

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

Zhi Da, Pengjie Gao and Ravi Jagannathan[†]

First Draft: May 10, 2007
This Draft: January 25, 2010

Abstract

We show that a mutual fund's stock selection skill can be decomposed into additional components that include liquidity absorbing impatient 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 impatient 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 after fee excess return (four factor-alpha) of 35 basis points per month in the future. Most of that superior performance comes from impatient trading. We also find that impatient 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, Robert Kosowski, Norris Larrimore, Dong Lou, Tim Loughran, Dermot Murphy, Rick Mendenhall, David Musto, Lübös Pastor, Christine Parlour, Paul Schultz, Clemens Sialm, Matthew Spiegel (the editor), Laura Starks, Sheridan Titman, Charles Trzcinka, Lance Young, two anonymous referees, 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, Financial Research Association 2007 Annual Meeting, NBER Asset Pricing Program Meeting (Spring, 2008), the Western Finance Association 2008 Annual Meeting, American Finance Association 2009 Annual Meeting, and the Oxford-Man Institute Hedge Fund Conference for comments. We thank Don Keim and Sunil Wahal for their assistance with institutional transaction data, Norris Larrimore for providing us with the bond factors data, and Dong Lou for providing us with his stock level mutual fund flow estimates data. We also thank Ken French and Antti Petajisto for making factor returns and active share data available through their websites. Patricia Andersen and Dermot Murphy provide excellent editorial assistance. This paper has been previously circulated under the title "When does a mutual fund's trade reveal its skill?"

[†]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 2008, US domestic equity mutual fund managers collectively had over \$2.8 trillion under their management. 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 trading, (b) recent liquidity providing trades, (c) positions in securities taken earlier, and (d) an adjustment term for inflows and outflows. We illustrate the use of our decomposition method for understanding the sources of superior performance of managed portfolios.

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. For example, 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 (Greenwald, Kahn, Sonkin and Biema, 2001).

Second, the manager can add value from medium-term trading by transacting in "mispriced"

¹These numbers are taken from Figures 2.1 and 2.9 in the Investment Company Fact Book (2009) published by the Investment Company Institute.

stocks expecting the market to agree with her view within, say, a quarter. For example, 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 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 (Blumenthal, 2007).

Third, the manager can add value from short-term trading. For example, it is well known that when index funds trade following index rebalancing, their trades tend to demand liquidity from the market during the few days surrounding index changes (see Blume and Edelen, 2004). Active fund managers taking the other side of those trades will benefit from liquidity provision. 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 from trading with those having an information advantage.² Another example of short-term liquidity-provision trading is pairs-trading strategies which are popular among technical traders. Engleberg, Gao, and Jagannathan (2009) demonstrate that a significant portion of the pairs-trading profit documented in Gatev, Goetzmann and Rouwenhorst (2006) represents compensation for short-term liquidity provision.

²Sometimes 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.

In the first case (long term value investing), 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 of what the manager holds. In the third case (short-term trading,) since we use mutual funds’ holdings reported at quarterly intervals, we can not say much about value added through active within-quarter trades.³ Therefore, in this paper our focus is on decomposing the value added by a manager from the second class of activities into different components.

We examine several empirical properties of the decomposition that lend support for its validity. First, we verify that the decomposition results for, (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.⁴ Second, we find that the impatient trading component is more important than the liquidity provision component in explaining cross-sectional variation in the Characteristic Selectivity measure (*CS* measure thereafter) — developed by Daniel, Grinblatt, Titman and Wermers (DGTW, 1997) — and impatient trading becomes relatively more important for growth-oriented funds while liquidity provision becomes relatively more important for income-oriented funds. Third, we find that funds with higher “return gaps” — defined in Kacperczyk, Sialm and Zheng (2008) 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

³As Kacperczyk, Sialm and Zheng (2008) show, “unobservable” actions (trades that cannot be inferred from quarterly holdings) by mutual funds could be important for some funds. However, we have little to say on that based on the data available to us.

⁴The findings in Keim (1999) 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) using the *CS* measure, after controlling for stock characteristics such as size, book-to-market ratio and return momentum. Easley, Hvidkjaer and O'Hara (2002) document that High-*PIN* stocks earn higher returns on average. They interpret this as compensation for risk associated with private information, i.e., *PIN*-risk. That does not explain our findings. 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 trading. In contrast, liquidity provision appears more important for funds trading in low-*PIN* stocks where there is little adverse selection risk. Although funds trading in high-*PIN* stocks outperform those trading in low-*PIN* stocks using the *CS* measure, both types of funds have after fee alphas that are either zero or negative, which is consistent with the findings in the early literature on portfolio performance evaluation.⁵

Several 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 and the when they trade.⁶ Interestingly, Mamaysky, Spiegel and Zhang (2007b) show that superior performance can be identified based on past performance alone by imposing restrictions implied by economic reasoning. However, during

⁵See Jensen (1968), Elton, Gruber, Das, and Hlavka (1993), Brown and Goetzmann (1995), Gruber (1996) and Carhart (1997).

⁶See Grinblatt and Titman (1989, 1993), Daniel, Grinblatt, Titman and Wermers (1997), Chen, Jegadeesh and Wermers (2000), Schultz (2007), and Alexander, Cici, and Gibson (2007).

the more recent sampling period from 1983-2004, Barras, Scaillet, and Wermers (2010), Kosowski, Timmermann, Wermers and White (2006), as well as Fama and French (2009) document that historical alphas alone do 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 decile portfolio of mutual funds that significantly outperforms its benchmark portfolio going forward at monthly frequency even during the more recent sampling period. In particular, the past winners among funds trading high-*PIN* stocks identified using the methods in Mamaysky, Spiegel and Zhang (2007b) has a statistically significant after fee four-factor alpha of 35 bps per month (t -value = 3.33.) Further, most of that outperformance comes from impatient trading. In contrast, past winners among all funds in our sample earn risk-adjusted returns that are not significantly different from zero. We confirm that these results are robust to the choice of subsample periods and models for risk-adjustments.

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 headquarters (Coval and Moskowitz, 2001); have more industry concentration in their holdings (Kacperczyk, Sialm and Zheng, 2005); have larger deviations from passive index or larger “active shares” (Cremers and Petajisto, 2009); 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 number of information events on the stock traded by a manager can be informative about the skill of that manager. In addition, we find that the impatient 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 *impatient trading* or *liquidity-providing*. 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 using quarterly holding changes 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 *impatient 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 “impatient trading” or “liquidity provision.” Nevertheless, this percentage of correct assignment is statistically significantly higher than 50% (corresponding to random assignment of trades to these categories) 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 “impatient 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 validate 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 (2008). 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 combine the use of *trade_PIN* and historical alpha to identify a decile portfolio that generates positive future alpha and analyze the sources of value. In section 6, with the help of actual fund transactions in the Plexus Group data, we analyze the noise associated with inferring mutual trades using quarterly holding data.

We conclude in section 7. The online appendices from the *Review of Financial Studies* website contain a numerical example on our decomposition, a brief discussion on the variance decomposition approach, and a short note on various measures of information events.

1 Decomposing A Mutual Fund’s Stock Selection Skills

In order to separate the value added via 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). 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:

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

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 online appendix A. Suppose mutual funds rebalance only at discrete points in time, $t = 1, 2, 3, \dots, T$. 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:

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

where the operator $\min()$ calculates the element-by-element minimum; 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:

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

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

$$H_t = P_t' N_t^H$$

$$B_t = P_t' N_t^B$$

$$S_t = P_t' N_t^S$$

At the end of quarter t , the mutual fund's stock holdings are a combination of the Hold portfolio and the Buy portfolio. 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:

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

The first component, the *old* component (CS_{t+1}^O), 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. 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 holding on to the positions for one more quarter.

