CFR-Working Paper NO. 06-09

# Forecasting stock returns through an efficient aggregation of mutual fund holdings

R. Wermers • T. Yao • J. Zhao

centre for financial research
Look deeper

# Forecasting Stock Returns through An Efficient Aggregation of Mutual Fund Holdings*

**Russ Wermers**

University of Maryland

**Tong Yao**

University of Iowa

**Jane Zhao**

PanAgora Asset Management

Current Draft: July 2012

*We thank the editor Laura Starks, an anonymous referee, Craig MacKinlay, Lubos Pastor, and seminar participants at the 2007 American Finance Association Annual Meetings, Barclays Global Investors, the 2006 BSI Gamma Foundation Annual Conference, Goldman Sachs Asset Management, the 2007 INQUIRE UK/Europe Joint Conference, PanAgora Asset Management, University of Arizona, University of British Columbia, University of Chicago, Chulalongkorn University (Bangkok), University of Kentucky, and the 2006 Wharton Conference on "Frontiers of Investing" for helpful comments. Financial support for this project was provided by INQUIRE Europe and the BSI Gamma Foundation. All errors are our own. Send correspondence to Russ Wermers, Department of Finance, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742. Email: rwermers@rhsmith.umd.edu. Phone: (301)405-0572. Fax: (301)405-0359.

# Forecasting Stock Returns through An Efficient Aggregation of Mutual Fund Holdings

## Abstract

We develop a stock return-predictive measure based on an efficient aggregation of the portfolio holdings of all actively managed U.S. domestic equity mutual funds, and use this model to study the source of fund managers' stock-selection abilities. This "generalized-inverse alpha" (GIA) approach reveals differences in the ability of managers to predict firms' future earnings from fundamental research. Notably, the GIA's return-forecasting power is not subsumed by publicly available quantitative predictors, such as momentum, value, and earnings quality, nor is it subsumed by methods shown in past research to forecast stock returns using fund holdings or trades.

The value of active investment management is a long-standing controversy. While it has been well documented that actively managed US equity funds, on average, underperform passive benchmarks, recent studies on fund returns find that a subset of funds exhibit some ability to outperform, especially when evaluated gross of fund expenses. Another body of literature uses fund portfolio holdings to evaluate the skills of fund managers. These studies suggest that fund portfolio holdings may reveal further information about the stock selection skills possessed by fund managers.[1]

This study proposes a systematic method to extract such stock selection information. The inputs to our model are fund portfolio holdings and proxies for fund stock selection skills (e.g., past fund performance). The output is a stock-level signal that aggregates this stock selection information across funds. This approach allows us to shift the analysis of the value of active investing from the fund level to its value in predicting individual stock returns. As an advantage demonstrated in the empirical part of the paper, the stock-level signal produced by the model serves as a "magnifying glass" on the collective stock picking wisdom of fund managers, and enables us to explore issues not straightforward to address at the fund level.[2]

Prior studies (e.g., Chen, Jegadeesh, and Wermers, 2000) have attempted to obtain stock-level return predictive information from aggregate fund holdings or aggregate fund trades. While intuitive, this simple aggregation approach may not be an optimal way to extract stock selection information. Specifically, since fund manager skills are heterogeneous with only a small group of managers actually possessing stock selection information at any given time, a simple aggregation of fund holdings or trades may be noisy because it mixes information from skilled managers with

[1]Recent studies using fund returns to evaluate performance include, for example, Kosowski et al. (2006), Barras et al. (2010), and Fama and French (2010). Studies using fund holdings to evaluate performance include, for example, Grinblatt and Titman (1989 and 1993), and Daniel, Grinblatt, Titman, and Wermers (1997). Recent studies also show that fund holdings are informative in understanding various aspects of fund performance and strategies; see, e.g., Cohen, Coval, and Pastor (2005), Kacperczyk, Sialm, and Zheng (2005, 2008), Kacperczyk and Seru (2007), Alexander, Cici, and Gibson (2007), Jiang, Yao, and Yu (2008), Cremers and Petajisto (2009), Huang, Sialm, and Zhang (2010), Kacperczyk, van Nieuwerburgh, and Veldkamp (2011).

[2]From the perspective of offering relevant information for investment decisions, our interest in the stock-level analysis is also motivated by the fact that many factors limit the potential of direct fund-level investment strategies. Berk and Green (2004) show that, due to a competitive supply of capital, fund managers or fund companies are able to extract a large portion of the economic rents their talents produce (e.g., via expanding the asset base in the presence of diseconomies-of-scale). As a consequence, if a fund manager generates any return surplus for investors net of fees and other costs, it tends to be short-term and fleeting. Further, due to frictions such as fund loads, short-term trading restrictions, and the inability to short-sell funds, it is difficult for investors to implement a mutual fund selection strategy to take advantage of short-term fund return predictability.

that from unskilled ones. In this paper, we extract stock selection information from fund holdings by taking into account cross-sectional differences in fund skills as well as the commonality in stock holdings across funds. For instance, we might expect a stock to outperform its benchmark in the future if it is heavily held in common by several fund managers exhibiting past skills; however, if this stock is also held by several fund managers exhibiting poor past skills, we might not be so optimistic about its outlook.

Our model starts with the notion that the stock selection skill of an equity fund is the weighted average of the alphas of stocks held by the fund. Combining this relation with an observed measure of fund stock selection skill (e.g., past fund alpha), we turn the problem of extracting information about future stock alphas into a regression, where the dependent variable is the past fund skill measure, the explanatory variables are fund portfolio weights, and the coefficients to be estimated are the future stock alphas. However, in solving for the stock alphas, there is a dimensionality issue: the number of mutual funds in any given period is smaller than the number of stocks they collectively hold; hence, the regression is underspecified. We resort to the generalized-inverse regression, instead of OLS, to solve this technical problem. The resulting estimator for stock alphas – which we term the "generalized-inverse alpha" (GIA) – has a form similar to the OLS estimator, and can be viewed as fund skill smoothed by functions of fund portfolio weights. Alternatively, the GIA can be viewed as a cross-sectional aggregate of private information about the value of a particular stock, where the weight on a given fund manager's private signal (which is revealed through her portfolio weight on a stock) is based on our assessment of the precision of her signal (relative to that of other managers). In a mean-variance optimization setting, this insight can be traced to Treynor and Black (1973). However, our approach does not require fund portfolios to be mean-variance efficient.

We show that the GIA exhibits significant and consistent power in predicting cross-sectional stock returns. In the baseline case, we assume no delays in fund portfolio disclosure to the public, and measure fund stock selection ability with a rolling, lagged 12-month four-factor Carhart (1997) alpha. Using this approach, the decile portfolio of stocks ranked highest by the GIA outperforms the bottom-ranked decile by 1.53% (with a t-statistic of 5.11) during the first quarter after portfolio formation. This performance difference remains significant for two more quarters; over a four-

quarter holding period, the return difference between top and bottom deciles is over 4% (t=3.99) per year, using the overlapping portfolio approach of Jegadeesh and Titman (1993). Notably, these return spreads remain significant after controlling for size, book-to-market ratio, and momentum, using both the Daniel, Grinblatt, Titman, and Wermers (1997) characteristic-based benchmark approach and the four-factor regression model of Carhart (1997).

To simulate the strategy in real-time, we also construct the GIA using fund portfolio holdings lagged by one quarter to take into account the portfolio reporting delays of funds (currently, a maximum of 60 days after the end of the fiscal quarter). The return-predictive performance of the GIAs is only slightly weakened when we use fund holdings that are lagged to reflect the time delay in their availability to the public. Further, we find that the GIA delivers significant return-predictive performance of stocks across the spectra of size, book-to-market ratio, past returns, idiosyncratic volatility, and breadth of mutual fund ownership. The fact that stock alphas are not concentrated among small and illiquid stocks further highlights the success of our approach in uncovering mutual fund manager skills.

We further explore the relation of the GIA to other studies that use mutual fund portfolio holdings to predict stock returns. Brown, Wei, and Wermers (2011) document an initial continuation, followed by a reversal of stock prices following mutual fund herding trades due to the price impact of unskilled fund herding. Chen, Jegadeesh, and Wermers (2000) show that aggregate (dollar-weighted) mutual fund trades positively predict stock returns, suggesting that mutual funds, in aggregate, have better stock selection abilities than other investors (e.g., individuals). Chen, Hong, and Stein (2002) find that stocks owned by fewer mutual funds tend to have low returns, and attribute this to the effect of (often self-imposed) short-sale constraints by mutual funds. Finally, Frazzini and Lamont (2008) document a "dumb money" effect – mutual fund flows are largely driven by investor sentiment, and high-sentiment stocks tend to generate low returns at long horizons. We find that the stock selection information produced by the GIAs has little correlation with these effects, and its return-predictive power remains significant after controlling for such effects.

Our new methodology also enables us to characterize the stock selection information fund managers possess, especially as it relates to the tendency of most fund managers to use "fundamental strategies" in picking stocks. A majority of active mutual fund managers claim that their stock

selection power stems from private information produced via fundamental research. We use the GIA approach to evaluate two aspects of this long-standing claim. First, we examine the standardized unexpected earnings, earnings surprises relative to the latest consensus forecast, revisions of analysts' earnings forecasts, and stock returns during short windows around earnings announcements, across the GIA ranked stock deciles. In all cases, we find that stocks with high GIAs are substantially more likely to have improving corporate fundamentals during the four subsequent quarters.

Second, we compare the GIA-based alphas with those earned by quantitative investment signals that are based on publicly available financial and accounting information. The 11 quantitative signals we examine are those analyzed by Jegadeesh et al. (2004), based on firm characteristic measures such as liquidity, momentum, value, earnings quality, and corporate investments. We find that the GIAs have a momentum tilt, but are not highly correlated with other types of indicators. Further, after controlling for the 11 quantitative signals, the GIAs continue to significantly predict stock returns. Therefore, we conclude that the stock selection information produced by fund managers through fundamental research is distinct from the information contained in publicly available quantitative signals.

Importantly, we demonstrate that our GIA approach can be extended to incorporate alternative proxies of fund skills in addition to past fund alphas. If additional fund characteristics help to predict fund performance, our approach enables us to determine whether such predictability is due to the stock selection ability captured by the characteristics, or to other economic reasons (e.g., liquidity provision or skills in trade execution). Further, if they are indeed related to funds' stock selection ability, we would like to know whether skilled funds suggested by different characteristics can identify different sets of mispriced stocks, or whether they tend to have similar stock selection information.

We find evidence of significant return predictive power by our GIA using a wide variety of skill proxies, including fund fees, TNA, turnover, the industry concentration index of Kacperczyk et al. (2005), the fund return gap of Kacperczyk et al. (2008), the similarity-based performance measure of Cohen, Coval, and Pastor (2005), the measure of earnings-announcement window returns of fund trades (Baker et al., 2009), and the R-squared of the Carhart four-factor model (Amihud and

Goyenko, 2009). Further, we find that the GIAs based on many of these alternative skill proxies and the baseline GIA (using the four-factor fund alpha) share common return-predictive components, suggesting that these fund skill measures tend to capture similar sources of fund managers' stock selection information.

Our stock-level study is related to the fund-level study of Cohen, Coval, and Pastor (2005). Based on the intuition that funds with similar ability may have similar holdings, they combine fund holdings and fund returns into a new performance measure, and show that the new measure predicts fund performance better than traditional fund alphas. At an intermediate step, they develop a "stock-quality" measure – the weighted average of the alphas of funds holding a stock. As discussed later in this paper, their stock-quality measure can be viewed as an intuitive and simplified version of the generalized inverse alpha by ignoring the cross-sectional covariance of fund portfolio holdings. Since their study focuses on predicting fund performance, it does not provide a formal derivation of the stock quality measure or direct empirical evidence on the return predictive power of the stock quality measure.

It is also worth mentioning that our method of aggregating holdings information across funds is distinct from a few recent studies that primarily distill stock selection information from the holdings within a given fund. Cohen, Polk, and Silli (2010) focus on "best-idea" stocks, i.e., stocks with the highest fund portfolio weights relative to benchmarks, and show that these stocks outperform other stocks held by funds. Based on the assumption that fund portfolios are mean-variance efficient, Shumway, Szefler, and Yuan (2009) reverse-engineer fund managers' beliefs about expected stock returns. However, these papers do not focus on developing an efficient way to aggregate stock selection information across funds.

The remainder of our paper proceeds as follows. Section 1 introduces model assumptions and develops the GIA estimators. Section 2 describes the data and empirical methodology, while Section 3 presents the main empirical results on the performance of the GIAs. Section 4 uses the GIA approach to perform additional analysis on the stock selection information possessed by fund managers. Section 5 concludes.

# 1. The Generalized-Inverse Alpha: Methodology

## 1.1 Assumptions

Assume that M mutual funds jointly hold N unique stocks. We start with the definition that a fund's stock selection ability is the weighted average of abnormal returns of the stocks the fund holds:

$$S_{jt+1} = \sum_{i=1}^{N} \omega_{ijt} \alpha^s_{it+1} \tag{1}$$

where $S_{jt+1}$ is the stock selection ability of fund j for the period from t to t+1, $\omega_{ijt}$ is the fund's portfolio weight on stock i at time t, $\alpha^s_{it+1}$ is the alpha, or abnormal return, of stock i from t to t+1. This definition follows the characteristic selectivity measure (CS) of Daniel, Grinblatt, Titman, and Wermers (1997), which uses characteristic-adjusted individual stock return as $\alpha^s_{it+1}$ in the above definition.

Our second assumption is that the fund stock selection skill $S_{jt+1}$ can be predicted, with noise, by information known at time t (i.e., past fund alpha or other fund characteristics that are potentially predictive of future fund stock selection ability). Let $\hat{S}_{jt}$ denote the expected stock selection skill from t to t+1 based on the information available at time t. Our assumption is, $\hat{S}_{jt} = S_{jt+1} + e_{jt+1}$, where $e_{jt+1}$ is the information noise, or an error term. Combining this with (1), we have, $\hat{S}_{jt} = \sum_{i=1}^{N} \omega_{ijt} \alpha^s_{it+1} + e_{jt+1}$.

Now, let $\hat{\boldsymbol{S}} = (\, \hat{S}_{1t} \quad \hat{S}_{2t} \quad ... \quad \hat{S}_{Mt} \,)'$, $\boldsymbol{\alpha} = (\, \alpha^s_{1t+1} \quad \alpha^s_{2t+1} \quad ... \quad \alpha^s_{Nt+1} \,)'$, and denote the error vector as $\boldsymbol{e} = (\, e_{1t+1} \quad e_{2t+1} \quad ... \quad e_{Mt+1} \,)'$. Further, let W be the M by N matrix of portfolio weights, with its (j, i) element being fund j's portfolio weight on stock i at time t, $\omega_{ijt}$. Then, the equation at the end of the previous paragraph can be written in the matrix form as

$$\hat{\boldsymbol{S}} = W\boldsymbol{\alpha} + \boldsymbol{e} \tag{2}$$

Here the time subscript is dropped for notational convenience. $\boldsymbol{e}$ is assumed to be white noise.

## 1.2 The Generalized-Inverse Alpha

Equation (2) describes the relation of stock alphas ( $\boldsymbol{\alpha}$) with observed fund skill measures, observed portfolio weights, and random error terms. Our goal is to obtain the expected value of future stock alphas, conditional on observed fund stock selection skills and portfolio weights, $E_t(\boldsymbol{\alpha}|\hat{\boldsymbol{S}}, W)$.

