ρ is the autocorrelation of the order flow, and θ captures the sensitivity of beliefs to unexpected order flows or the degree of private information. θ is therefore known as the information asymmetry component of the bid-ask spread and serves as an alternative measure of private information events. ϕ, ρ and θ will be jointly estimated with transaction level data using GMM on a quarterly basis.

To the extent that significant information events usually lead to abnormal trading in a stock, our last alternative measure is a measure of abnormal turnover ($aturn$) calculated in a similar fashion as in Chordia, Huh, and Subrahmanyam (2007). At the end of month t, for each stock, we estimate a regression in a 36-month rolling window $[t - 35, t]$:

$$turn = a + bx + \varepsilon$$

where $turn$ is monthly stock turnover defined as the ratio between total number of shares traded during the month and total number of shares outstanding, and x is a vector of adjustment regressors including 11 monthly dummy variables for months (January - November) as well as the linear and quadratic time-trend variables. The residual term for month t, ε_t, after standardization is the measure of abnormal turnover ($aturn$).

We find that past performance predicts future performance better among funds trading in stocks affected more by information events: past winners earn a risk adjusted excess return (four factor alpha) of 0.35% per month in the future. Most of that superior performance comes from impatient informed trading. We also find that informed trading is more important for growth-oriented funds while liquidity provision is more important for younger funds with income orientation.

References

[1] Amihud, Yahov, 2002, Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets* 5, 31-56.

[2] Baks, Klass, Jeffrey A. Busse, and T. Clifton Green, 2006, Fund Managers Who Take Big Bets: Skilled or Overconfident, *Working Paper*, Emory University

[3] Berk, Jonathan, and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112(6), 1269–1295.

[4] Bharath, Sreedhar, Paolo Pasquariello and Guojun Wu, 2006, Does asymmetric information drive capital structure decisions?, *Review of Financial Studies*, forthcoming.

[5] Blume, Marshall E., and Roger M. Edelen, 2004, On replicating the S&P 500 index, *Working Paper*, Wharton School.

[6] Blumenthal, Karen, 2007, Grande expectations: A year in the life of Starbuck's stock, Crown Publishing, first edition.

[7] Brown, Stephen, and William Goetzmann, 1995, Performance persistence, Journal of Finance 50(2), 679 - 698.

[8] Brown, Stephen, Steve Hillegeist and Kin Lo, 2004, Conference calls and information asymmetry, J*ournal of Accounting and Economics* 37, 343-366.

[9] Campbell, John Y., Tarun Ramadorai, and Allie Schwartz, 2007, Caught On Tape:Institutional Trading, Stock Returns, and Earnings Announcements, *Journal of Financial Economics*, Forthcoming.

[10] Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

[11] Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, 2000, An examination of the stockholdings and trades of fund managers, *Journal of Financial and Quantitative Analysis* 35, 343–368.

[12] Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey Kubik, 2004, Does fund size erode performance? Liquidity, organizational diseconomies and active money management, *American Economic Review* 94(5), 1276–1302.

[13] Chen, Qi, Itay Goldstein and Wei Jiang, 2006, Price informativeness and investment sensitivity to stock prices, *Review of Financial Studies* 20, 619-650.

[14] Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

[15] Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007, The cross-section of expected trading activity, *Review of Financial Studies* 20, 709-740.

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

[17] Christoffersen, Susan Kerr, Donald B. Keim, and David K. Musto, 2006, Valuable information and costly liquidity: evidence from individual mutual fund trades, *Working Paper*, Wharton School.

[18] Cohen, Randolph B., 2002, Dimensional fund advisors, *Harvard Business School Case* 294-025.

[19] Conrad, Jennifer, Kevin M. Johnson, and Sunil Wahal, 2003, Institutional trading and alternative trading systems, *Journal of Financial Economics* 70, 99–134.

[20] Coval, Joshua D., and Tobias J. Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109(4), 811–841.

[21] Coval, Joshua D., and Erik Stafford, 2006, Asset Fire Sales (and Purchases) in Equity Markets, *Journal of Financial Economics*, Forthcoming.

[22] Cremers, Martijn, and Antti Petajisto, 2006, How Active is Your Fund Manager? a New Measure that Predicts Performance, *Review of Financial Studies*, Forthcoming.

[23] Da, Zhi and Pengjie Gao, 2008, Clientele Change, Liquidity shock, and the Return on Financially Distressed Stocks, *Journal of Financial and Quantitative Analysis*, forthcoming.

[24] Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic based benchmarks, *Journal of Finance* 52, 1035–1058.

[25] Duarte, Jefferson, and Lance Young, 2007, Why is PIN priced? *Journal of Financial Economics*, forthcoming.

[26] Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57, 2185-2221.

[27] Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2004, Factoring information into returns, *Journal of Financial and Quantitative Analysis*, forthcoming.

[28] Easley, David, Nicholas M. Kiefer, and Maureen O'Hara, 1997, One day in the life of a very common stock, *Review of Financial Studies* 10, 805–835.

[29] Easley, David, Nicholas M. Kiefer, Maureen O'Hara, and Joseph Paperman, 1996, Liquidity, information, and less-frequently traded stocks, *Journal of Finance* 51, 1405–1436.

[30] Elton, Edwin J., Martin J. Gruber, Sanjiv Das, and Matthew Hlavka, 1993, Efficiency with costly information: a reinterpretation of evidence from managed portfolios, *Review of Financial Studies* 6, 1-22.

[31] Elton, Edwin J., Martin J. Gruber, Yoel Krasny and Sadi Ozelge, 2006, The effect of the frequency of holding data on conclusions about mutual fund behavior, *Journal of Financial and Quantitative Analysis*, forthcoming.

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

[33] Fama, Eugene F., and Kenneth French, 2008, Mutual fund performance, *Working Paper*.

[34] French, Kenneth, 2008, The cost of active investing, *Journal of Finance* 63(4), 1537-1573.

[35] Glosten, Lawrence R., and Lawrence E. Harris, 1988, Estimating the components of the bid / ask spread, *Journal of Financial Economics* 21, 123-142.

[36] Greenwald, Bruce C.N., Judd Kahn, Paul.D.Sonkin, and Michael Biema, 2001, Value Investing: From Graham to Buffett and Beyond, Wiley Publishing, first edition.

[37] 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.

[38] Grossman, S., and J. Stiglitz, 1980, On the impossibility of informational efficient markets, *American Economic Review* 70, 393-408.

[39] Gruber, Martin J., 1996, Another puzzle: the growth in actively managed mutual funds, , Journal of Finance, Vol. 51, No. 3, July 1996.

[40] Huang, Roger. and Hans R. Stoll, 1996, Dealer versus Auction markets: a paired comparison of execution costs on Nasdaq and the NYSE, *Journal of Financial Economics* 41, 313-358.

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

[42] Kacperczyk, Marcin, and Amit Seru, 2007, Fund Manager Use of Public Information: New Evidence on Managerial Skills, *Journal of Finance* 62, 485-528.

[43] Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2012.

[44] Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2007, Unobesrved actions of mutual funds, *Review of Financial Studies*, forthcoming.

