Comments Welcome

Informed Trading, Liquidity Provision and Stock Selection by Mutual Funds

Zhi Da, Pengjie Gao and Ravi Jagannathan*

First Draft: May 10, 2007
This Draft: September 3, 2007

Abstract

Active mutual fund managers with superior information-processing skills can add value through both "informed trading" and "liquidity provision." As information loses value over time, informed trading tends to be liquidity-absorbing. We conjecture that value enhancing informed trading is more likely in stocks during times when they are associated with more information events. In contrast, liquidity provision is more likely to add value for stocks associated with few information events and little adverse selection risk. We decompose the stock selection skill of a manager into several components including a liquidity absorbing informed trading component and a liquidity providing component and identify times when there are more information events associated with a stock by its Probability of Informed Trading (PIN, Easley et al., 1996) measure. We provide empirical support for our conjecture using quarterly mutual fund holdings data for the period from 1983 to 2004.

*We would like to thank Robert Battalio, Craig Holden, Rick Mendenhall, Paul Schutlz, Charles Trzcinka and seminar participants at IU/ND/Purdue Finance Symposium and Northwestern University for comments. 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 2006, US mutual fund managers collectively have over $10 trillion under their management with almost $6 trillion of it equity funds. A significant portion of this amount is actively managed. For example, in 2006 alone, US mutual funds bought and sold common stocks worth over $6 trillion.[1] Naturally, investors would like to understand how active fund managers add sufficient value to justify their higher fees. The early literature on portfolio performance evaluation (see: Jensen (1968), Gruber (1996) and Carhart (1997)) find that most managed portfolios earn close to zero or negative risk-adjusted returns especially after taking fees into account. In contrast, the more recent studies that make use of quarterly reports of mutual funds disclosing their stock holdings (e.g. Grinblatt and Titman (1989, 1993), Wermers (1997), Daniel, Grinblatt, Titman and Wermers (1997), Chen, Jegadeesh and Wermers (2000), Schultz (2007)) find active managers to possess considerable stock-picking abilities. On average, after adjusting for the stock characteristics but before deducting fees and expenses, stocks held by mutual funds outperform their benchmarks and stocks bought by mutual funds tend to outperform those sold by mutual funds. Further several mutual fund characteristics appear to be positively related to superior stock-selection skills. These characteristics include: (1) aggressive growth and growth fund stye (Daniel, Grinblatt, Titman and Wermers, 1997); (2) shorter average geographical distance between holdings and mutual funds headquarters (Coval and Moskowitz, 2001); (3) smaller fund size (Chen, Hong, Huang and Kubik, 2004); (4) higher industry concentration of fund's holdings (Kacperczyk, Sialm and Zheng, 2004; Lubomira, 2005); (5) less diversification in fund's holdings (Baks, Busse and Green, 2006); (6) larger deviations from passive index or larger "active shares" (Cremers and Petajisto, 2006) and (7) less dependency on analyst's recommendation (Kacperczyk and Seru, 2007).

Ultimately, a mutual fund manager's skill comes from her ability to better process valuation-relevant information on a stock which thereby enables her to identify potential mispricing. There are at least three different channels in which the superior information processing skill of a manager can add value, depending on how long it takes for the market to realize that the manager is right. First, the manager may be a patient long-term investor taking a position in a stock expecting the market to eventually agree with her view in, say, a few years.[2] In that case, the exact timing of

[1] These numbers are taken from Table 3 and 30 of the Mutual Fund Facts Book (2007) published by the Mutual Fund Institute.

[2] Using fundamental analysis, Mario Gabelli, a money manager, realized that the stock of Hudson General Corp (HGC) was heavily undervalued at around $25 in earlier 1994 and started to accumulate shares of HGC for his Gabelli funds (see Figure 1A). The investment started to payoff after two years when the stock price increased to $40. The

trades would not be critical. Evaluating the stock selection skill of such a portfolio manager who makes a few concentrated long term bets will be difficult based only on quarterly observations on what the manager holds. Second, the manager may be a medium-term trader expecting the market to agree with her view within, say, a quarter or two.[3] In that case the value of the information is likely to erode quickly over time and trades may have to be executed by paying a substantial price concession for immediacy. Third, the manager may be a short-term liquidity provider who provides liquidity by taking the other side of a trade thereby making a profit through market making activities. Since fund managers often hold an inventory of several of the stocks in their universe in order to track their performance benchmarks, they have a natural advantage when making a market in those stocks.[4] The superior knowledge about the stocks covered by a manager will help in any market making activities by minimizing potential losses that may arise by trading with those with an information advantage. Since portfolio managers in the latter two classes trade sufficiently frequently, examining how the stocks they bought and sold following their trades should help in assessing their stock selection skills with more confidence.

While in theory knowledge of what the manager holds should help evaluate a portfolio manager's skill better, the fact that mutual fund stock holdings data are available only at infrequent intervals (quarterly at most) makes it difficult to assess a manager's abilities when the manager

market eventually agreed with Mr. Gabelli when Lufthansa took over HGC at $76 per share. See Greenwald, Kahn, Sonkin and Biema (2001) for details on this case.

[3]For Starbucks, the year-to-year same store sales growth reported by the company every month is a widely watched number and matters a lot for stock price (probably more important than earnings announcement). From Jan to Sep 2005, Starbucks was reporting sales growth in the range of 7% to 9%. However, many people attribute a large part of it to a 3% sales price increase took place in Oct 2004. The 3% will not help for the sales growth starting Oct 2005, leading to much smaller expected growth rate (analyst's consensus is 3.6%) and could trigger price drop. However, a careful analysis of sales breakdown would suggest that the 3% price increase plays a less important role so the Oct sales growth figure should still be good. While most mutual funds decreased their holdings of Starbucks stock during Q3 of 2005, in anticipation of a drop of same-store sales growth announcement for Oct, Putnam Voyager Fund actually accumulated more shares (see Figure 1B). In Nov 3, Starbucks reported a more-than-solid sales growth of 7% for Oct and share price jumped. Details on this case can be found in Blumenthal (2007).

[4]In early 2006, a change in accounting rule requires companies to count stock-option-based compensation as an expense in reporting earnings. On Feb 7, 2006, Vertex Pharmaceuticals Inc. reported 2005Q4 earnings and issued the guidance for 2006's earnings forecast to be a loss of $205m to $225m. This forecast includes option expense. Thomson Financial which collects earnings forecasts, however, made a mistake and recorded the "ex-option-expense" version earnings estimate of a loss from $195m to $200m. On the next day, Vertex's share price jumped by 2.3% amid heavy buying pressure (trades are on average buyer-initiated) as some market participants thought analysts had revised up their estimates. Vertex noticed the mistake and notified Thomson Financial. The mistake was then corrected. On Feb 9, the stock price fell back by 1.4% amid selling pressure (trades are on average seller-initiated). Mutual fund who held the stocks on Feb 7 could benefit from liquidity provision on Feb 8. See Wall Street Journal article "Moving the market – tracking the numbers / street sleuth: Options expensing Jars consensus" (dated 4 April 2006) for details.

trades actively in between two holdings reporting points in time.[5] In spite of this limitation, a mutual fund's recent trades inferred from their quarterly holding changes, will contain interesting information about a manager's abilities if we can separate the value-added from the second and the third channel discussed above in order to reduce the noise in the data. That becomes possible when we recognize that (a) since information loses value over time, informed trading tends to be liquidity-absorbing on average; (b) as informed trading adds value from superior information-processing skills, value enhancing informed trading is more likely to take place in stocks during times when they are associated with more information events; (c) liquidity provision is more likely to add value for stocks associated with few information events and therefore little adverse selection risk (Glosten and Harris (1998)); and (d) it is possible to identify times when information events affect a stock at a higher frequency and intensity using the Easley, Kiefer, O'Hara and Paperman (1996) Probability of Informed Trading (PIN) measure.

We therefore decompose the stock selection skill of a manager into a liquidity absorbing informed trading component, a liquidity providing component, and other components using quarterly mutual fund holdings data from 1983 to 2004. We start with a popular holding-based measure of mutual fund stock selection skill - the "Characteristic Selectivity" (CS) - proposed by Daniel, Grinblatt, Titman and Wermers (DGTW, 1997), which measures the extent to which managers can select stocks that outperform the average stocks having the same characteristics.[6] We first decompose the CS measure into three components: a passive "buy-and-hold" component (CS^P), a small adjustment component due to fund flows (CS^{adj}) and an active component due to trading in the previous quarter (CS^A). The active component (CS^A) which captures return to managers' recent trades are more informative about their stock selection skills. CS^A can be further decomposed into an informed trading component (CS^{inf}) and a liquidity provision component (CS^{liq}), respectively. The last decomposition is motivated by the evidence that the stock level aggregate order imbalance serves as a good measure of the direction of liquidity needs on the underlying stock (see Chordia and Subrahmanyam (2004) among others). When managers trade in the same direction as the

[5] For instance, Kacperczyk, Sialm and Zheng (2007) and Elton et al. (2006) show that the "unobservable" actions (or high-frequency turnovers) by mutual funds could be important for some funds. Campbell, Ramadorai and Schwartz (2007) attempts to infer insitutional transactions within a quarter by selecting trade sizes to best match quarterly holding changes.

