Comments Welcome

When Does a Mutual Fund's Trade Reveal its Skill?

Zhi Da, Pengjie Gao, and Ravi Jagannathan*†

First Draft: May 10, 2007
This Draft: April 2, 2008

Abstract

We conjecture that a mutual fund manager with superior stock selection ability is more likely to benefit from trading in stocks affected by information-events. Taking the probability of informed trading (PIN, Easley, Kiefer, O'Hara, and Paperman, 1996) to measure the amount of informed trading in a stock, and inferring mutual fund trades from a large sample of mutual fund holdings, we provide empirical support for the conjecture. Funds trading high-PIN stocks exhibit superior performance on average, and superior performance that is more likely to persist. The findings are not due to price momentum or the higher returns earned by high-PIN stocks on average. Conclusions remain the same after testing for alternative measures for the amount of informed trading. Decomposing a fund's stock selection ability into "informed trading" and "liquidity provision" adds further insight into fund's underlying strengths. Impatient informed trading is a significant source of alpha for funds trading high-PIN stocks, while liquidity provision is more important as a source of alpha for funds trading low-PIN stocks.

---

*We thank Robert Battalio, Roger Edelen, Craig Holden, Paul Irvine, Tim Loughran, Rick Mendenhall, David Musto, Ľuboš Pástor, Paul Schultz, Clemens Sialm, Laura Starks, Sheridan Titman, Charles Trzcinka, Lance Young and seminar participants at Barclays Global Investors, FRA 2007 Annual Meeting, IU/ND/Purdue Finance Symposium, Northwestern University, University of Illinois at Urbana-Champaign, University of Michigan, Tilburg University, and the 4th Vienna Symposium on Asset Management for comments. An earlier version of this paper was titled "Informed Trading, Liquidity Provision and Stock Selection by Mutual Funds." Patricia Andersen provides excellent editorial support.

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

As of 2006, US mutual funds collectively managed \$10 trillion, and almost \$6 trillion of that in equity funds. A significant portion of this amount is actively managed. According to a survey by the Mutual Fund Institute, in 2006 alone, US mutual funds bought and sold common stocks worth over \$6 trillion. Over 1980 to 2006, investors paid at least 0.67 percent of portfolio value per year to the active portfolio managers (French, 2008). Naturally, investors want to identify fund managers with superior skills and to understand how these managers add sufficient value to justify their fees. Ultimately, an active mutual fund manager's success derives from his or her superior skill in processing valuation-relevant information about a stock, a skill that should allow the identification of potential mispricing. Thus, it is reasonable to expect such skills to be more valuable when stocks the manager can invest in are affected by more value-relevant information events. To the extent that rational managers have the option not to trade such stocks when they know that they do not have an advantage in analyzing the information affecting a stock, we should expect to find that managers who choose to trade in this case earn higher returns on average.

To examine this conjecture we need (1) a measure of the number of information events affecting a stock during a given period; and (2) information that facilitates identifying a manager's trades in those stocks during those periods. We use the probability of informed trading (PIN) - proposed by Easley, Kiefer, O'Hara, and Paperman (1996) - to proxy for the extent of information events affecting a stock. We infer a mutual fund's trades by comparing its quarter-end holdings with its holdings at the end of the quarter before.

Using quarterly holdings data from 1983 to 2004 for a large sample of US active equity mutual funds, we find that funds trading high-PIN stocks indeed outperform those trading low-PIN stocks. That is, if we sort funds according to the average PIN of stocks they trade during quarter t (defined as $trade\_PIN$), funds in the top $trade\_PIN$ decile outperform those in the bottom decile by more than 50 basis points per quarter on a risk-adjusted basis during quarter $t+1$. The risk-adjusted performance does not depend on whether we adjust for risk using a factor model or a holding-based characteristics model.

Easley, Hvidkjaer, and O'Hara (2002) document that high-PIN stocks earn higher returns – for convenience we will refer to this phenomenon as compensation for PIN risk. Our findings, however, are not driven by this phenomenon for the several reasons. First, we obtain very similar results after we control for PIN risk when we examine risk adjusted-returns. Second, our conclusions do not change when we replace PIN with several alternative measures of the amount of information affecting a stock. To account for the illiquidity associated with high-PIN stocks, we examine the information asymmetry component of the PIN (*adjPIN*; see Duarte and Young, 2007). To ensure that our results are not driven by the correlation between trading volume and PIN, we directly examine abnormal turnover in stock trading (*aturn*; see Chordia, Huh, and Subrahmanyam, 2006). We also consider the information asymmetry component of the bid-ask spread (*theta*; see Madhavan, Richardson and Roomans, 1997) which captures the impact of information events on stock price. Third, stocks recently bought by mutual funds outperform those sold by mutual funds, although they have very similar PIN measures.

We find that the ability of $trade\_PIN$ to predict future mutual fund performance is not driven by its correlation with other fund characteristics. In a cross-sectional regression framework, $trade\_PIN$ provides additional explanatory power on the next-quarter risk-adjusted mutual fund return even in the presence of many other fund level characteristics. We show that such predictability is not due to potential momentum trading by mutual funds.

While fund managers who choose to trade high-PIN stocks on average outperform the rest, we would expect outperformance to be concentrated among managers who in fact have superior skills. Using a four-factor-adjusted fund return or alpha as a proxy for manager skill, we provide supporting evidence for such conjecture. Sorting on $trade\_PIN$ variable generates significant spreads in future fund performance only among funds with large alpha. In particular, amongst funds associated with high alphas, only those that also trade high-PIN stocks during quarter t produce positive and significant alpha of 33 basis points during quarter $t+1$ after accounting for fees and expenses. Moreover, to the extent that a manager's skills are likely to persist for some time, past superior performance is more likely to be an indication of future performance for a manager who attained that performance by trading high-PIN stocks. Consistent with this conjecture, we find that sorting on fund alphas in quarter t within funds with high $trade\_PIN$ generates a much wider spread in fund risk-adjusted returns in quarter $t+1$ (91 basis points on average). Controlling for PIN risk does not alter our results in any significant way.

We also investigate the specific channels through which a fund manager's trades add value. In general, how a manager with superior skill trades to add value will depend on how long it takes for the market to realize that the manager is correct. Based on how long the informational advantage lasts, a manager's trades can be classified into three types: long-term value investing, medium-term informed trading, and short-term liquidity provision. Medium-term informed trading, in contrast to long-term value investing and short-term liquidity provision, is likely to demand liquidity because the value of information erodes quickly and the timely execution becomes important.

To attribute mutual fund performance to these different styles of investment, we use a characteristics-based performance measure, characteristic selectivity (CS), proposed by Daniel, Grinblatt, Titman, and Wermers (DGTW, 1997).[1] CS, computed using quarterly reported stock holdings by mutual funds, measures the extent to which managers can select stocks that outperform the average stocks with similar characteristics. Because mutual fund stock holdings data are available only at infrequent intervals (quarterly in most cases), it is difficult to assess a manager's skills when that manager trades actively in between two holding reporting dates.[2] We find nevertheless that a mutual fund's recent trades inferred from its quarterly holdings changes still reveals useful information about how a manager adds value.

We decompose the CS measure for a given quarter into three components: (1) an *old* component that captures value-added from the fund's (previous) investments prior to the preceding quarter; (2) a *trade* component that captures the value-added from the fund's trades during the previous quarter; and (3) a small *adjustment* component that results from preceding quarter's net fund flow. The trade component can then be decomposed further into an *informed trading* component and a *liquidity provision* component. This latter decomposition is motivated by the evidence that 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 aggregate market order imbalance on a stock, they demand liquidity. Such trades are therefore likely to be driven by information and are classified as informed trading. When fund managers trade in the opposite direction of the aggregate market order imbalance, they effectively supply liquidity, and such trades are classified as liquidity provision. We validate

[1] We abstract away from measuring the market timing ability of mutual funds in this paper. See Ferson and Schadt (1996), and Mamaysky, Spiegel, and Zhang (2007) for the discussion on empirical methods to detect mutual fund market timing.

[2] For instance, Kacperczyk, Sialm and Zheng (2007) and Elton, et al. (2006) show that "unobservable" actions (or high-frequency turnovers) by mutual funds could be significant for some funds. Campbell, Ramadorai and Schwartz (2007) attempts to infer institutional transactions within a given quarter by selecting trade sizes that best match quarterly holdings changes. Relying on a unique regulation governing mutual fund trade disclosure in Canada, Christoffersen, Keim and Musto (2006) investigate essentially all trades of 210 Canadian mutual funds between 2001 and 2003.

our decomposition approach by applying it to two specific cases before applying it to portfolios of active mutual funds sorted by the $trade\_PIN$ variable.