[45] Keim, Donald B., 1999, An analysis of mutual fund design: the case of investing in small-cap stocks, *Journal of Financial Economics* 51, 173-194.

[46] Keim, Donald B., and Ananth Madhavan, 1995, Anatomy of the trading process: Empirical evidence on the behavior of institutional traders, *Journal of Financial Economics* 37, 371-398.

[47] Korajczyk, Robert, and Ronnie Sadka, 2005, Are momentum trading robust to trading costs? *Journal of Finance* 59, 1039-1082.

[48] Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551-2595.

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

[50] Madhavan, Ananth, Matthew Richardson and Mark Roomans, 1997, Why do security prices change? A transaction-level analysis of NYSE stocks, *Review of Financial Studies* 10, 1035-1064.

[51] Mamaysky, Harry, Matthew Spiegel, and Hong Zhang, 2007a, Estimating the dynamic of mutual fund alphas and betas, Review of Financial Studies 21(1), 233 - 264.

[52] Mamaysky, Harry, Matthew Spiegel, and Hong Zhang, 2007b, Improved forecasting of mutual fund alphas and betas, Review of Finance 11, 359-400.

[53] Pástor, Ľuboš, and Robert F. Stambaugh, 2002a, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 63,315-349.

[54] Pástor, Ľuboš, and Robert F. Stambaugh, 2002b, Investing in equity mutual funds, *Journal of Financial Economics* 63,351-380.

[55] Pástor, Ľuboš, and Robert F. Stambaugh, 2003,Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

[56] Schultz, Paul, 2007, When is stock picking by mutual funds successful?, *Working Paper*, University of Notre Dame.

[57] Sirri, Erik R., and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *Journal of Finance* 53, 1589-1622.

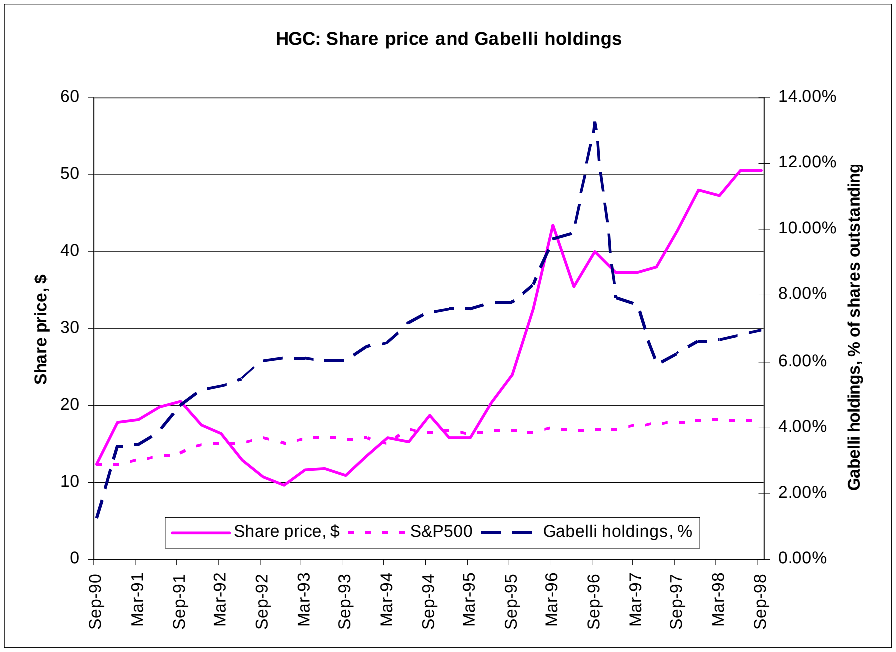
[58] Vega, Clara, 2006, Stock price reaction to Public and private information, *Journal of Financial Economics* 82, 103-133.

[59] Wermers, Russ, 1999, Mutual Fund Herding and the Impact on Stock Prices, *Journal of Finance* 54, 581-622.

Figure 1: Share price and mutual fund holdings

Panel A plots the share price of Hudson General Corp (HGC) and Gabelli Fund's holdings of HGC (as a percentage of total number of shares outstanding) from September 1990 to September 1998. Panel B plots the share prices of Starbucks (SBUX) from June to December 2005 (price is normalized so that the end-of-July-price is 1) and Putnam Voyager Fund's holdings of Starbucks (as a percentage of total number of shares outstanding) at the end of June, September and December.

A: Share price of Hudson General Corp (HGC) and Gabelli's Holdings



B: Share Price of Starbuck (SBUX, normalized) and Putnam Voyager Fund's Holdings

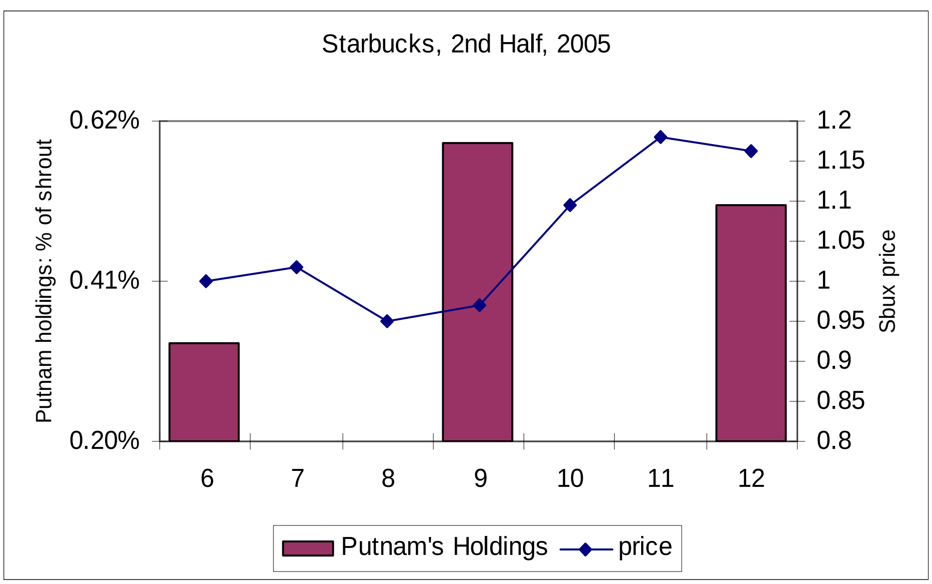


Figure 2: Graphic illustration of the Buy, Sell and Hold portfolios

This figure illustrates the three portfolios defined in Section 1 for the purpose of Characteristics Selectivity (CS) decomposition. Let the smaller pie represent the fund's holding at the end of quarter t-1 and the bigger pie represent the fund's holding at the end of quarter t. The intersection of the two pies thus represents the Hold portfolio which contains stocks untouched by the fund during quarter t. The portion of the smaller pie excluding the Hold portfolio represents the Sell portfolio which contains stocks sold by the fund during quarter t. The portion of the bigger pie excluding the Hold portfolio represents the Buy portfolio which contains stocks recently bought by the fund during quarter t.