The second component, the *trade* component (CS_{t+1}^T), measures the characteristics-adjusted returns on the most recent mutual fund stock trades during quarter t . Finally, the *adjustment* component (CS_{t+1}^{adj}) 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+1}^T) measures value added from both impatient trading and as patient liquidity provision. We therefore decompose the trade component CS_{t+1}^T further into the impatient trading and liquidity provision 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+1}^T :

$$\begin{aligned}
CS_{t+1}^T &= CS_{t+1}^{imp} + CS_{t+1}^{liq} \\
CS_{t+1}^{imp} &= \frac{B_t^+}{H_t + B_t} CS_{B,t+1}^+ - \frac{S_t^+}{H_t + S_t} CS_{S,t+1}^+ \\
CS_{t+1}^{liq} &= \frac{B_t^-}{H_t + B_t} CS_{B,t+1}^- - \frac{S_t^-}{H_t + S_t} CS_{S,t+1}^-
\end{aligned} \tag{2}$$

Given that the aggregate order imbalance is a good measure of the direction of liquidity needs of a stock (see Chordia and Subrahmanyam, 2004), CS_{t+1}^{imp} measures the characteristics-adjusted return on mutual fund trades that on average absorb market liquidity. Such trades are likely to be driven by informational advantage that decays relatively faster over time requiring “impatient trading.” CS_{t+1}^{liq} , 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_{t+1}^{liq} 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:

$$\begin{aligned}
CS_{t+1} &= CS_{t+1}^O + CS_{t+1}^{adj} + CS_{t+1}^T, \\
CS_{t+1}^T &= CS_{t+1}^{imp} + CS_{t+1}^{liq}.
\end{aligned} \tag{3}$$

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 which are hybrid funds. Following

standard practice in 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.⁷ Finally, we only include fund / quarter observations for which the fund holdings *at the end of the previous quarter* are also available so holding changes can be computed over consecutive quarters.⁸ 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) mutual fund database.

The CDA/Spectrum mutual fund holding data are matched to CRSP Mutual Fund data using the MFLINKS database. 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

⁷As we explicitly exclude asset allocation funds and lifecycle funds, which comprises most of the hybrid funds in our sample, equity holdings account for at least 90% of total asset values (based on comparison between equity holding values computed from Thomson Finance/CDA database and the corresponding period's total asset values reported in the CRSP mutual fund database). Comer, Larrymore and Rodriguez (2009) demonstrate that the inclusion of fixed-income factors helps to produce more consistent estimates of mutual fund alphas. Including bond factors should have little impact on our empirical analysis since bond holdings account for only 2% of holdings among funds in our sample (based on the numbers reported by CRSP mutual fund database).

⁸We select a group of funds with consecutive reporting because we need to compare holdings at consecutive quarters. These funds on average do better, as documented by Ge and Zheng (2006). Since our ranking based on past alpha and trade_pin is done at a given point in time, using funds revealing their holdings information for the current quarter and the previous quarter, there is no look-ahead bias in our empirical analysis. On the one hand, because of this our results may not speak about funds with semiannual reporting. On the other hand, now that mutual funds are required to report again on a semiannual basis, the methodology developed here is still of value to mutual fund selection.

average, there are about 701 distinct funds every quarter. The number of funds per quarter increases from about 132 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);⁹ *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\$); *act_share* (the active share variable proposed by Cremers and Petajisto (2009) to capture the deviations of a manager’s portfolio holdings from the underlying benchmark index holdings); 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 the fund level by value-weighting 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 billions of 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).¹⁰

3 Validating the Decomposition

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.¹¹ To the extent

⁹We use percentile age ranks to remove a time-series (increasing) trend in the age variable.

¹⁰Amihud illiquidity measure is defined as the average ratio of absolute daily return to 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 problems.

¹¹Campbell, 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.

that short-term liquidity provision occurs within a calendar quarter, by using quarter-end holdings only, we may underestimate the benefit from liquidity provision.

Second, we need to understand the forces that drive impatient trades and liquidity providing trades. When trading relatively large quantities quickly to take advantage of the time value of information, it is hard not to absorb liquidity. As a result, liquidity-demanding impatient trades are likely to be driven by short lived informational advantage on average. On the other hand, not all liquidity-demanding impatient trades are information driven. For example, sale of distressed stocks by mutual funds (see Da and Gao, 2008) and fire sale of assets triggered by extremely large unanticipated withdrawals flows (see Coval and Stafford, 2007) are likely to absorb liquidity. 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 a value-weighted average. When we leave out value-adding informed trading that is not liquidity demanding and include non value-adding distressed trading that is liquidity demanding, our impatient trading component underestimates the benefit of informed trading.

Finally, our classification of impatient trading and liquidity provision depends on quarterly data, which could also be noisy. Collectively, these noises may prevent us from finding any significant results. We therefore formally evaluate the direction and magnitude of the biases in Section 6 using a limited sample of high-frequency institutional transaction data. Our findings suggest that in spite of these challenges, the quarterly holdings data contain valuable information about the contribution of impatient 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 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.¹² If it is the case, using our decomposition procedure, one would expect to find a positive liquidity

¹²See the case studies by Keim (1999) and Cohen (2002).

provision component in DFA's *CS* measure and an impatient 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.

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 Panel A of Table 2 (lower 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 the 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.

3.2 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) so we should expect the impatient trading component to be negative. Index funds may also add value from securities lending that will show up in the return. Other than that we should expect little of interest happening in the case of 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 Panel A of Table 2 (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 impatient trading component (C^{imp} = -35 bps) which is statistically significant 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.

3.3 Fund Styles

Panel B of Table 2 suggests that, overall, active fund managers seem to have some stock selection skill. 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 magnitude (-1.8 basis points) but significant, potentially driven by fund flow to skilled managers as empirically documented by Chevalier and Ellison (1997), and Sirri and Tufano (1998), among others, and theoretically analyzed by Berk and Green (2004).¹³ Finally, although both the impatient trading component (C^{imp}) and the liquidity provision component (CS^{liq}) are positive, neither is significant.

Across fund styles, we expect growth-oriented funds to have more shorter-term impatient 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 the impatient informed trading component within growth funds and relatively more variation in the liquidity provision component within income-oriented funds. The results in Panel B of Table 2 confirm our conjectures. On average, growth-oriented funds have a larger impatient trading component and income-oriented funds have a larger liquidity provision component.

We examine the relative importance of different components of CS measures using the variance decomposition procedure detailed in online appendix B. In a nutshell, the variance decomposition delineates how much the cross-sectional variation in the total CS measure can be attributed to

¹³When managers have skill (CS^O 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).

the cross-sectional variation in each of its four components. The results are reported in Panel B of Table 2 for the full sample of all active US equity funds, and across three style subsamples. As we expected, the impatient informed trading component, CS^{imp} , varies relatively more across growth-oriented funds while the liquidity provision component, CS^{liq} , varies more across funds in the income-oriented category.

3.4 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 (2008). 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 where 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 (2008) show that funds with high return gaps tend to have higher alphas in the future. They conjecture that the higher positive alpha could in part be 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 contributing to the fund’s alpha. In that case, 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).¹⁴

¹⁴Avramov and Wermers (AW, 2006) show that it is possible to identify time variations in mutual fund managers’

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 C1 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 (2008). We document that the return gap is strongly negatively correlated with the Old CS component (CS^O) and the Informed Trading component (C^{imp}). 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 conduct interim trading within a quarter.

We then examine the relation between the current return gap and future CS measures. Following Kacperczyk, Sialm and Zheng (2008), 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 C2 of Table 2. We first confirm the findings in Kacperczyk, Sialm and Zheng (2008) that a 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

skill based on information in macroeconomic variables – aggregate dividend yield on stocks, default spread, term spread, and short term interest rate. The Agnostic PA-3 fund they construct has an after fee four-factor alpha of almost 13% per year in their sample (Table 5, AW). However, the CS measure of the Agnostic PA-3 fund is much lower: only 8.3% annualized (Table 6, AW). That suggests that a part of the superior performance of their Agnostic PA-3 fund may come from the return gap and macroeconomic variables help predict time variations in the return gap of fund managers.

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.

To summarize, our decomposition method for classifying a mutual fund's trades during a quarter into liquidity demanding and liquidity providing components based only on the fund's quarterly holdings data appears reasonable in several settings. In the next section, we will analyze a larger sample of mutual funds using our decomposition method in order to provide a better understanding of 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 should 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 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..

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:

$$trade_PIN_{j,t} = \frac{\sum_{i=1}^N PIN_{i,t} \times d_{i,j}}{\sum_{i=1}^N d_{i,j}} \quad (4)$$

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).