[6] Chan, Dimmock and Lakonishok (2006) discuss a variaties of performance benchmarks, and conclude that characteristic matching method may generate better tracking ability than regression-based procedures.

aggregate market order imbalance, they demand liquidity. Such trades are therefore likely driven by information and are classified as "informed trading". On the other hand, when managers trade in the opposite direction of the aggregate market order imbalance, they supply liquidity and such trades are classified as "liquidity provision". We do not mean to say that the trades motivated by informed trading and liquidity provision are orthogonal to each other - it may well be the case that these two types of transactions can be complimentary – we only aim at capturing the average effects of these transactions given our definition of different trading motives.

We document that when fund managers open new positions and close out or increase existing positions, they are likely to absorb market liquidity. When they decrease existing positions, they are likely to provide liquidity, consistent with our earlier conjecture that it is easier to provide liquidity on stocks currently in one's possession. We also demonstrate the effectiveness of this decomposition approach in two specific cases: (1) Dimensional Fund Advisors (DFA) and (2) a group of index funds.[7] We confirm that the decomposition results in those two cases are largely consistent with what one would expect.

We then apply the decomposition to portfolios of active mutual funds sorted by the trade-value-weighted-average-PIN of stocks they recently traded ($trade\_PIN$). Several interesting patterns emerge. First, funds trading high-PIN stocks outperform those trading low-PIN stocks by 53 bps per quarter (t-value = 2.87) after controlling for stock characteristics such as size, book-to-market ratio and return momentum. Using after-fee mutual fund returns from CRSP mutual fund database, we obtain similar results. Specifically, funds trading high-PIN stocks outperform those trading low-PIN stocks by 50 bps per quarter (t-value = 3.27) after four-factor risk adjustment (Fama and French, 1993 and Carhart, 1997), indicating that the better performance is unlikely to be driven by the window dressing actions of mutual funds. Second, a large part of the CS measure for high-$trade\_PIN$-funds (CS = 50 bps) indeed comes from active trading during the previous quarter ($CS^A = 30.7$ bps). Although both the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) are positive for high-$trade\_PIN$-funds, only the the

[7] For example, Keim (1999) shows that a passive index fund, the small-cap stock fund - "9-10 fund" - launched by the Dimensional Fund Advisors (DFA), outperforms its benchmark by roughly 2.2% during the period between 1982 to 1995. His case study reemphasizes how skillful trade execution enhances fund performance. Cohen (2002) provides a thorough description of the operation of DFA. He documents that managers at DFA make money by systematically providing liquidity to those who want to trade small stocks for non-information based reasons. We verify that most of DFA's stock selection skills indeed come from the liquidity provision component (CS^{liq}).

informed trading component is significant (20.4 bps with t-value of 2.25) and its size is twice that of the liquidity provision component (10.4 bps). The liquidity provision component is positive – certain skilled managers, by judiciously choosing their trades, could potentially benefit from the price impact working to their advantage, which could be sizable for high-PIN stocks. However, for all funds trading high-PIN stocks as a group, the positive liquidity provision is not significant (t-value = 1.37), probably because for most managers in this group this is a smaller fraction of the trades they execute. Moreover, some difficulty comes with detecting high-frequency liquidity events using quarterly holdings data. Third, we document a positive significant liquidity provision component (CS^{liq} = 16.2 bps per quarter with t-value of 2.57) for funds trading low-PIN stocks. For low-PIN stocks, fund managers are less likely to encounter informed traders when they trade. Consequently, when they trade against the market order imbalance, they are more likely to benefit from the price impact.

Easley, Hvidkjaer and O'Hara (2002) document that High-PIN stocks earn higher returns and interpret this as compensation for risk associated with private information. This PIN-risk should not drive our results – although stocks bought and sold by mutual funds are of similar PINs, the stocks bought by mutual funds still outperform those sold by mutual funds. In addition, we show that the results are not driven by momentum trading using mechanical rules described in the momentum literature.[8] Finally, our results are not driven by the choice of PIN as a measure of private information which assumes specific information arrival process. An easy-to-compute and direct measure of order imbalances (Aktas, Bodt, Declerck and Oppens (2007)) delivers very similar results.

In a cross-sectional regression framework, we attempt to identify fund characteristics that are associated with informed trading and liquidity provision. We first confirm that funds trading high-PIN stocks have a larger informed trading component. In addition, the informed trading component is associated with growth-oriented investment style. In contrast, the liquidity provision component is associated with younger funds and funds with income-oriented investment style. Finally, we document that the informed trading component of the skill of a mutual fund manager appears to persist for a while from one quarter into the next.

[8] For the related discussion of institutional momentum and feedback trading, see Lakonishok, Shliefer and Vishny (1992), Grinblatt, Titman and Wermers (1995), Nofsinger and Sias (1999), Badrinath and Wahal (2002), and Sias, Starks and Titman (2006), among others.

The remainder of this paper is organized as follows. Section 1 introduces our data sources and sample constructions. Section 2 introduces the decomposition approach and applies it to two specific examples; Section 3 presents the main empirical results when we apply the decomposition approach to portfolios of active funds sorted on $trade\_PIN$; Section 4 concludes and the Appendices contain a numerical example on skill decomposition, a brief discussion on the variance decomposition approach, and brief descriptions of the PIN measure and the approximate PIN measure.

1 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[9] and lifecylce funds as these are hybrid funds. In addition, following the standard practice in the mutual fund literature, we omit international funds, sector funds, bond funds, and domestic hybrid funds based on the self-reported fund style in the CDA/Spectrum database. Thus, we only keep funds that are self-reported as Aggressive Growth (AGG), Growth or Growth and Income (GNI). To ensure that the funds we examine are reasonably active, we only include fund / quarter observations if the fund trades at least 10 stocks and turns over at least 10% of its holdings during that quarter. Finally, we only include fund / quarter observations for which the fund holdings at the end of previous quarter are also available so holding changes can be computed over consecutive quarters. We obtain the information on the after-fee performance of the fund and other fund characteristics from the CRSP survivor-bias-free mutual fund database. The CDA/Spectrum mutual fund holding data are matched to CRSP Mutual Fund data using the MFLINKS database produced by the Wharton Research Data Service (WRDS) and Russ Wermers. An appealing feature of MFLINKS database is that it allows us to map different funds or different share classes of funds as recorded by CRSP mutual fund database which share the same underlying portfolio as recorded by CDA/Spectrum mutual fund holding database. For multiple funds or multiple share classes with the same holding data, we aggregate these funds or

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

share classes into one by equal-weighting or value-weighting (using the total net asset values). We report results from equal-weighting although the value-weighting results are similar.

The stock data come from the Center for Research in Security Prices (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 4654 distinct funds in our sample during the period from 1983 to 2004. On average, there are about 701 distinct funds every quarter. The number of funds per quarter increases from about 134 in 1983 to about 1700 more recently as in Table 1. About 61% of the funds in our sample are self-reported as "Growth" funds, about 26% are reported as "Growth and Income (GNI)" and the remaining 13% are reported as "Aggressive Growth (AGG)".

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

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

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

2 A Decomposition of Fund Stock Selection Skill

Daniel, Grinblatt, Titman and Wermers (DGTW, 1997) and Wermers (2004) develop a "Characteristic Selectivity" (CS) measure to detect whether managers can select stocks that outperform the average stocks with the same characteristics. By examining the actual stock holdings of the mutual fund, the CS measure during quarter $t+1$ is computed as

$$CS_{t+1} = \sum_j w_{j,t} \left[R_{j,t+1} - BR_{t+1}(j,t) \right], \tag{1}$$

where $R_{j,t+1}$ is the return on stock j during quarter $t+1$, $BR_{t+1}(j,t)$ is the benchmark portfolio return during quarter $t+1$ to which the stock j is matched at the end of quarter t based on its size, book-to-market equity ratio, and past 12-month return; and $w_{j,t}$ is the dollar value weight of stock j held by the mutual fund at the end of quarter t. In this section, we propose a further decomposition of the CS measure. A numerical example is provided in Appendix A. To better understand the intuition behind the decomposition, we make the simplified assumption that funds only trade at the end of the quarter.

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, three stock portfolios can be defined:

(1) "Hold" portfolio, whose stock holdings are:

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

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

(2) "Buy" portfolio, whose stock holdings are:

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

"Buy" portfolio contains stocks bought by the fund during quarter t.

(3) "Sell" portfolio, whose stock holdings are:

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

"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 and denote the market value of "Hold", "Buy" and "Sell" portfolios as H_t, B_t and S_t accordingly, we have:

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

At the end of quarter t, the mutual fund stock holding is a combination of the "Hold" portfolio and "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 $t+1$, respectively.