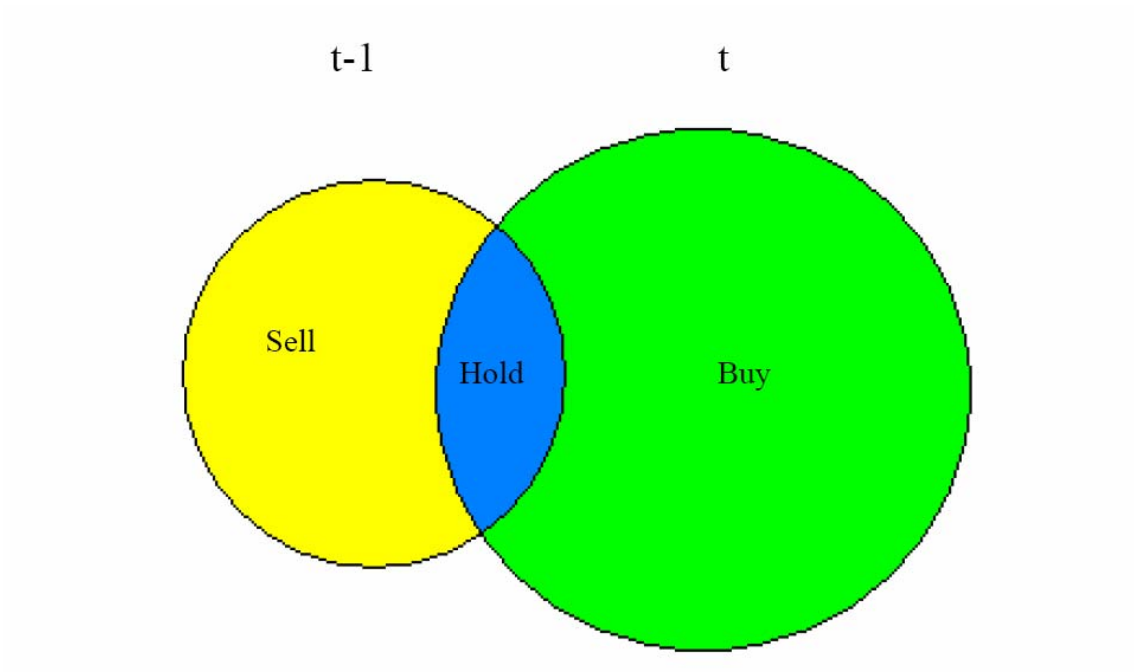


Table 1: Breakdown of mutual fund sample over time

We report the breakdown of our mutual fund sample by the self-reported investment objectives. Consistent with prior literature on actively managed mutual funds, we exclude all index funds, lifecycle mutual funds, bond funds, hybrid funds, sector funds, and international funds. We only keep funds that are self-reported as aggressive growth (AGG), growth (GROWTH) or growth and income (GNI). To ensure our sample of mutual funds are relatively active, we also exclude fund / quarter observations with quarterly turnover less than 10% or if the fund trades less than 10 during that quarter. Finally, we only include fund / quarter observations for which the fund holdings at the end of previous quarter are also available so that holding changes can be computed over consecutive quarters. The CDA/Spectrum mutual fund holding data are matched to CRSP Mutual Fund data via the MFLINKS database.

year	# of funds per qtr	AGG	GROWTH	GNI
1983	132	35	57	40
1984	163	38	73	52
1985	201	44	98	59
1986	234	43	125	66
1987	291	59	156	76
1988	328	73	173	82
1989	283	57	151	75
1990	293	59	157	77
1991	327	73	172	82
1992	397	84	217	96
1993	438	95	242	102
1994	353	65	208	80
1995	353	59	194	100
1996	468	54	271	142
1997	557	64	337	157
1998	913	88	586	238
1999	1291	125	856	310
2000	1843	190	1182	472
2001	1431	159	913	359
2002	1775	201	1106	468
2003	1776	181	1116	480
2004	1459	130	911	419
All	696	90	423	183

Table 2: Empirical validations of the Characteristics Selectivity (CS) decomposition

Panel A reports type of mutual fund trades and the average order imbalances. For each fund in our sample, we examine changes in its holdings over two consecutive quarters and categorize them into four groups: (1) "Open" (defined as holdings increase from zero to positive); (2) "Close" (defined as holdings decrease from positive to zero); (3) "Increase" (defined as holdings increase but not from zero); and (4) "Decrease" (defined as holdings decrease, but not to zero). For each group, we report the average dollar holding change as a percentage of the total dollar holding change of the fund (computed using prices at quarter-end), the average order-imbalance measure and the associated *t*-value. The average order-imbalance measure is computed as the difference between total numbers of buyer-initiated shares bought and total numbers of seller-initiated shares sold, divided by total number of shares traded during the quarter; the resulting number is then cross-sectionally demeaned. The sampling period is from 1983 to 2004.

Panel B provides two examples to illustrate the decomposition of the mutual fund stock selection skill. We decompose the mutual fund Characteristics Selectivity (CS) measure (Daniel et al., 1997) for DFA US Micro-Cap fund (FUNDNO=16500 in CDA/Spectrum S-12 mutual fund holding database) and Index funds a group (fund whose name contains any of the following: "INDEX," "INDE," "INDX," "S&P," "DOW JONES," "MSCI" or "ISHARE"). Specifically, the CS measure is decomposed into: $CS = CS^{O} + CS^{adj} + CS^{inf} + CS^{liq}$, where CS^{O} is the old component; CS^{adj} is an adjustment component due to fund inflows; CS^{inf} and CS^{liq} are the informed trading and liquidity provision components, respectively. The sampling period is from 1983 to 2004. *t*-values associated with the average measures are reported in *italic*s.

Panel C reports the percentage of total cross-sectional variation in the total "Characteristic Selectivity" (*CS*) measure (DGTW, 1997) explained by its four components: the old component (CS^{O}), the adjustment component (CS^{adj}), the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) in a variance decomposition framework outlined in the paper. We perform the variance decomposition on the full sample and on each style subsample. The *t*-values associated with the percentages are reported in *italic*s, using the weighted least squares (WLS) method. The sampling period is from 1983 to 2004.

Panel D reports the average components of CS measure (both contemporaneous and next-quarter) on mutual fund portfolios sorted on funds' return gaps. The return gaps are computed based on Kacperczyk, Sialm and Zheng (2007).