We find that a greater proportion of the CS measure for high $trade\_PIN$ funds (50 basis points) comes from positions taken through trades made during the previous quarter; the trade component is 31.2 basis points (t-value of 2.83). We also find that the stocks a fund bought as well the stocks it sold had similar PIN values and both added value in the next quarter. Further, we document that informed trading is more likely to add value when the fund is trading stocks associated with more information events. The informed trading component in general increases with $trade\_PIN$, and is positive and significant (20.4 basis points with t-value of 2.25) only for funds in the top $trade\_PIN$ decile. Liquidity provision, however, is more likely to be detected when the fund is trading stocks associated with fewer information events. The liquidity provision component is positive and significant (16.2 basis points with t-value of 2.57) only for funds in the bottom $trade\_PIN$ decile, where the risk of adverse selection in trading is low.

Our results are relevant for two important strands of literature in finance: the literature on mutual fund performance, and the literature on market microstructure. The early literature on mutual fund performance evaluation finds that most managed portfolios earn close to zero or negative risk-adjusted returns, especially after taking fees into account.[3] This is consistent with Berk and Green's (2004) observation that competition among investors for a given supply of managerial talents of fund managers will drive after-fee, risk adjusted performance to zero. More recent studies that use of quarterly reports of mutual fund stock holdings find that active managers possess considerable stock-picking abilities.[4] On average, after adjusting for stock characteristics but before deducting fees and expenses, stocks held by actively managed mutual funds outperform their benchmarks, and stocks bought by mutual funds tend to outperform those sold by mutual funds. Further, several mutual fund holding characteristics appear to correlate with better fund performance, including funds that follow aggressive growth and growth styles (Daniel, Grinblatt, Titman, and Wermers, 1997); funds that hold the stocks of firms whose headquarters are located geographically close to the fund's headquarters (Coval and Moskowitz, 2001); funds that have greater industry concentration in their holdings (Kacperczyk, Sialm, and Zheng, 2005); funds that have less diversification in their

[3] See Jensen (1968), Brown and Goetzmann (1995), Gruber (1996), and Carhart (1997).

[4] See Grinblatt and Titman (1989, 1993), Daniel, Grinblatt, Titman and Wermers (1997), Wermers (1999), Chen, Jegadeesh, and Wermers (2000), Sias, Starks and Titman (2006), and Schultz (2007)

holdings (Baks, Busse, and Green, 2006); funds that show greater deviations from passive indexes or greater proportions of "active shares" (Cremers and Petajisto, 2006); and funds that have less dependency on analyst recommendations (Kacperczyk and Seru, 2007). In addition, funds that are smaller in size perform better after controlling for fund family size (Chen, Hong, Huang, and Kubik, 2004).

We contribute to the literature by building on the insights in the market microstructure literature that trades reveal information associated with a stock, and that such information can help to differentiate various trading motives and ultimately identify skillful managers. Moreover, as providing liquidity requires specialization and may well require a particular type of talent, we decompose fund managers' stock selection ability into a liquidity-demanding informed trading component and a liquidity provision component. We believe that such a decomposition will help us better understand the strengths of an active portfolio manager, and the extent to which such strengths will continue to be of value in the future.

The rest of the paper is organized as follows. We describe our data sources and the sample in section 1. We provide evidence that the amount of information events associated with stocks traded by fund managers reveals their skills in section 2. We then show how to decompose the stock selection ability of a mutual fund manager into its informed trading, liquidity provision, and residual components in section 3. We conclude in section 4. Appendix A contain brief descriptions on the various measures of information events we use. Appendix B illustrates our decomposition approach with a numerical example and validates the approach empirically.

1 Data and Sample Construction

We use data from several sources. The mutual fund holding data are from the CDA/Spectrum $S12$ mutual fund holding database, which collects holding information from N30-D filings with the US Securities and Exchange Commission (SEC). A detailed description of this database can be found in Wermers (1999). We exclude index funds from most of our analysis.[5] Between 1983 and 2004, there were about 11 domestic equity only index funds identified each quarter on average in the holding database. Following the standard practice in the mutual fund literature, we also exclude all closed-

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

end funds, international funds, sector funds, bond funds, and domestic hybrid funds (including life-cycle and life-style funds), based on self-reported fund style in the CDA/Spectrum database. Thus, we retain only those funds that self-report as aggressive growth (AGG), growth (GRO), or growth and income (GNI). To ensure that the funds we examine are reasonably active, we examine funds that traded at least ten stocks with a turnover of at least ten percent of its holdings during the *previous* quarter. This filter eliminates about ten percent of all fund/quarter observations. Finally, we include only those fund/quarter observations for which the fund holdings at the end of the previous two quarters are available, so that holding changes can be computed over consecutive quarters. We obtain information on after-fee performance and other fund characteristics from the Center for Research in Security Prices (CRSP) Survivor-Bias Free US Mutual Fund database.

The CDA/Spectrum mutual fund holding data are matched to the CRSP Mutual Fund data using the MFLINKS database developed by the Wharton Research Data Service (WRDS) and Professor Russ Wermers. An appealing feature of the MFLINKS database is that it allows mapping different share classes of the same fund, which are recorded as distinct funds in the CRSP Mutual Fund database, to the corresponding mutual fund holdings data in the CDA/Spectrum database. For multiple share classes in CRSP that correspond to the same fund in the CDA/Spectrum database, we aggregate returns of those share classes into one large portfolio's return by equal-weighting or value-weighting (using total net asset values). The results for equal-weighting and value-weighting are similar, so we report the results for the former for brevity.

Stock data are from the CRSP. We include all common stocks (CRSP share codes 10 and 11) traded on the NYSE, the AMEX, and the Nasdaq (CRSP exchange code 1, 2 and 3). Accounting information is from the COMPUSTAT database. To link COMPUSTAT and CRSP, we use CRSP-LINK produced by CRSP. The tick-by-tick stock transaction data are from the ISSM (1983 - 1992) and TAQ (1993 - 2004) databases.

Overall, there are $4,654$ distinct funds in our sample for the period 1983 through 2004. On average, there are 701 distinct funds every quarter. The number of funds per quarter increases from 134 in 1983 to about $1,700$ toward the end of the sample, as shown in Figure 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 for every quarter. First, we obtain common fund characteristics from the CRSP mutual fund database. These characteristics include: *age* (age of the fund in months since inception, in terms of percentile rank in the cross section); *turnover* (turnover rate of the fund); *expense* (expense ratio of the fund); TNA (total net assets under management by the fund in millions US$); and pct_flow (net fund flows in percent, defined as $\frac{TNA(t)-TNA(t-1)\times[1+Ret(t-1,t)]}{TNA(t-1)}$).[6]

Second, we aggregate stock characteristics at the fund level by value-weighting them for stocks held by the fund using the quarter-end dollar values of the holdings. These characteristics include: $fund\_holding$ (average percentage of total number of shares outstanding of stocks held by the fund); $fund\_size$ (average market capitalization of stocks held by the fund, in billions of dollars); $fund\_bm$ (average book-to-market ratio of stocks held by the fund); $fund\_mom$ (average past one-year return on stocks held by the fund); and $fund\_amihud$ (average Amihud (2002) illiquidity measure, in terms of percentile rank in the cross section, of stocks held by the fund).[7]

To evaluate mutual fund performance, we use both factor-adjusted returns and holding-based characteristics-adjusted returns performance. Our first performance measure is based on the Carhart (1997) four-factor model, which adds to the Fama-French three-factor model (1993) a momentum factor. In particular, for each mutual fund each month, we estimate the following rolling-window regression using monthly data from month $(t-36)$ to $(t-1)$:

$$R_{i,t} - R_{rf,t} = \alpha + \beta_1 MKTRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \epsilon_{i,t} \qquad (1)$$

where R_i is the mutual fund i's return *after fees and expenses*; R_{rf} is the risk-free rate; $MKTRF$ is the CRSP value-weighted market excess return over the risk-free rate; SMB is the small-minus-big (SMB) factor; HML is the high-minus-low (HML) factor; MOM is the momentum factor.[8]

To ensure our estimation is relatively precise, we impose the constraint that there must be at least 24 months' worth of mutual fund return data available for estimation prior to month (t). That is, we require the fund must have survived at least two years before we estimate the regression.

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

[7] The Amihud (2002) 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; and second, the Amihud measure may be extreme and subject to outliers. Using percentile ranking alleviates these issues.

[8] The return factors are taken from the Kenneth French's Web site: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data Library.