The frequentist approach considers $\boldsymbol{\alpha}$ in (2) as a vector of coefficients. If $W'W$ is invertible, the OLS estimator for $\boldsymbol{\alpha}$ is $\hat{\boldsymbol{\alpha}}_{OLS} = (W'W)^{-1}W'\hat{\boldsymbol{S}}$. In empirical implementation, an issue renders the OLS estimator impractical – the number of stocks (N) is usually larger than the number of funds (M). As a consequence, $W'W$, an N×N matrix, is singular and not invertible. To overcome this problem, we develop an alternative approach based on generalized inversion, which is a statistical method to deal with the singularity or near-singularity problem in regressions (Moore 1920; Penrose 1955).

Let V be the N×N matrix consisting of all N eigenvectors for $W'W$, and $D$ be the N×N diagonal matrix of eigenvalues ranked in descending order. By definition, $W'W = VDV'$. When $W'W$ is non-singular, it is known that $(W'W)^{-1} = VD^{-1}V'$. When $W'W$ is singular, some diagonal elements of $D$ are zero, and D is not invertible. Now, let K be the number of leading eigenvalues that are substantially above zero, and $d_{ii}$ be the i-th diagonal element of D. Define $\hat{D}$ as a diagonal matrix with the i-th diagonal element $\hat{d}_{ii}$, where $\hat{d}_{ii} = d_{ii}^{-1}$ if $d_{ii}$ is substantially higher than zero (i.e., i≤K) and $\hat{d}_{ii} = 0$ if $d_{ii}$ is zero or close to zero (i.e., i>K). The generalized inverse of $W'W$ is then $V\hat{D}V'$, and the generalized-inverse estimator for $\boldsymbol{\alpha}$ is

$$\hat{\boldsymbol{\alpha}}_{GI} = (V\hat{D}V')W'\hat{\boldsymbol{S}} \tag{3}$$

We call this alpha estimator the "Generalized-Inverse Alpha" (henceforth, GIA).

An implementation issue is to decide the number (K) of nonzero eigenvalues of $W'W$. Since the dimension of $W$ is M×N (M<N), among the N eigenvalues of $W'W$, at least N-M are, theoretically, zero. Additionally, within the first M eigenvalues, some may be quite small and close to zero. In our baseline analysis, we set K equal to M/2, and treat the remaining (N-K) eigenvalues as zero. We have performed a sensitivity check by varying K, and find that the results are robust except for extreme choices of K (i.e., when K is close to zero or M).

Note that one major role of $(V\hat{D}V')$ in Equation (3) is to account for the correlation in portfolio holdings across funds. It is interesting to contemplate a case where this correlation is ignored. This is equivalent to the assumption that $\hat{\boldsymbol{\alpha}} \propto W'\hat{\boldsymbol{S}}$, which leads to the "stock-quality" measure ($\hat{\alpha}_{it+1}$) of Cohen et al. (2005):

$$\hat{\alpha}_{it+1} = \frac{\sum_{j=1}^{M} \omega_{ijt}\hat{S}_{jt}}{\sum_{j=1}^{M} \omega_{ijt}} \tag{4}$$

The measure has an intuitive interpretation that a stock's alpha is the weighted average of fund alphas, where the weights are proportional to fund portfolio weights. The portfolio weight, $\omega_{ijt}$, measures the size of the "bet" (i.e., the amount of information possessed by the manager), while the past fund alpha, $\hat{\alpha}^{f}_{jt+1}$, measures the precision of private information. The GIA approach carries the same intuition, and additionally takes into account the correlation of stock holdings across funds via the term $(V\hat{D}V')$.

## 1.3 Extension: GIA Based on Lagged Portfolio Weights and Weight Changes

Mutual funds report net asset values and returns at the end of each trading day. By contrast, fund holdings are reported at a quarterly frequency, and funds are required to disclose their holdings within 60 days after their fiscal quarter ends. Therefore, typically information about quarter-$t$ fund portfolio holdings is not publicly available at the end of quarter t. To take into account this reporting delay, we consider an alternative GIA using lagged fund holdings. Let $\mathrm{W}_L$ be the quarter-$t$ fund portfolio-weight matrix based on holdings at the end of quarter $t$-1. The resulting estimator is,

$$\hat{\boldsymbol{\alpha}}_{LGI} = (V_L\hat{D}_L V_L')W_L'\hat{\boldsymbol{S}} \tag{5}$$

where $V_L$ and $\hat{D}_L$ are counterparts of $V$ and $\hat{D}$ in (3) but defined on $W_L'W_L$ instead of $W'W$. Note that since there is essentially no delay in fund return reporting, there is no time lag in estimated fund alphas $\hat{\boldsymbol{S}}$. The resulting estimator, $\hat{\boldsymbol{\alpha}}_{LGI}$, (henceforth, LGIA), is fully feasible to implement using information publicly available at the end of quarter $t$.[3]

We further quantify the return-predictive information revealed by recent fund *trades* using two

[3]We thank Craig MacKinlay for this insight.

trade-based GIA estimators. They result from replacing fund portfolio weights $W$ in Equation (2) with either $\Delta W^+$ or $\Delta W^-$. $\Delta W^+$ is the matrix of positive portfolio weight changes from t-k to t (i.e, weight changes due to recent mutual fund buys), and $\Delta W^-$ is the matrix of negative portfolio weight changes (i.e., weight changes due to recent fund sells). The two trade-based GIA estimators are:

$$\hat{\boldsymbol{\alpha}}_{BUY} = V^+ \hat{D}^+ (V^+)' (\Delta W^+)' \hat{\boldsymbol{S}} \tag{6}$$

$$\hat{\boldsymbol{\alpha}}_{SELL} = V^- \hat{D}^- (V^-)' (\Delta W^-)' \hat{\boldsymbol{S}} \tag{7}$$

where $V^+$ and $\hat{D}^+$ ($V^-$ and $\hat{D}^-$) are the counterparts of $V$ and $\hat{D}$ in (3), but defined on $(\Delta W^+)'\Delta W^+$ ($(\Delta W^-)'\Delta W^-$), respectively.

# 2. Data and Methods

## 2.1 Data

Mutual fund data are from two sources – Thomson Reuters and CRSP. Funds in these two databases are linked using the MFLINKS data set (available from Wharton Research Data Services, WRDS). The sample period for our study starts in 1980Q1 and ends in 2006Q4. We start with domestic equity funds with investment objectives of aggressive growth, growth, or growth and income in the Thomson data, and then exclude index funds from the sample. However, we find some errors in the Thomson-reported investment objectives. Judged by fund names and by the investment styles from the CRSP data, some funds with Thomson-reported investment objectives in the above three categories are foreign-domiciled funds, US-domiciled international funds, fixed-income funds, real estate funds, precious metal funds, balanced funds, closed-end funds, or variable annuities. We take additional steps to manually screen out such funds that have misreported investment objectives. For the CRSP fund data, we combine different share classes of the same fund, so that monthly fund returns are computed as the weighted-average returns across share classes, with weights proportional to the beginning-of-month total net assets of each share class. In addition, we obtain stock returns from CRSP, financial statement variables from Compustat, and analyst earnings forecasts

from IBES.

Table 1 provides summary statistics for the mutual fund sample. We report the number of funds and the number of stocks that they hold, in aggregate, at the end of each year from 1980 to 2006. For funds that do not report portfolio holdings at year-end, we use their last reported portfolio snapshot during the second half of the year, and assume they hold this portfolio until year end (on a split-adjusted basis). The table shows that there are 284 actively managed domestic equity funds in our sample in 1980. The median fund holds 56 stocks, purchases 17 stocks and sells 18 stocks during the final six month period of 1980, with an average annual turnover ratio of 71%. They collectively hold 2,102 unique common stocks, accounting for 44.47% of all common stocks in the CRSP universe. The market value of the aggregate fund equity holdings is $33 billion, accounting for 2.54% of the total market capitalization of common stocks in the CRSP universe. The number of funds, the number of unique stocks held by these funds, and the market value of their equity holdings increase quickly during the sample period, except for the last several years. By 2006, there are 1,820 actively managed domestic equity funds in the sample. The median fund holds 91 stocks, purchases 40 stocks and sells 40 stocks during the final six month period of 2006, with an average annual turnover ratio of 92%. The funds collectively hold 4,129 unique common stocks, accounting for 87.02% of common stocks in the CRSP universe. The aggregate market value of fund equity holdings is $2.2 trillion, accounting for 13.01% of the total market capitalization of common stocks in the CRSP universe.

During each year, the number of funds is typically much lower than the number of stocks in the sample. During the sample years, the number of funds fluctuates between 270 and 2001, while the number of distinct common stocks these funds hold collectively ranges from 2,102 to 6,260.

## 2.2 Portfolio Weights and Weight Changes

Since our interest is in the persistent stock selection ability of mutual fund managers, we focus on the weights of stocks in the equity portion of a fund portfolio.[4] The fund portfolio weight is

[4]The results are robust when we define portfolio weights as the value of stock holdings divided by fund TNA.

computed as:

$$\omega_{ijt} = \frac{s_{ijt} p_{it}}{\sum_{i=1}^{N} s_{ijt} p_{it}} \tag{8}$$

where $s_{ijt}$ is the number of shares of stock i held by fund j in quarter t, and $p_{it}$ is the price of stock i at the end of quarter t. Note that funds report holdings for their fiscal quarter-ends, which may not coincide with calendar quarter-ends. In such a case, we assume that fund holdings reported at a fiscal quarter-end are valid for the immediately following calendar quarter-end, after adjusting for stock splits using the CRSP share adjustment factor. In addition, some funds in the sample report holdings semiannually instead of quarterly. For these funds, if they do not report holdings at the end of a given quarter $t$, we use their reported holdings at the end of quarter $t-1$. We find that including these lagged fund holdings, rather than omitting them, improves the return-predictive power of the resulting GIA estimator by increasing the size of the fund cross-section.

We compute the fund portfolio weight change, due to active trading as:

$$\Delta\omega_{ijt} = \frac{(s_{ijt} - s_{ijt-2}) p_{it}}{\sum_{i=1}^{N} s_{ijt} p_{it}} \tag{9}$$

where $s_{ijt-2}$ is the number of shares of stock $i$ held by the fund in quarter $t-2$. To make $\Delta\omega_{ijt}$ invariant to stock splits, we adjust the lagged holding, $s_{jit-2}$, using the share adjustment factor from CRSP to an equivalent shareholding at the end of quarter $t$. We compute the weight change over a two-quarter period instead of a one-quarter period so that funds reporting holdings semiannually can be seamlessly included. For funds that do not report holdings at the end of quarter $t$, we use holding changes from quarter $t-3$ to quarter $t-1$ as a proxy for holding changes from $t-2$ to $t$. Note that we have also performed an analysis using quarterly portfolio weight changes and have found similar results.

## 2.3 Measuring Fund Stock Selection Skills

In the baseline GIA model, we measure fund stock selection skills by the estimated alpha from the Carhart (1997) four-factor model:

$$r_t - r_{ft} = \alpha + \beta \cdot (r_{mt} - r_{ft}) + s \cdot \text{SMB}_t + h \cdot \text{HML}_t + u \cdot \text{UMD}_t + e_t \tag{10}$$

where the *pre-expense* monthly fund return $r_t$ is the net fund return plus $\frac{1}{12}$ times the annual expense ratio. The riskfree rate $r_{ft}$ is the yield on treasury bills with one-month maturity at the beginning of month t, obtained from CRSP. The market return $r_{mt}$ is the monthly CRSP value-weighted index return, while $\text{SMB}_t$, $\text{HML}_t$, and $\text{UMD}_t$ are the monthly size, book-to-market, and momentum factors obtained from Ken French's website. The regression is performed at the end of each calendar quarter on a rolling basis, using data for the prior 12 months.

We note that the four-factor alpha may capture skills of funds other than stock selection. For example, some funds may possess skills to time their trades and execute trades at low trading costs (Kacperczyk et al. 2008; Puckett and Yan 2011). Such skills can contribute to fund alphas but are not considered part of the stock selection skills. Therefore, the fund alpha is a noisy proxy for fund stock selection skills (as are any other proxies for fund stock selection measures we consider in this paper). If the fund alpha has nothing to do with the ability of fund managers to locate mispriced stocks, or if such skills do not persist from one year to the next, then by construction the resulting GIA will not predict future abnormal stock returns. Given the above, the validity of our baseline version of GIA depends on the persistence of stock selection skills over time. This assumption can be made concrete by considering the following process for fund manager stock selection skills: $S_{jt+1} = \rho\hat{\alpha}^F_{jt} + e_{jt+1}$, where $\hat{\alpha}^F_{jt}$ is the past four-factor alpha (assuming that $\hat{\alpha}^F_{jt}$, $e_{jt+1}$, and $S_{jt+1}$ have zero expected values). The GIA becomes $\hat{\boldsymbol{\alpha}}_{GI} = (V\hat{D}V')W'\rho\hat{\boldsymbol{\alpha}}^F_t$, where $\hat{\boldsymbol{\alpha}}^F_t$ is the vector of $\hat{\alpha}^F_{jt}$. Note that if the persistence parameter $\rho$ equals zero, all stocks are predicted to have the same alpha of zero.[5]

Mamaysky, Spiegel, and Zhang (2007) propose a backtesting procedure that substantially improves the power of the estimated fund alphas in predicting future fund performance. However this procedure would also result in a substantial reduction of fund sample size, which is crucial for the power of the GIA model. We therefore do not implement this procedure in our paper.

[5]In addition, $\rho$ could be different across funds. In untabulated analysis, we perform quarterly cross-sectional regressions of fund 12-month four-factor alpha on lagged 12-month four-factor alpha as well as the interaction of the lagged alpha with seven fund-level characteristics: the inverse of residual fund volatility, dummy variables for top- and bottom-decile fund alpha ranks, percentile rank of fund TNA, fees, turnover, and age. Also included are the levels of these seven characteristics. We use the estimated coefficients to form conditional fund persistence parameters, which are then used to compute "conditional GIAs." These GIAs do not forecast stock alphas significantly better than the unconditional approach to be presented in the next section.

## 2.4 Evaluating Return-predictive Performance

We rely on both the sorted-portfolio approach and Fama-MacBeth regressions to evaluate the performance of the GIA estimators. The details of the portfolio approach are as follows.

During the period from 1980 to 2006, at the end of each portfolio formation quarter (Q0), we sort stocks into equal-weighted decile portfolios based on one of the stock alpha estimators developed in Section 1, and hold them during the subsequent four quarters (Q1 to Q4, the holding quarters). During the holding period, we rebalance the portfolios quarterly, so that they have equal weights at the beginning of each holding quarter. To avoid market microstructure issues in measuring stock returns as well as to limit the impact of transaction costs, we require a stock to have a minimum price of \$5 at the beginning of a holding quarter to be included in any decile portfolio for that quarter.

In order to provide a summary performance measure over the entire four-quarter holding period, we further adopt an overlapping portfolio approach. In any given quarter, there are four portfolios with the same decile ranking, but formed during each of the prior four quarters. We further combine these four portfolios in equal weights into a single portfolio and hold it during the next quarter. This portfolio formation procedure is similar to the overlapping momentum portfolio procedure of Jegadeesh and Titman (1993), and is labeled "JT4" in the tables.

We evaluate portfolio performance using portfolio returns, characteristic-adjusted portfolio returns, and Carhart four-factor alphas. When computing the portfolio returns, we include CRSP-reported delisting returns for stocks delisted during a holding quarter. Following Shumway (1997), when the CRSP delisting return is missing, we replace it with -30% if the delisting is performance related, and zero, otherwise.