Panel A: Type of mutual fund trades and average order imbalances

trade type	ALL			AGG			GROWTH			GNI		
	% of all trades	oimb	t-value	% of all trades	oimb	*t*-value	% of all trades	oimb	*t*-value	% of all trades	oimb	*t*-value
Open	30.6%	0.31%	*4.09*	34.5%	0.36%	*3.19*	31.2%	0.14%	*1.61*	27.7%	0.61%	*7.21*
Close	26.7%	-0.27%	*-4.73*	30.8%	-0.15%	*-1.32*	27.4%	-0.26%	*-3.63*	24.0%	-0.30%	*-3.75*
Increase	22.8%	0.48%	*9.27*	17.5%	0.48%	*5.86*	22.0%	0.55%	*8.37*	26.4%	0.37%	*5.07*
Decrease	19.9%	1.27%	*18.06*	17.3%	1.69%	*14.51*	19.4%	1.34%	*15.67*	21.9%	0.84%	*11.17*

Panel B: Characteristics Selectivity (CS, measured in bps per quarter before fees) decomposition for DFA US Micro-Cap fund and index funds as a group

	Total CS (=1+2+3)	Old CS^O (1)	Adj CS^{adj} (2)	Trade CS^T (3=3a+3b)	Info trading CS^{inf} (3a)	Liquidity Prov CS^{liq} (3b)	CS^{inf} - CS^{liq} (3a)-(3b)
DFA US Micro-Cap:							
Value (bps)	36.1	19.3	-4.2	21	0.5	20.5	-20.0
t-value	*1.72*	*0.89*	*-0.64*	*1.3*	*0.06*	*1.84*	*-1.62*
Index Funds:							
Value (bps)	0	24.9	3.2	-28.1	-34.6	6.4	-41.0
t-value	*0*	*0.93*	*0.5*	*-1.11*	*-2.19*	*0.36*	*-1.86*

Panel C: *CS* decomposition across styles

	Total CS	Old CS^O	Adj CS^{adj}	Info trading CS^{inf}	Liquidity Prov CS^{liq}	Old CS^O	Adj CS^{adj}	Info trading CS^{inf}	Liquidity Prov CS^{liq}
	average (bps per quarter)					% of variation in total *CS* explained			
All	23.5	13.9	-1.8	3.6	8.8	56.8%	-2.5%	37.2%	8.4%
	1.91	*1.19*	*-2.38*	*0.55*	*1.5*	*127.2*	*-15.3*	*120.9*	*24.4*
AGG	54.7	30.8	-0.9	17.4	7.3	52.1%	-1.0%	44.9%	4.0%
	1.74	*1.22*	*-0.42*	*1.02*	*0.66*	*44.7*	*-2.7*	*55.2*	*4.2*
Growth	25.7	14.8	-2.7	5.6	8.0	55.7%	-3.0%	37.0%	10.2%
	1.80	*1.24*	*-2.34*	*0.64*	*1.27*	*96*	*-14.9*	*95*	*22.3*
GNI	12.0	3.0	0.1	-1.7	10.7	54.1%	-2.2%	37.0%	11.1%
	1.03	*0.23*	*0.03*	*-0.29*	*1.70*	*56.8*	*-5.4*	*55.9*	*16.6*

Panel D: Return Gaps and components of *CS* measure (in bps before fees per quarter)

Return Gap	D1. Contemporaneous					D2. Next Quarter				
	RG (monthly)	CS^O	CS^{adj}	CS^{inf}	CS^{liq}	Four-factor alpha	CS^O	CS^{adj}	CS^{inf}	CS^{liq}
Low	-120.3	141.2	-8.4	62.0	-13.8	-47.8	13.3	-0.9	10.0	5.9
		7.31	*-3.48*	*2.63*	*-1.39*	*-3.88*	*0.82*	*-0.43*	*0.83*	*0.62*
2	-30.3	62.2	-3.2	15.6	2.2	-31.2	16.1	-2.1	0.8	6.5
		4.72	*-2.52*	*2.07*	*0.34*	*-3.07*	*1.26*	*-1.08*	*0.12*	*0.94*
3	-7.3	11.0	-3.0	-0.1	11.5	-21.7	10.7	0.3	5.7	6.7
		0.87	*-1.84*	*-0.01*	*1.83*	*-2.19*	*0.86*	*0.22*	*0.82*	*1.07*
4	14.5	-32.6	1.4	-11.6	17.8	-19.8	16.6	-2.6	4.4	5.5
		-2.64	*0.66*	*-1.92*	*2.71*	*-2.10*	*1.24*	*-1.79*	*0.48*	*0.86*
High	74.9	-101.2	3.9	-34.7	22.9	2.0	23.2	-4.1	8.6	16.9
		-7.03	*2.03*	*-3.43*	*2.86*	*0.14*	*1.37*	*-1.97*	*0.76*	*1.97*
High - Low	195.2	-242.4	12.4	-96.7	36.7	49.9	9.9	-3.2	-1.4	10.9
		-15.10	*4.13*	*-4.51*	*4.15*	*4.84*	*0.81*	*-1.04*	*-0.21*	*1.81*

Table 3: Risk-adjusted *quarterly* fund returns across *trade_PIN* sorted fund deciles

In each quarter and for each fund, we compute a *trade_PIN* variable by value-weighing PIN of stocks traded by the fund during the quarter using the dollar value of the trade. At the end of each quarter from 1983 to 2004, we sort all mutual funds in our sample into deciles according to their *trade_PIN*s and examine the risk-adjusted fund portfolio return in the next four quarters within each decile. We consider two methods for risk adjustment. The first method uses either the Fama-French three-factor model or the Four-factor model (Fama-French three factors augmented by Carhart's momentum factor). The factor loadings are computed in a rolling window using fund returns during the past 36 months. The second method uses a characteristics-based risk adjustment. For each stock held by the fund at the end of each quarter, we compute its future excess returns over the returns of a characteristics-based benchmark portfolio that is matched to the stock along size, book-to-market and past return characteristics. These excess returns are then value-weighed across stocks at the fund level using dollar value of the stock holding to arrive at a characteristics-adjusted pseudo fund return.

trade_PIN	Four-factor adjusted returns (after fee)				Three-factor adjusted returns (after fee)				Chracteristic-adjusted returns (CS) (before fee)			
in Qtr t	Qtr t+1	Qtr t+2	Qtr t+3	Qtr t+4	Qtr t+1	Qtr t+2	Qtr t+3	Qtr t+4	Qtr t+1	Qtr t+2	Qtr t+3	Qtr t+4
Low	-0.0038	-0.0037	-0.004	-0.0038	-0.0040	-0.0044	-0.0048	-0.0044	-0.0003	0.0004	-0.0002	0.0002
	-3.45	-3.51	-3.45	-3.75	-3.73	-4.22	-4.04	-4.41	-0.29	0.42	-0.22	0.25
2	-0.0040	-0.0039	-0.0044	-0.0033	-0.0035	-0.0034	-0.0040	-0.0037	0.0011	0.0015	0.0007	0.0005
	-4.01	-4.34	-4.25	-3.47	-3.60	-3.94	-3.85	-3.78	0.92	1.4	0.7	0.52
3	-0.0032	-0.003	-0.0039	-0.0041	-0.0030	-0.0027	-0.0033	-0.0037	0.0012	0.0007	0.0007	-0.0003
	-2.87	-3.03	-4.46	-3.81	-3.01	-2.80	-3.66	-3.43	1.04	0.74	0.72	-0.32
4	-0.0031	-0.0036	-0.0037	-0.0037	-0.0023	-0.0027	-0.0031	-0.0030	0.0010	0.0012	0.0008	0.000
	-3.09	-3.66	-4.31	-3.38	-2.24	-2.96	-3.48	-2.84	1.01	1.19	0.84	0.01
5	-0.0017	-0.0026	-0.0035	-0.0034	-0.0006	-0.0015	-0.0025	-0.0021	0.0029	0.0014	0.0006	0.0007
	-1.57	-2.39	-3.24	-3.21	-0.61	-1.42	-2.50	-2.01	2.17	1.15	0.51	0.64
6	-0.0027	-0.0017	-0.0019	-0.0031	-0.0018	-0.0005	-0.0004	-0.0017	0.0032	0.0023	0.0016	-0.0005
	-2.18	-1.33	-1.36	-2.68	-1.48	-0.41	-0.34	-1.59	2.07	1.58	1.26	-0.37
7	-0.0025	-0.0022	-0.0016	-0.0022	-0.0008	-0.0006	-0.0001	-0.0009	0.0028	0.0016	0.0016	0.0006
	-1.8	-1.67	-1.24	-1.58	-0.61	-0.46	-0.08	-0.66	1.52	1.04	1.19	0.41
8	-0.0019	-0.0003	-0.0008	-0.0018	-0.0007	0.0003	-0.0002	-0.0010	0.0031	0.0024	0.0024	0.0014
	-1.16	-0.22	-0.52	-1.14	-0.45	0.20	-0.10	-0.67	1.73	1.41	1.53	0.85
9	-0.0018	-0.0014	-0.0018	-0.0018	-0.0007	-0.0002	-0.0005	-0.0005	0.0035	0.0016	0.0006	-0.0004
	-1.06	-0.82	-1.11	-1.18	-0.42	-0.14	-0.31	-0.34	1.75	0.84	0.35	-0.21
High	0.0011	0.0009	-0.0005	-0.0003	0.0015	0.0012	0.0002	0.0003	0.0050	0.0044	0.0017	0.0011
	0.85	0.67	-0.36	-0.21	1.27	0.91	0.13	0.23	2.7	2.75	1.2	0.68
High -	0.0048	0.0046	0.0035	0.0035	0.0056	0.0056	0.0050	0.0047	0.0053	0.0040	0.0019	0.0009
Low	3.15	2.98	2.24	2.26	3.42	3.47	3.06	2.89	2.87	2.55	1.29	0.52