We then decompose the CS measure into three components:

$$\begin{aligned} CS_{t+1} &= CS_{t+1}^P + CS_{t+1}^A + CS_{t+1}^{adj}, \\ CS_{t+1}^P &= \frac{H_t}{H_t + S_t} CS_{H,t+1} + \frac{S_t}{H_t + S_t} CS_{S,t+1}, \\ CS_{t+1}^A &= \frac{B_t}{H_t + B_t} CS_{B,t+1} - \frac{S_t}{H_t + S_t} CS_{S,t+1}, \\ CS_{t+1}^{adj} &= \frac{H_t}{H_t + B_t} \frac{S_t - B_t}{H_t + S_t} CS_{H,t+1}. \end{aligned} \tag{2}$$

The first component, CS^P_{t+1}, can be interpreted as the CS measure on the fund as if the fund adopts a passive strategy. If nothing happens to the fund during quarter t, its stock holding would remain unchanged ($N_t = N_{t-1}$) and would be comprised 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. The second component, CS^A_{t+1}, which measures the characteristics-adjusted returns to the recent mutual fund stock trades, captures the value-added from active fund trading during quarter t. As most fund managers are evaluated by comparing their performance against a performance benchmark, a large part of their holdings are chosen to minimize benchmark tracking errors. As a result, the active component (CS^A) is often more informative with regard to their stock selection skills. Finally, CS^{adj}_{t+1} represents an adjustment term whenever $S_t \neq B_t$, which could happen when there is inflow or outflow to the fund for example.

We then further decompose CS^A_{t+1} into two components by comparing the sign of quarterly mutual fund holding change and the sign of market order imbalance for each stock traded by the fund (the stocks in the "Buy" or "Sell" portfolio) during quarter t. The market order imbalance is defined as the total number of buyer-initiated trades minus the total number of seller-initiated trades in the quarter. Consistent with the literature, the trade classification is done using the standard algorithm in Lee and Ready (1991). We then classify stock trades where the two signs are identical into one group denoted using superscript "+" and those where the two signs are different into another group denoted using superscript "−". As a result, the characteristics-adjusted returns on trades from these groups sum up to CS^A_{t+1}:

$$\begin{aligned}
CS^A_{t+1} &= CS^{inf}_{t+1} + CS^{liq}_{t+1}, \\
CS^{inf}_{t+1} &= \frac{B^+_t}{H_t + B_t} CS^+_{B,t+1} - \frac{S^+_t}{H_t + S_t} CS^+_{S,t+1}, \\
CS^{liq}_{t+1} &= \frac{B^-_t}{H_t + B_t} CS^-_{B,t+1} - \frac{S^-_t}{H_t + S_t} CS^-_{S,t+1}.
\end{aligned} \tag{3}$$

Given that the aggregate market order imbalance is a good measure of the direction of liquidity needs on the stock (see Chordia and Subrahmanyam (2004)), CS^{inf}_{t+1} measures the characteristics-adjusted return on mutual fund trades that on average absorb market liquidity. Such trades are likely driven by information and therefore classified as "informed 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 classified as "liquidity provision". In an extreme case where the fund manager only trades one stock and when the time interval is a minute rather than a quarter, CS_{t+1}^{liq} will then closely resemble the realized spread of Huang and Stoll (1996) which measures the reward to market makers' liquidity provision activities. With quarterly holdings data, CS_{t+1}^{liq} is likely to underestimate the true reward for liquidity provision as it only captures liquidity-shock-induced price reversals that persist over a calendar quarter end. To summarize, we decompose the fund CS measure as:

$$\begin{aligned} CS_{t+1} &= CS_{t+1}^{P} + CS_{t+1}^{adj} + CS_{t+1}^{A}, \qquad (4) \\ CS_{t+1}^{A} &= CS_{t+1}^{inf} + CS_{t+1}^{liq}. \end{aligned}$$

As mutual funds' trades can only be inferred from changes in mutual fund holdings at quarterly frequency, we would therefore miss out high-frequency turnovers by mutual funds as studied in Kacperczyk, Sialm and Zheng (2007) and Elton et al. (2006).[12] These data limitations, combined with possible trade misclassifications in the order imbalance calculation using Lee and Ready's (1991) algorithm, introduce noises to the decomposition empirically. These noises should bias us against finding any significant results. In addition, there are also scenarios where our decomposition might be inappropriate. For example, a typical mutual fund, governed by its prospectus, usually avoids holding financially distressed stocks. Da and Gao (2006) show that when a stock suddenly becomes financially distressed, the fund is forced to sell it. Such a trade is likely to be seller-initiated and will therefore be misclassified as "iInformed trading" by our methodology. Given the fact that financially distressed stocks usually have very small market values and each component of the CS measure is computed using value-weighted average, such misclassification should have a relatively small impact on our results.

For active funds in our sample, we examine their mutual fund holding changes over two consec-

[12] Preliminary analysis suggests that results in our paper are not driven by such "unobservable actions" of mutual funds. We obtain very similar results after removing fund / quarter observations associated with extreme "return gaps" (top and bottom 20%) defined in Kacperczyk, Sialm and Zheng (2007). In addition, the return gap is not significant in explaining CS measure and its component in cross-sectional regressions.

utive quarters and categorize them into four groups: (1) "Open" (holdings increase from zero to positive); (2) "Close" (holdings decrease from positive to zero); (3) "Increase" (holdings increase but not from zero) and (4) "Decrease" (holdings decrease but not to zero). For each group, we then compute (1) the average dollar holding change as a percentage of the total dollar holding change of the fund (computed using prices at quarter end); (2) the average order-imbalance measure (defined as the difference between total numbers of buyer-initiated shares brought and total numbers of seller-initiated shares sold divided by total number of shares traded during the quarter, the resulting number is then cross-sectionally demeaned) and (3) the associated t-value. The results are provided in Table 2. The average order imbalance measure for each trade type tells us whether the trade is on average absorbing liquidity or demanding liquidity. We document that when fund managers open new positions and close out existing positions, they are likely to absorb market liquidity. In those cases, these trades are likely motivated by large "mispricings" perceived by fund managers who are willing to pay for the price of immediacy. When mutual funds adjust their holdings, they are likely to provide liquidity on average only when they decrease their holdings, consistent with our conjecture that it is easier to provide liquidity on stocks that one currently owns.

Before applying the decomposition to the entire sample of active US equity mutual funds in the next section, we first demonstrate the effectiveness of our decomposition methodology using two specific examples: (1) Dimensional Fund Advisors and (2) a group of index funds.

2.1 Illustrative examples

2.1.1 Dimensional Fund Advisors (DFA)

Dimensional Fund Advisor (DFA) is an asset management firm founded in 1981. It firmly believes in efficient markets (Fama, 1965) and does not pick stocks via fundamental analysis. Instead, the firm helps its clients get exposure to certain segments of the asset markets. It is well documented that funds managed by DFA create value by systematically providing liquidity to those who want to trade small stocks for non-information related reasons.[13] Consequently, one would expect to find a positive liquidity provision component in DFA's CS measure and an informed trading component close to zero. We examine the quarterly stock holdings of DFA's flagship fund, US Micro Cap

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

Portfolio, during the period from 1983 to 2004 and decompose its CS measure. The results are provided in Table 3. The overall CS measure for the fund is 36.1 bps per quarter and is close to being statistically significant (t-value = 1.72), indicating that the fund does possess some ability 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 the firm's investment philosohpy.

2.1.2 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). These trades would be incorrectly classified as "Informed Trading" within our decomposition framework, and the Informed Trading component, if deviating from zero, is likely to be negative. It is therefore less appropriate to apply the decomposition to index funds. We will then be focusing only on actively managed funds for the remaining parts of the paper. Nevertheless, the examination of 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 funds identified each quarter on average from the holding database, starting from 1 fund each quarter in 1983 to about 25 funds each quarter after 2000. Using their stock holdings, we apply our decomposition to each fund and the results are then equally-weighted across funds during every quarter. The results are again presented in Table 3. 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^P 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^P and CS^{liq} are offset by a negative Informed Trading component (CS^{inf}= -35 bps) which is statistically significant, indicating a sizable price for liquidity paid by the index funds, most likely during the period of index rebalancing to minimize the tracking errors.

3 Decomposing Stock Selection Skills for Active Fund Managers

We implement the decomposition for all active equity funds in our sample. To examine the relative importance of each component of the total CS measure, we carry out a variance decomposition exercise similar to that used in Vuolteenaho (2002). The details are provided in Appendix B. In a nutshell, the variance decomposition delineates how much the cross-sectional variations of the total CS measure can be attributed to the cross-sectional variation in each of its four components. The results are reported in Table 4 for the full sample of all active US equity managers and across three style-subsamples. Overall, the passive component (CS^P) explains about 57% of the total cross-sectional variation in the total CS measure. The informed trading component (CS^{inf}) explains about 37% of the total variation, more important than the liquidity provision component (CS^{liq}) which explains slightly more than 8%. In addition, CS^{inf} becomes relatively more important for growth-oriented funds while CS^{liq} becomes relatively more important for income-oriented funds.

The average magnitude of each component is summarized in the top portion of Table 6. Overall, the active fund managers seem to possess some stock selection skill. The average CS measure is 23.5 bps per quarter (t-value = 1.91), indicating stocks selected by the fund managers outperform those with similar characteristics. Out of the 23.5 bps, 13.9 bps come from the passive "buy-and-hold" strategy and 14.2 bps come from stocks recently traded by the funds. The adjustment component is small (-1.9 bps) in absolute term but significant, potentially driven by fund flow to managers with skills as empirically documented by Chevalier and Ellison (1997) among others and theoretically analyzed by Berk and Green (2004).[14] Finally, although both the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) are positive, neither is significant.