The characteristic-adjusted return is computed using the characteristic benchmark approach of Daniel, Grinblatt, Titman, and Wermers (1997; DGTW) and Wermers (2003).[6] In each quarter we sort all common stocks with price no less than \$5 into 125 ($5\times5\times5$) benchmark portfolios using a sequential triple-sorting procedure based on size, book-to-market ratio (BM), and momentum.

[6]The DGTW benchmarks are available via http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm. This paper uses a modified version of the DGTW approach, where the benchmark portfolios are updated quarterly instead of annually to provide a more precise control for both active and passive momentum returns achieved by the funds.

Size is the market cap at the end of the quarter (using NYSE breakpoints when sorting). BM is computed using the book value of equity for the most recently reported fiscal year and the quarter-end market cap. Momentum is the 12-month return prior to the quarter-end. An additional group is created to include stocks that cannot be assigned into any of the above 125 groups due to missing characteristics. The characteristic-adjusted return for a stock is its quarterly return in excess of the respective quarterly value-weighted benchmark portfolio return.

Finally, the four-factor alpha for a ranked portfolio is estimated using Equation (10), based on quarterly portfolio returns. The monthly factor returns obtained from Ken French's website are compounded into quarterly returns before they are used in the regression. When interpreting empirical results based on regression alphas, we keep in mind two known issues. First, the factor-regression approach may lack power to detect abnormal returns (e.g., DGTW 1997; Kothari and Warner, 2001). Second, Cremers, Petajisto, and Zitzewitz (2009) point out that the regression results may be sensitive to factor construction.

# 3. Main Empirical Results

## 3.1 GIA Based on Current Fund Holdings

Table 2 reports the characteristics and performance of equal-weighted portfolios formed on the GIA defined in (3). As described in Section 1, in the baseline case, we set K=M/2, and use a rolling (lagged) 12-month four-factor alpha to measure fund ability. Fund portfolio weights are measured at the end of the portfolio formation quarter, Q0.

The portfolio characteristics reported in the table are the cross-sectional average quintile ranks of stocks in the size, book-to-market ratio, and momentum dimensions (with 1 being the lowest and 5 the highest quintile). These quintile ranks are those used to construct the DGTW characteristic benchmarks. The table shows that stocks in both the top and bottom GIA-ranked deciles (D1 and D10) have larger market capitalizations relative to stocks in the middle deciles. Stocks in the two extreme deciles also tend to have slightly lower book-to-market ratios. In addition, there is a generally monotonic relation between the portfolio decile rank and momentum quintile rank, with the exception of the two bottom deciles. Since these stock characteristics are predictive of

stock returns, it is important to control for them in order to properly measure abnormal portfolio performance.

During the quarter (Q1) after portfolio formation, portfolio returns generally increase with their decile ranks, although not in a perfectly monotonic way. The D1 portfolio return is 3.02%, while that of D10 is 4.55%. The return difference is 1.53%, with a $t$-statistic of 5.11. The return gap between D10 and D1 remains significantly positive for the next two quarters, Q2 and Q3, but becomes insignificant during Q4. For the JT4 portfolio (which takes positions in four portfolios formed during four past quarters), the top-bottom return spread is 1.10%, with a $t$-statistic of 3.99. The decreasing D10-D1 return spreads as we move from Q1 to Q4 suggest that the return-predictive information possessed by fund managers with persistent skills is relatively short-lived.

The two performance measures that control for stock characteristics, the characteristic-adjusted return and the four-factor alpha, also tend to be higher for D10 stocks than for D1 stocks. For example, the characteristic-adjusted return spreads between D10 and D1 are 1.14%, 0.83%, and 0.53% from Q1 to Q3, and 0.65% for the JT4 portfolio, all highly significant. The four-factor alpha spreads between D10 and D1 are 1.15%, 1.00%, 0.73% from Q1 to Q3, and 0.86% for the JT4 portfolio. Note that in terms of the magnitude, the four-factor alpha spreads are comparable with or slightly higher than the characteristic-adjusted return spreads. However, the t-statistics are lower, possibly due to the lower statistical power for regression-based performance measures relative to portfolio-matching techniques (as discussed in Section 2.4).

The high statistical significance for the D10-D1 performance spreads suggests a high economic value of the return-predictive information, as measured by the Sharpe ratio. The $t$-statistic for the JT4 return spread of GIA portfolios is 3.99 during the 111 holding quarters (1980Q2 to 2007Q4), implying an annualized Sharpe ratio of 0.76 (=3.99/$\sqrt{111/4}$). By comparison, during the same period, the return spread between the top and bottom momentum decile portfolios of stocks, formed in JT4 style (i.e., based on past 12-month returns and skipping one month between formation and holding periods, following Jegadeesh and Titman (1993)), is 1.68%, with a t-statistic of 2.09 and an annualized Sharpe ratio of 0.40.

In sum, the GIA exhibits statistically and economically significant power to predict stock returns. This is consistent with a strong persistence in stock selection ability across funds.

## 3.2 GIA Based on Lagged Fund Holdings and Fund Trades

Due to reporting delays, the GIA strategy analyzed in Table 2 is not fully implementable in practice. Therefore we turn to the estimator (5), which is based on lagged fund holdings. Except for the use of lagged portfolio holdings, our procedure remains the same as that of Table 2.

Panel A of Table 3 reports the results. To save space we only report the D10-D1 performance differences, in terms of return spreads, the corresponding characteristics-adjusted spreads and the four-factor alphas of the spreads. The patterns of the portfolio performance are similar to those in Table 2, although with slightly lower magnitude. During Q1, the D10-D1 differences in net returns, characteristic-adjusted returns, and four-factor alphas are 1.52%, 0.95%, and 0.99% respectively, all significantly positive. The corresponding D10-D1 spreads for the JT4 portfolios are 1.00%, 0.48%, and 0.75%, also significant. Therefore, even with the delayed observation of fund holdings in practice, the return-predictive information produced by our model is highly valuable to investors.

The performance of GIA estimator based on fund buys (6) is reported in Panel B of Table 3. Again, to save space, we only report the D10-D1 performance differences. The spread in net returns, characteristic-adjusted returns, and four-factor alphas are significantly positive for Q1, Q2, and for the JT4 portfolios. They are not significant for Q3 and Q4, though. Even when the performance spreads are significant, they are generally lower in magnitude and have lower $t$-statistics relative to the corresponding numbers in Table 2, which are based on fund holdings.

There are several factors that affect the relative performance of holding- vs. trade-based GIAs. First, recent fund trades contain fresh information about stock values, whereas fund holdings may contain some stale information. Second, funds may make short-term trades for non-performance reasons, such as accommodating flows. A third effect is that, as reported in Table 1, the number of stocks held by funds is twice that of the stocks bought or sold by a fund. The smaller sample of fund buys may also limit the power of the buy-based GIA. Our results indicate that the advantage of fresh information contained in fund trades does not offset the two negative effects.

As reported in Panel C of Table 3, the GIA based on fund sales (7), fares much worse in predicting returns. Under all three portfolio performance measures, and across all holding horizons (Q1 to Q4, and JT4), there are no significant performance differences between D10 and D1 portfolios, with some performance differences being negative.

In addition to the fact that the number of fund sell observations is much lower than the number of fund holdings, there are three potential reasons why the sell-based signals perform poorly. First, most mutual funds have a self-imposed constraint on short-selling, which may limit the information revealed by selling of stocks. Second, a skilled fund manager buys stocks that are temporarily undervalued. Assuming the fund manager has talent, these stocks will subsequently be sold after a price run-up, and thus will no longer have prospects for further positive alphas. If true, then we cannot look to stock sales by skilled managers as a signal of these stocks being overpriced in an absolute sense. Finally, some funds may be forced to sell stocks due to large money outflows. Such forced sells are typically not related to fund managers' stock selection information (e.g., Alexander, Cici, and Gibson, 2007).

## 3.3 Robustness and Variations

### 3.3.1 Subperiod and Subsample Analysis

Figure 1 plots the performance of GIA and LGIA in nine subperiods that span the entire 27-year sample. The subperiods are based on the portfolio formation quarters; i.e., each three-year subperiod consists of 12 portfolio formation quarters occurring within the three-year span. The first subperiod is 1980-1982 and the last subperiod is 2004-2006. The performance measure here is the D10-D1 spread in characteristic-adjusted returns, for portfolios formed on GIA and LGIA respectively, using the JT4 approach. The plot shows that with the exception of the first subperiod (1980-1982), GIA and LGIA generate positive performance during all subperiods. The performance during the 7th subperiod (1998-2000, during the internet bubble) is notably strong. However, in untabulated analysis, we find that, after removing this subperiod, the performance during the remaining sample period is still significantly positive. In addition, we have performed a time-trend analysis by regressing the characteristic-adjusted JT4 return spreads onto a consecutively-increasing quarter index (taking a value of 1 for 1980Q1 and a value of 108 for 2006Q4). Using the data during the entire sample period from 1980 to 2006, the coefficients for the quarter index are 0.0046 (t=1.70) and 0.0048 (t=1.30) for GIA and LGIA respectively. During the subperiod from 1990 to 2006, the coefficients for the quarter index are -0.0089 (t=-0.48) and -0.0024 (t=-0.24) for

GIA and LGIA respectively. Thus, although there is a slight improvement of performance during the entire sample period and a slight deterioration of performance during the subperiod of 1990 to 2006, such time trends tend to be statistically insignificant.[7]

Recent studies have found improved stock market market efficiency over time. For example, Chordia, Subrahmanyam, and Tong (2010) show that increased liquidity has substantially weakened the magnitude of various stock market anomalies. Interestingly, we do not detect a visible declining trend of performance for the GIA and LGIA strategies. However, we note that the investment strategies analyzed by Chordia et al. (2010) are based on publicly available information that can be obtained at a relatively low cost. Thus the profitability of such strategies may be vulnerable to increased competition accompanied by improved liquidity in the financial markets. As discussed further in Section 4.3 of the paper, the stock selection information extracted using our approach is likely the result of fundamental research by fund managers and such information is of a different nature than the return-predictive information produced by easily exploited market anomalies. Further, to the extent that fundamental-based information is costly to produce, we would expect that fund managers producing such costly information will still be rewarded, even in a competitive market (Grossman and Stiglitz 1980).

Another factor that affects the time trend in the performance of GIA and LGIA is the increased fund sample size. Over time as more funds enter our sample, the power of the generalized inverse approach improves, which may somewhat offset a decline in the intensity of stock selection information each fund manager possesses. This can be easily understood because of the increased number of rows (M) in the MxN weight matrix of Equation (2). The GIA is able to extract more information from fund weights as this matrix becomes closer to a full-rank square matrix. The increase in model power due to an enlarged fund sample size may also reconcile our findings with Barras, Scaillet, and Wermers (2010), who report a shrinking proportion of funds with positive pre-fee alphas but increasing *numbers* of total funds and skilled funds.[8] In addition, the GIA model

[7] In addition, to see if a time trend exists but is confounded by the internet bubble period, we further compare the performance during the pre-bubble period of 1990-1997 with the post-bubble period of 2002 to 2006. We find that the performance during these two periods is at comparable magnitude, for GIA as well as for LGIA.

[8] ecifically, according to their Internet Appendix, based on pre-expense fund alphas, the proportion of skilled funds decreases from around 28% in 1989 to 9.6% in 2006; meanwhile, the number of funds in their sample increases from around 400 in 1989 to 1836 in 2006. Thus, the entire fund cross-section has dramatically expanded, and the number of skilled funds also increases (from 112 to 176).

draws its power from the dispersion of fund stock selection skills, not just the proportion of funds with positive alphas. Specifically, the growth in the numbers of other (unskilled) managers also helps, since the weight matrix makes use of the holdings of unskilled managers, too.

We also examine the performance of the GIA in subsamples of stocks formed based on various stock characteristics. For brevity the results are not tabulated but we summarize them here. We find that, based on the JT4 spread, the GIA exhibits strong return-predictive power among small-cap and mid-cap stocks; however, its return predictive power among large-cap stocks remains significant. Further, based on the JT4 spread, the predictive power of GIA is stronger among value stocks, past losers, and stocks with high idiosyncratic return volatility, relative to growth stocks, past winners, and stocks with low idiosyncratic return volatility. Finally, the GIA has the strongest predictive power among stocks with a medium level of breadth of fund ownership (measured by the number of funds holding the stock, orthogonalized with firm size). Intuitively, this might be due to a combination of the following two offsetting effects. On the one hand, the GIA approach does not have much statistical power for stocks owned by very few funds. On the other hand, the return-predictive information may be quickly competed away and reflected in stock prices when too many funds own the stocks.

### 3.3.2 Variations

We have performed additional analyses to ensure the robustness of the results and to deepen our understanding on the properties of the GIAs. For brevity we summarize the results below without tabulating them.

We first check the impact of fund alpha estimation on the performance of the GIAs. In addition to the Carhart four-factor model, we use the CAPM and the Fama-French three-factor model to estimate fund alphas. The resulting GIAs continue to exhibit significant return predictive power; however, since the CAPM and three-factor models do not control for the price momentum effect, the resulting GIAs tend to generate short-term momentum-like patterns. Moreover, we extend the rolling window for fund alpha estimation from 12 months to 60 months, at increments of 12 months. Regardless of the factor model used, as we extend the rolling window, the return predictive performance of the resulting version of the GIA weakens. This indicates that the stock picking

ability of fund managers is fleeting, consistent with the theoretical prediction of Berk and Green (2004) and the empirical findings of several studies (Carhart 1997; DGTW; Hendricks, Patel, and Zeckhauser 1993; and Kosowski, Timmermann, Wermers, and White 2006).

Second, we implement a benchmark-adjusted version of GIA, by replacing portfolio weights W with the portfolio weights in excess of the weights of stocks in the CRSP value-weighted index in the GIA implementation. We find slightly improved results relative to those reported in Table 2 in terms of the statistical significance of the top-bottom decile return spreads. However, the return spreads per se are slightly smaller. Therefore, the improvement due to benchmark adjustment is not clear-cut.

Third, we evaluate the performance of GIAs at a monthly frequency instead of quarterly frequency, and obtain similar results. The GIAs significantly predict stock returns over horizons of five to seven months after portfolio formation, depending on performance measures. In addition, the Jegadeesh and Titman (1993) style 12-month overlapping portfolios produce significantly positive top-bottom decile return spread in net returns, characteristic-adjusted returns as well as four-factor alphas during the holding period.

Fourth, we use the liquidity-augmented five-factor model to evaluate the performance of the GIAs. The fifth-factor (in addition to the Carhart four factors) is either the Pastor and Stambaugh (2003) liquidity factor or the Sadka (2006) liquidity factor. The return predictive power of the GIAs remains significant with the additional control for the exposure to the liquidity factor. In other words, the performance of the GIAs is not sensitive to the liquidity-provision or liquidity-demanding role of mutual funds.

Finally, Gaspar, Massa, and Matos (2005) document that some funds receive favorite IPO allocations from fund families and such allocations substantially benefit their performance. In addition, Cohen, Polk, and Silli (2010) show that stocks held with the largest portfolio weights (i.e., the "best-ideas") by funds tend to outperform other stocks. To see how important the IPO allocations and the "best-idea" stocks is to the performance of the GIAs, we exclude stocks with less than 1 year of history in the CRSP data, and exclude the "best-idea" stocks when constructing the GIAs. The resulting performance of the GIAs is robust to these variations. Therefore, favorable IPO allocations received by funds are not a key contributor to the performance of the GIAs, and

that the GIAs can extract stock return information effectively on stocks that are not the "best ideas" of fund portfolios.