Table 4: Robustness checks

Panel A first reports the next quarter PIN-risk-adjusted returns on decile portfolios of mutual funds sorted on *trade_PIN*. To control for the systematic risk associated with high-PIN stocks, we compute a five-factor-adjusted mutual fund return by augmenting the benchmark four-factor model with a PIN risk factor. The PIN risk factor is constructed as the high-PIN decile portfolio return minus the low-PIN decile portfolio return. To control for PIN characteristics risk, we construct characteristics benchmark portfolios by matching along size, book-to-market, past return and PIN simultaneously.

Panel A also reports the next quarter risk-adjusted returns on decile portfolios of mutual funds constructed using alternative measures of information events. These measures include: the information asymmetry component of the PIN (adjPIN, Duarte and Young, 2007); the information asymmetry component of the bid-ask spread (theta, Madhavan, Richardson and Roomans,1997) and the abnormal turnover in stock trading (aturn, Chordia, Huh and Subrahmanyam, 2006). Trade_adjPIN, trade_theta and trade_aturn are then computed in the same fashion as trade_PIN to measure the average amount of information events on stocks traded by the mutual funds. While the factor-adjusted returns (alpha) are before fee, the characteristics-adjusted returns (CS) are after fee.

Panel B reports the average PIN and past one-year return of stocks bought / sold by mutual funds across *trade_PIN* sorted deciles. For each fund, we compute the value-weighted average PIN and past one-year return of stocks in the "Buy" portfolio (stocks recently bought by the fund) and the "Sell" portfolio (stocks recently sold by the fund). These PIN and past returns are then averaged across funds and across time. *t*-values associated with the average measures are reported in *italic*s.

Panel A: PIN-risk-adjusted fund returns and alternative measure of information events. CS measures are before fees and alphas are after fees, per quarter

Portfolio	Control for PIN risk: Sorted on trade_PIN in Qtr t		Alternative measures of information events: Sorted on trade_adjPIN in Qtr t		Sorted on trade_Theta in Qtr t		Sorted on trade_aturn in Qtr t	
	5f alpha Qtr t+1	CS* Qtr t+1	4f alpha Qtr t+1	CS Qtr t+1	4f alpha Qtr t+1	CS Qtr t+1	4f alpha Qtr t+1	CS Qtr t+1
Low	-0.0034	-0.0008	-0.0033	0.0010	-0.0035	0.0005	-0.0045	0.0003
	-2.97	-0.75	-3.20	0.79	-3.18	0.44	-3.46	0.20
2	-0.0032	0.0010	-0.0034	0.0002	-0.0031	0.0022	-0.0031	0.0012
	-2.98	0.79	-3.09	0.13	-3.11	1.71	-3.10	0.92
3	-0.0026	0.0006	-0.0038	0.0010	-0.0029	0.0016	-0.0032	0.0021
	-2.28	0.51	-3.44	0.79	-2.88	1.41	-3.18	1.78
4	-0.0026	0.0014	-0.0035	0.0014	-0.0033	0.0011	-0.0035	0.0011
	-2.57	1.10	-2.85	1.05	-3.64	0.97	-3.95	1.03
5	-0.0012	0.0021	-0.0022	0.0033	-0.0029	0.0013	-0.0030	0.0010
	-1.09	1.49	-1.86	2.26	-2.44	1.09	-2.82	0.85
6	-0.0021	0.0028	-0.0011	0.0028	-0.0022	0.0020	-0.0024	0.0023
	-1.60	1.63	-0.88	1.81	-1.81	1.26	-1.98	1.55
7	-0.0019	0.0032	-0.0019	0.0030	-0.0017	0.0030	-0.0026	0.0016
	-1.28	1.59	-1.33	1.85	-1.18	1.84	-1.76	0.99
8	-0.0019	0.0022	-0.0015	0.0039	-0.0006	0.0049	-0.0007	0.0055
	-1.05	1.23	-1.02	2.05	-0.39	2.55	-0.53	2.97
9	-0.0015	0.0019	-0.0012	0.0037	0.0002	0.0043	-0.0008	0.0031
	-0.86	0.88	-0.81	2.16	0.14	2.17	-0.50	1.60
High	0.0011	0.0035	0.0015	0.0058	0.0006	0.0055	0.0004	0.0051
	0.86	1.97	1.14	3.13	0.40	2.70	0.29	2.56
High-Low	0.0045	0.0043	0.0048	0.0048	0.0041	0.0050	0.0049	0.0048
	2.85	2.11	3.13	2.97	2.48	2.66	2.70	2.23