3.1 Stock selection and PIN

As informed trading adds value through superior information-processing skills, value enhancing informed trading is more likely to take place in stocks during times when they are associated with

[14] When managers have skill (CS^P is likely to be positive), fund inflow is more likely ($B > S$); When managers have no skill (CS^P is likely to be negative), fund outflow is more likely ($S > B$). Both effects lead to a negative CS^{adj} as in equation (2).

more information events. To identify the occurrence of information events, we make use of the Probability of Informed Trading measure (PIN), which is a market microstructure based measure developed by 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 (primarily for liquidity reasons) and the order is equal likely to be a Buy or Sell, resulting in an order imbalance measure close to zero on average and a low PIN measure. On the other other hand, when there are significant information events and informed traders also trade, there will be large amount of Buy orders *or* Sell orders (depending on the nature of the information), resulting a large order imbalance and a high PIN measure.[15]

To the extent that PIN indeed captures the frequency and intensity of information events, we expect that funds trading high-PIN stocks to have a large and significant informed trading component (CS^{inf}), since the benefit from their information must be higher than the cost for demanding liquidity for them to initiate the trades (see Grossman and Stiglitz, 1980). Funds trading high-PIN stocks may also trade against the order imbalance. The return on these trades will be considered as the liquidity provision component. On one hand, these funds may face the danger of trading against informed traders. On the other hand, by judiciously choosing their trades, they could potentially benefit from the price impact working to their advantage, which could be sizable for high-PIN stocks. The net effect of the two will determine the sign and magnitude of the liquidity provision component for these funds. Funds trading low-PIN stocks are less likely to encounter informed traders when they trade. Consequently, when they trade against the market order imbalance, they are more likely to benefit from the price impact, resulting in a positive liquidity provision component. When they initiate the trade however, they suffer from the price impact which may outweigh the information advantage, and informed trading component will as a result be negative.

To estimate PIN , we use the tick-by-tick transaction data at the quarterly frequency from 1983 to 2004 using the entire three-month data to ensure the precision of estimation. A breakdown of our stock sample is provided in Panel A of Table 5. Overall, we have on average 4110 stocks with PIN measures in a quarter. The sample is initially dominated by NYSE/AMEX stocks. NASDAQ

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

stocks enter the sample in 1987 and account for a large portion of the sample. The mean of PIN measures in our sample is 25.8% with an associated standard deviation of 12.1%. The correlations between PIN and other stock characteristics are tabulated in Panel B of Table 5. Consistent with Easley, Hvidkjaer and O'Hara (2002), we find that High-PIN stocks are likely to be smaller and less liquid stocks. There is also some positive correlation between PIN and book-to-market ratio.

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

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

where $PIN_{i,q}$ is the estimated PIN measure of the i-th stock traded by the 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 high PIN stocks would have higher $trade\_PIN$ measures.

We then sort all funds in our sample into deciles at the end of each quarter from 1983 to 2004 according to their $trade_PIN$s and decompose the CS measure within each decile. Results are presented in Table 6. The CS measure and its components are winsorized at 1^{st} and 99^{th} percentiles to alleviate the effect of outliers. Several interesting patterns emerge from this table. First, funds trading high-PIN stocks (High $trade\_PIN$) outperform those trading low-PIN stocks (low $trade\_PIN$) by almost 53 bps per quarter on the dimension of stock selection. The 53 bps spread is highly significant with a t-value of 2.87. We also find similar results using actual after-fee mutual fund returns from the CRSP mutual fund database. The after-fee mutual fund return spread between the two deciles is 50 bps per quarter with a t-value of 3.27 after four-factor risk adjustment, which would suggest that window-dressing by mutual fund does not drive our result. Second, the spread is mainly driven by high-$trade\_PIN$-funds with an average CS measure of almost 50 bps (t-value = 2.70). In contrast, the CS measure of the low-$trade\_PIN$-fund is small and negative (-2.9 bps). Third, a large part of the CS measure for high-$trade\_PIN$-funds comes from active trading during the quarter ($CS^A = 31.2$ bps with a t-value of 2.83). Fourth, although both the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) are positive for high-$trade\_PIN$-funds, only the the informed trading component is significant (20.4 bps with a t-value of 2.25) and its size is twice that of the liquidity provision component (10.4 bps). This is consistent with our conjecture. When skillful managers absorb liquidity trading high-PIN stocks, they are likely to possess valuation-relevant information and therefore make money from informed trading. For them, the added cost of demanding immediacy in the market must be smaller than the benefit from superior information. In terms of liquidity provision, not all of them can perform as well as DFA. As a result, although the liquidity provision component is positive on average, it is much smaller and not significant, potentially due to the possibility of trading against informed traders. Finally, low -$trade\_PIN$-funds, having almost zero stock selection skill on average, seem to possess some skill in liquidity provision. The liquidity provision component (16.2 bps) is significant (t-value = 2.57). This is again consistent with our conjecture. 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. The positive liquidity provision component is partly offset by a negative informed trading component, resulting in a close-to-zero CS measure.

Although PIN is motivated in a structural model of informed trading, the empirical estimates should not be taken too literally. According to Hasbrouck (2007), PIN by construction is a meaningful measure of order flow one-sidedness. Independent of specific assumptions imposed on the trade arrival processes, frequent and large information events would result in order imbalances. Consistent with this interpretation, Aktas, Bodt, Declerck and Oppens (2007) consider an approximate PIN measure: relative order imbalance (rel_OIB), defined as:

$$rel\_OIB = \frac{E\left[|B - S|\right]}{E\left[B + S\right]},$$

where B and S denote the daily number of buyer-initiated trades and seller-initiated trades, respectively. Like PIN, rel_OIB is also a number between 0 and 1 and can therefore be interpreted as a probability. Aktas et al. (2007) show that rel_OIB is exactly equal to PIN on a daily basis and serves as a very good approximation during a longer time window.[16] rel_OIB is clearly a measure of order flow one-sidedness. Compared to PIN, it is extremely easy to compute and does not suffer from the problem of poor convergence during the maximum likelihood estimation of PIN. On the other hand, extreme daily order imbalance could introduce large noise to rel_OIB but have little impact on the estimation of PIN since a structural model is imposed for PIN which alleviates the effect of outliers.

To show the robustness of the results, we repeat the entire exercise by replacing PIN with rel_OIB. Empirically, we exclude stocks that trade less than 15 days within a quarter and estimate rel_OIB using simple daily average within each quarter. We first verify that rel_OIB is an reasonable approximation of PIN. The average cross-sectional correlation between these two measures is above 0.75. In addition, we are able to compute rel_OIB for a larger number of stocks since we avoid the convergence problem associated with PIN. The decomposition results within rel_OIB-sorted fund decile can be found in Table 7. The main results are almost identical qualitatively using the alternative PIN measure.

How can we discern which funds are more likely to trade high-PIN stocks? We tabulate the average fund-level characteristics across $trade\_PIN$-sorted fund deciles in Panel A of Table 8. All characteristics are winsorized at 1^{st} and 99^{th} percentile to alleviate the effect of outliers. We find that high-$trade\_PIN$ funds are typically associated with smaller fund size, smaller fund age, higher expense ratios, higher percentage fund inflow. In addition, high-$trade\_PIN$ funds tend to hold more stocks, smaller and more illiquid stocks. Their stock holding as a percentage of total number of share outstanding is also higher on average. These patterns are confirmed in Panel B of Table 8 which reports the correlations among these variables. Finally, their investment style more likely belongs to the AGG or Growth categories. In contrast, the investment style of the low-$trade\_PIN$ funds leans more towards the GNI spectrum.

3.1.1 Stock selection and momentum trading

Grinblatt, Titman and Wermers (1995) document that the majority of mutual funds use momentum as stock selection criterion and thus the well-known momentum effects can significantly influence the mutual fund performance. Panel A of Table 8 indeed shows that funds trading high-PIN stocks

[16] See Appendix C for a brief discussion on this issue.

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, the CS measure and its components throughout the paper are computed after adjusting for book-to-market, size and momentum characteristics following Daniel, Grinblatt, Titman and Wermers (DGTW, 1997). Second, when we regress the CS measure on several fund characteristics in a cross-sectional regression in the next subsection, we find $fund\_mom$ to be insignificant while $trade\_PIN$ is still highly significant, confirming that the large CS measure 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, sold and held by the funds separately in Table 9. Specifically, 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), the "Sell" portfolio (stocks recently sold by the fund) and the "Hold" portfolio (stocks held by the fund throughout the quarter). These past returns are then averaged across funds in the same $trade\_PIN$ decile and across time. Although high $trade\_PIN$ funds do 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), they are selling 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) and therefore 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 CS measures on average. That is clearly not the case as in Table 9.

3.2 Informed trading, liquidity provision and fund characteristics

To examine the relation between fund characteristics and the CS measures, we use a Fama-MacBeth (1973) cross-sectional regression approach. Specifically, we regress the next quarter CS measure and its components on several fund-level characteristics during each quarter from 1983 to 2004. Variables are winsorized at 1% and 99% to alleviate the effect of outliers. The effect of a fund's style is captured by two dummy variables that correspond to "Growth" and "AGG", respectively.