# 4. Further Analysis

In this section, we take advantage of the GIA approach to evaluate four issues on the stock selection information possessed by fund managers.

## 4.1 Comparison with Other Return-predictive Signals from Mutual Fund Holdings

Existing studies have discovered several other ways to use fund holdings or trades to predict stock returns. We compare the return-predictive power of the GIAs with four such effects.

The first is the effect of herding. Although early studies such as Lakonishok, Shleifer, and Vishny (1992) and Grinblatt, Titman, and Wermers (1995) find little evidence of institutional herding, Wermers (1999) and Brown, Wei, and Wermers (2011) document a significant price impact by mutual fund herds. Wermers (1999) reports return continuation following mutual fund herding trades, while Brown at al. (2011) find that, in recent periods, herding trades generate initial return continuation followed by return reversal. To control for the effect of herding we follow Brown et al. (2011) to construct a variable HERD. We first construct a buy herding measure, BHM, for stocks that mutual funds "buy-herd," and a sell herding measure, SHM, for stocks that funds "sell-herd." The details of these two measures can be found in their paper. We then rank buy-herded stocks (by BHM) into quintiles, and separately rank sell-herded stocks (by SHM) into quintiles. Finally, we stack the buy-herding quintiles and sell-herding quintiles into deciles (denoted as HERD), with the stocks having the largest buy herding measure, BHM, in the top decile, and those with the top sell herding measure, SHM, in the bottom decile.

The second effect is the aggregate stock-picking ability by mutual funds vs other investors. Chen, Jegadeesh, and Wermers (2000) document that aggregate mutual fund trades have significant power to predict stock returns. They argue that this is because mutual funds, on average, are better stock-pickers than unsophisticated individual investors. Following their paper, we measure the

aggregate fractional mutual fund trades (TRADE) as the one-quarter change in total mutual fund holdings (in dollars) of a stock divided by the market capitalization of that stock.

The third effect is related to short-sale constraints. Chen, Hong, and Stein (2002) find that a decrease in the number of mutual funds holding a stock (reduced breadth of ownership) is associated with a lower future return for that stock, as the negative outlook of many funds is not fully expressed through their portfolio holdings due to the self-imposed short-sale constraint of most funds. Following their study, we measure the change in breadth of ownership of a stock ($\Delta$BRD) as the one-quarter change in the number of mutual funds who hold a long position in the stock divided by the total number of mutual funds that exist during both the formation quarter and the prior quarter.

The fourth is the "dumb money" effect of fund flows documented by Frazzini and Lamont (2008). They find that mutual fund flows are largely driven by investor sentiment, and high sentiment stocks tend to generate low returns at long horizon(s). To capture this effect, we follow Equation (8) of their paper to construct a measure, FLOW. Fund flows used in measuring the variable are over the portfolio formation quarter, Q0.[9]

The performance of equal-weighted decile portfolios formed in HERD, TRADE, $\Delta$BRD, and FLOW is reported in Panel A of Table 4. For brevity, we only report the spreads in characteristic-adjusted returns between the top and bottom decile portfolios. A common pattern across the four variables is that their top-bottom return spreads decrease monotonically as the holding period moves from Q1 to Q4 (with the exception of $\Delta$BRD during Q3), and the spreads are positive during the first two quarters, Q1 and Q2, but turn negative during Q3 and Q4 (again with the exception of $\Delta$BRD during Q3). Therefore, herding, aggregate fund trades, change in breadth of ownership, and fund flows all generate short-term return continuation that is subsequently at least partially reversed. Consequently, only $\Delta$BRD produces a significant spread for the JT4 portfolio. These results are consistent with those separately reported for each variable by various studies.

We further compare the return-predictive power of the GIAs with these variables using Fama-

[9]We measure fund flows during a quarter to capture the return continuation effect induced by short-term autocorrelated fund flows. By contrast, Frazzini and Lamont (2008) use fund flows at longer horizons (e.g., 1 to 3 years), in order to capture the return reversal effect. However, such long-run reversal is unlikely to explain the short-term return-predictive power of the GIA.

MacBeth regressions. The successive dependent variables in the four cross-sectional regressions are the characteristic-adjusted stock returns during Q1, Q2, Q3, and Q4. The regressors include GIA (or LGIA), HERD, TRADE, $\Delta$BRD, and FLOW. The regressions are performed quarterly, and we compute the time-series averages of the estimated coefficients. To obtain a summary measure of predictive performance over the four quarters, we also compute JT4 regression coefficients that are similar to the JT4 portfolio procedure. Specifically, during each quarter t, there are four sets of cross-sectional regressions performed which use the characteristic-adjusted stock return during that quarter as the dependent variable. These four sets of cross-sectional regressions involve explanatory variables measured during the previous four quarters (t-4 to t-1). We first average the four sets of coefficients on the same explanatory variable across the four lags, and then further compute their time-series averages. Finally, $t$-statistics are computed using a Newey-West covariance estimator with a two-quarter lag.

The regression results are reported in Panel B of Table 4. Note that, even in the presence of HERD, TRADE, $\Delta$BRD, and FLOW, the coefficients for GIA (LGIA) are significantly positive during the first three (two) quarters after Q0, and for the JT4-style regression. Further, the coefficients for GIA and LGIA decline from Q1 to Q4, consistent with the pattern on their top-bottom decile spreads reported in Table 2 and Panel A of Table 3. Overall, the results suggest that the GIAs reflect an aspect of fund managers' stock selection information that is different from that captured by the effects already examined in existing studies.

## 4.2 Fundamental Information Contained in GIAs

Next, we delve into a deeper issue regarding the economic source of the return-predictive power of the GIAs, or equivalently, the stock selection information implied by the cross-sectional difference in fund skills. One conjecture is that the GIAs are related to future stock returns due to mere price pressure mechanisms. The pure price pressure effect could arise in various forms. For example, mutual funds tend to herd on (mimic) the prior trades of top-performing funds (Sias 2004). As mentioned in the prior section, Wermers (1999) documents that fund herding moves stock prices. In our context, herding might push up the prices of stocks held by winning funds even when there is no private information conveyed by winning fund trades or subsequent herding trades. The second

type of price pressure is induced by fund flows. It is well-documented that fund flows chase past performance (e.g., Sirri and Tufano 1998). If winning funds respond to new money inflows by adding to already-held stock positions, they may push up the prices of these stocks. Wermers (2003) finds some evidence of price-pressure in stock trades that are motivated by fund flows.

An alternative explanation is that the GIAs predict stock returns because they aggregate unique fundamental information produced by fund managers. Most mutual funds claim to pick stocks via fundamental research, a process that may result in the production of non-public information about the business fundamentals (e.g., sales and earnings) of firms. Perhaps because fund managers' fundamental research efforts are not directly observed, the effectiveness of fundamental analysis has, to date, rarely been scrutinized by academic studies. However, if fund managers do possess fundamental information about stocks, we expect that it should result in a significant relation between the GIAs and measures of firms' future operating performance.

Our analysis in Section 4.1 already provides some information to evaluate the price pressure hypothesis. The variables HERD and FLOW capture specific forms of price pressure due to fund herding and sentiment-driven fund flows. One can also argue that the relation between TRADE and subsequent stock returns may be in part due to the effect of price pressure. However, none of these variables explains the return-predictive power of the GIAs. In addition, we check the long-run performance of the stock portfolios. If the Q1 through Q3 price increases of high GIA (or LGIA) stocks are due to mere price pressure, their performance would reverse at longer horizons. In untabulated analysis, we do not find any significant reversal when we trace the performance up to three years after portfolio formation.

Further, we provide direct evidence that the GIAs contain predictive information about firms' fundamentals. We measure information shocks to corporate fundamentals using four variables – standardized unexpected earnings (SUE), earnings surprise relative to consensus forecast (SUR), analyst forecast revision (FRV), and earnings-announcement-window return (EAR). SUE is defined in Appendix A. SUR is the quarterly EPS reported during a given quarter minus the last consensus forecast for that EPS prior to the earnings announcement. FRV is the consensus EPS forecast for the (as yet unreported) current fiscal year measured during the last month of the given quarter, minus the consensus EPS forecast for the same fiscal year measured three months ago,

then scaled by the IBES-reported stock price during month m. EAR is the buy-and-hold return during the five trading days (day -2 to day 2) around the date of a given earnings announcement. If earnings are announced after the market close or during a non-trading day, we treat the next immediate trading day as the announcement date. SUE and EAR are based on data from Compustat, while SUR and FRV are based on IBES data. To alleviate the effect of outliers, in each quarter, we cross-sectionally winsorize SUE, SUR and FRV at the top and bottom 1 percentile before using them in our analysis. Note that these four variables measure unexpected information about firms' fundamentals from different perspectives. SUE and SUR measures earnings surprises relative to past earnings and relative to analysts' forecasts. FRV measures changes in earnings expectations throughout a quarter, presumably due to new information arrival during the quarter. EAR measures the magnitude of investors' earnings surprises in terms of stock returns.

Table 5 reports the averages of the four fundamental information measures for each decile portfolio sorted on GIA and LGIA, respectively, during the four quarters (Q1 to Q4) after portfolio formation. We also compute the JT4-style averages for these portfolio-level fundamental-information measures. For brevity we only report the D10-D1 differences in these measures. The table shows that both GIA and LGIA have significant power to predict firms' fundamentals. All D10-D1 spreads in the fundamental measures are significantly positive, with the only exception being the EAR spreads for Q4. The predictive power of the GIA and LGIA degrades through time, matching the degrading return spread of D10-D1 that we document earlier. In untabulated analysis, we obtain similar results using the four fundamental information measures in excess of those of their DGTW benchmarks.

Overall, the evidence is consistent with the hypothesis that the stock selection information contained in the GIAs reflects the persistent ability by mutual funds to perform research on corporate fundamentals, rather than being driven by pure price pressure.

## 4.3 Public or Private Information? Relation with Quantitative Signals

Given our evidence in the last section on the fundamental information revealed by the GIAs, a follow-up issue is the relation between such information and the return-predictive information contained in various publicly available firm characteristics. A large body of academic studies has

found that stock returns are predictable based on firm-specific financial and accounting variables, which are sometimes referred to as "quantitative characteristics" or "quantitative investment signals" (e.g, Jegadeesh et al., 2004). There is evidence that mutual funds trade on at least some of these variables, such as price momentum (e.g., Grinblatt, Titman, and Wermers 1995). One might question whether the return-predictive information contained in the GIAs is truly private in nature. For instance, if top performing managers simply trade on previously documented market anomalies using publicly available financial information, then we might not conclude that our approach truly uncovers fund managers' private information on stocks. Therefore, we next examine the relation of the GIA signals with prior-documented stock return anomalies.

We consider 11 quantitative investment signals documented in the prior literature. These signals are used by Jegadeesh et al. (2004) to assess the value of analyst stock recommendations. We follow their definitions of these variables, and provide a detailed description of the variables in Appendix A. The first two variables are related to liquidity. SIZE is the log of stock market capitalization and TURN is the exchange-specific percentile ranking of stock trading turnover. The next four variables are momentum signals. RETP and RET2P measure price momentum by stock returns during months -6 through -1 and months -12 through -7, respectively. FREV and SUE measure earnings momentum of stocks using analysts' EPS forecast revision and standardized unexpected earnings. There are three contrarian or value signals. EP is the earnings-to-price ratio, while BP is the log book-to-market ratio. SG is the past sales growth rate. Further, the earnings quality indicator, ACC, is the total accounting accruals. Finally, a measure of corporate investment, CAPEX, is capital expenditure intensity. All the 11 variables are cross-sectiontally winsorized at the top and bottom 1 percentile. We further standardize these variables by subtracting the cross-sectional means and then dividing by cross-sectional standard deviations.

Based on untabulated Fama-MacBeth regressions, we find that all these variables exhibit a significant ability to predict returns during at least one of the four evaluation quarters (Q1 to Q4) during the sample period, and the signs of the estimated coefficients are generally consistent with those documented in previous studies. We further examine the correlation between the GIAs and the quantitative investment signals (again, untabulated). The GIAs are positively correlated with SIZE, but have an insignificantly negative correlation with the second liquidity variable, TURN.

They also have a significantly negative correlation with BP, but are insignificantly correlated with other signals on value, earnings quality, or corporate investments. However, the GIAs have an apparent momentum tilt, as they have significantly positive correlations with RETP, RET2P, FREV, and SUE. This finding appears to be consistent with previous literature in that momentum is an important factor in explaining performance persistence.

To confirm that the GIA contains unique private information, we run a "horserace" between the GIAs and quantitative signals using Fama-MacBeth regressions. The successive dependent variables are cross-sectional stock returns during each of the four evaluation quarters (Q1 through Q4, respectively), and the explanatory variables include the GIA (LGIA) and the 11 quantitative signals. We also compute the JT4-style average coefficients, following the procedure described in Section 4.1.

A technical issue we encounter is the missing observations for explanatory variables. During any given quarter (especially in the early sample period), a significant number of stocks have at least one missing quantitative signal. Excluding these stocks from multivariate regressions would result in a substantial reduction in the sample size. Instead, we use a multiple imputation method to replace missing observations with simulated values. This method is developed in the statistical literature for regressions where explanatory variables have a subset of observations that are missing at random (e.g., Rubin 1987; Yuan 2000). We adopt this method for the Fama-MacBeth regressions with Newey-West covariance estimates. Briefly, we assume joint normal distributions for the dependent and explanatory variables, and use the Markov chain Monte Carlo (MCMC) procedure to generate simulated values whenever the actual observations are missing. We then perform Fama-MacBeth regressions using the data augmented by simulated values, and adjust the $t$-statistics for the parameter estimates to take into account the effect of simulated regressors. The details of the procedure are described in Appendix B.

In Table 6, we report time-series average estimated coefficients and the corresponding time-series $t$-statistics computed with Newey-West standard errors with a two-quarter lag. The table shows that, notably, after controlling for the 11 quantitative signals, loadings on the GIA and LGIA are significantly positive for the first three quarters, Q1 to Q3, and for the JT4-style averages. In addition, consistent with the information decay pattern reported in Table 2 and Panel A of Table 3,

the coefficients for GIA and LGIA in Table 6 monotonically decline as we move from Q1 to Q4.

Therefore, the stock selection information revealed by the GIA approach cannot be explained by fund trading on momentum or other public quantitative investment signals. This result reflects an interesting difference between the fundamental analysis pursued by most mutual funds and the quantitative investment analysis that has become popular among hedge funds and some institutional money managers. Fundamental analysis may enable funds to obtain private information about stock values, which could be quite different from publicly available investment signals used in quantitative research. This confirms the value of our approach of using fund holdings to extract stock selection information.

## 4.4 GIAs Based on Alternative Fund Skill Proxies

Our last set of analyses uses the GIA approach to explore a further issue. The existing literature has identified various fund characteristics and fund skill measures, which are alternatives to the return-based fund alphas, and, yet, are indicative of future fund performance. An important question is whether these alternative skill proxies predict fund performance by capturing the stock selection ability of fund managers, or because of other effects. Such other effects range from simply charging lower fees (Carhart 1997), engaging in market timing or making active asset allocation decisions (Jiang, Yao, and Yu, 2007), being more skillful in executing trades (Anand, Irvine, Puckett, and Venkataraman, 2012), to being more successful in interim (intra-quarter) trading (Kacperczyk, Sialm, and Zheng, 2008; Puckett and Yan, 2011). Another issue of interest is whether the above skill proxies generally impart the same stock selection information, or whether they capture different and unique information.