Panel B: Average PIN and past one-year return of stocks bought / sold by mutual funds across trade_PIN sorted deciles

trade_PIN	PIN: Buy	Sell	Buy-sell	t-value	Past One-year Return: Buy	Sell	Buy-sell	t-value
Low	0.1123	0.1144	-0.0021	*-3.39*	20.90%	24.30%	-3.40%	*-5.74*
2	0.1235	0.1250	-0.0015	*-2.50*	23.20%	26.40%	-3.20%	*-5.18*
3	0.1298	0.1323	-0.0026	*-4.26*	25.10%	26.80%	-1.70%	*-2.96*
4	0.1372	0.1383	-0.0011	*-1.82*	26.70%	29.60%	-2.90%	*-4.08*
5	0.1446	0.1455	-0.0009	*-1.26*	28.50%	32.10%	-3.60%	*-4.48*
6	0.1529	0.1547	-0.0018	*-1.37*	32.10%	36.20%	-4.10%	*-3.66*
7	0.1631	0.1642	-0.0012	*-1.22*	36.20%	39.60%	-3.40%	*-3.13*
8	0.1767	0.1779	-0.0012	*-1.15*	39.20%	43.80%	-4.60%	*-3.6*
9	0.1936	0.1953	-0.0017	*-1.45*	40.20%	46.10%	-5.90%	*-4.65*
High	0.2294	0.2301	-0.0007	*-0.41*	34.30%	46.60%	-12.30%	*-8.93*
High-Low	0.1171	0.1157			13.47%	22.36%		
	64.53	*53.82*			*6.37*	*8.42*		

Table 5: Characteristic Selectivity (CS) measure decomposition
across *Trade_PIN* sorted fund deciles

In each quarter and for each fund, we compute a *trade_PIN* variable by value-weighing the probability of information trading (PIN) of stocks traded by the fund during the quarter using the dollar values of the trade as weights. At the end of each quarter from 1983 to 2004, we sort all mutual funds in our sample into deciles according to their *trade_PIN*s and decompose the Characteristic Selectivity (CS) measure within each decile. The CS measure and its component are reported in before fee basis points (bp) per quarter. *t*-values associated with the average measures are reported in *italic*s.

Trade_PIN Qrt t	Total CS (=1+2+3) Qtr t+1	Old CS^{O} (1) Qtr t+1	Adj CS^{adj} (2) Qtr t+1	Trade CS^{T} (3=3a+3b) Qtr t+1	Info trading CS^{inf} (3a) Qtr t+1	Liquidity Prov CS^{liq} (3b) Qtr t+1	CS^{inf} - CS^{liq} (3a)-(3b) Qtr t+1
Low	-2.9	-7.6	-0.4	3.4	-12.1	16.2	-28.3
	-0.29	*-0.70*	*-0.20*	*0.42*	*-2.02*	*2.57*	*-2.71*
2	11.4	10.4	-0.7	2.6	-6.4	8.9	-15.3
	0.92	*0.87*	*-0.47*	*0.40*	*-0.93*	*1.28*	*-1.48*
3	11.8	9.2	-1.0	5.5	-5.5	9.8	-15.3
	1.04	*0.81*	*-0.69*	*0.88*	*-0.78*	*1.65*	*-1.20*
4	10.3	8.3	-1.1	5.8	-3.5	6.0	-9.5
	1.01	*0.76*	*-0.81*	*0.76*	*-0.54*	*0.89*	*-0.91*
5	28.6	23.3	-2.3	6.4	2.5	5.5	-2.9
	2.17	*1.80*	*-1.52*	*0.72*	*0.31*	*0.74*	*-0.16*
6	31.9	19.2	-2.4	18.7	5.6	9.4	-3.8
	2.07	*1.20*	*-1.52*	*1.69*	*0.62*	*1.10*	*-0.10*
7	28.4	19.4	-0.9	17.2	9.7	0.2	9.5
	1.52	*1.17*	*-0.56*	*1.18*	*0.89*	*0.03*	*0.38*
8	30.9	14.6	-3.4	25.0	8.6	13.7	-5.1
	1.73	*0.87*	*-2.05*	*2.26*	*0.84*	*1.64*	*-0.10*
9	35.1	15.7	-2.7	26.6	16.8	7.6	9.2
	1.75	*0.82*	*-1.35*	*2.38*	*1.70*	*0.77*	*0.88*
High	50.0	26.5	-3.2	31.2	20.4	10.4	10.0
	2.70	*1.43*	*-1.40*	*2.83*	*2.25*	*1.37*	*0.28*
High - Low	52.9	34.1	-2.8	27.8	32.5	-5.8	38.4
	2.87	*1.94*	*-0.93*	*2.26*	*3.50*	*-0.68*	*2.16*

Table 6: Performance of fund deciles sorted on previous alphas

At the end of each month from 1983 to 2004, we sort all mutual funds in our sample into deciles according to their OLS four-factor alphas ("4f", Fama-French three factors augmented by Carhart's momentum factor) estimated using the previous 5 years data. Each fund decile portfolio is then held for 1 month and rebalanced next month. We then compute the three-factor ('3f") and four-factor ('4f") monthly alphas (after fee) on these rebalanced fund decile portfolios. We also repeat the exercise after including the three filters described in Mamaysky, Spiegel and Zhang (2007). For any fund to be included in any decile the following must be true: (1) the absolute value of alpha must be less than 2% per month, (2) the CAPM beta must be in between 0 and 2, and (3) in the previous month the forecasted alpha the difference between the realized return and the market return must have the same sign. *t*-values associated with the alphas and factor loadings are reported in *italic*s.

Sort on previous	Without filters						With filters described in MSZ (2007)					
4f alpha	4f alpha	MKTRF	SMB	HML	UMD	3f alpha	4f alpha	MKTRF	SMB	HML	UMD	3f alpha
Low alpha	-0.0032	0.8877	0.3081	0.0749	-0.0130	-0.0034	-0.0063	0.9238	0.2760	0.1656	-0.0465	-0.0067
	-3.48	*38.26*	*10.81*	*2.18*	*-0.64*	*-3.69*	*-5.50*	*32.45*	*7.89*	*3.93*	*-1.88*	*-5.98*
2	-0.0011	0.8650	0.1570	0.0679	0.0180	-0.0010	-0.0037	0.9047	0.1352	0.1409	-0.0093	-0.0038
	-2.53	*77.15*	*11.39*	*4.10*	*1.85*	*-2.19*	*-5.22*	*51.72*	*6.29*	*5.45*	*-0.61*	*-5.47*
3	-0.0012	0.8561	0.1185	0.0820	0.0264	-0.0009	-0.0023	0.8869	0.1152	0.1207	0.0104	-0.0022
	-3.09	*90.94*	*10.24*	*5.89*	*3.22*	*-2.44*	*-4.05*	*61.54*	*6.51*	*5.66*	*0.83*	*-3.97*
4	-0.0008	0.8356	0.0911	0.1047	0.0109	-0.0007	-0.0021	0.9003	0.0817	0.1066	0.0089	-0.0020
	-1.95	*84.43*	*7.49*	*7.15*	*1.26*	*-1.73*	*-4.21*	*73.22*	*5.41*	*5.86*	*0.83*	*-4.13*
5	-0.0011	0.8385	0.0587	0.1088	0.0071	-0.0010	-0.0008	0.8828	0.0487	0.0967	-0.0091	-0.0009
	-2.75	*85.63*	*4.88*	*7.51*	*0.83*	*-2.64*	*-1.55*	*70.05*	*3.14*	*5.19*	*-0.83*	*-1.76*
6	-0.0008	0.8087	0.0549	0.0927	0.0199	-0.0006	0.0000	0.8889	0.0560	0.0934	-0.0082	-0.0001
	-1.97	*84.88*	*4.69*	*6.58*	*2.40*	*-1.50*	*-0.06*	*77.35*	*3.97*	*5.50*	*-0.82*	*-0.24*
7	-0.0007	0.8354	0.0677	0.0872	0.0062	-0.0007	0.0004	0.8641	0.0744	0.0529	-0.0135	0.0002
	-1.89	*87.35*	*5.76*	*6.17*	*0.75*	*-1.77*	*0.67*	*65.99*	*4.62*	*2.73*	*-1.18*	*0.44*
8	-0.0007	0.8599	0.0777	0.0674	-0.0044	-0.0007	0.0000	0.8731	0.1290	0.0620	-0.0064	-0.0001
	-1.62	*81.67*	*6.01*	*4.32*	*-0.48*	*-1.76*	*0.01*	*62.52*	*7.52*	*3.00*	*-0.53*	*-0.10*
9	-0.0007	0.9014	0.1378	0.0372	-0.0094	-0.0008	-0.0001	0.8365	0.1718	0.0357	-0.0108	-0.0002
	-1.44	*73.44*	*9.14*	*2.05*	*-0.88*	*-1.66*	*-0.17*	*44.86*	*7.50*	*1.29*	*-0.67*	*-0.31*
High alpha	0.0002	0.9470	0.2641	-0.1135	-0.0085	0.0001	0.0015	0.8192	0.2456	-0.0525	-0.0265	0.0012
	0.21	*49.29*	*11.19*	*-3.99*	*-0.51*	*0.11*	*1.19*	*26.70*	*6.52*	*-1.16*	*-0.99*	*1.01*
High - Low	0.0034	0.0594	-0.0440	-0.1884	0.0045	0.0034	0.0078	-0.1046	-0.0304	-0.2181	0.0201	0.0080
	2.93	*2.06*	*-1.24*	*-4.41*	*0.18*	*3.04*	*4.12*	*-2.23*	*-0.53*	*-3.14*	*0.49*	*4.32*