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 one standard deviation change in the variable. In addition, the regression intercept can be interpreted as the average effect of having a "GNI" fund style. Finally, the regression coefficients are averaged across time and the associated t-values are computed using Newey-West formula of lead / lag of 8 quarters to account for the autocorrelations in the error terms. The regression results are reported in Table 10.

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

3.3 Persistence of the informed trading and liquidity provision component

We examine the persistence in the CS measure and its component. To do that, at the end of each quarter from 1983 to 2004, we sort funds into deciles based on their CS measure during the quarter. We then tabulate the average CS measure across the deciles during the next quarter. If the manager's stock selection skill is persistent, we would expect funds with the highest CS

measures this quarter to continue to have significantly higher CS next quarter relative to funds with the lowest CS measures. We repeat the sorting exercise for the components of the CS measure: CS^P, CS^{inf} and CS^{liq}. The results are reported in Table 11. Overall, there is weak evidence of persistence in the active fund managers' stock selection skills. The average CS measure of funds in the highest CS-decile during the prior quarter is 73 bps higher than that of funds in the lowest CS-decile, although the spread is only marginally significant at 10% level. Interestingly, when we look at the components of CS, only the informed trading component (CS^{inf}) seems to be persistent. The insignificant persistence in the liquidity provision may be consistent with the notion that liquidity-based trading at quarterly frequency is episodic since the opportunity of low-frequency liquidity provision is sporadic. For example, Coval and Stafford (2007) investigates forced mutual funds transactions due to fund inflows and outflows, and identify economically important but statistically noisy profits when market participants are able to trade against the mutual funds "fire" sales and purchases.

4 Conclusion

This paper explores how active U.S. mutual funds add values in their stock selection through the channels of informed trading and liquidity provision. We develop a decomposition procedure to evaluate the relative contribution of informed trading and liquidity provision to the overall investment value generated in the stock selection process. To empirically identify these two types of value creation trading mechanisms, we make several identification assumptions. First, we assume that informed trading activities of mutual funds are likely to occur during the time period when information events of the underlying assets are more prevalent. Second, we assume that the informed trading on average is likely to absorb rather than provide liquidity because the value of information diminishes over time. Third, we notice that liquidity provision more likely occurs among stocks with fewer information events since that minimizes the adverse effects due to trading with those with an information advantage. To characterize the direction of liquidly needs of the fund, we use the stock-level order imbalance. When the fund trade against the aggregate stock-level order imbalance, we classify such trades as liquidity absorbing trades; otherwise, we classify such trades as liquidity provision trades. To quantify the frequency of informational events, we adopt the measure of average Probability of Informed Trading measure developed in Easley et al. (1996).

Although the decomposition procedure developed in this paper separates the relative contribution of informed trading and liquidity provision to the investment value of stock selection, the identification assumptions as well as empirical implementations are imperfect. First, among all the limitations, availability of data at quarterly frequency only allows us to examine long-lived liquidity shocks and the role of liquidity provision arisen from such shocks, whereas we cannot capture high-frequency liquidity events and associated liquidity provision in the traditional market microstructure sense. As the availability of high frequency fund transaction data, our procedure in theory could be used to evaluate the manager's contribution on these two dimensions with greater precision.[17] Second, PIN and the alternative measure are imperfect measures of the frequency of informational event affecting a stock during a quarter.

Bearing these challenges to our empirical exercise, we have made the first attempt to bring both informed trading and liquidity provision into the evaluation of mutual funds stock selectivity. We find mutual funds seem to possess skills to engage informed trading and liquidity provision. During the period from 1983 to 2004 on average, funds trading high-PIN stocks – those that are most affected by information events – add value via informed buys as well as sells. Funds trading low-PIN stocks – those that are subject to lower adverse selection costs due to trading with informed traders – add value via liquidity provision. Overall, funds trading high-PIN stocks outperform those trading low-PIN stocks by 53 bps per quarter and the 53 bps spread can be largely attributed to the difference in the informed trading components of the two fund groups. In addition, the informed trading component seems to be quite different from the liquidity provision component. While the informed trading component is associated with growth-oriented investment objective and is persistent, the liquidity provision component, on the other hand, is associated with younger fund age, income-oriented investment objective and is not persistent.

[17] As a practical matter, this may not be an issue for fund of funds and large institutional investors – they get almost daily reports of the holdings of the managers who manage their funds.

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

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

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

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

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

The portfolio values H_t, B_t and S_t are determined using the prices at the end of quarter t (P_t). Notice $B_t > S_t$, and the difference is likely financed by fund inflows, or a decrease in cash position or sale of other non-stock assets held by the fund. The "Hold", "Buy" and "Sell" can be treated as three separate funds whose CS measures can be computed using equation (1) and holdings as:

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

With the above information, equation (2) then decomposes the total CS measure into three components:

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

If we further assume that the fund traded B and F in the same direction as the aggregate order imbalance and traded A and E against the direction of aggregate order imbalance, then equation (3) further decomposes the active component (CS^A_{t+1}) into a "informed trading" component (CS^{inf}_{t+1}) and a "liquidity provision" component (CS^{liq}_{t+1}):

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

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

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

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

Consequently, we have

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

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

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

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

The term $\beta_{(\cdot)}$ then measures the contribution of component $(\cdot)$ to the cross-sectional variations of CS. The sum of the contribution from the four components is equal to one by construction. β can be measured by regression. For instance, β_P is estimated by regressing CS^P on CS cross-sectionally. Empirically, we have a panel data of cross-sectionally demeaned CS, CS^P, CS^{adj}, CS^{inf} and CS^{liq}. To estimate β, we run a Weighted Least Squares (WLS) regression. In practice, this means deflating the data for each fund-quarter by the number of funds in the corresponding cross-section (see Vuolteenaho (2002)).

Appendix C: Probability of Informed Trading (PIN) and its Approximation— A Brief Description

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

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

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

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

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

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

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

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

hence we estimate a modified version of (5),

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

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

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

It is interesting to note that the probability that an information event occurs (α) and the rate of informed trade arrivals (μ) enter PIN as a product term ($\alpha\mu$). Although α and μ may be individually estimated rather imprecisely, since estimation errors in these two parameters are usually strongly negatively correlated, the resulting PIN estimate is quite precise. In addition, the variation in α and μ are offsetting, making PIN a much stable measure bounded between 0 and 1. This is an important reason why PIN is chosen over alternative measures for private information such as the price non-synchronicity measure (see Roll, 1988 and Morock, Yeung and Yu, 2000, Durnev, Morck, Yeung and Zarowin, 2003 and Durnev, Morck and Yeung, 2004) and the adverse-selection component of bid-ask spread (Glosten and Harris, 1988 and Huang and Stoll, 1996).

In the economy of Easley et al. (2001), the total number of trades $B+S$ and the order imbalance $B-S$ are related to parameters of the model. as:

$$\begin{aligned} E\left[B+S\right] &= \alpha\mu + 2\epsilon, \\ E\left[B-S\right] &= \alpha\mu(1-2\delta). \end{aligned}$$

Since each day is either a good day ($\delta = 0$), a bad day ($\delta = 1$), or a no-event day ($\alpha = 0$), the expected daily absolute OIB is then:

$$E\left[|B-S|\right]=\alpha\mu.$$

Aktas, Bodt, Declerck and Oppens (2007) show that a relative order imbalance measure $rel\_OIB = E\left[|B-S|\right]/E\left[B+S\right]$ serves as a very good approximation to PIN. In fact, on a daily basis, rel_OIB is equivalent to PIN.

References

[1] Aktas, Nihat, Eric de Bodt, Fany Declerck, and Herve Van Oppens, 2007, The pin anomaly around M&A announcements, *Journal of Financial Markets* 10, 169-191.

[2] Badrinath, S.G. and Sunil Wahal, 2002, Momentum trading by institutions, *Journal of Finance* 57, 2449 - 2478.

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

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

[5] Blume, Marshall E. and Roger M. Edelen, 2004, On Replicating the S&P 500 Index, *Working Paper*.

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

[7] Campbell, John Y., Tarun Ramadorai, and Allie Schwartz, 2007, Caught On Tape:Institutional Trading, Stock Returns, and Earnings Announcements, *Working Paper*, Harvard University.

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

[9] Chan, Louis K. C., Stephen G. Dimmock, and Josef Lakonishok, 2006, Benchmarking money manager performance: Issues and evidence, *Working Paper*.

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

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

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

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

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

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

[16] 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.

[17] Cremers, Martijn, and Antti Petajisto, 2006, How Active is Your Fund Manager? a New Measure that Predicts Performance, *Working Paper*, Yale School of Management, Yale University.

[18] Da, Zhi and Pengjie Gao, 2006, Clientele Change, Liquidity shock, and the Return on Financially Distressed Stocks, *Working Paper*, University of Notre Dame.

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

[20] Durnev, A., R. Morck, and B. Yeung, 2004, Value Enhancing Capital Budgeting and Firm-Specific Stock Return Variation, *Journal of Finance* 59, 65–105.