To address these issues, we extend the GIA model by substituting the alternative skill proxies in place of the four-factor fund alpha as $\hat{\boldsymbol{S}}$ in Equations (3) and (5). The alternative skill proxies that we consider are the following. The first three are based on simple fund characteristics – fund fees, size, and turnover. Various studies, such as Carhart (1997), Gil-Bazo and Ruiz-Verdú (2009), Chen, Hong, Huang, and Kubik (2004), Pollen and Wilson (2008), and Yan (2008), have shown that these characteristics are related to fund skills and fund performance. We define FEE as the annual fund expense ratio plus amortized load (one seventh of total load; see Sirri and Tufano,

1998), TNA as the total net assets, and TURN as fund trading turnover, all using information available prior to the end of the portfolio formation quarter, Q0.

The next four measures are based on information regarding fund portfolio holdings. Kacperczyk, Sialm, and Zheng (2005) find that the industry concentration of a fund portfolio signals fund ability and is positively related to future fund performance. Kacperczyk, Sialm, and Zheng (2008) show that unobserved actions of mutual funds, measured by the gap between before-expense fund net return and hypothetical buy-and-hold return based on beginning-of-period holdings, predict fund performance. They suggest that the gap reflects the joint effects of fund manager skills and trading costs. Cohen, Coval, and Pastor (2005) show a model that combines past fund performance with the similarity of fund stock holdings produce superior power in predicting future fund performance. Baker, Litov, Wachter, and Wurgler (2010) use the relation between fund trading and stock returns around subsequent earnings announcements to measure fund managers' stock selection ability. Following these studies, we define measures on industry concentration index (ICI), return gap (GAP), similarity-based fund performance measure (SIM), and earnings-announcement returns of a fund (EAR). The details for constructing these variables are provided in Appendix C. Finally, Cremers and Petajisto (2009) and Amihud and Goyenko (2009) find that measures of fund activeness are positively related to future fund performance. We follow Amihud and Goyenko (2009) to construct a return-based (in)activeness measure, which is the R-square from the Carhart four-factor model using a rolling window of 12 months.

In Panel A of Table 7, we report the performance of these alternatively constructed GIAs. For brevity, we only tabulate their D10-D1 spreads in characteristic-adjusted returns. Notably, all eight alternative alpha estimators exhibit significant return predictive power during at least one of the four holding quarters (Q1 to Q4). The return spreads for the GIAs based on FEE, TNA, TURN, and R2 are negative. On the other hand, those based on ICI, GAP, SIM, and EAR generate positive return spreads. Therefore, all these GIAs capture some aspects of mutual fund stock-selection ability, and the directions in which they predict stock returns are consistent with the directions in which the corresponding fund characteristics or fund skill measures predict fund performance, as documented in existing studies.

The GIAs based on return gap ($\alpha_{GAP}$) and portfolio similarity ($\alpha_{SIM}$) exhibit particularly

strong predictive power. At 0.85% and 1.18%, the characteristic-adjusted return spreads they generate for Q1 are at the same magnitude as those for the baseline GIA (Table 2). The next two are $\alpha_{EAR}$ and $\alpha_{R2}$. However, the predictive power of $\alpha_{EAR}$ seems to be rather short-lived, significant only for Q1, while the predictive power of $\alpha_{R2}$ appears to last longer.

Panel B of Table 7 reports the correlations across the eight alternative GIAs and the baseline GIA estimator. The correlations are first calculated across stocks during each quarter, then averaged over time. A noted pattern is the high correlations among the baseline GIA and the two alphas based on GAP and SIM, suggesting that the stock selection abilities captured by fund alpha, return gap, and the similarity-based performance measure may be closely related. In addition, the correlations of $\alpha_{R2}$ with $\alpha_{TNA}$ and $\alpha_{ICI}$, and the correlation between $\alpha_{FEE}$ and $\alpha_{TURN}$, are modestly high.

Given such correlation patterns, we further explore the distinctiveness of the return predictive information contained in these GIAs. We do so using multivariate Fama-MacBeth regressions. Similar to the regressions performed in Table 6, we use the multiple imputation procedure to replace missing observations among explanatory variables with simulated values, and adjust $t$-statistics to account for simulated regressors. As it turns out, due to the high correlations among the baseline GIA, $\alpha_{GAP}$, and $\alpha_{SIM}$, the regressions exhibit symptoms of multi-collinearity when they are jointly included as explanatory variables. In many quarters, the variance inflation factors (VIF) are high, and all the coefficients often become statistically insignificant or exhibit opposite signs.[10] To deal with this issue, we replace the three variables by their principal components as regressors.

Table 8 reports the regression results. Out of the three principal components for the baseline GIA, $\alpha_{GAP}$, and $\alpha_{SIM}$, only the first one exhibits strong and persistent return predictive power, suggesting that these three skill proxies capture similar types of stock picking skills. In addition, the return-predictive power of many other variables is subdued in joint regressions, suggesting that the stock selection information contained in these variables also shares common sources.

[10]The high correlation between the baseline GIA and $\alpha_{GAP}$ is consistent with the notion that unobserved action is an important determinant of fund performance (Kacperczyk, Sialm, and Zheng, 2008). The predictive power of $\alpha_{GAP}$ further suggests that the "unobserved actions" of funds are positively correlated with the stock selection ability of the funds. The high correlation between the baseline GIA and $\alpha_{SIM}$ is consistent with the intuition of Cohen, Coval, and Pastor (2005) that their similarity-based performance measure is a smoothed version of fund alpha.

To take into account the reporting delays for fund holdings, we further perform regressions using one-quarter lagged measures of $\alpha_{ICI}$, $\alpha_{GAP}$, $\alpha_{SIM}$, and $\alpha_{EAR}$, together with LGIA (but keeping the current values of explanatory variables that do not involve fund holdings). The predictive pattern of the three principal components remains similar, although with a slightly weaker magnitude. In addition, the one-quarter lagged $\alpha_{ICI}$ and $\alpha_{EAR}$ are no longer predictive of returns, in contrast to the significant predictive power when their current values are used in the regressions. This indicates that the stock selection information contained in these two variables is relatively short-lived.

# 5. Conclusions

We provide a model to efficiently aggregate stock selection information across stock portfolios held by mutual funds with different skills. We accomplished two objectives. First, we provide strong evidence that fund managers (or a subset of them) possess stock selection information. The statistical and economic magnitude of the predictive performance of the GIA strategies dwarfs that of a popular price momentum strategy. Further, the return-predictive power of the GIAs is not concentrated among small and illiquid stocks. They deliver significant predictive performance across stocks with a wide spectra of characteristics, such as size, book-to-market ratio, past returns, idiosyncratic volatility, and breadth of mutual fund ownership.

Second, the GIA approach serves as a "magnifying glass" on the stock selection information possessed by fund managers, and thus enables us to look at issues related to such stock selection information in a way that we may not be able to at the fund level. Using this approach, we address four such issues. First, we show that the stock selection information captured by the GIA is distinct from several effects documented in existing literature via which fund holdings or trades affect stock returns, namely, the effects of herding, aggregate fund trading, short-sale constraints, and fund trading driven by fund flows. Second, we provide evidence consistent with the notion that fund managers' stock selection ability stems from fundamental analysis, via which fund managers may uncover information about corporate fundamentals not reflected in current stock prices. Third, we find that the return-predictive information possessed by fund managers cannot be ex-

plained away by well-known quantitative investment signals and, therefore, is truly private, which highlights the distinction between the two popular stock selection approaches – fundamental analysis and quantitative research. Finally, we find that many alternative measures of mutual fund skills documented in existing literature can also be converted into signals that predict cross-sectional stock returns, suggesting that these fund skill measures indeed reflect valuable fund manager stock selection information.

# Appendix A: Quantitative Investment Signals

This appendix describes the eleven quantitative signals. [text] refers to the data source, where D# is the item number from the Quarterly Compustat file. t refers to the portfolio formation quarter Q0. m is the last month of Q0. q refers to the most recently reported fiscal quarter prior to the end of Q0, based on Compustat fiscal quarter reporting dates. If the reporting date in Compustat is missing we assume a two-month reporting lag from the fiscal quarter-end.

| Variable | Description | Computation Details |
|---|---|---|
| 1. SIZE | Natural log of market capitalization | Size$_t$ = LOG ($P_t$ * Shares Outstanding$_t$)<br>= LOG (price at the end of the quarter t [D14], multiplied by common shares outstanding at the end of quarter t [D61]) |
| 2. TURN | Average daily volume turnover | Percentile rank $\frac{\sum_{i=1}^{n} \text{Daily Volume/Shares Outstanding}}{n}$ by exchange, where n = number of days available for 6 months preceding the end of quarter t (months m-6 though m-1) [CRSP] |
| 3. RETP | Cumulative market adjusted return for the preceding six months (months -6 through -1) | $[\Pi_{i=m-6}^{m-1}(1 + \text{monthly return}_i)] - 1$<br>-$[\Pi_{i=m-6}^{m-1}(1 + \text{value-weighted market monthly return}_i] - 1$,<br>where m is the last month of quarter t [CRSP] |
| 4. RET2P | Cumulative market-adjusted return for the second preceding six months (months -12 through -7) [CRSP] | $[\Pi_{i=m-12}^{m-7}(1 + \text{monthly return}_i)] - 1$<br>-$[\Pi_{i=m-12}^{m-7}(1 + \text{value-weighted market monthly return}_i] - 1$, |
| 5. FREV | Analyst forecast revisions to price | $\sum_{i=0}^{5}(\frac{f_{m-i}-f_{m-1-i}}{P_{m-1-i}})$, where<br>$f_m$ = mean consensus analyst FY1 forecast at month m [IBES]<br>$P_{m-1}$ = price at the end of month m-1 [CRSP]. Thus,<br>$\sum_{i=0}^{5}(\frac{f_{m-i}-f_{m-1-i}}{P_{m-1-i}})$ = rolling sum of preceding six months revisions to price ratios |
| 6. SUE | Standardized unexpected earnings | $\frac{(\text{EPS}_q - \text{EPS}_{q-4})}{s_q}$, where<br>EPS$_q$ − EPS$_{q-4}$= unexpected earnings for quarter q, with EPS defined as earnings per share (diluted) excluding extraordinary items [D9], adjusted for stock distributions [D17]<br>$s_q$ = standard deviation of unexpected earnings over eight preceding quarters (quarters q-7 though q) |
| 7. EP | Earnings to price | $\frac{\sum_{i=0}^{3}\text{EPS}_{q-i}}{P_t}$, where<br>$EPS_q$= earnings per share before extraordinary items for quarter q[D19]<br>$P_t$ = price at the end of the quarter t [D14]<br>Thus, $\frac{\sum_{i=0}^{3}\text{EPS}_{q-i}}{P_t}$=rolling sum of EPS for preceding four quarters, deflated by price |
| 8. BP | Natural log of book to price ratio | LOG($\frac{\text{Book value of common equity}}{\text{Mktcap}}$), where<br>Book value of common equity$_q$ = book value of total common equity at the end of quarter q [D59]<br>Mktcap$_t$ = $P_t$ * Shares Outstanding$_t$<br>= price at the end of the quarter t [D14], multiplied by common shares outstanding at the end of quarter t [D61] |
| 9. SG | Sales growth | $\frac{\sum_{i=0}^{3}\text{Sales}_{q-i[D2]}}{\sum_{i=0}^{3}\text{Sales}_{q-4-i[D2]}}$ where<br>Thus, $\sum_{i=0}^{3}\text{Sales}_{q-i}$=rolling sum of sales for preceding four quarters and $\sum_{i=0}^{3}\text{sales}_{q-4-i}$=rolling sum of sales for second preceding set of four quarters |
| 10. ACC | Total accruals to total assets | {(△Current Assets$_q$ [D40] -△Cash$_q$[D36])<br>-(△Current Liabilities$_q$ [D49]- △Current LTD$_q$ [D45])<br>-△Deferred Taxes$_q$ [D35] - Depreciation and Amortization$_q$ [D5] }<br>/{$(TA_q + TA_{q-4})/2$ [D44] }<br>$\triangle X_q = X_q - X_{q-4}$, e.g., △Current Assets$_{t-1}$ = Current Assets$_{t-1}$ −Current Assets$_{t-5}$ |
| 11. CAPEX | Capital expenditures to total assets | $\frac{\text{CAPEX}_q}{(TA_q+TA_{q-4})/2[D44]}$<br>CAPEX$_q$ = rolling sum of four quarters (quarters q-3 through q) of Capital Expenditures [D90] (As D90 is fiscal-year-to-date, adjustments are made as needed to calculate the rolling sum of the preceding four quarters.) |

# APPENDIX B: Multiple Imputation Procedure for Missing Observations

Consider the multivariate regressions performed in Table IV. Let $y_{it}$ denote a dependent variable for stock i at time t and $X_{it}$=$\{x_{1it}, x_{2it}, ..., x_{Kit}\}$ denote K explanatory variables for stock i at time t. For a given time t and a given stock i, a subset of the explanatory variables are randomly missing.

We perform P imputations. In empirical implementation P is set to 20. Each imputation p consists of the following three steps.

- Step 1: replacing missing data via simulations. We use the Monte Carlo Markov Chain (MCMC) method to simulate a set of observations to replace the missing observations at each time t. Details of the procedure are explained below. Denote the resulting data (with missing variables replaced) $\{x^p_{kit}\}$, k=1, ..., K.

- Step 2: performing cross-sectional regressions. Using the replaced data generated in Step 1, we perform the following cross-sectional regressions across all stocks at each time t:

$$y_{it} = \theta^p_{0t} + \sum_{k=1}^{K} \theta^p_{kt} x^p_{kit} + e_{it} \tag{11}$$

  where $\theta^p_{kt}$ (k=0, ..., K) is a coefficient. Denote the resulting vector of estimated coefficients $\hat{\theta}^p_t$.

- Step 3: computing statistics for each imputation. We compute the time series average of the estimated coefficient for the p-th imputation as $\bar{\theta}^p = \frac{1}{T}\sum_{t=1}^{T} \hat{\theta}^p_t$. And let the Newey-West estimator of the covariance matrix of $\hat{\theta}^p_t$ be $V^p$.

Based on the average coefficients $\bar{\theta}^p$ and the covariances $V^p$ (p=1, ..., P), we compute the average coefficients and average covariances across P imputations: $\hat{\bar{\theta}} = \frac{1}{P}\sum_{p=1}^{P} \bar{\theta}^p$ and $\hat{V} = \frac{1}{P}\sum_{p=1}^{P} V^p$. We further compute the between-imputation covariance as $\hat{S} = \frac{1}{P-1}\sum_{p=1}^{P}(\bar{\theta}^p - \hat{\bar{\theta}})$. Finally, the imputation-adjusted covariance for $\hat{\bar{\theta}}$ is $\hat{V}^* = \hat{V} + (1 + \frac{1}{P})\hat{S}$. And the imputation-adjusted t-statistic for a coefficient $\theta_k$ is $t^* = \hat{\bar{\theta}}_k / \hat{V}^*_{k,k}$, where $\hat{V}^*_{k,k}$ is the k-th diagonal element of $\hat{V}^*$.