For month (t), we use the estimated factor loadings in the prior 36-month rolling window and the current factor realizations to adjust the mutual fund excess return over the risk-free rate and obtain the four-factor adjusted-alpha, or:

$$\widehat{\alpha}_{i,t} = (R_{i,t} - R_{rf,t}) - \left[\widehat{\beta}_1 MKTRF_t + \widehat{\beta}_2 SMB_t + \widehat{\beta}_3 HML_t + \widehat{\beta}_4 UMD_t\right] \tag{2}$$

We sum across the four-factor adjusted-alphas within a quarter to obtain the mutual fund's quarterly four-factor adjusted-return.[9]

Our second performance measure is based on the mutual fund holdings. Daniel, Grinblatt, Titman and Wermers (DGTW, 1997) and Wermers (2004) develop a characteristic selectivity (CS) measure to detect whether managers are able to select stocks that outperform average stocks with similar characteristics. We examine the actual stock holdings of the mutual fund and compute its CS measure during quarter $t+n$ as:

$$CS_{t+n} = \sum_j w_{j,t} \left[R_{j,t+n} - BR_{t+n}(j,t)\right] \tag{3}$$

where $R_{j,t+n}$ is the return on stock j during quarter $t+n$; $w_{j,t}$ is the dollar value weight of stock j held by the mutual fund at the end of quarter t; and $BR_{t+n}(j,t)$ is the benchmark portfolio return during quarter $t+n$ matching stock j at the end of quarter t according to its size, book-to-market equity ratio, and past 12-month return.

2 Information Events and Mutual Fund Performance

An active mutual fund manager will be successful if he or she has superior skill in processing valuation-relevant information on a stock, which helps the manager identify potential mispricing. It is reasonable to expect that such skills add more value when stocks the manager can invest in are affected by many value-relevant information events. Because rational managers can choose not to trade such stocks if they do not have an advantage in analyzing information, we should expect

[9] We obtain the quarterly risk-adjusted return by aggregating the monthly risk-adjusted returns because the regressions over quarterly return using the past three years of data have only 12 data points, making estimation imprecise. This is also the approach used widely in the asset management industry (see, the Style Advisor's User Guide developed by Zephyr's Associates).

those managers who choose to trade them to earn superior returns on average.

To identify the occurrence of information events, we first 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 this 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 equally likely to be a buy or a sell, resulting in an order imbalance measure close to zero on average, and a low PIN measure. On the other hand, when there are significant information events and informed traders also trade, there will be large amounts of buy orders *or* sell orders (depending on the nature of the information), resulting in a large order imbalance and a high PIN measure.[10] Empirically, PIN decreases with trading volume, size and analyst coverage, but increases with bid-ask spread, and insider and institutional ownership, consistent with it being a reasonable measure of private information events.

To estimate PIN, we use tick-by-tick transaction data for each quarter from 1983 to 2004, employing the entire three-month data to ensure precision of the estimates. A breakdown of our stock sample is provided in Panel A of Table 1. Overall, we have on average 4110 stocks with PIN measures in a quarter. Due to data availability from ISSM, NASDAQ stocks enter the sample in 1987 and account for a large portion of the sample afterwards. The mean of PIN measures in our sample is 25.8% with an associated standard deviation of 12.1%. The correlations between PIN and other stock characteristics are tabulated in Panel B of Table 1. 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

[10] A more detailed description of the PIN measure and its estimation procedure is contained in Appendix B. Recently, the PIN measure has been widely used in the empirical finance literature, for instance, in Brown, Hillegeist and Lo (2004), Vega (2006), Bharath, Pasquariello and Wu (2006), and Chen, Goldstein and Jiang (2006).

as:

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

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

At the end of each quarter over 1983 - 2004, we sort all mutual funds in the sample into deciles according to their $trade_PIN$s and examine the four-factor adjusted and characteristics-adjusted mutual fund portfolio return in the next four quarters after portfolio formation within each decile. The results are summarized in Table 2.

The central message in Table 2 is that funds trading more high-PIN stocks outperform the funds trading low-PIN stocks. Using the four-factor model for risk adjustment, we find that funds in the top $trade\_PIN$ decile outperform funds in the bottom $trade\_PIN$ decile by 48 basis points in the next quarter with a t-value of 3.15. The return spread is 46 basis points in the second quarter with a t-value of 2.98. Return spreads are 35 basis points in the third quarter (t-value = 2.24) and 35 basis points in the fourth quarter (t-value = 2.26). Thus, within a one-year horizon after portfolio formation, funds within the highest $trade\_PIN$ decile outperform the funds within the lowest trade-PIN decile by roughly 1.6 percentage points. In general, we see a positive relation between future risk-adjusted fund returns and the $trade\_PIN$ variable: the lowest five trade-PIN decile portfolios usually have large negative and statistically significant factor-adjusted return during four quarters after portfolio formation; in contrast, the highest five trade-PIN decile portfolios have small negative and in most cases statistically insignificant factor-adjusted returns.

The results are similar for the characteristics-based risk adjustment. On average, the top trade-PIN decile portfolio of funds outperform the bottom decile portfolio by 53 and 40 basis points per quarter in the first and second quarter after portfolio formation. These differences are significant at 1% significance level. The return spreads between the top and bottom deciles attenuate to 19 to 9 basis points in the third and fourth quarter.

The factor-adjusted mutual fund returns differ from the characteristics-adjusted mutual fund returns in several important aspects. First, while the four-factor adjusted-returns are after fees and expenses, the characteristic-adjusted returns are before fees and expenses. This difference explains why the characteristic-adjusted returns are largely positive while the factor-adjusted returns are mostly negative. Second, the characteristics-adjusted returns are calculated from the reported mutual fund stock holdings at portfolio formation. Therefore, they do not account for the fact that mutual funds can change their stock holdings afterward. Perhaps the best way to interpret our characteristics-adjusted returns is to ask when the "mispricing" at the stock portfolio level, if any, dissipates. Interestingly, our evidence shows such dissipation occurs within two quarters. Finally, unlike the factor-adjusted fund returns, the characteristics-adjusted fund returns ignore possible cash, stock holdings below reporting threshold and other non-stock holdings of the mutual fund. However, such holdings are usually small, accounting for less than 5% of the fund holdings on average in our sample. In addition, factor-adjustment and characteristics-adjustment generate similar return spreads between high $trade\_PIN$ fund and low $trade\_PIN$ fund deciles. These results indicate that non-stock holdings by mutual funds are unlikely to introduce any systematic biases.

2.1 Fund Characteristics

How can we tell which funds are more likely to trade high-PIN stocks? In Panel A of Table 3, we tabulate the average fund-level characteristics across $trade\_PIN$-sorted fund deciles. All characteristics are winsorized at the 1st and 99th percentiles to mitigate the effect of outliers. We find that high-$trade\_PIN$ funds are typically associated with smaller fund size, younger fund age, higher expense ratio, and higher percentage fund inflow. High-$trade\_PIN$ funds tend to hold more stocks, and a greater amount of smaller and less liquid stocks. Their stock holdings as a percentage of total number of shares outstanding are also higher on average. Finally, their investment style more likely belongs to the aggressive growth or growth funds categories, compared to the low low-$trade\_PIN$ funds, which lean more toward the growth and income.

Some fund characteristics have been documented as associated with superior fund performance. For instance, Chen, Hong, Huang, and Kubik (2004) show that smaller funds are more likely to outperform. Schultz (2007) documents that small growth stocks held by mutual funds produce abnormal returns on a characteristics-adjusted basis. The positive relation between the fund return and the $trade\_PIN$ variable could be entirely driven by other correlated fund characteristics. We examine this possibility in a cross-sectional regression framework. For each quarter in 1983 - 2004, we regress the next-quarter factor-adjusted or characteristics-adjusted fund returns on several fund-level characteristics. All right-hand side variables (except the style dummies) are measured as deviations from their corresponding cross-sectional means, standardized to have unit variance. The regression intercept can be interpreted as the average effect of the growth and income (GNI) fund style. Finally, the regression coefficients are averaged across time, and the associated t-values are computed using Newey-West correction with lag terms of eight quarters to account for autocorrelations in the error terms. The regression results are reported in Panel B of Table 3. We find $trade\_PIN$ to be significant even with the presence of other fund-level characteristics.

2.2 Robustness

We consider two alternative explanations for the empirical findings in the previous section, and we discuss them in turn.

PIN Risk?

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

To address this concern, we first directly control for PIN risk in the risk adjustment. In the case of the factor-risk adjustment, we augment the benchmark four-factor model in (1) and (2) with a PIN risk factor - PIN_t. Similar to Easley, Hvidkjaer, and O'Hara (2004), we construct the PIN risk factor as the high-PIN decile portfolio return minus the low-PIN decile portfolio return. The resulting five-factor-adjusted mutual fund return thus controls for any systematic PIN risk.

In the case of characteristics-based risk adjustment, we construct characteristics benchmark portfolios by matching along size, book-to-market, past return, and PIN characteristics simultaneously. At the end of each quarter, we sort all stocks into 81 portfolios using a 3 by 3 by 3 by 3 sequential sort based their sizes, book-to-market ratios, past 12 month returns and PIN measures (in that order). We then compute a new characteristics-adjusted fund return or characteristic selectivity measure (CS^*) using the 81 benchmark portfolio returns in (3).