[21] Durnev, A., R. Morck, B. Yeung, and P. Zarowin, 2003, Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing, *Journal of Accounting Research* 41, 797–836.

[22] Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2002, Is Information Risk a Determinant of Asset Returns?, *Journal of Finance* 57, 2185-2221.

[23] 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.

[24] 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.

[25] Elton, Edwin J., Martin J. Gruber , Yoel Krasny and Sadi Ozelge, 2006, The effect of the frequency of holding data on conclusions about mutual fund behavior, *Working Paper*.

[26] Fama, Eugene F., 1965, The behaviour of stock market prices, *Journal of Business*, 34-105.

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

[28] Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

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

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

[31] 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.

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

[33] 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.

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

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

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

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

[38] 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.

[39] Lakonishok, J., A. Shleifer, and R. W. Vishny, 1992, The Impact of Institutional Trading on Stock Prices, *Journal of Financial Economics* 32, 23-43.

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

[41] Lubomira, Ivanova, 2005, Essays on mutual funds performance, *unpublished Ph.D. Dissertation*, Northwestern University.

[42] Morck, Randall, Bernard Yeung and Wayne Yu, 2000, The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58, 215-260.

[43] Nofsinger, J., and R. W. Sias, 1999, Herding and Feedback Trading by Institutional and Individual Investors, *Journal of Finance* 54, 2263-2295.

[44] Roll, Richard, 1988, R^2, *Journal of Finance* 43, 541–566.

[45] Sias, Richard, Laura T. Starks, and Sheridan Titman, 2006, The price impact of institutional trading, *Journal of Business* 79, 2869-2910.

[46] Schultz, Paul, 2007, When is Stock Picking by Mutual Funds Successful?, Working Paper, University of Notre Dame.

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

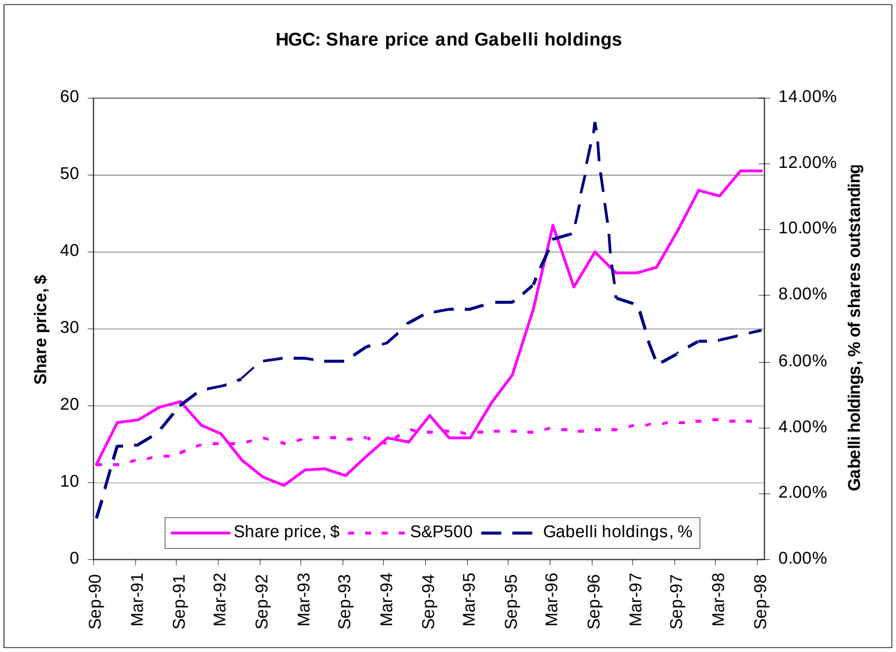
[48] Vuolteenaho, T. (2002): "What drives firm-level stock returns," *Journal of Finance*, Vol 57, 233-264.

[49] 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

Figure 1 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 Sep 1990 to sep 1998. Figure 2 plots the share prices of Starbucks (SBUX) from June to Dec 2005 (price is normalized so 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, Sep and Dec.

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

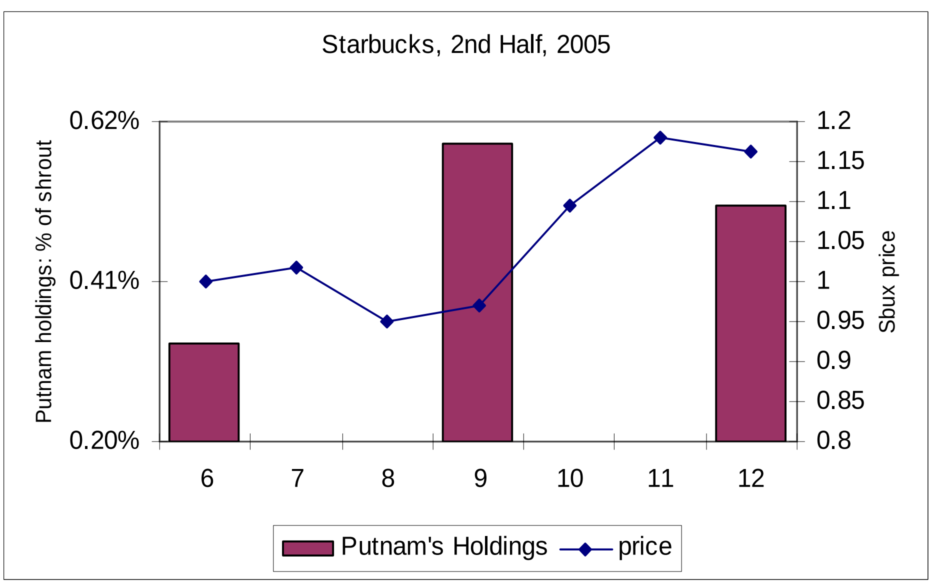


Table 1: Breakdown of mutual fund sample over time

We report the breakdown of our mutual fund sample by investment objective. We exclude index fund and lifecylce funds that are likely passively managed. In addition, we only keep funds that are self-reported as aggressive growth (AGG), growth (GROWTH) or growth and income (GNI). We also exclude fund / quarter observations with quarterly turnover less than 10% or if the fund trades less than 10 during that quarter. Finally, we only include fund / quarter observations for which the fund holdings at the end of previous quarter are also available so holding changes can be computed over consecutive quarters. The CDA/Spectrum mutual fund holding data are matched to CRSP Mutual Fund data using the MFLINKS tables developed by WRDS.

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

Table 2: Type of mutual fund trades and the average order imbalances

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

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

Table 3: Decomposition of the mutual fund "Characteristics selectivity" (CS) measure for DFA US Micro-Cap fund and index funds as a group

We provide two examples to illustrate the decomposition of the mutual fund stock selection skill. We decompose the mutual fund "Characteristics selectivity" (CS) measure (Daniel et al., 1997) for DFA US Micro-Cap fund (fundno=16500 in CDA/Spectrum S-12 mutual fund holding database) and Index funds a group (fund whose name contains any of the following: "INDEX", "INDE", "INDX", "S&P", "DOW JONES", "MSCI" or "ISHARE"). Specifically, the CS measure is decomposed into:

$$CS = CS^{P} + CS^{adj} + CS^{inf} + CS^{liq},$$

Where CS^{P} is the passive component; CS^{adj} is an adjustment component due to fund inflows; CS^{inf} and CS^{liq} are the informed trading and liquidity provision components, respectively. The sampling period is from 1983 to 2004. *t*-values associated with the average measures are reported in *italic*s.

	Total CS (=1+2+3)	Passive CS^{P} (1)	Adj CS^{adj} (2)	Active CS^{A} (3=3a+3b)	Info trading CS^{inf} (3a)	Liquidity Prov CS^{liq} (3b)
DFA US Micro-Cap:						
Alpha (bps)	36.1	19.3	-4.2	21.0	0.5	20.5
t-value	*1.72*	*0.89*	*-0.64*	*1.30*	*0.06*	*1.84*
Index Funds:						
Alpha (bps)	0.0	24.9	3.2	-28.1	-34.6	6.4
t-value	*0.00*	*0.93*	*0.50*	*-1.11*	*-2.19*	*0.36*

Table 4: Variance Decomposition of the *CS* measure

This table reports the percentage of total cross-sectional variation in the total "Characteristic Selectivity" (*CS*) measure (DGTW, 1997) explained by its four components: the passive component (CS^{P}), the adjustment component (CS^{adj}), the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) in a variance decomposition framework. We have performed the decomposition on the full sample and on each style-subsample. WLS *t*-values associated with the percentages are reported in *italic*s. The sampling period is from 1983 to 2004. Details on the variance decomposition can be found in Appendix A.

Passive CS^{P}	Adj CS^{adj}	Info trading CS^{inf}	Liquidity Prov CS^{liq}
	All		
56.8%	-2.5%	37.2%	8.4%
127.2	*-15.3*	*120.9*	*24.4*
	Aggressive Growth (AGG)		
52.1%	-1.0%	44.9%	4.0%
44.7	*-2.7*	*55.2*	*4.2*
	Growth (Growth)		
55.7%	-3.0%	37.0%	10.2%
96.0	*-14.9*	*95.0*	*22.3*
	Growth and Income (GNI)		
54.1%	-2.2%	37.0%	11.1%
56.8	*-5.4*	*55.9*	*16.6*

Table 5: Descriptive statistics of *PIN*

The Probability of Informed trading (PIN) is estimated at quarterly frequency from 1983 to 2004 using the entire three months trade and quote data from TAQ. A breakdown of our stock PIN sample over time is provided in Panel A. The correlations among PIN and other stock characteristics are reported in Panel B.