We now describe the MCMC simulation procedure in the above Step 1. The simulation is based on the assumption of joint multivariate normal distribution for all the variables $y_{it}$ and $X_{it}$, the joint normal distribution for the mean of the variables, and the inverted Whishart distribution for the covariances of the variables. The priors on these distributional parameters are either non-informative or based on sample estimates, whenever appropriate. The MCMC method iterates over the following two steps. In the simulation step, we simulate the data from the current prior distribution for the variables with missing observations, and then replace the missing observations with simulated ones. In the posterior step, we update our posterior estimates based on the replaced data. The simulation and posterior steps are iterated by a maximum of 100 times or until the posterior estimates converge, whichever comes first (plus first 200 burn-in iterations). This is implemented using the "proc mi" procedure in SAS.

# APPENDIX C: Fund Skill Measures: ICI, GAP, SIM, and EAR

**ICI**. The industry concentration index of fund i follows Kacperczyk, Sialm, and Zheng (2005):

$$\text{ICI}_{i,t} = \sum_{j=1}^{10} (\omega_{i,j,t} - \tilde{\omega}_{j,t})^2 \tag{12}$$

where $j = 1$ to 10, representing 10 industries. $\omega_{i,j,t}$ is fund i's portfolio weight in industry j, and $\tilde{\omega}_{j,t}$ is the weight of industry j in the CRSP market portfolio. The 10-industry classification is in Appendix B of Kacperczyk et al. (2005).

**GAP**. Following Kacperczyk, Sialm, and Zheng (2006), we define fund return gap as:

$$\text{GAP}_{i,t} = \frac{1}{4}\sum_{k=0}^{3}(R_{i,t-k} - \text{GR}_{i,t-k}) \tag{13}$$

where $R_{i,t-k}$ is the before-expense net return for fund i in quarter t-k (k=0, 1, 2, 3). $\text{GR}_{i,t}$ is quarterly fund gross return, computed as the buy-and-hold return on the portfolio at beginning of quarter t:

$$\text{GR}_{i,t} = \omega_{i,s,t}\sum_{j=1}^{N} w_{i,j,t}R_{j,t} + \omega_{i,pb,t}R_{pb,t} + \omega_{i,c,t}R_{ft} \tag{14}$$

where $w_{i,j,t}$ is the portfolio weight of fund i on stock j, calculated as the number of shares of stock j held by the fund multiplied by its market price, and then divided by the total stock holding value of the fund, with all variables involved measured at the beginning of quarter t. $R_{j,t}$ is the buy-and-hold return of stock j during quarter t. $\omega_{i,s,t}$ is the value of stock holding as a fraction of fund TNA, $\omega_{i,pb,t}$ is the value of holdings of preferred stocks and bonds as a fraction of fund TNA, and $\omega_{i,c,t}$ is the value of cash holding as a fraction of fund TNA. $\omega_{i,s,t}$, $\omega_{i,pb,t}$, and $\omega_{i,c,t}$ are obtained from the CRSP. We set them missing if CRSP reports 0 for all these variables. $R_{pb,t}$ is the total return of the Lehman Brothers Aggregate Bond Index and $R_{ft}$ is the risk-free return (yield on three-month Treasury bills, from CRSP) during quarter t.

**SIM**. Based on Cohen, Coval, and Pastor (2005), we construct a variable SIM to gauge the skill of a fund manager by the extent to which his stock holding resemble those of funds with superior past performance. In each quarter, SIM for fund i is $\text{SIM}_{i,t} = \sum_{j=1}^{N} w_{i,j,t}\delta_{j,t}$, where there are M funds (i=1, ..., M) and N stocks (j = 1, ..., N). $w_{i,j,t}$ is the weight on stock j by fund i and $\delta_{j,t}$ is the quality of stock j:

$$\delta_{j,t} = \frac{\sum_{i=1}^{M} w_{i,j,t}\alpha_{i,t}}{\sum_{i=1}^{M} w_{i,j,t}} \tag{15}$$

where $\alpha_{i,t}$ is a fund's four-factor alpha estimated using the past twelve months fund return prior to the quarter end.

**EAR**. Baker, Litov, Watchter, and Wurgler (2010) use the returns realized around the subsequent earnings announcements of stocks traded by mutual funds to measure fund managers' stock picking ability. Consistent with their approach, we estimate the earnings announcement returns (EAR) of fund i in quarter t as the difference between weighted earnings announcement returns of all stocks purchased and those of all stocks sold:

$$\text{EAR}_{i,t} = \sum_{j\epsilon J+}\frac{w_{i,j,t-1} - \tilde{w}_{i,j,t-k-1}}{\sum(w_{i,j,t-1} - \tilde{w}_{i,j,t-k-1})}R_{j,t} - \sum_{j\epsilon J-}\frac{w_{i,j,t-1} - \tilde{w}_{i,j,t-k-1}}{\sum(w_{i,j,t-1} - \tilde{w}_{i,j,t-k-1})}R_{j,t} \tag{16}$$

where k is either 1 (for funds reporting holdings quarterly) or 2 (for funds reporting holdings semi-annually). $w_{i,j,t-1}$ is the portfolio weight of fund i on stock j and $\tilde{w}_{i,j,t-k-1}$ is the portfolio weight measured at the end of quarter t-1 under the assumption that the fund holds its positions, without trading, from its previous

reporting date. $J^+ = \{j : w_{i,j,t-1} - \tilde{w}_{i,j,t-k-1} > 0\}$ denotes all stocks purchased by fund i in quarter t-1 and $J^- = \{j : w_{i,j,t-1} - \tilde{w}_{i,j,t-k-1} < 0\}$ denotes all stocks sold by fund i in quarter t-1. $R_{j,t}$ is the three-day event-window announcement return of stock j during quarter t, from one trading day before to one trading day after the earnings announcement.

# References


Alexander, G., G. Cici, and S. Gibson. 2007. Does Motivation Matter When Assessing Trade Performance? An Analysis of Mutual Funds. *Review of Financial Studies* 20:125-150.

Almazan, A., K. Brown, M. Carlson, and D. Chapman. 2004. Why Constrain Your Fund Managers? *Journal of Financial Economics* 73:289-321.

Amihud, Y., and R, Goyenko. 2009. Mutual Fund's R2 As Predictor of Performance. Working Paper, New York University.

Anand, A., P. Irvine, A. Puckett and K. Venkataraman. 2012. Performance of Institutional Trading Desks: An Analysis of Persistence in Trading Costs. *Review of Financial Studies* 25:557-598.

Baker, M., L. Litov, J. Watchter, and J. Wurgler. 2010. Can Mutual Fund Managers Pick Stocks? Evidence from Their Trades Prior to Earnings Announcements. *Journal of Financial and Quantitative Analysis* 45:1111-1131.

Barras L, O. Scaillet, and R. Wermers. 2010. False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas. *Journal of Finance* 65:179-216.

Berk, J., and R. Green. 2004. Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy* 112:1269-1295.

Brown, N., K. Wei, and R. Wermers. 2011. Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices. Working Paper, University of Maryland.

Carhart, M. 1997. On Persistence in Mutual Fund Performance. *Journal of Finance* 52:57-82.

Chen, H. L., N. Jegadeesh, and R. Wermers. 2000. An Examination of the Stockholdings and Trades of Fund Managers. *Journal of Financial and Quantitative Analysis* 35:43-68.

Chen, J., H. Hong, and J. Stein. 2002. Breadth of Ownership and Stock Returns. *Journal of Financial Economics* 66:171-205.

Chordia, T., A. Subrahmanyam, and Q. Tong. 2010. Trends in the Cross-Section of Expected Stock Returns. Working Paper, University of California at Los Angeles.

Cohen, R., J. Coval, and L. Pastor. 2005. Judging Fund Managers by the Company They Keep. *Journal of Finance* 60:1057-1096.

Cohen R., C. Polk, and B. Silli. 2010. Best Ideas. Working Paper, London School of Economics.

Cremers, M., A. Petajisto, and E. Zitzewitz. 2009. Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation. Working Paper, Yale University.

Cremers, M. and A. Petajisto. 2009. How Active Is Your Fund Manager? A New Measure that Predicts Performance. *Review of Financial Studies* 22:3329-3365.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring Mutual Fund Performance with Characteristic-based Benchmarks. *Journal of Finance* 52:1035-1058.

Fama, E. F. and K. French. 1992. The Cross-section of Expected Stock Returns. *Journal of Finance* 47:427-465.

Fama, E. F. and K. French. 2010. Luck versus Skill in the Cross Section of Mutual Fund Returns. *Journal of Finance* 65:1915-1947.

Frazzini, A., and O. Lamont. 2008. Dumb Money: Mutual Fund Flows and the Cross-section of Stock Returns. *Journal of Financial Economics* 88:299-322.

Gaspar, J., M. Massa, and P. Matos. 2005. Favoritism in Mutual Fund Families? Evidence on Strategic Cross-fund Subsidization. *Journal of Finance* 61:73-104.

Gil-Bazo, J., and P. Ruiz-Verdú. 2009. The Relation between Price and Performance in the Mutual Fund Industry. *Journal of Finance* 64:2153-2183.

Grinblatt, M. and S. Titman. 1989. Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings. *Journal of Business* 62:393-416.

Grinblatt, M. and S. Titman. 1993. Performance Measurement Without Benchmarks: An Examination of Mutual Fund Returns. *Journal of Business* 66:47-68.

Grinblatt, M., S. Titman, and R. Wermers. 1995. Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. *American Economic Review* 85:1088-1105.

Grossman, S., and J. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *American Economic Review* 70:393-408.

Hendricks, D., J. Patel, and R. Zeckhauser. 1993. Hot Hands in Mutual Funds: The Persistence of Performance. 1974-88. *Journal of Finance* 48:93-130.

Huang, J., C. Sialm, and H. Zhang. 2010. Risk Shifting and Mutual Fund Performance. *Review of Financial Studies*, forthcoming.

Jegadeesh, N., J. Kim, S. Krische, and C. Lee. 2004. Analyzing the Analysts: When Do Recommendations Add Value? *Journal of Finance* 59:1083-1124.

Jiang, G., T. Yao, and T. Yu. 2007. Do Mutual Funds Time the Market? Evidence from Portfolio Holdings. *Journal of Financial Economics* 86:724-758.

Kacperczyk, M., and A. Seru. 2007. Fund Manager Use of Public Information: New Evidence on Managerial Skills. *Journal of Finance* 62:485-528.

Kacperczyk, M., C. Sialm, and L. Zheng. 2005. On the Industry Concentration of Actively Managed Equity Mutual Funds. *Journal of Finance* 60:1983-2011.

Kacperczyk, M., C. Sialm, and L. Zheng. 2008. Unobserved Actions of Equity Mutual Funds. *Review of Financial Studies* 21:2379-2416.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp. 2011. Time-Varying Fund Manager Skill. Working Paper, New York University.

Kosowski, R., A. Timmermann, R. Wermers, and H. White. 2006. Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from A Bootstrap Analysis. *Journal of Finance* 56:2551-2595.

Kothari, S. and Warner, J.. 2001. Evaluating mutual Fund Performance. *Journal of Finance* 56:1985-2010.

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

Mamaysky H., M. Spiegel, and H. Zhang. 2007. Improved Forecasting of Mutual Fund Alphas and Betas. *Review of Finance* 11:359-400.

Moore, E. H. 1920. On the Reciprocal of the General Algebraic Matrix. (Abstract). *Bulletin of the American Mathematical Society* 26:394-5.

Newey, W. and K. West. 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55:703-708.

Penrose, R. A. 1955. A Generalized Inverse for Matrices. *Proceedings of the Cambridge Philosophical Society* 51:406-13.

Pollet, J. M., and M. Wilson. 2008. Does Size Affect Mutual Fund Behavior? *Journal of Finance* 63:2941-2969.

Puckett A., and X. Yan. 2011. The Interim Trading Skills of Institutional Investors. *Journal of Finance* 66:601-633.

Rubin, D.B. 1987. Multiple Imputation for Nonresponse in Surveys. New York, NY: John Wiley & Sons, Inc.

Shumway, T. 1997. The Delisting Bias in CRSP Data. *Journal of Finance* 52:327-340.

Shumway, T., M. Szefler, and K. Yuan. 2009. The Information Content of Revealed Beliefs in Portfolio Holdings. Working Paper, London School of Economics

Sias, R. 2004. Institutional Herding. *Review of Financial Studies* 17:165-206.

Sirri, E., and P. Tufano. 1998. Costly Search and Mutual Fund Flows. *Journal of Finance* 53:1589-1622.

Treynor, J. L. and F. Black. 1973. How to Use Security Analysis to Improve Portfolio Selection. *Journal of Business* 46:66-86.

Wermers, R. 1999. Mutual Fund Herding and the Impact on Stock Prices. *Journal of Finance* 53:581-622.

Wermers, R. 2003. Is Money Really "Smart"? New Evidence on the Relation between Mutual Fund Flows, Manager Behavior, and Performance Persistence. Working Paper, University of Maryland.

Yan, X. 2008. Liquidity, Investment Style, and the Relation between Funds Size and Fund Performance. *Journal of Financial and Quantitative Analysis* 43:741-768.

Yuan, Y. 2000. Multiple Imputation for Missing Values: Concepts and New Development. in *SAS Users Group International Conference Proceedings*. Cary, NC: SAS Institute.

## Table 1. Summary Statistics

This table provides summary statistics on the sample of mutual funds and their stock holdings each year from 1980 to 2006. To obtain the sample, we merge the Thomson mutual fund holdings data with CRSP mutual fund returns data. We include actively managed U.S. domestic equity mutual funds with a self-declared investment objective of aggressive growth, growth, or growth-and-income. We report the number of funds at the end of each sample year, the average annual fund turnover ratio (turnover ratio is missing in the CRSP data for 1991; they are replaced by the 1990 observations), the median numbers of stocks held per fund, as well as the median numbers of stocks bought and sold per fund during past six months, the total number of distinct stocks held by funds, and the aggregate market value of stock holdings by sample funds. For comparison we also report the total number of common stocks in the CRSP universe and the total market capitalizations of these stocks. The last two columns report the total number of distinct stocks held by sample funds and the total market value of fund stock holdings as percentages of the total number of common stocks and the total market capitalization in the CRSP universe. For funds that do not report portfolio holdings at a year-end, we use their last reported portfolio snapshots during the second half of the year and assume that they hold the reported shares until the year-end on a split-adjusted basis. The last three rows display the average, minimum, and maximum of the annual statistics across the sample years.