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

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

To measure the average number of information events associated with stocks traded by mutual funds during a quarter, we compute $trade\_adjPIN$, $trade\_theta$, and $trade\_aturn$ in the same fashion as $trade\_PIN$ by replacing PIN with $adjPIN$, *theta*, and *aturn* in (4) accordingly. The results appear in Table 4, Panel A.

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

Momentum Trading?

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

First, factor-adjusted and characteristics-adjusted fund returns are computed throughout after adjusting for momentum effects. Second, when we regress the risk-adjusted fund returns on several fund characteristics in a cross-sectional regression, we find $fund\_mom$ to be insignificant, while $trade\_PIN$ is still highly significant (see Panel B of Table 3), confirming that the higher return associated with funds trading high-PIN is not driven by the momentum effect. Finally, we directly examine the average past return characteristics of stocks bought, sold, and held by the funds separately in Panel B of Table 4. In each quarter and for each fund, we first compute the value-weighted average past one-year return of stocks in the "buy" portfolio (stocks recently bought by the fund), 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 seem to buy more recent winners than low-$trade\_PIN$ funds (the average past one-year return in the "buy" portfolio is 34.3% for high-$trade\_PIN$ funds *vs.* 20.9% for low-$trade\_PIN$ funds), high-$trade\_PIN$ funds also sell more extreme recent winners at the same time (the average past one-year return in the "sell" portfolio is 46.6% for high-$trade\_PIN$ funds); they thus are not momentum traders in the traditional sense. In addition, funds in $trade\_PIN$ deciles 7 to 9 seem to buy or hold even more winners than funds in the top $trade\_PIN$ decile. If the momentum effect drives the high CS measure, we would expect funds in $trade\_PIN$ deciles 7 to 9 to have higher risk-adjusted returns on average. This is clearly not the case.

2.3 PIN and Fund Manager Skill

While fund managers who choose to trade high-PIN stocks outperform the other managers on average, we would expect outperformance to be concentrated among managers who in fact have superior skills. To test this conjecture, we use the four-factor-adjusted fund return or alpha as a proxy for manager skill. At the end of each quarter from 1983 through 2004, we first sort all mutual funds in the sample into three portfolios according to their alpha during the quarter. Within the high alpha fund portfolio (top one-third) and the low alpha fund portfolio (bottom one-third), we then further sort funds according to their $trade\_PIN$ variables during the quarter into three portfolios. The average next quarter four-factor-adjusted and characteristic-adjusted mutual fund returns are presented in Table 5, Panel A. We observe that funds with higher alphas this quarter continue to perform better in the next quarter. Both the factor-adjusted returns and the characteristics-adjusted returns are higher than returns of funds with low-alphas this quarter.

Perhaps more interestingly, sorting on the $trade\_PIN$ variable generates positive and significant spreads in future fund performance measures only among funds with high past alphas. Among the funds with high alphas, those trading high-PIN stocks outperform funds trading low-PIN stocks both during the current quarter and the next quarter. Their risk-adjusted return spread is 48 basis points (t-value = 3.64) based on factor-adjusted returns and by 43 basis points (t-value = 2.43) based on characteristics-adjusted returns. Funds associated with high alphas that also trade high-PIN stocks produce positive and significant alpha of 33 basis points next quarter, even after accounting for fees and expenses. In sharp contrast, among funds with low alphas, sorting on the $trade\_PIN$ variable does not generate much dispersion in future fund performance measures at all. In addition, during the current quarter, funds trading high-PIN stocks actually significantly underperform funds trading low-PIN stocks, reflecting the possibility that fund managers who are overconfident have more to lose in trading high-PIN stocks. Overall, these results provide additional support for our main hypothesis that skillful managers who can better analyze value-relevant information can add more value when trading stocks associated with more information events.

Panel B of Table 5 presents the results when we sort on $trade\_PIN$ first and fund alpha second. We find that sorting on fund alphas this quarter for funds with high $trade\_PIN$ generates a much wider spread in fund risk-adjusted returns during both this quarter and the next quarter than within funds with low $trade\_PIN$. Among funds that trade high-PIN stocks, funds that earn higher quarter t alphas have higher quarter $t+1$ alphas than funds that earn low quarter t alphas (91basis points on average with t-value of 4.36), even after accounting for fees and expenses. Among funds that trade low-PIN stocks, the quarter $t+1$ alpha spread between funds in the high- and low-quarter-t-alpha groups is only 38 basis points. The t-value of 2.43 is also much smaller, although still very significant. Results in Panel B suggest that past fund alpha does a better job in separating talented managers from ordinary managers if we also examine the stocks they trade.

Table 5 also presents the risk-adjusted mutual fund returns after directly controlling for PIN risk, and the results are very similar. Overall, the results indicate that talented managers indeed add more value by trading stocks associated with more information events. The implication of this finding is that analysis of the number of information events associated with stocks traded by the managers helps us to identify the talented ones.

3 Decomposing Mutual Fund's Stock Selection Skills

To better understand the channels through which talented fund managers can add value, we decompose the stock selection skill of a mutual fund. In general, how a manager with superior skill trades to add value will depend on how long it takes for the market to realize that the manager is correct. Analyzing how long the informational advantage lasts, a manager's trades can be classified into the following three types:

1. A manager can add value from long-term "value investing" by taking a position in a stock with the expectation that the market will eventually agree with his or her view in, say, a few years.[11]

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

2. A manager can add value from medium-term informed trading by transacting in "mispriced" stocks, expecting the market to agree with his or her view within, say, a quarter or two.[12] In this case, the value of the information is likely to erode quickly, and successful and timely trade execution may require paying a substantial price concession.

3. A fund manager can add value using short-term liquidity provision by taking the other side of a trade when liquidity is most needed.[13] Since fund managers often hold an inventory of stocks in order to track their performance benchmarks, they have a natural advantage in making a market in those stocks. Superior knowledge about the stocks covered by a manager will help in any market-making activities by minimizing potential losses that would arise from trading with market participants with an information advantage.[14]

To separate the value-added in these three different types of trading, we focus on the characteristics-based performance measure, characteristic selectivity (CS). Recall that the CS measure of a mutual fund during quarter $t+1$, based on its actual stock holdings at the end of quarter t, can be computed as:

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

where $R_{j,t+1}$ is the return on stock j during quarter $t+1$; $w_{j,t}$ is the dollar value weight of stock j held by the mutual fund at the end of quarter t; and $BR_{t+1}(j,t)$ is the benchmark portfolio

[12] The year-to-year same store sales growth reported by Starbucks every month is a widely watched number, and is considered about as important as the company's quarterly earnings announcements for valuation purposes. For January to September 2005, Starbucks' reported sales growth rates were in the range of 7% to 9%. Most analysts were of the view that a large part of that growth rate was attributable to the 3% sales price increase that took effect in October 2004, and that this price increase would not help with respect to same-month year-to-year sales growth rates beginning with October 2005. That probably explains the much smaller anticipated growth rate (analyst consensus was 3.6%). However, a careful analysis of sales breakdown would have indicated that the 3% price increase in October 2004 explained little of the sales growth during January-September 2005. So, the October sales growth figure should be more like that for the early months of 2005. While most mutual funds decreased their holdings of Starbucks stock during Q3 2005 in anticipation of an announcement of a drop in same-store sales growth for October, Putnam Voyager Fund actually accumulated more shares (see Figure 2B). On November 3, 2005 Starbucks reported unexpectedly strong sales growth of 7% for October, and its share price jumped. Details on this case can be found in Blumenthal (2007).

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

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

return during quarter $t+1$ to which stock j is matched at the end of quarter t based on its size, book-to-market equity ratio, and past 12-month return.

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

1. Hold portfolio, which has stock holdings:

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

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

2. Buy portfolio, which has stock holdings:

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

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

3. Sell portfolio, which has stock holdings:

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

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

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

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

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

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

At the end of quarter t, the mutual fund's stock holdings are a combination of the hold portfolio and the buy portfolio. The fund CS measure for quarter $t+1$ is therefore the value-weighted average of CS measures on the hold portfolio and buy portfolio for quarter $t+1$:

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

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

We then decompose the CS measure into three components:

$$
\begin{aligned}
CS_{t+1} &= CS^O_{t+1} + CS^T_{t+1} + CS^{adj}_{t+1} \\
CS^O_{t+1} &= \frac{H_t}{H_t + S_t} CS_{H,t+1} + \frac{S_t}{H_t + S_t} CS_{S,t+1} \\
CS^T_{t+1} &= \frac{B_t}{H_t + B_t} CS_{B,t+1} - \frac{S_t}{H_t + S_t} CS_{S,t+1} \\
CS^{adj}_{t+1} &= \frac{H_t}{H_t + B_t} \frac{S_t - B_t}{H_t + S_t} CS_{H,t+1}
\end{aligned}
\tag{5}
$$

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

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

The trade component (CS^T_{t+1}) measures value-added from both medium-term informed trading and short-term liquidity provision. Since mutual fund holdings are typically reported at quarterly frequency at most, in order to make a reasonable attempt to separate them, we rely on a key difference between these two types of trades. Informed trading, unlike liquidity provision trade, is likely to demand liquidity since the value of information erodes quickly over time, so timely execution becomes important. Given this intuition, we can further decompose the trade component CS^T_{t+1} into two components by comparing the sign of quarterly mutual fund holding change and the sign of market order imbalance for each stock traded by the fund (the stocks in the buy or sell portfolios) during quarter t. The market order imbalance is defined as the total number of buyer-initiated trades minus the total number of seller-initiated trades in the quarter.