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. Whereas momentum trading involves price impact that has to be taken into account, the UMD factor does not. Therefore, a

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

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 funds’ quarter end 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 holdings of stocks below reporting thresholds, and cash and other non-stock holdings of a mutual fund. However, such holdings are usually small, accounting for less than 10% 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 affecting our conclusions.

4.1 Information Events and Informed Trading

At the end of each quarter in 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 during the four quarters following portfolio formation. 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 after fees in the next quarter with a *t*-value of 3.15. The return spread is 46 basis points after fees 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 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 adjustment – with the caveat that characteristics-based adjustment is done to pre-fee returns not after fee returns. On average, the top *trade_PIN* decile portfolio of funds outperforms the bottom decile portfolio by 53 and 40 basis points per quarter in the first and second quarter after portfolio formation, respectively. These differences are significant at the 1% level. The return spreads between the top and bottom deciles drop steeply to 19 and 9 basis points in the third and fourth quarter respectively, primarily due to the drop in the *CS* measure of the top decile. Notice that the *CS* measure exhibits less persistence than the three- and four-factor alphas. The fact that the three-factor after fees alpha of the top decile declines little over the four quarters following portfolio formation (remains at essentially zero), while the *CS* measure of quarter t falls steeply after quarter $t + 2$, indicates that most of the convergence occurs within two quarters after taking a position, and new positions are created to maintain the alpha. The *CS* measures of bottom decile managers remain stable over the four quarters at essentially zero, and the after fee alphas are stable and hover between -40 and -48 bp/quarter. That indicates that the trading activities of managers in the lowest decile do little to add value to cover the fees.

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 (2002), 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 return thus controls for any systematic PIN risk. In the case of characteristics-based 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 online 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 (2007).

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. 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 *PIN*s 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 *PIN*s (0.2294 *vs.* 0.2301, not statistically significantly different.) 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 *PIN*s. This finding again suggests that the alpha generated by funds trading high-*PIN* stocks comes from superior stock selection skill rather than *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); thus they 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.

Impatient 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) come from active trading during the previous quarter ($CS^T = 31.2$ basis points with a t -value of 2.83.) CS^O that measures the contribution of positions put in earlier is 26.5, indicating that those positions contributed equally, though not statistically significant ($t = 1.43$.) The funds trading medium- $trade_PIN$ stocks (deciles 5 and 6) also earn significantly positive before-fee 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 that their before-fee 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 impatient trading component (C^{imp}) and the liquidity provision component (CS^{liq}) are positive for high- $trade_PIN$ -funds, only the impatient 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 impatient 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 impatient trading component, resulting in a close-to-zero *CS* measure. As can be seen from the last column the difference between the impatient trading component and liquidity provision component is negative and statistically significant for funds trading low *trade_PIN* stocks.

To summarize, impatient informed trading is more likely to add value at times when the stocks that a mutual fund trades are affected by information events. Liquidity provision is more likely to add value when the stocks that the fund trades are associated with few information events.

5 Identifying Funds With Positive Future Alphas

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.

5.1 Sorting on Historical Alpha

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 estimated using the previous 5 years of monthly mutual fund return data with at least 3 years of monthly 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 benchmark 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 benchmark 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 these three filters. Specifically, for any fund to be included in any decile, the following criteria must be met: (1) the absolute value of alpha must be less than 2% per month, (2) the CAPM beta must be between 0 and 2, and (3) in the previous month the forecasted alpha and the difference between the realized fund 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 benchmark 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 benchmark 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 (2009).¹⁶ 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 expect stronger performance persistence among funds that traded in high-*PIN* stocks 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

¹⁶However, we also want to point out that 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.

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 traded 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 a 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 benchmark going forward at monthly frequency using *trade_PIN* variable alone. The portfolio of past winners within the high-*trade_PIN* group does outperform the benchmark after fees (the 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 benchmark going forward at monthly frequency. The past winner within the high-*trade_PIN* group now has a statistically significant after-fee four factor alpha of 35 bps per month (t -value = 3.33). In fact, among the high-*trade_PIN* funds, the performance persistence is as strong in the past winners as in the past losers. We obtain similar results using the Fama-French three-factor alphas as in Panel B of Table 7.

¹⁷In the tests presented in this section, we have also verified that controlling for the PIN risk directly using either a fifth PIN risk factor or using PIN characteristic-adjustment produces very similar results.

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 other 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 of Table 7, 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%.

Robustness

The Carhart (1997) four-factor model has been used extensively in mutual fund performance studies. Several recent papers identify some limitations associated with the four-factor model, and propose improvement and extension. For example, Cremers, Petajisto, and Zitzewitz (2008) show passive benchmark indices, such as S&P 500 and Russell 2000, have non-zero alphas when the three factors in Fama and French (1993) augmented by the momentum factor in Carhart (1997) are used to control for systematic risk. Cremers, Petajisto, and Zitzewitz (2008) provide a modification of factor construction, and suggest a four-factor model and a seven-factor model that effectively resolve these issues. Other issues are related to additional factors to be included in the factor model. Comer, Larrymore and Rodriguez (2009) demonstrate that the inclusion of fixed-income factors produces more consistent estimates of mutual fund alphas, especially for those funds with substantial fixed-income holdings.

Motivated by these recent studies, we repeat our analysis on fund performance persistence using alternative risk adjustment models and in different subsample periods. We examine the four-factor model and the seven-factor model of Cremers, Petajisto, and Zitzewitz (2008) (labeled as CPZ4 and CPZ7) and the factor model of Comer, Larrymore and Rodriguez (2009) which includes bond factors (labeled as CLR). For comparison purpose, we also include the Carhart (1997) four-factor model (labeled as FF4). We split our sample into two subsample periods: 1983 – 1993 and 1994 – 2004. The bond factors required by CLR only start from 1994 so we estimate the CLR-alphas only in the second subsample period. We report the results in Panel C of Table 7.

Overall, our findings on fund performance persistence is robust to the choice of sampling periods

and risk-adjustment models. During either the subsample period or the full-sample period, using any of the alternative factor models, we find very little evidence of performance persistence among funds trading low-*PIN* stocks. If there is any performance persistence, it is among those with low past alphas. In contrast, there is strong performance persistence among funds trading high-*PIN* stocks. The portfolio of past high past alpha funds always outperforms the portfolio of past low past alpha funds. Further, the portfolio of past high alpha funds has a statistically significant after-fee alpha going forward. The exact magnitude of the alpha depends on the factor model we choose. During the full sample period, the portfolio of high-*trade_PIN* and high past alpha funds has an after-fee alpha of 37 to 47 bps per month. During the earlier period from 1983 to 1993, the alpha ranges from 36 to 40 bps per month. During the recent period from 1994 to 2004, the alpha ranges from 37 to 54 bps per month. In all cases, these alphas are statistically significant at the one percent level.

Finally, Kosowski, Timmermann, Wermers, and White (2006) suggest that bootstrapped p -values are potentially better when conducting statistical inference for the returns on the extreme portfolios of funds, since these fund returns may violate the standard normality assumptions underlying the calculation of t -statistics. To generate bootstrapped p -values of the four-factor alphas and their associated t -statistics, we bootstrap fund portfolio excess returns using factor loadings estimated during the prior five-year period under the null of a zero true alpha. This process is repeated for each month in our sample to build a full time series of fund portfolio excess returns. We then estimate the alpha and its t -statistic for each fund portfolio. We repeat the above procedure 1000 times to produce the empirical distribution of the alpha and its t -statistic. Comparing the actual alpha and t -statistic estimates to their empirical distributions allows us to calculate the bootstrapped p -values. These results are reported in the last two columns of Panel C, Table 7. Our main conclusions remain robust to the bootstrap procedure.

Sources of Future Performance

Table 8 describes the decomposition of the before-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 impatient (informed) trading, and the rest mostly come from positions taken from earlier quarters. The positive and significant C^{imp} and CS^O explain why these past winners are able to outperform the other funds in the future –

they are indeed very good at impatient informed trading. Interestingly, the past loser funds in the low *trade_PIN* class have a significantly positive liquidity provision component (CS^{liq}) and a significantly negative impatient trading component (C^{imp}) – only that the two are about equal (in absolute terms) and cancel each other out. 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 impatient 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.