Panel A: Summary Statistics on PIN

Year	# of stocks per quarter	% of NYSE/AMEX stocks	% of NASNAQ stocks	mean	std dev
1983	1915	99.3%	0.7%	22.5%	10.2%
1984	1747	99.5%	0.5%	25.2%	13.0%
1985	1812	99.1%	0.9%	24.1%	11.8%
1986	1828	99.2%	0.8%	23.4%	11.1%
1987	3732	46.7%	53.3%	27.0%	12.1%
1988	3399	50.0%	50.0%	28.1%	13.4%
1989	3373	49.7%	50.3%	27.4%	13.3%
1990	3321	49.4%	50.6%	27.7%	13.6%
1991	3362	50.4%	49.6%	26.7%	12.7%
1992	4117	43.4%	56.6%	27.2%	13.0%
1993	4106	53.8%	46.2%	25.4%	12.0%
1994	5258	36.3%	63.7%	27.4%	12.9%
1995	5500	35.1%	64.9%	27.2%	12.5%
1996	6028	33.7%	66.3%	26.6%	12.1%
1997	6473	32.5%	67.5%	25.8%	11.8%
1998	6453	32.6%	67.4%	25.6%	11.8%
1999	5879	33.9%	66.1%	26.0%	12.1%
2000	5526	33.1%	66.9%	26.3%	12.6%
2001	4842	32.3%	67.7%	28.0%	13.6%
2002	4476	36.4%	63.6%	25.3%	11.6%
2003	3999	39.4%	60.6%	22.7%	9.8%
2004	3727	42.0%	58.0%	21.1%	9.5%
All	4130	51.3%	48.7%	25.8%	12.1%

Panel B: Cross-correlation

	PIN	log(Size)	log(BM)	Mom
log(Size)	-0.536			
log(BM)	0.169	-0.193		
Mom	-0.066	0.058	-0.148	
Amihud	0.557	-0.872	0.190	-0.198

Table 6: CS measure decomposition across *trade_PIN* sorted fund deciles

In each quarter and for each fund, we compute a *trade_PIN* variable by value-weighing 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 decompose the CS measure within each decile. The last column reports the average 4-factor (Fama-French three factors and the momentum factor) risk adjusted mutual fund returns. *t*-values associated with the average measures are reported in *italic*s. The CS measure and its components are winsorized at 1st and 99th percentiles to alleviate the effect of outliers.

Trade_PIN	Total CS (=1+2+3)	Passive CS^P (1)	Adj CS^{adj} (2)	Active CS^A (3=3a+3b)	Info trading CS^{inf} (3a)	Liquidity Prov CS^{liq} (3b)	4f-adj MF return
All stocks	23.5	13.9	-1.8	14.2	3.6	8.8	-23.6
	1.91	*1.19*	*-2.38*	*2.09*	*0.55*	*1.50*	*-2.55*
Low	-2.9	-7.6	-0.4	3.4	-12.1	16.2	-42.2
	-0.29	*-0.70*	*-0.20*	*0.42*	*-2.02*	*2.57*	*-3.97*
2	11.4	10.4	-0.7	2.6	-6.4	8.9	-35.7
	0.92	*0.87*	*-0.47*	*0.40*	*-0.93*	*1.28*	*-3.85*
3	11.8	9.2	-1.0	5.5	-5.5	9.8	-28.6
	1.04	*0.81*	*-0.69*	*0.88*	*-0.78*	*1.65*	*-2.59*
4	10.3	8.3	-1.1	5.8	-3.5	6.0	-33.8
	1.01	*0.76*	*-0.81*	*0.76*	*-0.54*	*0.89*	*-3.52*
5	28.6	23.3	-2.3	6.4	2.5	5.5	-15.5
	2.17	*1.80*	*-1.52*	*0.72*	*0.31*	*0.74*	*-1.36*
6	31.9	19.2	-2.4	18.7	5.6	9.4	-22.3
	2.07	*1.20*	*-1.52*	*1.69*	*0.62*	*1.10*	*-1.65*
7	28.4	19.4	-0.9	17.2	9.7	0.2	-28.1
	1.52	*1.17*	*-0.56*	*1.18*	*0.89*	*0.03*	*-1.93*
8	30.9	14.6	-3.4	25.0	8.6	13.7	-20.5
	1.73	*0.87*	*-2.05*	*2.26*	*0.84*	*1.64*	*-1.22*
9	35.1	15.7	-2.7	26.6	16.8	7.6	-17.4
	1.75	*0.82*	*-1.35*	*2.38*	*1.70*	*0.77*	*-1.06*
High	50.0	26.5	-3.2	31.2	20.4	10.4	8.3
	2.70	*1.43*	*-1.40*	*2.83*	*2.25*	*1.37*	*0.65*
High - Low	52.9	34.1	-2.8	27.8	32.5	-5.8	50.5
	2.87	*1.94*	*-0.93*	*2.26*	*3.50*	*-0.68*	*3.27*

Table 7: CS measure decomposition across *trade_rel_OIB* sorted fund deciles

In each quarter and for each fund, we compute a *trade_rel_OIB* variable by value-weighing *rel_OIB* 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_rel_OIB*s and decompose the CS measure within each decile. *t*-values associated with the average measures are reported in *italic*s. The CS measure and its components are winsorized at 1st and 99th percentiles to alleviate the effect of outliers.

Trade_rel_OIB	Total CS (=1+2+3)	Passive CS^P (1)	Adj CS^{adj} (2)	Active CS^A (3=3a+3b)	Info trading CS^{inf} (3a)	Liquidity Prov CS^{liq} (3b)
Low	1.9	-7.9	-2.0	14.9	-11.8	23.2
	0.12	*-0.59*	*-1.12*	*1.15*	*-1.09*	*2.62*
2	7.7	3.2	-0.4	6.3	-7.4	12.5
	0.59	*0.29*	*-0.30*	*0.64*	*-0.81*	*1.73*
3	7.1	5.9	-1.3	7.2	-5.2	8.7
	0.56	*0.48*	*-0.88*	*0.99*	*-0.69*	*1.29*
4	10.4	6.0	-2.4	9.6	-4.0	12.0
	0.86	*0.49*	*-1.57*	*0.91*	*-0.47*	*1.57*
5	30.2	23.1	0.5	7.9	-0.3	7.2
	2.24	*1.71*	*0.31*	*0.88*	*-0.04*	*0.99*
6	30.1	16.4	-2.6	17.1	12.2	3.1
	1.90	*1.06*	*-1.50*	*2.23*	*1.37*	*0.39*
7	31.9	22.9	-2.3	13.3	7.0	4.9
	1.96	*1.37*	*-1.47*	*1.29*	*0.76*	*0.56*
8	40.8	19.7	-2.6	26.3	16.7	7.8
	2.32	*1.21*	*-1.75*	*2.40*	*1.73*	*0.84*
9	37.8	20.8	-6.0	29.0	20.4	7.1
	1.87	*1.10*	*-2.62*	*2.97*	*1.95*	*0.78*
High	55.1	23.7	-1.8	34.4	26.9	6.3
	2.89	*1.21*	*-0.71*	*3.69*	*3.26*	*0.82*
High - Low	53.2	31.6	0.2	19.5	38.7	-16.9
	2.00	*1.25*	*0.08*	*1.22*	*2.83*	*-1.35*

Table 8: Fund-level characteristics

Panel A reports the average fund-level characteristics across the *trade_PIN* sorted deciles. Fund-level stock characteristics are computed by value-weighing the stock characteristics of stocks held by the fund at quarter end using the dollar value of the holding. All characteristics are winsorized at 1st and 99th percentile to alleviate the effect of outliers. The correlations among the characteristics are reported in Panel B.