Following custom in the literature, we implement the trade classification using the standard algorithm in Lee and Ready (1991). We then classify stock trades where the two signs are identical into one group, denoted by superscript "+"; and where the two signs are different into another group, denoted by superscript "−". As a result, the characteristics-adjusted returns on trades from these groups sum up to CS^T_{t+1}:

$$\begin{aligned} CS^T_{t+1} &= CS^{inf}_{t+1} + CS^{liq}_{t+1} \\ CS^{inf}_{t+1} &= \frac{B^+_t}{H_t + B_t} CS^+_{B,t+1} - \frac{S^+_t}{H_t + S_t} CS^+_{S,t+1} \\ CS^{liq}_{t+1} &= \frac{B^-_t}{H_t + B_t} CS^-_{B,t+1} - \frac{S^-_t}{H_t + S_t} CS^-_{S,t+1} \end{aligned} \tag{6}$$

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

$$
\begin{aligned}
CS_{t+1} &= CS^{O}_{t+1} + CS^{adj}_{t+1} + CS^{T}_{t+1}, \\
CS^{T}_{t+1} &= CS^{inf}_{t+1} + CS^{liq}_{t+1}.
\end{aligned} \tag{7}
$$

3.1 Discussion of the Empirical Implementation

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

Second, the division of informed trading and liquidity provision is imprecise. On the one hand, not all informed trading is liquidity demanding, especially when the trader is very patient and trades in small quantities over a relatively long period of time. In the trading of relatively large quantities quickly to take advantage of the time value of information, however, it is extremely hard not to absorb liquidity. As a result, liquidity-demanding trades are still likely information-driven on average. On the other hand, not all liquidity-demanding trades are information driven. For example, distress stocks sales by mutual funds (see Da and Gao, 2006) and assets fire-sales due to extreme flows (see Coval and Stafford, 2007) are likely to absorb liquidity but have nothing to do with mispricing trading motives. As distressed stocks are typically of small market capitalization, the impact of transactions will be alleviated, as each component of the CS measure is computed using the value-weighted average. When we leave out value-adding informed trading that is not liquidity demanding and including value-destorying distressed trading that is liquidity demanding, we are underestimating the benefit of informed-trading, and overestimating the benefit of liquidity provision. Finally, our classification of informed trading and liquidity provision depends on quarterly data, which could also be noisy. These noises may prevent us from finding any significant results.

Recognizing these challenges in the empirical exercise, we have all the same made the first attempt to bring both informed trading and liquidity provision into the evaluation of mutual fund's stock selectivity. In Appendix B, we examine several empirical properties of the decomposition that lend support for its validity. First, we show that mutual funds are likely to provide liquidity on average only when they reduce their holdings, consistent with our conjecture that it is easier to provide liquidity on stocks that one currently owns. Second, we demonstrate the effectiveness of our decomposition methodology using two specific examples: (1) Dimensional Fund Advisors, and (2) a group of domestic index funds. Third, we find that the informed trading component is more important than the liquidity provision component in explaining cross-sectional variation in the CS measures. In addition, informed trading becomes relatively more important for growth-oriented funds while liquidity provision becomes relatively more important for income-oriented funds, consistent with what one would expect.

3.2 Empirical Decomposition Results

We show for all US domestic active equity funds in the sample, the average size of each component in the first line of Table 6. Overall, the active fund managers seem to have some stock selection skill that requires trading with the order imbalance in the market. The average character selectivity measure is 23.5 basis points per quarter (t-value = 1.91), indicating the stocks selected by fund managers outperform stocks with similar characteristics. Of the 23.5 basis points, 13.9 basis points come from the passive buy-and-hold strategy and 14.2 basis points come from stocks recently traded by the funds. The adjustment component is small in absolute term (-1.8 basis points) but significant, potentially driven by fund flow to managers with skills as empirically documented by Chevalier and Ellison (1997), and Sirri and Tufano (1998), among others, and theoretically analyzed by Berk and Green (2004).[15] Finally, although both the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) are positive, neither is significant.

Table 6 also shows the decomposition applied to the decile portfolios of funds sorted on $trade\_PIN$. This reveals interesting differences in value-added between funds trading high-PIN stocks and funds trading low-PIN stocks. For high-$trade\_PIN$-funds, a large part of the character selectivity (CS)

[15] 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 (5).

measure comes from active trading during the previous quarter ($CS^T = 31.2$ basis points with a t-value of 2.83). We can confirm that the stocks bought by mutual funds (the buy portfolio) and the stocks sold by mutual funds (the sell portfolio) have very similar average PIN measures. The trade component, measuring their return difference, should be less subject to PIN risk. The fact that it is positive and significant indicates that we are not capturing just the PIN risk. Although both the informed trading component (CS^{inf}) and the liquidity provision component (CS^{liq}) are positive for high-$trade\_PIN$-funds, only the informed trading component is significant (20.4 basis points with a t-value of 2.25), and it is twice the size of the liquidity provision component (10.4 basis points). This is consistent with our conjecture. When skillful managers absorb liquidity by trading high-PIN stocks, they are likely to have valuation-relevant information, and thus make money on informed trading. For them, there is less of an added cost of demanding immediacy in the market than there is a benefit from superior information, as Grossman and Stiglitz (1980) would predict. In terms of liquidity provision, not all of them can perform well consistently. As a result, although the liquidity provision component is positive on average, it is much smaller and not significant, perhaps because of the possibility of trading against informed traders and the noise associated with identifying liquidity provision using quarterly mutual fund holdings data..

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

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

4 Conclusion

The folk wisdom is that portfolio managers can profitably take advantage of their talent when the stocks they follow are affected by information events. To the extent that economic conditions that favor one type of talent over another tend to persist for a while, managers who were able to generate value at a certain time will continue to do so for a while. It follows that the past superior performance of a portfolio manager is less likely to be due to just plain luck when that superior performance is attributable to trading in stocks that were most affected by information events.

We provide empirical support for this view. Funds that trade in stocks most affected by information events on average earned 50 basis points per quarter more after risk adjustment than funds trading in stocks least affected by information events. Among funds with a high positive alpha in one quarter, only those that traded in stocks the most affected by information events earned a significantly positive alpha (33 basis points per quarter on average) in the next quarter.

When we decompose a manager's stock selection ability further into different components, we find that impatient informed trading is more important for alpha generation for the funds trading stocks affected most by information events. Liquidity provision, on the other hand, is more important for the funds trading stocks affected least by information events. Decomposition of holding like this further facilitate the understanding of the strengths of an active portfolio manager and the extent to which such strengths will persist into the future.

Appendix A: Measures of Private Information Events

Easley and O'Hara with several coauthors develop a measure - probability of informed trading (PIN) - to capture the probability of information-based trading. Let α denote the probability that an information event occurs; δ denote the low value of underlying asset, conditioning on the occurrence of an information event; μ the rate of informed trade arrivals; ϵ_b the arrival rate of uninformed buy orders; and ϵ_s the arrival rate of uninformed sell orders.

Easley, Hvdkjaer, and O'Hara (2002) propose an maximal likelihood estimation of the parameter vector $\Theta \equiv \{\alpha, \mu, \epsilon_b, \epsilon_s, \delta\}$

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

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

$$PIN = \frac{\alpha\mu}{\alpha\mu + \epsilon_b + \epsilon_s} \tag{A-2}$$

With some independence assumptions across trading days, the likelihood function (A-1) becomes

$$L\left(\Theta|\,(B_i,S_i)_{i=1}^{i=N}\right) = \prod_{i=1}^{N} L(\Theta|B_i,S_i)\,. \tag{A-3}$$

The problem with estimation of PIN measure is that numbers of buy and sell orders have risen considerably since 2001, particularly for some Nasdaq stocks. One way to alleviate this problem, as in Vega (2006), is to impose the constraint that the informed and uninformed orders arrive at the same rate:

$$\epsilon_b = \epsilon_s = \epsilon \tag{A-4}$$

Then we estimate a modified version of (A-1):

$$L(\Theta|B,S) = (1-\alpha)\, e^{-2\epsilon}\frac{\epsilon^{B+S}}{B!S!} + \alpha\delta e^{-(\mu+2\epsilon)}\frac{\epsilon^B(\mu+\epsilon)^S}{B!S!} + \alpha(1-\delta)\, e^{-(\mu+2\epsilon)}\frac{\epsilon^S(\mu+\epsilon)^B}{B!S!} \tag{A-5}$$

The probability of informed trading, PIN, is

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\epsilon}. \tag{A-6}$$

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

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

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

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 order imbalance (OIB) is then:

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

Aktas, Bodt, Declerck, and Oppens (2007) and Kaul, Lei, and Stoffman (2007) show that a relative order imbalance measure $rel\_OIB = E[|B-S|]/E[B+S]$ is a good approximation of PIN. In fact, on a daily basis, rel_OIB is equivalent to PIN. In addition, the average cross-sectional correlation between these two measures is above 0.75. rel_OIB is clearly a measure of order flow one-sidedness. Our results are almost identical when we replace PIN with rel_OIB.