One may be concerned that the flow-induced trading from mutual funds temporarily move the underlying asset prices, and the performance persistence results could be due to such temporary price pressure (Wermers, 2003; Lou, 2008). In the last column of Table 8, we directly control for the impact of fund flows on underlying stock prices when we evaluate the characteristics-adjusted returns. Lou (2008) explores mutual funds’ trading decisions when they face inflows and outflows. He estimates the stock-level holding changes within a fund in the presence of fund level flows. Aggregating flows from all mutual funds into a stock, he provides a measure of flow-induced trading at the stock level, and shows that flow-induced trading affects future stock returns.

Using Lou’s estimates of quarterly stock-level flows from mutual funds, we construct 81 four-way sorted benchmark portfolios. All stocks are first sorted into tercile portfolios, based on market capitalization. Within each market capitalization sorted tercile portfolio, stocks are further sorted into tercile portfolios based on the industry-adjusted book-to-market equity ratios. Then, within each of these 9 size and book/market sorted portfolios, stocks are sorted into tercile portfolios based on past twelve-month returns. Finally, within each of these 27 size, book/market, and past return sorted portfolios, stocks are sorted based on the aggregate flow-induced trading measure proposed in Lou (2008). The composition of size and book/market portfolios changes annually, while the composition of past return and flow-induced trading portfolios changes quarterly. A stock’s four-characteristic adjusted return is computed as the stock return in excess of the benchmark portfolio return, where the stock-specific benchmark is chosen based on information that was available during the previous quarter.

Among the high *trade_PIN* portfolios, funds with high past alpha continue to earn a statisti-

cally significant four-characteristic adjusted return of 97 basis points per quarter (t-value = 2.57), which decreases only by about 10 basis points from three-characteristic adjusted return of 107 basis points per quarter. In sharp contrast, among the low *trade_PIN* portfolios, funds with high past alpha do not earn any statistically significant returns, using either four-characteristic adjusted returns (31 basis points per quarter, t-value = 1.61), or three-characteristic-adjusted returns (28 basis points per quarter, t-value = 1.55). Overall, the four-characteristic-adjusted return results suggest that flow induced trading does not seem to explain performance persistence among high *trade_PIN* past winners identified in this paper.

5.3 Fund Characteristics and Performance Components

We examine the relation between fund characteristics and *CS* measures using Fama-MacBeth (1973) cross-sectional regressions. 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 8 lags 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. Cremers and Petajisto (2009) propose the active share (*act_share*) concept, which captures the deviations of a manager’s portfolio holdings from the underlying benchmark index holdings. They show that funds with higher active shares tend to perform better in the future, which we confirm in the regression. *act_share* is significantly positively related to the future total *CS* measure. 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 impatient trading component (CS^{imp}), and the liquidity provision component (CS^{liq}). Interestingly, fund characteristics associated with impatient trading, and with liquidity provision are quite different. When we regress CS^{imp} 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 impatient trading. act_share ceases to be significant, which suggests that not all actively managed funds can benefit from impatient trading. In addition, $flow$ and $dummy_AGG$ remain significant, indicating that impatient trading is more prevalent in funds with an “Aggressive Growth (AGG)” investment style and money is chasing fund managers who are good at impatient 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, the 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. Finally, $dummy_AGG$ and $dummy_growth$ have negative and significant coefficients, indicating that aggressive growth funds do not specialize in liquidity provision.¹⁸

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 calculation of $trade_PIN$. Second, inferring mutual fund trades using their quarterly holdings also adds noise to the classification of trades as belonging to the *impatient-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

¹⁸It is also useful to consider the evidence in Wermers (2003; Table IX). He relates the next quarter’s characteristics-adjusted returns to the past characteristics-adjusted returns. He finds the past characteristics-adjusted returns do not predict the next quarter’s characteristics-adjusted returns in the presence of fund flows. Therefore, contrasting regression results from these two papers, we interpret our evidence as being inconsistent with the idea that flow-induced trading explains the results in this paper.

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 1991Q2-1993Q1 and 1995Q4-1998Q2. Prior research that use the Plexus Group data include Keim and Madhavan (1995) and Conrad, Johnson, and Wahal (2003), among others.

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 manager in Plexus Group data, we are able to compute their *trade_PIN*, and accurately classify their trades as belonging to the “impatient-trading” or “liquidity-provision” category. However, note that Plexus Group data cover only a subset of the mutual funds we examined earlier, and for a shorter 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 affects our conclusions. Other things being equal, 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, which is defined as the ratio of the total value of all within-quarter round-trip trades to 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 on the net trades which exclude all 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 the “impatient–trading” category or “liquidity-provision” category using quarterly data is more noisy. Table 10 shows that on average, 66% of these trades are correctly assigned to either the “impatient–trading” category or “liquidity-provision” category even with the quarterly data. The percentage of correct assignment is higher when it is measured using the total dollar value of the trades. The percentage of correct assignment is significantly higher than 50% (i.e., random assignment of trades into these two groups) for all fund quintiles, suggesting that trade classification using quarterly data is still sufficiently informative. The accuracy of the assignment is better among funds that engage in fewer interim transactions. For these funds, 70% of their trades are correctly assigned to be either “impatient-trading” or “liquidity-provision.” Interestingly, these funds trade stocks with higher PIN s 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 has little impact on the calculation of a fund’s $trade_PIN$. Although the classification of trades into either the “impatient-trading” category or the “liquidity-provision” category is affected to a larger extent by ignoring interim trades, the classification scheme remains sufficiently informative. In summary, our general conclusions regarding impatient 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 (“impatient 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 impatient 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 impatient informed trading and liquidity provision dimensions with greater precision, thereby contributing to better allocation of resources across different money managers.

We confirm 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 has 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), has a statistically significant after-fee alpha of 35 bps per month (t -value = 3.33). Most of that alpha comes from impatient informed trading positions that converge on average within two quarters. In contrast, the alpha of the corresponding portfolio of past winners *among all funds* is neither statistically nor economically different from zero.

References

- [1] Alexander, Gordon J., Gjergji Cici, and Scott Gibson, 2007, Does Motivation Matter When Assessing Trade Performance? An Analysis of Mutual Funds, *Review of Financial Studies* 20, 125-150
- [2] Avramov, Doron, and Russ Wermers, 2006, Investing in mutual funds when returns are predictable, *Journal of Financial Economics* 81, 339-377.
- [3] Amihud, Yarov, 2002, Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets* 5, 31-56.
- [4] Laurent Barras, Olivier Scaillet, and Russ Wermers, 2010, False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas, *Journal of Finance*, Forthcoming.
- [5] Berk, Jonathan, and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112(6), 1269–1295.
- [6] Blume, Marshall E., and Roger M. Edelen, 2004, On replicating the S&P 500 index, *Working Paper*, Wharton School.
- [7] Blumenthal, Karen, 2007, Grande expectations: A year in the life of Starbuck's stock, Crown Publishing, first edition.
- [8] Brown, Stephen, and William Goetzmann, 1995, Performance persistence, *Journal of Finance* 50(2), 679 - 698.
- [9] Brown, Stephen, Stephen A. Hillegeist and Kin Lo, 2004, Conference calls and information asymmetry, *Journal of Accounting and Economics* 37, 343-366.
- [10] Campbell, John Y., Tarun Ramadorai, and Allie Schwartz, 2007, Caught On Tape: Institutional Trading, Stock Returns, and Earnings Announcements, *Journal of Financial Economics*, Forthcoming.
- [11] Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.
- [12] Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, 2000, The Value of Active Mutual Fund Management: An examination of the stockholdings and trades of fund managers, *Journal of Financial and Quantitative Analysis* 35, 343–368.
- [13] Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey Kubik, 2004, Does fund size erode mutual fund performance? The Role of Liquidity and Organization, *American Economic Review* 94(5), 1276–1302.