Panel A: Average fund-level characteristics across *trade_PIN* sorted fund deciles

Trade _PIN	num _stock	trade _PIN	fund _holding	fund _size	fund _bm	fund _mom	fund _amihud	age	turnover	expense	TNA	pct _flow	% of AGG	% of Growth	% of GNI
Low	64	11.2%	0.25%	32.9	0.56	0.232	4.5%	53.3%	0.680	1.14%	1020.9	2.74%	3.7%	50.9%	45.4%
2	72	12.3%	0.26%	30.1	0.56	0.253	5.1%	55.3%	0.772	1.12%	972.5	2.11%	4.5%	54.1%	41.5%
3	74	13.0%	0.27%	28.0	0.56	0.268	5.5%	55.9%	0.841	1.13%	848.4	1.98%	5.6%	54.9%	39.5%
4	74	13.7%	0.28%	25.4	0.55	0.282	6.3%	54.2%	0.855	1.16%	740.8	1.95%	7.9%	60.4%	31.7%
5	75	14.4%	0.32%	20.5	0.55	0.304	7.4%	53.8%	0.880	1.20%	719.1	1.82%	11.3%	61.5%	27.2%
6	75	15.3%	0.37%	15.6	0.55	0.329	8.8%	50.8%	0.909	1.22%	636.8	2.42%	17.2%	59.7%	23.1%
7	74	16.3%	0.44%	11.2	0.54	0.365	10.9%	48.1%	0.945	1.26%	557.5	3.23%	24.0%	59.8%	16.1%
8	73	17.7%	0.54%	6.6	0.55	0.381	14.3%	45.3%	0.970	1.30%	404.4	2.85%	28.6%	59.6%	11.9%
9	81	19.4%	0.63%	3.9	0.54	0.393	18.9%	42.6%	0.904	1.32%	353.9	3.92%	30.0%	61.8%	8.2%
High	97	22.7%	0.91%	2.1	0.62	0.339	28.4%	37.6%	0.725	1.34%	295.1	5.29%	29.2%	64.9%	5.9%
H-L	33	11.5%	0.66%	-30.8	0.06	0.107	24.0%	-15.7%	0.044	0.20%	-725.8	2.54%	25.5%	14.1%	-39.5%
t-value	11.5	73.2	53.5	-9.8	3.0	6.2	47.4	-14.5	1.6	12.9	-14.9	4.6	13.7	7.5	-26.8

Panel B: Correlations among fund-level characteristics

	trade_PIN	fund _holding	fund_size	fund_bm	fund_mom	fund _amihud	age	turnover	expense	TNA
fund_holding	0.336									
fund_size	-0.612	-0.285								
fund_bm	0.336	0.177	-0.461							
fund_mom	0.087	0.016	-0.095	-0.308						
fund_amihud	0.731	0.417	-0.516	0.311	0.013					
age	-0.142	0.100	0.084	-0.051	-0.041	-0.168				
turnover	0.018	-0.130	-0.089	-0.074	0.210	-0.009	-0.105			
expense	0.008	-0.146	0.037	-0.181	0.061	0.132	-0.256	0.220		
TNA	-0.209	0.392	0.155	-0.069	-0.015	-0.173	0.231	-0.127	-0.175	
pct_flow	0.082	0.004	-0.086	0.037	0.110	0.045	-0.135	0.010	0.035	-0.009

Table 9: Average past one-year return of stocks Bought / Sold / Held by mutual funds across *trade_PIN* sorted deciles

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. For each fund, we then compute the value-weighted average past one-year return of stocks in the "Buy" portfolio (stocks recently bought by the fund), the "Sell" portfolio (stocks recently sold by the fund) and the "Hold" portfolio (stocks held by the fund throughout the quarter). These past returns are then averaged across funds and across time. *t*-values associated with the average measures are reported in *italic*s.

trade_pin	Past One-year Return				
	Buy	Sell	Hold	Buy-sell	t-value
Low	20.9%	24.3%	22.2%	-3.4%	*-5.74*
2	23.2%	26.4%	24.6%	-3.2%	*-5.18*
3	25.1%	26.8%	25.6%	-1.7%	*-2.96*
4	26.7%	29.6%	27.2%	-2.9%	*-4.08*
5	28.5%	32.1%	29.2%	-3.6%	*-4.48*
6	32.1%	36.2%	32.4%	-4.1%	*-3.66*
7	36.2%	39.6%	35.2%	-3.4%	*-3.13*
8	39.2%	43.8%	37.5%	-4.6%	*-3.60*
9	40.2%	46.1%	37.8%	-5.9%	*-4.65*
High	34.3%	46.6%	34.0%	-12.3%	*-8.93*
H-L	13.47%	22.36%	11.73%		
	6.37	*8.42*	*6.47*		

Table 10: Cross-sectional regressions

We regress the next quarter components of CS measure on several fund-level characteristics during each quarter from 1983 to 2004. Variables are winsorized at 1st and 99th percentiles to alleviate the effect of outliers. 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 one standard deviation change in the variable. Finally, the regression coefficients are averaged across time and the associated t-values are computed using Newey-West formula of lead / lag of 8 to account for the autocorrelations in the error terms. *t*-values associated with the average measures are reported in *italic*s.
trade_pin is the (log) 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; *expense* is the expense ratio of fund; *turnover* is the turnover rate of the fund; *dummy_growth* is a dummy variable which assumes 1 if the self-reported investment objective is "growth" and 0 otherwise; *dummy_Agg* is a dummy variable which assumes 1 if the self-reported investment objective is "AGG" and 0 otherwise.

	Intercept	trade_pin	log_fund _size	log_fund _bm	fund _mom	fund _amihud	log_TNA	Age	expenses	turnover	dummy _growth	dummy _Agg	Average R^2
							LHS = CS						
coeff	0.0018	0.0020	0.0014	0.0006	0.0016	-0.0002	0.0000	-0.0001	-0.0004	0.0003	0.0006	0.0036	0.20
t-value	*1.25*	*2.79*	*1.42*	*0.52*	*1.69*	*-0.22*	*0.12*	*-0.64*	*-0.95*	*0.60*	*0.67*	*3.05*	
							LHS = CS^{inf}						
coeff	0.0001	0.0010	0.0008	0.0008	0.0011	0.0001	0.0000	-0.0003	-0.0004	0.0001	-0.0001	0.0015	0.12
t-value	*0.20*	*3.13*	*1.41*	*1.42*	*1.54*	*0.23*	*-0.21*	*-1.58*	*-1.57*	*0.26*	*-0.20*	*2.64*	
							LHS = CS^{liq}						
coeff	0.0013	-0.0003	0.0000	-0.0005	-0.0002	0.0004	0.0000	-0.0004	0.0001	0.0004	-0.0005	-0.0011	0.09
t-value	*2.56*	*-0.98*	*-0.07*	*-1.51*	*-0.56*	*1.09*	*0.03*	*-2.94*	*0.55*	*1.60*	*-1.37*	*-1.54*	

Table 11: Persistence of the informed trading component and the liquidity provision component of the mutual fund CS measure

At the end of each quarter from 1983 to 2004, we sort funds into deciles based on their CS measure during the quarter. We then tabulate the average CS measure across the deciles during the next quarter. We repeat the sorting exercise also for the components of the CS measure: CS^{p}, CS^{inf} and CS^{liq}. *t*-values associated with the average measures are reported in *italic*s.

	Total CS		Passive CS^{p}		Info trading Cs^{inf}		Liquidity Prov Cs^{liq}	
	Qtr t	Qtr t+1	Qtr t	Qtr t+1	Qtr t	Qtr t+1	Qtr t	Qtr t+1
Low	-6.51%	-0.08%	-6.79%	-0.02%	-4.49%	-0.09%	-4.60%	0.14%
	-28.15	*-0.39*	*-28.30*	*-0.10*	*-23.01*	*-0.71*	*-18.89*	*0.86*
2	-3.20%	0.17%	-3.26%	0.05%	-1.99%	0.03%	-1.85%	0.03%
	-22.54	*1.08*	*-22.35*	*0.37*	*-20.14*	*0.30*	*-16.23*	*0.34*
3	-1.90%	0.13%	-1.95%	0.09%	-1.16%	0.01%	-1.01%	0.04%
	-16.66	*0.99*	*-16.19*	*0.74*	*-16.70*	*0.11*	*-13.35*	*0.50*
4	-0.97%	0.13%	-1.05%	0.09%	-0.62%	0.04%	-0.50%	0.08%
	-9.38	*0.90*	*-9.39*	*0.62*	*-12.02*	*0.46*	*-9.18*	*1.29*
5	-0.21%	0.31%	-0.30%	0.26%	-0.22%	0.03%	-0.11%	0.07%
	-1.96	*1.91*	*-2.68*	*1.85*	*-4.73*	*0.34*	*-2.41*	*1.08*
6	0.51%	0.27%	0.43%	0.04%	0.17%	0.07%	0.26%	0.08%
	4.31	*1.71*	*3.70*	*0.30*	*3.46*	*0.79*	*5.93*	*1.33*
7	1.31%	0.27%	1.22%	0.17%	0.60%	0.12%	0.66%	0.12%
	9.34	*1.71*	*9.10*	*1.05*	*9.10*	*1.14*	*12.75*	*1.64*
8	2.27%	0.32%	2.17%	0.18%	1.16%	0.06%	1.20%	0.15%
	12.38	*1.61*	*13.50*	*1.00*	*11.65*	*0.69*	*17.24*	*2.09*
9	3.69%	0.41%	3.54%	0.32%	2.10%	0.08%	2.03%	0.00%
	13.25	*1.76*	*17.11*	*1.50*	*12.12*	*0.63*	*20.79*	*-0.02*
High	7.70%	0.65%	7.42%	0.15%	5.05%	0.29%	4.70%	0.16%
	15.40	*1.84*	*20.13*	*0.47*	*12.79*	*1.75*	*23.62*	*1.30*
High - Low	14.21%	0.73%	14.21%	0.17%	9.54%	0.38%	9.31%	0.02%
	22.87	*1.83*	*27.51*	*0.44*	*18.86*	*2.80*	*23.89*	*0.13*