| | Mutual Fund Sample | | | | | | | CRSP Universe | | Proportion | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Total Number of Funds | Average Annual Turnover | Median Number of Stocks Held | Median Number of Stocks Bought | Median Number of Stocks Sold | Number of Distinct Stocks Held | Market Value of Stocks Held ($B) | Number of Distinct Stocks | Total Market Value ($B) | Number of Stocks (%) | Value of Stocks (%) |
| 1980 | 284 | 0.71 | 56 | 17 | 18 | 2102 | 33 | 4727 | 1316 | 44.47 | 2.54 |
| 1981 | 280 | 0.67 | 57 | 17 | 17 | 2289 | 30 | 5090 | 1228 | 44.97 | 2.44 |
| 1982 | 270 | 0.74 | 57 | 23 | 18 | 2469 | 39 | 5060 | 1409 | 48.79 | 2.75 |
| 1983 | 297 | 0.73 | 65 | 24 | 20 | 3082 | 54 | 5673 | 1740 | 54.33 | 3.10 |
| 1984 | 313 | 0.73 | 69 | 24 | 20 | 3177 | 56 | 5797 | 1681 | 54.80 | 3.33 |
| 1985 | 342 | 0.79 | 73 | 26 | 24 | 3447 | 74 | 5774 | 2095 | 59.70 | 3.54 |
| 1986 | 403 | 0.79 | 72 | 27 | 25 | 3659 | 91 | 6058 | 2360 | 60.40 | 3.87 |
| 1987 | 461 | 0.90 | 70 | 26 | 28 | 3568 | 102 | 6363 | 2321 | 56.07 | 4.40 |
| 1988 | 500 | 0.84 | 65 | 22 | 24 | 3708 | 111 | 6102 | 2520 | 60.77 | 4.41 |
| 1989 | 545 | 0.79 | 64 | 24 | 25 | 3638 | 143 | 5909 | 3067 | 61.57 | 4.66 |
| 1990 | 589 | 0.91 | 67 | 25 | 25 | 3345 | 136 | 5765 | 2754 | 58.02 | 4.95 |
| 1991 | 691 | 0.90 | 70 | 29 | 22 | 3598 | 215 | 5810 | 3721 | 61.93 | 5.77 |
| 1992 | 800 | 0.79 | 75 | 34 | 23 | 3783 | 278 | 5927 | 4121 | 63.83 | 6.74 |
| 1993 | 1074 | 0.81 | 82 | 37 | 27 | 5095 | 392 | 6471 | 4682 | 78.74 | 8.37 |
| 1994 | 1212 | 0.83 | 83 | 36 | 28 | 5398 | 415 | 6785 | 4636 | 79.56 | 8.95 |
| 1995 | 1347 | 0.90 | 87 | 40 | 31 | 5782 | 674 | 7017 | 6329 | 82.40 | 10.65 |
| 1996 | 1502 | 0.94 | 92 | 41 | 32 | 6209 | 892 | 7479 | 7718 | 83.02 | 11.56 |
| 1997 | 1653 | 0.91 | 91 | 42 | 33 | 6260 | 1233 | 7467 | 10055 | 83.84 | 12.26 |
| 1998 | 1651 | 0.95 | 88 | 39 | 34 | 5870 | 1632 | 7040 | 12425 | 83.38 | 13.14 |
| 1999 | 1655 | 1.13 | 89 | 40 | 35 | 5771 | 2103 | 6685 | 15892 | 86.33 | 13.23 |
| 2000 | 1717 | 1.17 | 90 | 40 | 36 | 4842 | 1977 | 6381 | 14438 | 75.88 | 13.69 |
| 2001 | 1856 | 1.26 | 92 | 43 | 36 | 4486 | 1684 | 5678 | 12806 | 79.01 | 13.15 |
| 2002 | 1586 | 1.20 | 92 | 41 | 40 | 4402 | 1268 | 5254 | 9938 | 83.78 | 12.76 |
| 2003 | 2001 | 1.17 | 93 | 44 | 37 | 4283 | 1695 | 4925 | 12972 | 86.96 | 13.07 |
| 2004 | 1999 | 1.04 | 95 | 42 | 40 | 4248 | 1911 | 4867 | 14444 | 87.28 | 13.23 |
| 2005 | 1940 | 0.93 | 90 | 42 | 37 | 4124 | 2010 | 4796 | 14983 | 85.99 | 13.42 |
| 2006 | 1820 | 0.92 | 91 | 40 | 40 | 4129 | 2166 | 4745 | 16652 | 87.02 | 13.01 |
| Average | 1066 | 1 | 78 | 33 | 29 | 4176 | 793 | 5913 | 6974 | 70.11 | 8.26 |
| Min | 270 | 0.67 | 56 | 17 | 17 | 2102 | 30 | 4727 | 1228 | 44.47 | 2.44 |
| Max | 2001 | 1.26 | 95 | 44 | 40 | 6260 | 2166 | 7479 | 16652 | 87.28 | 13.69 |

## Table 2. Performance of GIA Estimated Using Fund Holdings

This table report the performance of decile stock portfolios formed on GIA, which is estimated using current fund holdings. From 1980 to 2006, in each quarter we form equal-weighted portfolios based on GIA and hold them during the next four quarters (Q1-Q4), rebalancing quarterly. SIZE, BM, and MOM are the average size, book-to-market, and momentum quintile ranks (1 the lowest and 5 the highest quintiles). We report the net returns, characteristic-adjusted returns, and the Carhart four-factor alphas for the decile portfolios. We also report the returns to a four-quarter overlapping portfolio strategy following Jegadeesh and Titman (1993) (referred to as *JT4*), where in each quarter, the four portfolios with the same decile ranks but formed during the past four different quarters are further combined in equal weights into a single portfolio. Returns and alphas are expressed in percentage points. The time-series t-statistics for the differences in the returns and alphas between the top and bottom decile portfolios are reported in parentheses.

| | Characteristic Quintile Ranks | | | Net Return (%) | | | | |
|---|---|---|---|---|---|---|---|---|
| | SIZE | BM | MOM | Q1 | Q2 | Q3 | Q4 | JT4 |
| D1 (Bottom) | 3.53 | 2.81 | 2.97 | 3.02 | 3.02 | 2.95 | 3.03 | 2.97 |
| D2 | 2.85 | 3.01 | 2.89 | 3.31 | 3.33 | 3.22 | 3.27 | 3.25 |
| D3 | 2.41 | 3.06 | 2.85 | 3.24 | 3.52 | 3.37 | 3.18 | 3.32 |
| D4 | 1.98 | 3.14 | 2.87 | 3.19 | 3.21 | 3.27 | 3.45 | 3.25 |
| D5 | 1.56 | 3.21 | 2.93 | 3.64 | 3.82 | 3.68 | 3.73 | 3.70 |
| D6 | 1.69 | 3.2 | 2.96 | 3.69 | 3.59 | 3.53 | 3.56 | 3.59 |
| D7 | 2.07 | 3.14 | 2.94 | 3.70 | 3.68 | 3.39 | 3.64 | 3.62 |
| D8 | 2.44 | 3.08 | 2.97 | 3.88 | 3.73 | 3.75 | 3.58 | 3.74 |
| D9 | 2.89 | 2.97 | 3.07 | 4.17 | 3.92 | 3.62 | 3.49 | 3.81 |
| D10 (Top) | 3.51 | 2.7 | 3.26 | 4.55 | 4.34 | 3.93 | 3.42 | 4.07 |
| D10-D1 | | | | 1.53 | 1.31 | 0.98 | 0.39 | 1.10 |
| | | | | (5.11) | (4.08) | (2.85) | (0.65) | (3.99) |

| | Characteristic-adjusted Return (%) | | | | | 4-factor Alpha (%) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q4 | JT4 | Q1 | Q2 | Q3 | Q4 | JT4 |
| D1 (Bottom) | -0.38 | -0.36 | -0.30 | -0.09 | -0.29 | -0.35 | -0.38 | -0.36 | -0.27 | -0.36 |
| D2 | 0.04 | -0.01 | -0.10 | -0.08 | 0.02 | -0.17 | -0.09 | -0.11 | -0.11 | -0.12 |
| D3 | 0.08 | -0.08 | 0.13 | -0.02 | 0.11 | -0.19 | -0.05 | -0.20 | -0.27 | -0.14 |
| D4 | -0.02 | -0.11 | -0.01 | 0.10 | 0.01 | -0.18 | -0.12 | -0.12 | -0.06 | -0.10 |
| D5 | 0.16 | 0.17 | 0.07 | 0.21 | 0.20 | -0.15 | -0.08 | 0.00 | -0.13 | -0.06 |
| D6 | 0.36 | 0.19 | 0.19 | 0.15 | 0.23 | 0.06 | -0.12 | -0.03 | -0.05 | -0.02 |
| D7 | 0.36 | 0.19 | 0.15 | 0.33 | 0.25 | 0.11 | -0.06 | 0.02 | 0.19 | 0.09 |
| D8 | 0.43 | 0.26 | 0.28 | 0.19 | 0.31 | 0.19 | 0.03 | 0.15 | 0.06 | 0.13 |
| D9 | 0.64 | 0.31 | 0.12 | 0.10 | 0.32 | 0.44 | 0.20 | 0.17 | 0.12 | 0.25 |
| D10 (Top) | 0.76 | 0.47 | 0.24 | -0.06 | 0.36 | 0.80 | 0.62 | 0.37 | 0.12 | 0.50 |
| D10-D1 | 1.14 | 0.83 | 0.53 | 0.04 | 0.65 | 1.15 | 1.00 | 0.73 | 0.40 | 0.86 |
| | (5.72) | (3.87) | (2.64) | (0.18) | (4.50) | (3.50) | (2.98) | (2.23) | (1.16) | (3.22) |

## Table 3. Performance of GIAs Estimated Using Fund Buys and Fund Sells

This table report the performance difference between the top and bottom deciles of stock portfolios formed on GIAs that are estimated using lagged fund holdings, recent fund buys, and recent fund sells respectively. From 1980Q3 to 2006, in each quarter we form equal-weighted portfolios based on one of the GIAs and hold them during the next four quarters (Q1-Q4), rebalancing quarterly. We report the net returns, characteristic-adjusted returns, and the Carhart four-factor alphas for the decile portfolios. We also report the returns to a four-quarter overlapping portfolio strategy following Jegadeesh and Titman (1993) (referred to as *JT4*), where in each quarter, the four portfolios with the same decile ranks but formed during the past four different quarters are further combined in equal weights into a single portfolio. Returns and alphas are expressed in percentage points. The time-series t-statistics are reported in parentheses.

| | Q1 | Q2 | Q3 | Q4 | JT4 |
|---|---|---|---|---|---|
| Panel A: GIA Based on Lagged Holdings (LGIA) | | | | | |
| Net RET | 1.52 | 1.27 | 0.79 | 0.24 | 1.00 |
| | (4.88) | (3.17) | (1.97) | (0.47) | (3.22) |
| Char-adj. RET | 0.95 | 0.62 | 0.31 | -0.02 | 0.48 |
| | (4.62) | (2.68) | (1.51) | (-0.08) | (3.23) |
| 4-factor Alpha | 0.99 | 0.95 | 0.73 | 0.26 | 0.75 |
| | (2.94) | (2.48) | (2.09) | (0.82) | (2.84) |
| Panel B: GIA Based on Recent Fund Buys ($\alpha_{BUY}$) | | | | | |
| Net RET | 0.79 | 0.71 | 0.14 | 0.22 | 0.51 |
| | (2.67) | (2.42) | (0.50) | (0.72) | (2.18) |
| Char-adj. RET | 0.62 | 0.61 | -0.05 | 0.05 | 0.35 |
| | (2.82) | (2.60) | (-0.21) | (0.24) | (2.15) |
| 4-factor Alpha | 0.97 | 0.83 | 0.18 | 0.49 | 0.68 |
| | (2.89) | (2.48) | (0.56) | (1.45) | (2.66) |
| Panel C: GIA Based on Recent Fund Sells ($\alpha_{SELL}$) | | | | | |
| Net RET | -0.15 | -0.37 | -0.14 | 0.11 | -0.16 |
| | (-0.53) | (-1.23) | (-0.47) | (0.35) | (-0.69) |
| Char-adj. RET | 0.16 | -0.16 | -0.02 | 0.26 | 0.04 |
| | (0.78) | (-0.69) | (-0.09) | (1.19) | (0.31) |
| 4-factor Alpha | -0.04 | -0.26 | -0.13 | -0.05 | -0.17 |
| | (-0.11) | (-0.72) | (-0.39) | (-0.17) | (-0.68) |

## Table 4. Comparison with Herding, Aggregate Fund Trading, Change in Breadth of Ownership, and Flow Effect

This table provides a comparison of the GIAs with measures of herding (HERD), aggregate fund trading (TRADE), change in breadth of fund ownership ($\Delta$BRD), and the fund flow effect (FLOW). Panel A reports the difference in characteristic-adjusted returns between the top and bottom decile portfolios formed on HERD, TRADE, $\Delta$BRD, and FLOW respectively. Panel B reports the results of Fama-MacBeth regressions. The dependent variables are characteristic-adjusted stock returns during the four subsequent quarters (Q1 to Q4) as well as the returns to the four-quarter overlapping portfolios (JT4). The explanatory variables include GIA (or LGIA), HERD, TRADE, $\Delta$BRD, and FLOW. Coefficients for HERD, TRADE, and FLOW are multiplied by 10. The regression intercepts are not reported. The t-statistics are reported in parentheses.

Panel A: Performance of HERD, TRADE, $\Delta$BRD, and FLOW

| | Q1 | Q2 | Q3 | Q4 | JT4 |
|---|---|---|---|---|---|
| HERD | 1.24 | 0.30 | -0.64 | -1.08 | -0.03 |
| | (3.40) | (1.00) | (-1.80) | (-3.09) | (-0.13) |
| TRADE | 0.41 | 0.37 | -0.10 | -0.62 | 0.06 |
| | (1.78) | (1.67) | (-0.47) | (-3.36) | (0.49) |
| $\Delta$BRD | 1.31 | 0.22 | 0.37 | -0.67 | 0.33 |
| | (3.97) | (0.77) | (1.68) | (-2.51) | (3.09) |
| FLOW | 0.51 | 0.26 | -0.02 | -0.12 | 0.20 |
| | (1.77) | (0.90) | (-0.06) | (-0.41) | (0.83) |

Panel B: Multivariate Regressions

| | GIA | | | | | LGIA | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q4 | JT4 | Q1 | Q2 | Q3 | Q4 | JT4 |
| GIA | 0.16 | 0.11 | 0.08 | 0.00 | 0.08 | | | | | |
| | (3.92) | (3.20) | (2.07) | (-0.02) | (2.78) | | | | | |
| LGIA | | | | | | 0.15 | 0.10 | 0.04 | 0.02 | 0.07 |
| | | | | | | (4.00) | (2.72) | (1.07) | (0.46) | (2.57) |
| HERD | 0.26 | -0.02 | -0.14 | -0.21 | -0.02 | 0.26 | -0.02 | -0.14 | -0.21 | -0.02 |
| | (3.87) | (-0.29) | (-1.93) | (-3.31) | (-0.47) | (3.85) | (-0.27) | (-1.95) | (-3.33) | (-0.50) |
| TRADE | 0.03 | 0.38 | 0.24 | -0.02 | 0.18 | 0.04 | 0.36 | 0.24 | -0.01 | 0.18 |
| | (0.06) | (1.78) | (1.29) | (-0.04) | (0.70) | (0.09) | (1.72) | (1.14) | (-0.02) | (0.70) |
| $\Delta$BRD | 0.14 | 0.12 | 0.11 | -0.14 | 0.03 | 0.12 | 0.10 | 0.10 | -0.14 | 0.02 |
| | (1.85) | (1.39) | (1.40) | (-1.38) | (0.73) | (1.92) | (1.51) | (1.34) | (-1.45) | (0.48) |
| FLOW | 0.09 | -0.14 | 0.05 | -0.27 | -0.05 | 0.08 | -0.15 | 0.09 | -0.26 | -0.04 |
| | (0.54) | (-0.85) | (0.43) | (-2.16) | (-0.53) | (0.49) | (-0.87) | (0.69) | (-2.07) | (-0.44) |

## Table 5. Fundamental Information Contained in GIAs

This table reports the average differences, between the top and bottom decile portfolios sorted on GIA and LGIA, in the following four measures of firms' future fundamentals: standardized unexpected earnings (SUE), earnings surprises relative to consensus forecasts (SUR), analyst forecast revision (FRV), and returns during a five-day window around earnings announcements (EAR). We report SUE, SUR, FRV, and EAR during the four quarters (Q1 to Q4) after portfolio formation. In addition, JT4 refers to a four-quarter overlapping portfolio strategy following Jegadeesh and Titman (1993), where, in each quarter, the four portfolios with the same decile rank but formed during the past four different quarters are further combined in equal weights into a single portfolio. SUE, SUR, and FREV are cross-sectionally winsorized at the 1 and 99 percentiles in each quarter. SUR, FRV, and EAR are expressed in percentage points. In the parentheses are the t-statistics computed using the Newey-West covariance estimates with a 4-quarter lag.