- [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] Comer, George, Norris Larrimore, and Javier Rodriguez, 2009, Controlling for Fixed-Income Exposure in Portfolio Evaluation: Evidence from Hybrid Mutual Funds, *Review of Financial Studies*, 22, 481-507
- [16] Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007, The cross-section of expected trading activity, *Review of Financial Studies* 20, 709-740.
- [17] Chordia, Tarun and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485 - 518.
- [18] 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.
- [19] Cohen, Randolph B., 2002, Dimensional fund advisors, *Harvard Business School Case* 294-025.
- [20] Conrad, Jennifer, Kevin M. Johnson, and Sunil Wahal, 2003, Institutional trading and alternative trading systems, *Journal of Financial Economics* 70, 99–134.
- [21] 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.
- [22] Coval, Joshua D., and Erik Stafford, 2007, Asset Fire Sales (and Purchases) in Equity Markets, *Journal of Financial Economics*, Issue 2, 479-512.
- [23] Cremers, Martijn, and Antti Petajisto, 2009, How Active is Your Fund Manager? A New Measure that Predicts Performance, *Review of Financial Studies*, 22, 3329-3365.
- [24] Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2008, Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation, Manuscript, Yale School of Management
- [25] Da, Zhi and Pengjie Gao, 2008, Clientele Change, Liquidity shock, and the Return on Financially Distressed Stocks, *Journal of Financial and Quantitative Analysis*, forthcoming.
- [26] Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic based benchmarks, *Journal of Finance* 52, 1035–1058.
- [27] Duarte, Jefferson, and Lance Young, 2009, Why is PIN priced? *Journal of Financial Economics*, Issue 2, 119-138.
- [28] Easley, David, Soeren Hvidkjaer, and Maureen O’Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57, 2185-2221.

- [29] 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.
- [30] 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.
- [31] 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.
- [32] Engleberg, Joseph, Pengjie Gao, and Ravi Jagannathan, 2009, An Anatomy of Pairs Trading: The role of idiosyncratic news, common information, and liquidity, *Working Paper*, University of North Carolina at Chapel Hill, University of Notre Dame and Northwestern University.
- [33] Fama, Eugene F., and Kenneth French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–56.
- [34] Fama, Eugene F., and Kenneth French, 2009, Mutual fund performance, *Journal of Finance*, Forthcoming.
- [35] French, Kenneth, 2008, The cost of active investing, *Journal of Finance* 63(4), 1537-1573.
- [36] Fama, Eugene F., and J.D. MacBeth, 1973, Risk Return and Equilibrium: Empirical Tests, *Journal of Political Economy*, 607-36.
- [37] Evan, Gatev, William N. Goetzmann, and K. Geert Rouwenhorst, 2006, Pairs Trading: Performance of a Relative-Value Arbitrage Rule, *Review of Financial Studies*, 19, 797-827
- [38] Ge, Weili, and Lu Zheng, 2006, The Frequency of Mutual Fund Portfolio Disclosure, *Working Paper*, University of Clifornia, Irvine.
- [39] Glosten, Lawrence R., and Lawrence E. Harris, 1988, Estimating the components of the bid / ask spread, *Journal of Financial Economics* 21, 123-142.
- [40] Greenwald, Bruce C.N., Judd Kahn, Paul D.Sonkin, and Michael van Biema, 2001, Value Investing: From Graham to Buffett and Beyond, Wiley Publishing, first edition.
- [41] 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.
- [42] Grossman, S., and J. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393-408.
- [43] Gruber, Martin J., 1996, Another puzzle: the growth in actively managed mutual funds, *Journal of Finance* 51, No. 3, July 1996.

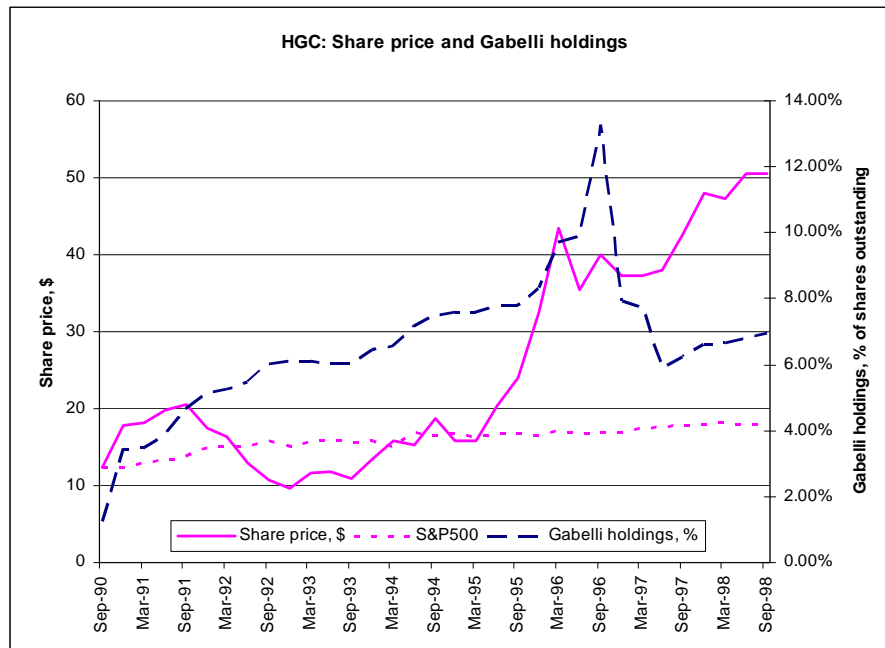
- [44] Huang, Roger and Hans R. Stoll, 1996, Dealer versus Auction markets: a paired comparison of execution costs on Nasdaq and the NYSE, *Journal of Financial Economics* 41, 313-358.
- [45] Jensen, Michael, 1968, The performance of mutual funds in the period 1945-1964, *Journal of Finance* 23, 389-416.
- [46] Kacperczyk, Marcin, and Amit Seru, 2007, Fund Manager Use of Public Information: New Evidence on Managerial Skills, *Journal of Finance* 62, 485-528.
- [47] Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2012.
- [48] Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies*, 21, 2379-2416
- [49] 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.
- [50] 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.
- [51] 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.
- [52] Lee, Charles M. C., and Mark J. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733-746.
- [53] Lou, Dong, 2008, A Flow-based Explanation for Return Predictability, *Working Paper*, London School of Economics.
- [54] 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.
- [55] 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.
- [56] Mamaysky, Harry, Matthew Spiegel, and Hong Zhang, 2007b, Improved forecasting of mutual fund alphas and betas, *Review of Finance* 11, 359-400.
- [57] Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.
- [58] Schultz, Paul, 2007, When is stock picking by mutual funds successful?, *Journal of Financial and Quantitative Analysis*, Forthcoming.

- [59] Sirri, Erik R., and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *Journal of Finance* 53, 1589-1622.
- [60] Wermers, Russ, 2003, Is Money Really “Smart”? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence, *Working Paper*, University of Maryland.
- [61] 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

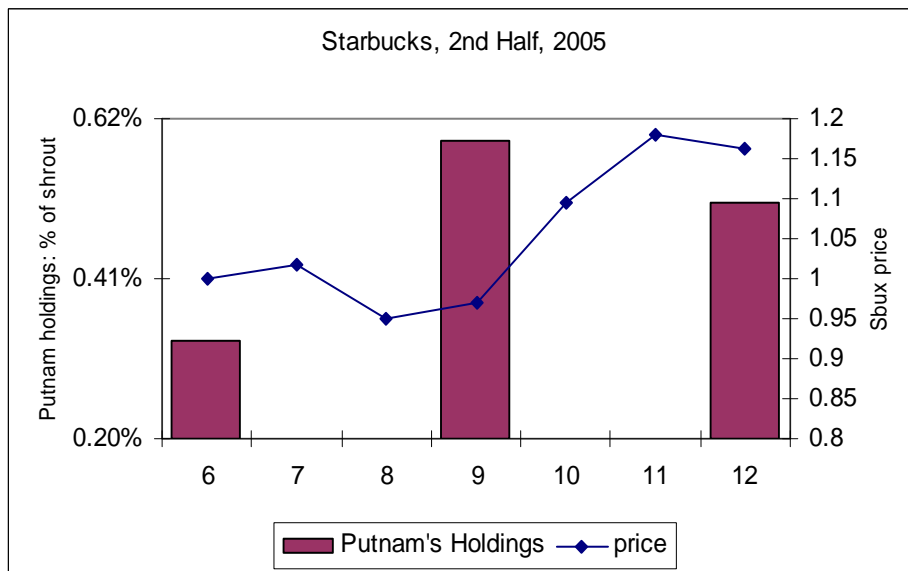


Table 1: Breakdown of mutual fund sample over time

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 days 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	All	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 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 (any 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: $CS = CS^O + CS^{adj} + CS^{imp} + CS^{liq}$, where CS^O is the old component; CS^{adj} is an adjustment component due to fund inflows; CS^{imp} and CS^{liq} are the impatient trading and liquidity provision components, respectively. The sampling period is from 1983 to 2004. t -values associated with the average measures are reported in parentheses.