Table 7: Persistence of mutual fund alphas and *trade_PIN*s

At the end of each quarter from 1983 to 2004, we conduct 2 by 5 double sorts based on *trade_PIN*s first and then on OLS four-factor alphas ("4f", Fama-French three factors augmented by Carhart's momentum factor) estimated using the previous 5 years data in the second sort. Each fund portfolio is then held for 1 month and rebalanced next month. In Panel A, we report the four-factor ('4f") monthly alphas (after fees) on these rebalanced fund portfolios. We also repeat the exercise after including the three filters described in Mamaysky, Spiegel and Zhang (2007). For any fund to be included in any decile the following must be true: (1) the absolute value of alpha must be less than 2% per month, (2) the CAPM beta must be in between 0 and 2, and (3) in the previous month the forecasted alpha the difference between the realized return and the market return must have the same sign. Panel B reports the three-factor alphas after fee, alphas before fees and the growth rates in dollar fees (we apply the three filters described in Mamaysky, Spiegel and Zhang, 2007). *t*-values are reported in *italic*s.

Panel A: Monthly four-factor alphas (after fee) and filters described in MSZ (2007)

Trade_Pin at Qtr t-1	Previous 4f alpha	Without filters						With filters described in MSZ (2007)					
		Trade_Pin	4f alpha	mktrf	smb	hml	umd	Trade_Pin	4f alpha	mktrf	smb	hml	umd
low	low	0.1295	-0.0009	0.9209	-0.0519	0.0217	-0.0035	0.1292	-0.0014	0.9343	-0.0654	0.0167	-0.0253
			-1.86	*78.36*	*-3.59*	*1.25*	*-0.34*		*-1.72*	*47.39*	*-2.69*	*0.57*	*-1.47*
low	medium	0.1280	-0.0011	0.9465	-0.0723	0.0705	-0.0002	0.1287	-0.0007	0.9350	-0.0799	0.0728	-0.0309
			-2.81	*94.40*	*-5.86*	*4.75*	*-0.02*		*-1.34*	*72.13*	*-5.01*	*3.79*	*-2.75*
low	high	0.1294	-0.0008	0.9743	-0.0004	-0.0050	-0.0803	0.1303	0.0006	0.9420	-0.0098	-0.0229	-0.0798
			-1.24	*63.09*	*-0.02*	*-0.22*	*-5.98*		*0.63*	*41.77*	*-0.35*	*-0.68*	*-4.07*
	high - low	-0.0001	0.0001	0.0534	0.0515	-0.0267	-0.0767	0.0011	0.0019	0.0077	0.0555	-0.0396	-0.0545
			0.17	*3.33*	*2.61*	*-1.13*	*-5.51*		*1.54*	*0.25*	*1.44*	*-0.86*	*-2.01*
High	low	0.1782	-0.0023	1.0184	0.3776	0.0303	0.0222	0.1792	-0.0034	1.0449	0.3176	0.0738	-0.0085
			-3.34	*59.89*	*18.05*	*1.20*	*1.50*		*-3.47*	*43.30*	*10.69*	*2.06*	*-0.41*
High	medium	0.1781	-0.0007	0.9679	0.2676	0.0938	0.0219	0.1801	-0.0009	1.0044	0.2853	0.0474	0.0096
			-1.09	*62.98*	*14.15*	*4.12*	*1.64*		*-1.24*	*54.41*	*12.56*	*1.73*	*0.60*
High	high	0.1821	0.0013	1.0092	0.3825	-0.0962	0.0408	0.1818	0.0035	0.9481	0.3384	-0.0412	0.0186
			1.65	*51.67*	*15.91*	*-3.33*	*2.40*		*3.33*	*36.15*	*10.49*	*-1.06*	*0.82*
	high - low	0.0039	0.0036	-0.0091	0.0049	-0.1265	0.0186	0.0026	0.0069	-0.0968	0.0208	-0.1149	0.0271
			5.10	*-0.52*	*0.23*	*-4.89*	*1.23*		*4.61*	*-2.61*	*0.46*	*-2.09*	*0.84*

Panel B: Alpha and Fee Growth -- With filters described in MSZ (2007)

Trade_Pin at Qtr t-1	Previous 4f alpha	3f alpha (after fee)	4f alpha (before fee)	3f alpha (before fee)	fee growth rate (next qtr)
low	low	-0.0016	-0.0004	-0.0006	2.36%
		-2.07	*-0.47*	*-0.78*	*1.90*
low	mde	-0.0010	0.0001	-0.0001	2.98%
		-1.92	*0.28*	*-0.29*	*2.90*
low	high	-0.0002	0.0014	0.0007	6.46%
		-0.20	*1.57*	*0.73*	*5.58*
	high - low	0.0014	0.0018	0.0013	4.10%
		1.15	*1.42*	*1.03*	*3.94*
High	low	-0.0034	-0.0022	-0.0023	1.42%
		-3.63	*-2.32*	*-2.45*	*1.15*
High	mde	-0.0008	0.0000	0.0001	4.76%
		-1.14	*0.01*	*0.13*	*4.15*
High	high	0.0037	0.0045	0.0047	8.49%
		3.57	*4.24*	*4.50*	*5.57*
	high - low	0.0071	0.0067	0.0070	7.08%
		6.94	*4.50*	*4.78*	*6.56*

Table 8: Characteristic Selectivity (CS) measure decomposition across *Trade_PIN* and alpha double sorted fund portfolios

At the end of each quarter from 1983 to 2004, we conduct 2 by 5 double sorts based on *trade_PIN*s and four-factor alphas of the mutual funds. We then decompose the next-quarter Characteristic Selectivity (CS) measure within each portfolio. The CS measure and its component are reported in basis points before fees (bp) per quarter. *t*-values associated with the average measures are reported in *italic*s.