Durate and Young (2007) extend (A-1) to take into account large buy and sell volatilities, and pervasive positive correlation between buy and sell orders. Their model allows the possibility of order flow shocks and different distributions of the number of the buyer-initiated informed trades and the number of the seller-initiated informed trades. With such an extension, one may estimate an adjusted version of the probability of informed trading (AdjPIN) as

$$AdjPIN = \frac{\alpha \times [(1-\delta) \times \mu_b + \delta \times \mu_s]}{\alpha \times [(1-\delta) \times \mu_b + \delta \times \mu_s] + (\Delta_b + \Delta_s) \times \big[\alpha \times \theta' + (1-\alpha) \times \theta\big] + \epsilon_b + \epsilon_s} \tag{A-7}$$

where the additional parameter θ denotes the probability of symmetric order flow shocks conditional on no arrival of private information event, and θ' denotes the probability of symmetric order flow shocks conditional on the arrival of private information. Δ_b and Δ_s denote the additional arrival rate of buy orders and sell orders conditional on the arrival of the symmetric order flow shocks.

Durate and Young (2007) simplify (A-7) by restricting $\theta = \theta'$. The most important feature of the model is the probability of symmetric order flow shocks could be non-negative ($\theta \geq 0$). Since we estimate the $AdjPIN$ measure quarterly using the daily trade imbalanced data aggregated from the tick-by-tick transactions, to reduce the sheer volume of calculations, and to estimate a relatively parsimonious model with fewer parameters, we further impose the constraints that $\mu_b = \mu_s = \mu$ and $\Delta_b = \Delta_s = \Delta$. According to Durate and Young (2007), the adjusted-PIN estimated with these constraints generate similar results to their full-fledged model.

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

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

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

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

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

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

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

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

Appendix B: Decomposition of Mutual Fund Stock Selection Skill - A Numerical Example and Empirical Validations

B.1: A numerical example

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

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

The hold, buy and sell are then defined by the holdings N_t^H, N_t^B and N_t^S:

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

The portfolio values H_t, B_t, and S_t are determined using the prices at the end of quarter t (P_t). Notice that $B_t > S_t$, and the difference is likely financed by fund inflows, or a reduction in cash position or the sale of other non-stock assets held by the fund. The hold, buy, and sell can be treated as three separate funds whose CS measures can be computed using equation (3) and holdings as:

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

Given this information, equation (5) then decomposes the total CS measure into three components:

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

If we further assume that the fund trades B and F in the same direction as the aggregate order imbalance and trades A and E against the direction of aggregate order imbalance, equation (6) further decomposes the trade component (CS_{t+1}^T) into an informed trading component (CS_{t+1}^{inf}) and a liquidity provision component (CS_{t+1}^{liq}):

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

B.2: Type of mutual fund trades and average order imbalances

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

The results are provided in Panel A of Table 7. The average order imbalance measure for each trade type tells us whether the trade is on average demanding liquidity. When fund managers open new positions and close out standing positions, they are likely to absorb market liquidity. In this case, these trades are likely motivated by considerable mispricing perceived by fund managers who are willing to pay for the price of immediacy. When fund managers adjust their holdings, they are likely to provide liquidity on average only when they reduce their holdings, consistent with our conjecture that it is easier to provide liquidity on stocks a fund currently owns.

B.3: Empirical decomposition examples

We demonstrate the effectiveness of our decomposition methodology using two specific examples: (1) Dimensional Fund Advisors, and (2) a group of domestic index funds.

Dimensional Fund Advisors (DFA)

Dimensional Fund Advisors (DFA) is an asset management firm founded in 1981. It is well known that the firm does not use fundamental analysis to pick stocks, but instead helps its clients gain exposure to certain segments of the asset markets through passive indexing or enhanced indexing. Anecdotal evidence suggests that a subset of the funds managed by DFA create value by systematically providing liquidity to those who want to trade small stocks for non-information-related reasons.[16]

If this is the case, our decomposition procedure would allow us to find a positive liquidity provision component in DFA's characteristic selectivity measure and an informed trading component

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

close to zero. Of course, since we examine one specific fund over a limited time period, statistical significance could be rather weak.

We examine the quarterly stock holdings of DFA's flagship fund, US Micro Cap Portfolio, over the period 1983 - 2004 and decompose its CS measure. The results are presented in Panel B of Table 7. The overall CS measure for the fund is 36.1 basis points per quarter, but this is not statistically significant (t-value = 1.72), indicating that the fund does not seem to exhibit any ability to select stocks that outperform those with similar characteristics. As expected, the largest component of the overall CS measure results from liquidity provision (20.5 basis points per quarter), which is significant at the 10% level (t-value = 1.84). The informed trading component, however, is very close to zero and statistically insignificant, which is consistent with the claims of the firm's investment policy.

Index funds

The majority of index funds are formed to track a market index or other broad index with the objective of minimizing tracking errors, so we would not expect them to have a high CS measure. Index funds are most likely to trade during index rebalancing and to demand liquidity in those trades (see Blume and Edelen, 2004). These trades would be incorrectly classified as informed trading within our decomposition framework, and the informed trading component, if different from zero, is likely to be negative. It is of course not useful to apply the decomposition to index funds. For that reason we focus only on actively managed funds in the text, but examining index funds provides another way to test the validity of our decomposition approach.

We identify index funds by their fund names as recorded in CDA/Spectrum S12 mutual fund holdings database. Between1983 and 2004, there are an average of 11 domestic equity-only index funds identified each quarter, starting from one 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 then equal-weight the results across funds for every quarter. These results are in the second part of Panel B, Table 7. The overall CS measure for index funds as a group is almost exactly zero. The index fund group has a positive although not significant CS^O component of about 25 basis points per quarter on average (t-value = 0.93), but this is not statistically significant. 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 basis points per quarter (t-value = 0.36). Interestingly, the positive CS^O and CS^{liq} are offset by a negative informed trading component (CS^{inf}= -35 basis points), which is statistically significant, indicating a sizable price for liquidity paid by the index funds for trades that arise due to index rebalancing, new money flowing in, and redemptions.

B.4: Variance decomposition results

To examine the relative importance of each component of the total characteristic selectivity measure, we carry out a variance decomposition exercise. We decompose the total characteristic selectivity measure into four components:[17]

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

Consequently, we have

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

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

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

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_O is estimated by regressing CS^O 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. The variance decomposition delineates how much the cross-sectional variation in the total CS measure can be attributed to the cross-sectional variation in each of its four components.

The results are reported in Panel C of Table 7 for the full sample of all active US equity funds, and across the three style subsamples. Overall, the old component (CS^O) explains about 57% of the

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

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 more important for growth-oriented funds, while CS^{liq} becomes relatively more important for income-oriented funds.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[27] Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2004, Factoring information into returns, *Working Paper*, Cornell University.

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

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

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

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

[32] Ferson, Wayne, and Richard Schadt. 1996. Measuring Fund Strategy and Performance in Changing Economic Conditions, *Journal of Finance* 51(2), 425- 462.

[33] French, Kenneth, 2008, The cost of active investing, *Working Paper*, Dartmouth College.

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

[35] 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.

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

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

[38] 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.

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

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

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

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

[43] Kaul, Gauta, Qin Lei, and Noah Stoffman, 2007, Information asymmetry or liquidity, *Working Paper*, University of Michigan.

[44] 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.

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

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

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

[48] Sias, Richard W., Laura T. Starks, and Sheridan Titman, 2006, Changes in institutional ownership and stock returns: assessment and and methodology, Journal of Business 79(6), 2869-2910.

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

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

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

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

Figure 1: Mutual fund sample according to investment objective

Index funds, lifecycle mutual funds, bond funds, hybrid funds, sector funds, and international funds are excluded. Only those funds that self-report as aggressive growth (AGG), growth (GROWTH), or growth and income (GNI) are included. Fund/quarter observations with quarterly turnover under 10% or under 10 stocks are excluded. Fund holdings at the end of the preceding quarter must also be available in order to calculate holding changes over consecutive quarters. The CDA/Spectrum mutual fund holding data are matched to CRSP Mutual Fund data via the MFLINKS database.