Panel B 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 impatient trading component (CS^{imp}) 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. t -values associated with the average measures are reported in parentheses, using the weighted least squares (WLS) method. The sampling period is from 1983 to 2004.

Panel C reports the average components of the 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: Characteristics Selectivity (CS, measured in bps per quarter) decomposition for DFA US Micro-Cap fund and index funds as a group. t -values in parentheses.

	Total	Old	Adj	Trade	Impatient trading	Liquidity Prov	$CS^{imp} - CS^{liq}$
	CS	CS^O	CS^{adj}	CS^T	CS^{imp}	CS^{liq}	
	(=1+2+3)	(1)	(2)	(3=3a+3b)	(3a)	(3b)	(3a)-(3b)
DFA US Micro-Cap:							
Value (bps)	36.1	19.3	-4.2	21	0.5	20.5	-20
	(1.72)	(0.89)	(-0.64)	(1.30)	(0.06)	(1.84)	(-1.62)
Index Funds:							
Value (bps)	0	24.9	3.2	-28.1	-34.6	6.4	-41
	(0.00)	(0.93)	(0.50)	(-1.11)	(-2.19)	(0.36)	(-1.86)

Panel B: CS decomposition across styles. t-values in parentheses.

	Total	Old	Adj	Impatient trading	Liquidity Prov	Old	Adj	Impatient trading	Liquidity Prov
	CS	CS ^O	CS ^{adj}	CS ^{imp}	CS ^{liq}	CS ^O	CS ^{adj}	CS ^{imp}	CS ^{liq}
	average (bps per quarter)					% of variation in total CS explained			
All	23.5 (1.91)	13.9 (1.19)	-1.8 (-2.38)	3.6 (0.55)	8.8 (1.50)	56.8% (127.20)	-2.5% (-15.30)	37.2% (120.90)	8.4% (24.40)
AGG	54.7 (1.74)	30.8 (1.22)	-0.9 (-0.42)	17.4 (1.02)	7.3 (0.66)	52.1% (44.70)	-1.0% (-2.70)	44.9% (55.20)	4.0% (4.20)
Grow	25.7 (1.80)	14.8 (1.24)	-2.7 (-2.34)	5.6 (0.64)	8 (1.27)	55.7% (96.00)	-3.0% (-14.90)	37.0% (95.00)	10.2% (22.30)
GNI	12 (1.03)	3 (0.23)	0.1 (0.03)	-1.7 (-0.29)	10.7 (1.70)	54.1% (56.80)	-2.2% (-5.40)	37.0% (55.90)	11.1% (16.60)

Panel C: Return Gaps and components of CS measure (in bps per quarter) t-values in parentheses.

Return Gap	C1. Contemporaneous					C2. Next Quarter				
	RG (monthly)	CS ^O	CS ^{adj}	CS ^{imp}	CS ^{liq}	Four-factor alpha	CS ^O	CS ^{adj}	CS ^{imp}	CS ^{liq}
Low	-120.3	141.2 (7.31)	-8.4 (-3.48)	62 (2.63)	-13.8 (-1.39)	-47.8 (-3.88)	13.3 (0.82)	-0.9 (-0.43)	10 (0.83)	5.9 (0.62)
2	-30.3	62.2 (4.72)	-3.2 (-2.52)	15.6 (2.07)	2.2 (0.34)	-31.2 (-3.07)	16.1 (1.26)	-2.1 (-1.08)	0.8 (0.12)	6.5 (0.94)
3	-7.3	11 (0.87)	-3 (-1.84)	-0.1 (-0.01)	11.5 (1.83)	-21.7 (-2.19)	10.7 (0.86)	0.3 (0.22)	5.7 (0.82)	6.7 (1.07)
4	14.5	-32.6 (-2.64)	1.4 (0.66)	-11.6 (-1.92)	17.8 (2.71)	-19.8 (-2.10)	16.6 (1.24)	-2.6 (-1.79)	4.4 (0.48)	5.5 (0.86)
High	74.9	-101.2 (-7.03)	3.9 (2.03)	-34.7 (-3.43)	22.9 (2.86)	2 (0.14)	23.2 (1.37)	-4.1 (-1.97)	8.6 (0.76)	16.9 (1.97)
High - Low	195.2	-242.4 (-15.10)	12.4 (4.13)	-96.7 (-4.51)	36.7 (4.15)	49.9 (4.84)	9.9 (0.81)	-3.2 (-1.04)	-1.4 (-0.21)	10.9 (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-weighting 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 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-weighted across stocks at the fund level using the dollar value of the stock holding to arrive at a characteristics-adjusted pseudo fund return. The returns are reported in basis points (bps per quarter) and t-values are in parenthesis.

trade_PIN in Qtr t	Four-factor adjusted returns (after fee)				Three-factor adjusted returns (after fee)				Chracteristic-adjusted returns (CS) (before fee)			
	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	-38 (-3.45)	-37 (-3.51)	-40 (-3.45)	-38 (-3.75)	-40 (-3.73)	-44 (-4.22)	-48 (-4.04)	-44 (-4.41)	-3 (-0.29)	4 (0.42)	-2 (-0.22)	2 (0.25)
2	-40 (-4.01)	-39 (-4.34)	-44 (-4.25)	-33 (-3.47)	-35 (-3.60)	-34 (-3.94)	-40 (-3.85)	-37 (-3.78)	11 (0.92)	15 (1.40)	7 (0.70)	5 (0.52)
3	-32 (-2.87)	-30 (-3.03)	-39 (-4.46)	-41 (-3.81)	-30 (-3.01)	-27 (-2.80)	-33 (-3.66)	-37 (-3.43)	12 (1.04)	7 (0.74)	7 (0.72)	-3 (-0.32)
4	-31 (-3.09)	-36 (-3.66)	-37 (-4.31)	-37 (-3.38)	-23 (-2.24)	-27 (-2.96)	-31 (-3.48)	-30 (-2.84)	10 (1.01)	12 (1.19)	8 (0.84)	0 (0.01)
5	-17 (-1.57)	-26 (-2.39)	-35 (-3.24)	-34 (-3.21)	-6 (-0.61)	-15 (-1.42)	-25 (-2.50)	-21 (-2.01)	29 (2.17)	14 (1.15)	6 (0.51)	7 (0.64)
6	-27 (-2.18)	-17 (-1.33)	-19 (-1.36)	-31 (-2.68)	-18 (-1.48)	-5 (-0.41)	-4 (-0.34)	-17 (-1.59)	32 (2.07)	23 (1.58)	16 (1.26)	-5 (-0.37)
7	-25 (-1.80)	-22 (-1.67)	-16 (-1.24)	-22 (-1.58)	-8 (-0.61)	-6 (-0.46)	-1 (-0.08)	-9 (-0.66)	28 (1.52)	16 (1.04)	16 (1.19)	6 (0.41)
8	-19 (-1.16)	-3 (-0.22)	-8 (-0.52)	-18 (-1.14)	-7 (-0.45)	3 (0.20)	-2 (-0.10)	-10 (-0.67)	31 (1.73)	24 (1.41)	24 (1.53)	14 (0.85)
9	-18 (-1.06)	-14 (-0.82)	-18 (-1.11)	-18 (-1.18)	-7 (-0.42)	-2 (-0.14)	-5 (-0.31)	-5 (-0.34)	35 (1.75)	16 (0.84)	6 (0.35)	-4 (-0.21)
High	11 (0.85)	9 (0.67)	-5 (-0.36)	-3 (-0.21)	15 (1.27)	12 (0.91)	2 (0.13)	3 (0.23)	50 (2.70)	44 (2.75)	17 (1.20)	11 (0.68)
High - Low	48 (3.15)	46 (2.98)	35 (2.24)	35 (2.26)	56 (3.42)	56 (3.47)	50 (3.06)	47 (2.89)	53 (2.87)	40 (2.55)	19 (1.29)	9 (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 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 after fee, the characteristics-adjusted returns (CS) are before fee. The returns are reported in basis points (bps) per quarter.