Trade_Pin at Qtr t	4f alpha	Total CS (=1+2+3a+3b) Qtr t+1	Old CS^{O} (1) Qtr t+1	Adj CS^{adj} (2) Qtr t+1	Info trading CS^{inf} (3a) Qtr t+1	Liquidity Prov CS^{liq} (3b) Qtr t+1	CS^{inf} - CS^{liq} (3a)-(3b) Qtr t+1
low	low	-7.89	-8.96	-1.92	-17.09	20.09	-37.17
		-0.62	*-0.69*	*-1.21*	*-2.33*	*2.64*	*-2.90*
low	mde	19.99	21.33	-1.10	-1.66	1.48	-3.14
		1.72	*1.64*	*-0.77*	*-0.26*	*0.22*	*-0.27*
low	high	27.98	17.78	-1.19	4.87	6.51	-1.63
		1.55	*1.13*	*-0.53*	*0.37*	*0.78*	*-0.09*
	high -	35.87	26.74	0.73	21.96	-13.58	35.54
	low	*1.75*	*1.42*	*0.28*	*1.79*	*-1.56*	*3.11*
high	low	-6.79	-11.17	-2.83	0.22	6.99	-6.77
		-0.33	*-0.59*	*-1.60*	*0.02*	*0.80*	*-0.38*
high	mde	33.41	16.59	1.61	5.51	9.70	-4.19
		2.02	*1.04*	*0.87*	*0.56*	*1.21*	*-0.28*
high	high	106.86	56.32	-5.35	48.04	7.90	40.14
		3.20	*2.09*	*-1.69*	*2.63*	*0.67*	*1.55*
	high -	113.65	67.49	-2.52	47.82	0.91	46.91
	low	*4.14*	*2.85*	*-0.76*	*2.99*	*0.08*	*2.91*

Table 9: Cross-sectional analysis

At the end of each quarter from 1983 to 2004, we regress the next quarter characteristics-adjusted before fee quarterly return (CS) on several fund-level characteristics in each quarter from 1983 to 2004. *trade_pin* is the average PIN of stocks recently traded by the funds; *log_fund_size* is the (log) average market cap of stocks held by the fund; *log_fund_bm* is the (log) average book-to-market ratio of stocks held by the fund; *fund_mom* is the average past one-year returns on stocks held by the fund; *fund_amihud* is the average Amihud illiquidity measure, in terms of percentile rank in the cross-section, of stocks held by the fund; *log_ TNA* is the (log) total net assets under management by the fund; *Age* is the age of the fund since inception, in terms of percentile rank in the cross-section; *flow* is the percentage fund flow; *turnover* is the turnover rate of the fund; *dummy_growth* is a dummy variable which assumes a value of 1 if the self-reported investment objective is "growth" and 0 otherwise; *dummy_Agg* is a dummy variable which assumes a value of 1 if the self-reported investment objective is "AGG" and 0 otherwise. All explanatory variables (except for the style dummy variables) are cross-sectionally demeaned and standardized so the corresponding coefficients can be interpreted as the impact on return of a one standard deviation change in the variable. Variables are winsorized at the 1st and 99th percentiles to alleviate the effect of outliers. Finally, the regression coefficients are averaged across time and the associated t-values are computed using Newey-West corrections with 8 lags to account for autocorrelations in the error terms. t-values associated with the average measures are reported in *italic*s. There are on average 320 funds in each cross-section.

	Intercept	trade_pin	log_fund _size	log_fund _bm	fund _mom	fund _amihud	log_TNA	Age	flow	turnover	dummy _growth	dummy _Agg	Average R^2
							LHS = CS						
Estimates	0.0018	0.0021	0.0017	0.0007	0.0013	-0.0001	0.0002	0.0000	0.0010	0.0001	0.0005	0.0033	0.20
t-value	*1.20*	*2.78*	*1.64*	*0.61*	*1.52*	*-0.13*	*0.79*	*0.10*	*3.22*	*0.24*	*0.60*	*2.38*	
							LHS = CS^{inf}						
Estimates	0.0001	0.0010	0.0009	0.0011	0.0010	0.0000	0.0001	-0.0001	0.0008	-0.0001	0.0000	0.0016	0.12
t-value	*0.17*	*3.03*	*1.48*	*1.52*	*1.30*	*0.09*	*0.52*	*-0.39*	*4.45*	*-0.24*	*-0.12*	*2.66*	
							LHS = CS^{liq}						
Estimates	0.0012	-0.0002	0.0000	-0.0004	-0.0001	0.0004	0.0000	-0.0004	0.0000	0.0004	-0.0004	-0.0010	0.10
t-value	*2.37*	*-0.76*	*0.04*	*-1.40*	*-0.37*	*1.14*	*-0.14*	*-3.55*	*0.07*	*1.78*	*-0.98*	*-1.36*	

Table 10: Noise associated with using quarterly data: evidence from the Plexus group data

We analyze the actual transactions of institutional money managers monitored by the Plexus Group, a consulting firm for institutional investors that monitors the cost of institutional trading. The Plexus Group dataset we use covers 1991Q2-1993Q1and 1995Q4-1998Q2. Each quarter, we sort manages in the Plexus Group data into quintiles based on the intensity of interim trading defined as the ratio between the total value of within-quarter roundtrip trades and the total value of all net trades in that quarter. For each quintile, we report the average number of managers per quarter and the average number of trades per manager per quarter. We then compute trade_PIN in two ways: either use the actual trades or use the net trades in the quarter which miss out interim trading. The difference between the two estimates captures the estimation error introduced by using quarterly data. We report the average estimation error and the average absolute estimation error in trade_PIN. Finally, we calculate the percentage of all trades that are correctly classified into either "informed-trading" or "liquidity-provision" group. The percentage is computed using either number of trades or total value of the trades. Although not reported, all percentages of correct classifications are significantly higher than 50%.

Portfolio	% of interim trading	avg # of managers	avg # of trades per manager	actual trade_PIN	avg est. err in trade_PIN	avg abs est. err in trade_PIN	% of correct trade classification (by # of trades)	% of correct trade classification (by value of trades)
1	0.2%	89	181	0.1559	0.0000	0.0001	69.8%	70.1%
2	1.8%	80	193	0.1501	0.0000	0.0003	67.2%	69.4%
3	5.2%	69	203	0.1451	0.0001	0.0012	66.4%	68.8%
4	12.3%	72	364	0.1455	-0.0003	0.0023	64.6%	68.7%
5	57.2%	71	683	0.1465	-0.0001	0.0043	61.8%	64.5%
All	15.3%	76	325	0.1486	-0.0001	0.0016	66.0%	68.3%