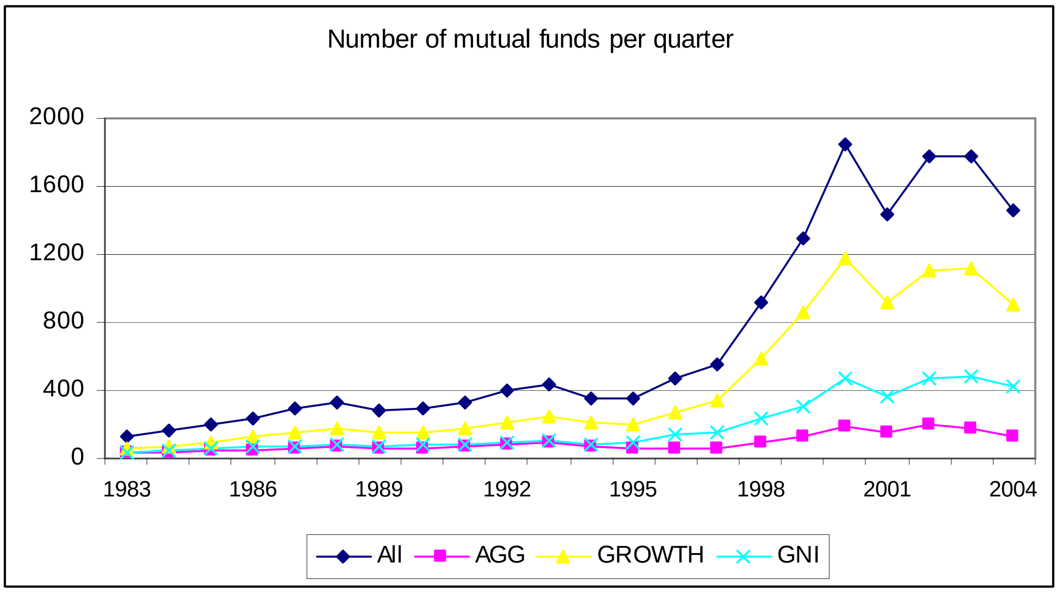
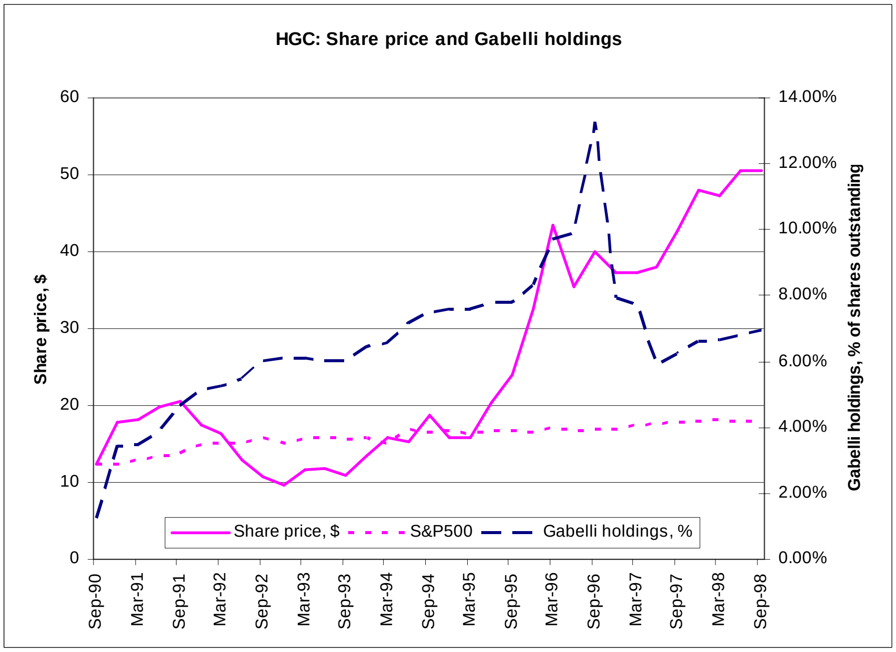


Figure 2: Share price and mutual fund holdings

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

A: Hudson General Corp (HGC) and Gabelli Holdings



B: Starbuck and Putnam Voyager Fund Holdings

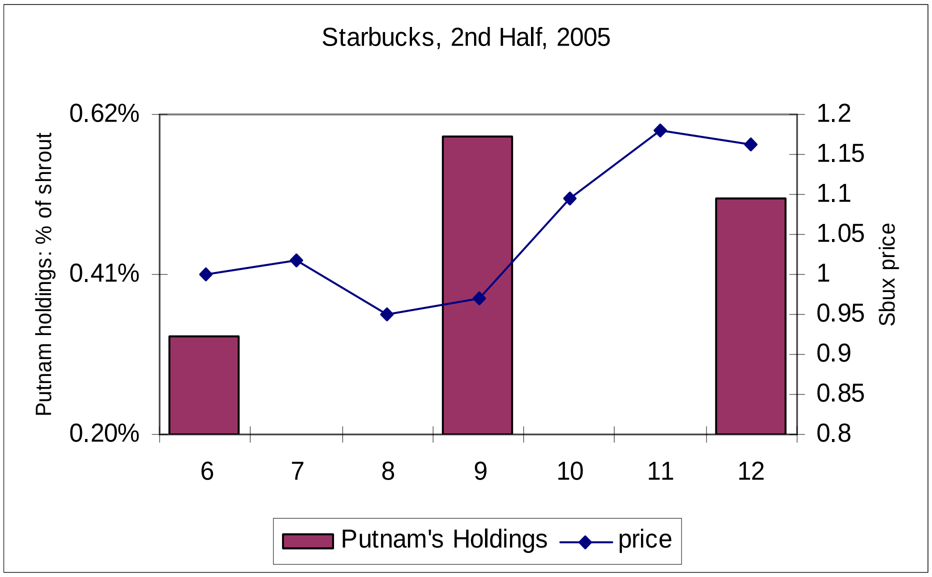


Table 1: Descriptive statistics of *PIN*

Probability of informed trading (PIN) is estimated at quarterly frequency over 1983-2004 using three-month trade and quote data from TAQ. Panel A describes the stock *PIN* sample over time. 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	100.0%	0.0%	22.5%	10.2%
1984	1747	100.0%	0.0%	25.2%	13.0%
1985	1812	100.0%	0.0%	24.1%	11.8%
1986	1828	100.0%	0.0%	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)	Momentum
log(Size)	-0.536			
log(BM)	0.169	-0.193		
Momentum	-0.066	0.058	-0.148	
Amihud	0.557	-0.872	0.190	-0.198

Table 2: Risk-adjusted *quarterly* fund returns across *trade_PIN* sorted fund deciles

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

trade_PIN in Qtr t	4-factor adj returns				Char-adj returns (CS)			
	Qtr t+1	Qtr t+2	Qtr t+3	Qtr t+4	Qtr t+1	Qtr t+2	Qtr t+3	Qtr t+4
Low	-0.0038	-0.0037	-0.0040	-0.0038	-0.0003	0.0004	-0.0002	0.0002
	-3.45	-3.51	-3.45	-3.75	-0.29	0.42	-0.22	0.25
2	-0.0040	-0.0039	-0.0044	-0.0033	0.0011	0.0015	0.0007	0.0005
	-4.01	-4.34	-4.25	-3.47	0.92	1.40	0.70	0.52
3	-0.0032	-0.0030	-0.0039	-0.0041	0.0012	0.0007	0.0007	-0.0003
	-2.87	-3.03	-4.46	-3.81	1.04	0.74	0.72	-0.32
4	-0.0031	-0.0036	-0.0037	-0.0037	0.0010	0.0012	0.0008	0.0000
	-3.09	-3.66	-4.31	-3.38	1.01	1.19	0.84	0.01
5	-0.0017	-0.0026	-0.0035	-0.0034	0.0029	0.0014	0.0006	0.0007
	-1.57	-2.39	-3.24	-3.21	2.17	1.15	0.51	0.64
6	-0.0027	-0.0017	-0.0019	-0.0031	0.0032	0.0023	0.0016	-0.0005
	-2.18	-1.33	-1.36	-2.68	2.07	1.58	1.26	-0.37
7	-0.0025	-0.0022	-0.0016	-0.0022	0.0028	0.0016	0.0016	0.0006
	-1.80	-1.67	-1.24	-1.58	1.52	1.04	1.19	0.41
8	-0.0019	-0.0003	-0.0008	-0.0018	0.0031	0.0024	0.0024	0.0014
	-1.16	-0.22	-0.52	-1.14	1.73	1.41	1.53	0.85
9	-0.0018	-0.0014	-0.0018	-0.0018	0.0035	0.0016	0.0006	-0.0004
	-1.06	-0.82	-1.11	-1.18	1.75	0.84	0.35	-0.21
High	0.0011	0.0009	-0.0005	-0.0003	0.0050	0.0044	0.0017	0.0011
	0.85	0.67	-0.36	-0.21	2.70	2.75	1.20	0.68
High-Low	0.0048	0.0046	0.0035	0.0035	0.0053	0.0040	0.0019	0.0009
	3.15	2.98	2.24	2.26	2.87	2.55	1.29	0.52

Table 3: Fund-level characteristics and cross sectional regressions

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 the 1st and 99th percentile to alleviate the effect of outliers.