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 parenthesis.

Panel A: PIN-risk-adjusted fund returns and alternative measures of information events

Portfolio	Control for PIN risk				Alternative measures of information events			
	Sorted on		Sorted on		Sorted on		Sorted on	
	trade_PIN in Qtr t		trade_adjPIN in Qtr t		trade_Theta in Qtr t		trade_aturn in Qtr t	
	5f alpha	CS*	4f alpha	CS	4f alpha	CS	4f alpha	CS
	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1
Low	-34	-8	-33	10	-35	5	-45	3
	(-2.97)	(-0.75)	(-3.20)	(0.79)	(-3.18)	(0.44)	(-3.46)	(0.20)
2	-32	10	-34	2	-31	22	-31	12
	(-2.98)	(0.79)	(-3.09)	(0.13)	(-3.11)	(1.71)	(-3.10)	(0.92)
3	-26	6	-38	10	-29	16	-32	21
	(-2.28)	(0.51)	(-3.44)	(0.79)	(-2.88)	(1.41)	(-3.18)	(1.78)
4	-26	14	-35	14	-33	11	-35	11
	(-2.57)	(1.10)	(-2.85)	(1.05)	(-3.64)	(0.97)	(-3.95)	(1.03)
5	-12	21	-22	33	-29	13	-30	10
	(-1.09)	(1.49)	(-1.86)	(2.26)	(-2.44)	(1.09)	(-2.82)	(0.85)
6	-21	28	-11	28	-22	20	-24	23
	(-1.60)	(1.63)	(-0.88)	(1.81)	(-1.81)	(1.26)	(-1.98)	(1.55)
7	-19	32	-19	30	-17	30	-26	16
	(-1.28)	(1.59)	(-1.33)	(1.85)	(-1.18)	(1.84)	(-1.76)	(0.99)
8	-19	22	-15	39	-6	49	-7	55
	(-1.05)	(1.23)	(-1.02)	(2.05)	(-0.39)	(2.55)	(-0.53)	(2.97)
9	-15	19	-12	37	2	43	-8	31
	(-0.86)	(0.88)	(-0.81)	(2.16)	(0.14)	(2.17)	(-0.50)	(1.60)
High	11	35	15	58	6	55	4	51
	(0.86)	(1.97)	(1.14)	(3.13)	(0.40)	(2.70)	(0.29)	(2.56)
High-Low	45	43	48	48	41	50	49	48
	(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				Past One-year Return			
	Buy	Sell	Buy-sell	t-value	Buy	Sell	Buy-sell	t-value
Low	11.23%	11.44%	-0.21%	(-3.39)	20.90%	24.30%	-3.40%	(-5.74)
2	12.35%	12.50%	-0.15%	(-2.50)	23.20%	26.40%	-3.20%	(-5.18)
3	12.98%	13.23%	-0.26%	(-4.26)	25.10%	26.80%	-1.70%	(-2.96)
4	13.72%	13.83%	-0.11%	(-1.82)	26.70%	29.60%	-2.90%	(-4.08)
5	14.46%	14.55%	-0.09%	(-1.26)	28.50%	32.10%	-3.60%	(-4.48)
6	15.29%	15.47%	-0.18%	(-1.37)	32.10%	36.20%	-4.10%	(-3.66)
7	16.31%	16.42%	-0.12%	(-1.22)	36.20%	39.60%	-3.40%	(-3.13)
8	17.67%	17.79%	-0.12%	(-1.15)	39.20%	43.80%	-4.60%	(-3.60)
9	19.36%	19.53%	-0.17%	(-1.45)	40.20%	46.10%	-5.90%	(-4.65)
High	22.94%	23.01%	-0.07%	(-0.41)	34.30%	46.60%	-12.30%	(-8.93)
	11.71%	11.57%			13.47%	22.36%		
High-Low	(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-weighting 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 basis points (bps) per quarter. *t*-values associated with the average measures are reported in parentheses.

Trade_PIN	Total	Old	Adj	Trade	Impatient	Liquidity	CS ^{imp} -
Qtr t	CS	CS ^O	CS ^{adj}	CS ^T	trading	Prov	CS ^{liq}
	(=1+2+3)	-1	-2	(3=3a+3b)	CS ^{imp}	CS ^{liq}	(3a)-(3b)
	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1	Qtr t+1	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	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	-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	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	26.5	-3.2	31.2	20.4	10.4	10
	(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 of 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”) alphas 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 and the difference between the realized return and the market return must have the same sign. *t*-values associated with the average measures are reported in parentheses. Alphas are reported in basis points (bps/month).

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
Low alpha	-32	0.8877	0.3081	0.0749	-0.0130	-34	-63	0.9238	0.2760	0.1656	-0.0465	-67
	(-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	-11	0.8650	0.1570	0.0679	0.0180	-10	-37	0.9047	0.1352	0.1409	-0.0093	-38
	(-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	-12	0.8561	0.1185	0.0820	0.0264	-9	-23	0.8869	0.1152	0.1207	0.0104	-22
	(-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	-8	0.8356	0.0911	0.1047	0.0109	-7	-21	0.9003	0.0817	0.1066	0.0089	-20
	(-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	-11	0.8385	0.0587	0.1088	0.0071	-10	-8	0.8828	0.0487	0.0967	-0.0091	-9
	(-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	-8	0.8087	0.0549	0.0927	0.0199	-6	0	0.8889	0.0560	0.0934	-0.0082	-1
	(-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	-7	0.8354	0.0677	0.0872	0.0062	-7	4	0.8641	0.0744	0.0529	-0.0135	2
	(-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	-7	0.8599	0.0777	0.0674	0.8599	-7	0	0.8731	0.1290	0.0620	-0.0064	-1
	(-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	-7	0.9014	0.1378	0.0372	-0.0094	-8	-1	0.8365	0.1718	0.0357	-0.0108	-2
	(-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	2	0.0594	-0.0440	0.1884	0.0045	1	15	0.8192	0.2456	-0.0525	-0.0265	12
	(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	34	594	-440	-1884	45	34	78	-0.1046	-0.0304	-0.2181	0.0201	80
	(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_PINs*

At the end of each quarter from 1983 to 2004, we conduct 2 by 5 double sorts based on *trade_PINs* 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 of 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’) alphas 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 fees (we apply the three filters described in Mamaysky, Spiegel and Zhang, 2007). Panel C reports results in subsample periods and uses alternative risk-adjustment models including the four-factor model and the seven-factor model of Cremers, Petajisto, and Zitzewitz (2008, CPZ4 and CPZ7) and the factor model of Comer, Larrimore and Rodriguez (2009, CLR). Following the procedures in Kosowski, Timmermann, Wermers, and White (2006, KTWW), we also report the *p*-values associated with both the alphas and their *t*-statistics corresponding to the four-factor model. *t*-values are reported in parentheses. Alphas are reported in basis points (bps/month).

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	-9	0.9209	-0.0519	0.0217	-0.0035	0.1292	-14	0.9343	-0.0654	0.0167	-0.0253
			(-1.86)	(78.36)	(-3.59)	(1.25)	(-0.34)						
low	medium	0.1280	-11	0.9465	-0.0723	0.0705	-0.0002	0.1287	-7	0.9350	-0.0799	0.0728	-0.0309
			(-2.81)	(94.40)	(-5.86)	(4.75)	(-0.02)						
low	high	0.1294	-8	0.9743	-0.0004	-0.0050	-0.0803	0.1303	6	0.9420	-0.0098	-0.0229	-0.0798
			(-1.24)	(63.09)	(-0.02)	(-0.22)	(-5.98)						
	high - low	-0.0001	1	0.0534	0.0515	-0.0267	-0.0767	0.0011	19	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	-23	1.0184	0.3776	0.0303	0.0222	0.1792	-34	1.0449	0.3176	0.0738	-0.0085
			(-3.34)	(59.89)	(18.05)	(1.20)	(1.50)						
High	medium	0.1781	-7	0.9679	0.2676	0.0938	0.0219	0.1801	-9	1.0044	0.2853	0.0474	0.0096
			(-1.09)	(62.98)	(14.15)	(4.12)	(1.64)						
High	high	0.1821	13	1.0092	0.3825	-0.0962	0.0408	0.1818	35	0.9481	0.3384	-0.0412	0.0186
			(1.65)	(51.67)	(15.91)	(-3.33)	(2.40)						
	high - low	0.0039	36	-0.0091	0.0049	-0.1265	0.0186	0.0026	69	-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: 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	-16	-4	-6	2.36%
		(-2.07)	(-0.47)	(-0.78)	(1.90)
low	medium	-10	1	-1	2.98%
		(-1.92)	(0.28)	(-0.29)	(2.90)
low	high	-2	14	7	6.46%
		(-0.20)	(1.57)	(0.73)	(5.58)
	high - low	14	18	13	4.10%
		(1.15)	(1.42)	(1.03)	(3.94)
High	low	-34	-22	-23	1.42%
		(-3.63)	(-2.32)	(-2.45)	(1.15)
High	medium	-8	0	1	4.76%
		(-1.14)	(0.01)	(0.13)	(4.15)
High	high	37	45	47	8.49%
		(3.57)	(4.24)	(4.50)	(5.57)
	high - low	71	67	70	7.08%
		(6.94)	(4.50)	(4.78)	(6.56)

Panel C: Subsample results, alternative risk adjustment models and bootstrap p -values

Trade_PIN Qtr t-1	Previous 4f alpha	FF4	CPZ4	CPZ7	FF4	CPZ4	CPZ7	FF4	CPZ4	CPZ7	CLR	KTWW p -values	
		1983 - 2004			1983 - 1993			1994 - 2004			alpha	t-stat	
low	low	-14	-16	-20	2	-5	-12	-24	-25	-30	-25	0.02	0.02
		(-1.72)	(-2.05)	(-2.80)	(0.20)	(-0.45)	(-1.23)	(-2.35)	(-2.38)	(-3.00)	(-2.31)		
low	medium	-7	-9	-8	1	-2	-3	-10	-14	-14	-9	0.07	0.07
		(-1.34)	(-1.84)	(-1.68)	(0.11)	(-0.37)	(-0.42)	(-1.50)	(-2.19)	(-2.14)	(-1.30)		
low	high	6	4	9	-2	-5	-6	13	9	17	18	0.26	0.26
		(0.63)	(0.40)	(0.89)	(-0.26)	(-0.58)	(-0.60)	(0.85)	(0.61)	(1.15)	(1.12)		
	high - low	19	20	29	-4	0	6	37	34	47	43	0.04	0.04
		(1.54)	(1.61)	(2.18)	(-0.32)	(0.04)	(0.45)	(1.80)	(1.67)	(2.27)	(1.97)		
High	low	-34	-25	-29	-25	-25	-27	-40	-26	-29	-41	<0.01	<0.01
		(-3.47)	(-2.78)	(-2.96)	(-2.91)	(-2.91)	(-2.81)	(-2.41)	(-1.64)	(-1.87)	(-2.44)		
High	medium	-9	-2	-2	2	2	8	-18	-6	-8	-17	0.19	0.21
		(-1.24)	(-0.34)	(-0.32)	(0.24)	(0.24)	(0.93)	(-1.73)	(-0.52)	(-0.85)	(-1.58)		
High	high	35	40	47	36	36	40	37	46	54	47	<0.01	<0.01
		(3.33)	(4.26)	(4.43)	(3.08)	(3.08)	(2.78)	(2.22)	(3.03)	(3.55)	(2.79)		
	high - low	69	66	76	61	61	67	76	72	84	89	<0.01	<0.01
		(4.61)	(4.46)	(4.55)	(4.11)	(4.11)	(3.56)	(2.95)	(2.80)	(3.18)	(3.30)		

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_PINs* and four-factor alphas of the mutual funds. We then decompose the next-quarter total Characteristic Selectivity (CS) measure within each portfolio. The DGTW CS measure (column 1) and its component (columns 2 to 5) are reported in basis points (bps). The size, book-to-market, momentum and mutual fund trading induced flow four-characteristic adjusted return is reported in the last column. The *t*-values associated with the average measures are reported in parentheses. The returns are reported in basis points (bps) per quarter.

Trade_PIN 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 ^{imp} (3a) Qtr t+1	Liquidity Prov CS ^{liq} (3b) Qtr t+1	CS ^{imp} - CS ^{liq} (3a)- (3b) Qtr t+1	Total CS (4-char.) Qtr t+1
low	low	-7.89 (-0.62)	-8.96 (-0.69)	-1.92 (-1.21)	-17.09 (-2.33)	20.09 (2.64)	-37.17 (-2.90)	-27.73 (-2.13)
	medium	19.99 (1.72)	21.33 (1.64)	-1.10 (-0.77)	-1.66 (-0.26)	1.48 (0.22)	-3.14 (-0.27)	0.53 (0.06)
	high	27.98 (1.55)	17.78 (1.13)	-1.19 (-0.53)	4.87 (0.37)	6.51 (0.78)	-1.63 (-0.09)	30.98 (1.61)
	high - low	35.87 (1.75)	26.74 (1.42)	0.73 (0.28)	21.96 (1.79)	-13.58 (-1.56)	35.54 (3.11)	58.71 (2.42)
high	low	-6.79 (-0.33)	-11.17 (-0.59)	-2.83 (-1.60)	0.22 (0.02)	6.99 (0.80)	-6.77 (-0.38)	-15.74 (-0.68)
	medium	33.41 (2.02)	16.59 (1.04)	1.61 (0.87)	5.51 (0.56)	9.70 (1.21)	-4.19 (-0.28)	33.86 (2.19)
	high	106.86 (3.20)	56.32 (2.09)	-5.35 (-1.69)	48.04 (2.63)	7.90 (0.67)	40.14 (1.55)	96.79 (2.57)
	high - low	113.65 (4.14)	67.49 (2.85)	-2.52 (-0.76)	47.82 (2.99)	0.91 (0.08)	46.91 (2.91)	112.54 (3.20)

Table 9: Cross-sectional analysis

At the end of each quarter from 1983 to 2004, we regress the next quarter characteristics-adjusted 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; *act_share* is a variable proposed by Cremers and Petajisto (2009) to capture the deviations of a manager’s portfolio holdings from the underlying benchmark index holdings; *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 parenthesis. There are on average 300 funds in each cross-section.

	Intercept	trade_PIN	log_fund_size	log_fund_bm	fund_mom	fund_amihud	log_TNA	Age	flow	turnover	act_share	dummy_growth	dummy_Agg	Average R ²
LHS = CS														
coeff	0.0020	0.0022	0.0019	0.0004	0.0011	-0.0005	0.0004	0.0001	0.0013	0.0000	0.0007	0.0004	0.0045	0.20
t-value	(1.04)	(1.90)	(1.24)	(0.28)	(0.83)	(-0.40)	(0.98)	(0.40)	(3.75)	(-0.07)	(2.22)	(0.38)	(2.92)	
LHS = CS ^{imp}														
coeff	0.0004	0.0011	0.0009	0.0014	0.0011	-0.0004	0.0001	0.0000	0.0010	0.0000	0.0003	0.0000	0.0017	0.13
t-value	(0.49)	(2.17)	(1.08)	(1.79)	(1.01)	(-0.62)	(0.76)	(-0.19)	(3.80)	(-0.02)	(1.20)	(-0.05)	(1.89)	
LHS = CS ^{liq}														
coeff	0.0018	-0.0004	-0.0001	-0.0008	-0.0001	0.0006	-0.0001	-0.0003	-0.0002	0.0003	0.0001	-0.0010	-0.0018	0.09
t-value	(3.19)	(-0.87)	(-0.16)	(-1.86)	(-0.29)	(1.15)	(-0.34)	(-2.06)	(-1.15)	(0.89)	(0.24)	(-2.02)	(-2.26)	

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 1991Q2-1993Q1 and 1995Q4-1998Q2. Each quarter, we sort fund managers in the Plexus Group data into quintiles based on the intensity of their interim trading defined as the ratio of the total value of within-quarter roundtrip trades to 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 the "impatient-trading" or the "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 number of managers	avg number 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%