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

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	*11.48*	*73.17*	*53.47*	*-9.76*	*3.05*	*6.19*	*47.36*	*-14.53*	*1.58*	*12.92*	*-14.94*	*4.59*	*13.74*	*7.50*	*-26.84*

Panel B: Cross-sectional regressions

	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 = Four-factor alpha in Qtr t+1													
coeff	-0.0037	0.0026	0.0016	-0.0008	-0.0013	-0.0003	-0.0002	-0.0004	-0.0014	0.0003	0.0011	0.0017	0.18
t-value	*-2.92*	*3.97*	*1.66*	*-1.21*	*-1.14*	*-0.21*	*-0.71*	*-1.86*	*-3.89*	*0.51*	*1.07*	*1.18*	
LHS = CS in Qtr t+1													
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*	

Table 4: Robustness

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

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

Panel B reports the average past one-year return of stocks bought / sold / held by mutual funds across *trade_PIN* sorted deciles. For each fund, we 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.

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

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

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

trade_PIN	Past One-year Return			Buy-sell	t-value
	Buy	Sell	Hold		
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 5: Persistence of mutual fund returns and *trade_PIN*s

At the end of each quarter from 1983 to 2004, we conduct three by three double sorts based on *trade_PIN*s and four-factor alphas of the mutual funds. We then report the next quarter risk-adjusted returns on different mutual fund portfolios constructed by the double sorts. For benchmark risk adjustment, we use both the four-factor model and the characteristics model for risk adjustment. To control for the systematic risk associated with high-PIN stocks, we compute a five-factor-adjusted mutual fund return by augmenting the benchmark four-factor model with a PIN risk factor. The PIN risk factor is constructed as the high-PIN-decile portfolio return minus the low-PIN-decile portfolio return. To control for PIN characteristics risk, we construct characteristics benchmark portfolios by matching along size, book-to-market, past return and PIN simultaneously. *t*-values are reported in *italic*s. Panel A reports the result on a double sort where we sort on mutual fund return alpha first and *trade_PIN* second. Panel B reports the result on a double sort where we sort on *trade_PIN* first and mutual fund return alpha second.

Panel A: Sorting on *trade_PIN* among mutual funds with high and low alphas

Sort on *trade_PIN* (Qtr t)	*trade_PIN* (Qtr t)	Benchmark risk adjustment				Controlling for PIN risk			
		4f alpha (Qtr t)	CS (Qtr t)	4f alpha (Qtr t+1)	CS (Qtr t+1)	5f alpha (Qtr t)	CS* (Qtr t)	5f alpha (Qtr t+1)	CS* (Qtr t+1)
High alpha, Qtr t (top 1/3):									
Low	0.1261	0.0279	0.0183	-0.0015	0.0013	0.0283	0.0188	-0.0013	0.0004
	62.63	*19.07*	*11.89*	*-1.07*	*0.88*	*18.21*	*11.17*	*-0.95*	*0.24*
Medium	0.1535	0.0329	0.0236	0.0003	0.0031	0.0333	0.0244	0.0008	0.0027
	71.92	*19.50*	*11.59*	*0.20*	*1.64*	*19.79*	*10.57*	*0.47*	*1.25*
High	0.2011	0.0365	0.0273	0.0033	0.0056	0.0365	0.0264	0.0034	0.0045
	88.08	*22.59*	*11.52*	*2.10*	*2.33*	*23.12*	*11.25*	*2.10*	*1.93*
High - Low	0.0750	0.0086	0.0090	0.0048	0.0043	0.0082	0.0075	0.0047	0.0041
		11.25	*5.30*	*3.64*	*2.43*	*9.73*	*4.47*	*3.43*	*2.15*
Low alpha, Qtr t (bottom 1/3):									
Low	0.1245	-0.0331	-0.0150	-0.0062	0.0004	-0.0325	-0.0160	-0.0052	0.0000
	62.81	*-24.67*	*-14.92*	*-4.56*	*0.33*	*-23.34*	*-15.82*	*-3.63*	*-0.03*
Medium	0.1509	-0.0377	-0.0185	-0.0041	0.0026	-0.0370	-0.0198	-0.0038	0.0016
	70.51	*-24.78*	*-12.90*	*-2.71*	*1.64*	*-22.62*	*-13.76*	*-2.34*	*1.08*
High	0.1965	-0.0408	-0.0202	-0.0048	0.0009	-0.0404	-0.0229	-0.0044	-0.0008
	86.80	*-26.08*	*-11.88*	*-2.66*	*0.46*	*-24.00*	*-12.82*	*-2.26*	*-0.38*
High - Low	0.0720	-0.0077	-0.0052	0.0014	0.0005	-0.0080	-0.0069	0.0008	-0.0008
		-9.72	*-3.58*	*0.89*	*0.30*	*-8.81*	*-4.58*	*0.49*	*-0.39*

Panel B: Sorting on alpha among mutual funds with high and low trade_PINs

Sort on 4f alpha (Qtr t)	*trade_PIN* (Qtr t)	Benchmark risk adjustment				Controlling for PIN risk			
		4f alpha (Qtr t)	CS (Qtr t)	4f alpha (Qtr t+1)	CS (Qtr t+1)	5f alpha (Qtr t)	CS* (Qtr t)	5f alpha (Qtr t+1)	CS* (Qtr t+1)
High *trade_PIN* funds, Qtr t (top 1/3):									
Low	0.1957	-0.0420	-0.0211	-0.0050	0.0009	-0.0416	-0.0236	-0.0047	-0.0007
	89.49	*-23.56*	*-11.83*	*-2.68*	*0.43*	*-22.50*	*-12.75*	*-2.36*	*-0.34*
Medium	0.1970	-0.0014	0.0030	0.0001	0.0043	-0.0015	0.0021	0.0000	0.0032
	89.49	*-1.07*	*1.82*	*0.05*	*2.31*	*-1.00*	*1.12*	*-0.03*	*1.65*
High	0.1983	0.0397	0.0298	0.0040	0.0058	0.0397	0.0287	0.0044	0.0046
	91.68	*20.61*	*11.73*	*2.44*	*2.33*	*20.83*	*11.38*	*2.55*	*1.90*
High - Low	0.0025	0.0816	0.0509	0.0091	0.0049	0.0813	0.0523	0.0091	0.0054
	2.81		*21.93*	*4.36*	*2.27*	*30.75*	*21.88*	*4.17*	*2.36*
Low *trade_PIN* funds, Qtr t (bottom 1/3):									
Low	0.1231	-0.0314	-0.0141	-0.0054	0.0005	-0.0307	-0.0148	-0.0045	0.0001
	62.96	*-23.38*	*-13.78*	*-4.24*	*0.41*	*-21.89*	*-14.43*	*-3.34*	*0.09*
Medium	0.1234	-0.0036	0.0007	-0.0036	0.0011	-0.0029	0.0001	-0.0030	0.0004
	61.68	*-4.14*	*0.70*	*-3.81*	*1.02*	*-3.21*	*0.09*	*-3.25*	*0.38*
High	0.1234	0.0238	0.0154	-0.0017	0.0022	0.0241	0.0157	-0.0014	0.0013
	61.87	*18.78*	*10.88*	*-1.33*	*1.49*	*17.56*	*10.13*	*-1.12*	*0.86*
High - Low	0.0003	0.0552	0.0295	0.0038	0.0017	0.0549	0.0304	0.0031	0.0012
	0.89		*20.79*	*2.43*	*1.15*	*26.85*	*21.49*	*1.97*	*0.77*

Table 6: Characteristic Selectivity (CS) measure decomposition across *Trade_PIN* sorted fund deciles

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

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

Table 7: Empirical validations of the Characteristics Selectivity (CS) decomposition

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

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

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

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

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

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

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

Panel B: Characteristics Selectivity (CS) decomposition for DFA US Micro-Cap fund and index funds as a group

	Total CS (=1+2+3)	Old CS^{O} (1)	Adj CS^{adj} (2)	Trade CS^{T} (3=3a+3b)	Info trading CS^{inf} (3a)	Liquidity Prov CS^{liq} (3b)
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*

Panel C: Variance Decomposition of *CS* measure

Old CS^{O}	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*