| | GIA | | | | | LGIA | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q4 | JT4 | Q1 | Q2 | Q3 | Q4 | JT4 |
| SUE | 0.14 | 0.14 | 0.10 | 0.08 | 0.12 | 0.21 | 0.21 | 0.17 | 0.13 | 0.19 |
| | (2.99) | (3.54) | (2.92) | (2.79) | (3.45) | (4.82) | (5.82) | (5.62) | (5.14) | (6.16) |
| SUR | 0.09 | 0.10 | 0.12 | 0.13 | 0.11 | 0.11 | 0.12 | 0.15 | 0.13 | 0.13 |
| | (2.16) | (2.07) | (2.07) | (2.19) | (2.26) | (2.28) | (3.13) | (2.48) | (2.63) | (2.80) |
| FRV | 0.20 | 0.16 | 0.14 | 0.11 | 0.16 | 0.26 | 0.20 | 0.15 | 0.13 | 0.19 |
| | (6.76) | (6.13) | (4.61) | (2.82) | (5.79) | (7.67) | (6.35) | (4.96) | (3.25) | (6.70) |
| EAR | 0.27 | 0.18 | 0.13 | 0.02 | 0.15 | 0.23 | 0.17 | 0.19 | 0.05 | 0.15 |
| | (4.03) | (2.68) | (1.94) | (0.30) | (3.19) | (2.89) | (2.28) | (3.01) | (0.66) | (3.12) |

## Table 6. GIAs, Quantitative Signals, and Stock Returns

This table reports the results of the following Fama-MacBeth regressions. The dependent variables are stock returns during the four quarters Q1 to Q4. The explanatory variables jointly include GIA (or LGIA) and the 11 quantitative signals. A multiple imputation procedure is used to generate values for missing observations of the quantitative signals. Coefficients reported under "JT4" are those averaged over four different regressions with the stock returns (the dependent variable) in the same quarter but the explanatory variables in past four different quarters. We report the time-series averages of the estimated coefficients as well as the time-series t-statistics computed using the Newey-West procedure with a 4-quarter lag. The t-statistics are adjusted to account for the generated explanatory variables. The regression intercepts are not reported.

| | GIA | | | | | LGIA | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q4 | JT4 | Q1 | Q2 | Q3 | Q4 | JT4 |
| GIA | 0.26 | 0.20 | 0.15 | 0.05 | 0.14 | | | | | |
| | (4.38) | (3.46) | (2.27) | (1.01) | (3.18) | | | | | |
| LGIA | | | | | | 0.21 | 0.17 | 0.13 | 0.06 | 0.13 |
| | | | | | | (3.97) | (2.34) | (2.04) | (1.27) | (2.93) |
| SIZE | -0.35 | -0.31 | -0.14 | -0.17 | -0.26 | -0.27 | -0.21 | -0.15 | -0.16 | -0.22 |
| | (-1.87) | (-1.44) | (-0.74) | (-0.85) | (-1.44) | (-1.57) | (-1.07) | (-0.82) | (-0.75) | (-1.28) |
| TURN | -0.48 | -0.54 | -0.91 | -0.64 | -0.63 | -0.61 | -0.73 | -0.95 | -0.58 | -0.66 |
| | (-0.94) | (-0.85) | (-1.31) | (-0.91) | (-0.98) | (-1.11) | (-1.08) | (-1.33) | (-0.90) | (-1.02) |
| RETP | 1.20 | 1.15 | 0.48 | -0.10 | 0.68 | 1.11 | 1.12 | 0.45 | -0.05 | 0.65 |
| | (5.01) | (5.44) | (1.69) | (-0.91) | (3.68) | (4.99) | (5.73) | (1.57) | (-0.58) | (3.71) |
| RET2P | 0.48 | -0.21 | -0.36 | -0.06 | -0.01 | 0.45 | -0.19 | -0.38 | -0.05 | -0.01 |
| | (1.66) | (-1.10) | (-2.36) | (-0.71) | (-0.35) | (1.68) | (-1.09) | (-2.50) | (-0.49) | (-0.36) |
| FREV | 0.40 | 0.19 | 0.15 | 0.06 | 0.20 | 0.41 | 0.21 | 0.13 | 0.04 | 0.20 |
| | (3.91) | (2.18) | (2.42) | (1.12) | (2.94) | (3.90) | (2.40) | (2.10) | (0.88) | (2.88) |
| SUE | 0.56 | 0.21 | 0.07 | -0.09 | 0.21 | 0.54 | 0.18 | 0.05 | -0.08 | 0.19 |
| | (7.10) | (2.30) | (0.39) | (-1.25) | (2.79) | (7.62) | (2.09) | (0.60) | (-1.27) | (2.74) |
| EP | 0.38 | 0.23 | 0.33 | 0.33 | 0.32 | 0.38 | 0.24 | 0.34 | 0.31 | 0.31 |
| | (2.19) | (1.76) | (1.28) | (1.35) | (1.66) | (2.30) | (2.14) | (1.32) | (1.12) | (1.78) |
| BP | 0.57 | 0.25 | 0.29 | 0.35 | 0.30 | 0.58 | 0.33 | 0.31 | 0.29 | 0.32 |
| | (2.31) | (1.05) | (1.27) | (1.48) | (1.32) | (2.44) | (1.42) | (1.35) | (1.28) | (1.40) |
| SG | -0.48 | -0.46 | -0.38 | -0.36 | -0.42 | -0.49 | -0.44 | -0.35 | -0.34 | -0.41 |
| | (-2.76) | (-2.85) | (-2.43) | (-2.67) | (-2.81) | (-2.79) | (-2.67) | (-2.32) | (-2.63) | (-2.79) |
| ACC | -0.33 | -0.27 | -0.27 | -0.35 | -0.29 | -0.30 | -0.27 | -0.27 | -0.39 | -0.28 |
| | (-2.40) | (-2.99) | (-3.02) | (-4.15) | (-3.50) | (-2.60) | (-3.32) | (-2.98) | (-4.45) | (-3.44) |
| CAPEX | -1.09 | -1.10 | -1.03 | -0.92 | -1.00 | -0.98 | -1.28 | -1.22 | -1.16 | -0.98 |
| | (-2.85) | (-2.15) | (-1.87) | (-1.51) | (-2.24) | (-2.51) | (-2.44) | (-2.05) | (-1.74) | (-2.21) |

## Table 7. Performance of GIA Estimators using Alternative Fund Skill Measures

Panel A of this table reports the differences in characteristic-adjusted returns between the top and bottom decile portfolios formed on alternatively-constructed GIAs, which are based on fund characteristics and various fund skill measures. These alternative measures include fund expense ratio (FEE), total net assets (TNA), turnover ratio (TURN), industry concentration index (ICI), the fund return gap (GAP), the similarity-based fund performance measure (SIM), returns around subsequent earnings announcements to fund trades (EAR), as well as the R-square of regressing fund returns onto the Fama-French three factors (R2). The portfolios are equal-weighted, held over four quarters (Q1 to Q4) and rebalanced quarterly. We also report the returns to a four-quarter overlapping portfolio strategy following Jegadeesh and Titman (1993) (referred to as *JT4*), where in each quarter, the four portfolios with the same decile rank but formed during the past four different quarters are further combined in equal weights into a single portfolio. Returns are expressed in percentage points. The time-series t-statistics are reported in parentheses. Panel B reports the correlations of these alternatively GIAs as well as the baseline GIA.

Panel A: Performance of Alternative Stock Alpha Estimators

| | Q1 | Q2 | Q3 | Q4 | JT4 |
|---|---|---|---|---|---|
| $\alpha_{FEE}$ | -0.19 | -0.21 | -0.26 | -0.20 | -0.23 |
| | (-1.25) | (-1.32) | (-1.95) | (-1.55) | (-2.17) |
| $\alpha_{TNA}$ | -0.36 | -0.32 | -0.43 | -0.26 | -0.31 |
| | (-2.33) | (-2.36) | (-3.45) | (-1.75) | (-3.05) |
| $\alpha_{TURN}$ | 0.00 | -0.12 | -0.13 | -0.35 | -0.12 |
| | (0.00) | (-0.74) | (-0.76) | (-2.06) | (-0.85) |
| $\alpha_{ICI}$ | 0.36 | -0.09 | 0.32 | 0.41 | 0.27 |
| | (1.85) | (-0.48) | (1.59) | (1.96) | (1.53) |
| $\alpha_{GAP}$ | 0.85 | 0.44 | 0.34 | -0.08 | 0.35 |
| | (3.90) | (2.07) | (1.70) | (-0.38) | (2.36) |
| $\alpha_{SIM}$ | 1.18 | 1.03 | 0.50 | 0.04 | 0.60 |
| | (4.12) | (3.24) | (1.69) | (0.13) | (2.39) |
| $\alpha_{EAR}$ | 0.43 | -0.07 | 0.08 | -0.05 | 0.08 |
| | (3.06) | (-0.49) | (0.75) | (-0.42) | (0.91) |
| $\alpha_{R2}$ | -0.42 | -0.27 | -0.51 | -0.34 | -0.37 |
| | (-2.54) | (-1.54) | (-2.89) | (-1.99) | (-2.55) |

Panel B: Correlations

| | GIA | $\alpha_{FEE}$ | $\alpha_{TNA}$ | $\alpha_{TURN}$ | $\alpha_{ICI}$ | $\alpha_{GAP}$ | $\alpha_{SIM}$ | $\alpha_{EAR}$ |
|---|---|---|---|---|---|---|---|---|
| $\alpha_{FEE}$ | 0.00 | | | | | | | |
| | (0.30) | | | | | | | |
| $\alpha_{TNA}$ | -0.01 | 0.03 | | | | | | |
| | (-1.19) | (2.67) | | | | | | |
| $\alpha_{TURN}$ | -0.01 | 0.30 | 0.08 | | | | | |
| | (-1.57) | (23.20) | (6.35) | | | | | |
| $\alpha_{ICI}$ | 0.02 | 0.11 | -0.04 | 0.00 | | | | |
| | (1.94) | (18.32) | (-7.81) | (0.20) | | | | |
| $\alpha_{GAP}$ | 0.41 | 0.00 | -0.01 | -0.06 | 0.02 | | | |
| | (25.79) | (0.21) | (-0.98) | (-5.05) | (1.90) | | | |
| $\alpha_{SIM}$ | 0.75 | -0.01 | 0.00 | 0.01 | -0.03 | 0.32 | | |
| | (83.20) | (-0.80) | (-0.19) | (1.37) | (-3.02) | (17.97) | | |
| $\alpha_{EAR}$ | 0.07 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.04 | |
| | (8.85) | (4.02) | (2.88) | (2.83) | (2.72) | (3.53) | (5.00) | |
| $\alpha_{R2}$ | -0.03 | 0.11 | 0.35 | 0.16 | -0.27 | 0.01 | 0.00 | 0.01 |
| | (-2.77) | (9.19) | (30.06) | (12.79) | (-35.39) | (0.50) | (-0.43) | (1.87) |

## Table 8. Fama-MacBeth Regressions with Alternatively Estimated GIAs

This table reports the results of the following Fama-MacBeth regressions. The dependent variables are the stock returns during the four quarters Q1 to Q4. The explanatory variables include the baseline GIA (or LGIA) and the alternative GIAs ($\alpha$) based on various fund skills measures, including FEE, TNA, TURN, ICI, GAP, SIM, R2, and EAR. PRIN1, PRIN2, and PRIN3 are the three principle components of the baseline GIA, $\alpha_{GAP}$, and $\alpha_{SIM}$ (or their variations based on lagged holdings). We report the time-series averages of the estimated coefficients as well as the time-series t-statistics (in parentheses). Coefficients reported under "JT4" are those averaged over four different regressions with the stock returns (the dependent variable) in the same quarter but the explanatory variables in past four different quarters. The regression intercepts are not reported.

| | Based on Current Fund Holdings | | | | | Based on Lagged Holdings | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q4 | JT4 | Q1 | Q2 | Q3 | Q4 | JT4 |
| PRIN1 | 0.26 | 0.19 | 0.09 | -0.01 | 0.12 | 0.24 | 0.14 | 0.06 | 0.00 | 0.10 |
| | (5.46) | (3.79) | (1.88) | (-0.25) | (3.13) | (4.85) | (2.52) | (1.73) | (0.03) | (2.25) |
| PRIN2 | -0.06 | -0.03 | 0.00 | -0.06 | -0.04 | 0.02 | -0.03 | -0.05 | -0.05 | -0.02 |
| | (-0.75) | (-0.36) | (0.04) | (-0.97) | (-0.72) | (0.54) | (-0.79) | (-1.57) | (-1.30) | (-0.82) |
| PRIN3 | 0.20 | 0.17 | 0.11 | 0.08 | 0.16 | -0.06 | -0.02 | -0.04 | -0.03 | -0.04 |
| | (1.68) | (1.45) | (0.83) | (0.57) | (1.63) | (-1.74) | (-0.71) | (-1.05) | (-0.81) | (-1.55) |
| $\alpha_{FEE}$ | 0.00 | -0.04 | -0.04 | -0.03 | -0.02 | 0.03 | -0.03 | -0.02 | -0.08 | -0.02 |
| | (0.06) | (-1.07) | (-1.22) | (-0.90) | (-0.76) | (0.68) | (-0.71) | (-0.55) | (-1.99) | (-0.62) |
| $\alpha_{TNA}$ | -0.10 | -0.05 | -0.03 | 0.00 | -0.04 | 0.00 | 0.01 | 0.06 | 0.04 | 0.03 |
| | (-2.86) | (-1.35) | (-0.87) | (0.09) | (-1.55) | (-0.02) | (0.18) | (1.10) | (0.88) | (0.69) |
| $\alpha_{TURN}$ | -0.01 | -0.03 | 0.00 | -0.08 | -0.02 | 0.05 | 0.08 | -0.02 | 0.10 | 0.05 |
| | (-0.22) | (-0.70) | (0.01) | (-1.91) | (-0.69) | (0.56) | (1.04) | (-0.31) | (1.63) | (1.08) |
| $\alpha_{ICI}$ | 0.09 | 0.01 | 0.03 | 0.05 | 0.06 | 0.01 | 0.02 | 0.05 | 0.03 | 0.03 |
| | (1.97) | (0.20) | (0.62) | (0.92) | (1.13) | (0.14) | (0.42) | (1.08) | (0.60) | (0.74) |
| $\alpha_{EAR}$ | 0.10 | 0.03 | 0.02 | -0.02 | 0.03 | 0.01 | 0.01 | -0.02 | -0.01 | -0.00 |
| | (3.05) | (1.01) | (0.97) | (-0.56) | (1.48) | (0.30) | (0.59) | (-0.86) | (-0.31) | (-0.25) |
| $\alpha_{R2}$ | -0.06 | -0.04 | -0.10 | -0.07 | -0.06 | -0.04 | -0.04 | -0.10 | -0.10 | -0.07 |
| | (-1.21) | (-0.82) | (-2.07) | (-1.66) | (-1.62) | (-0.79) | (-0.86) | (-2.27) | (-2.21) | (-1.78) |

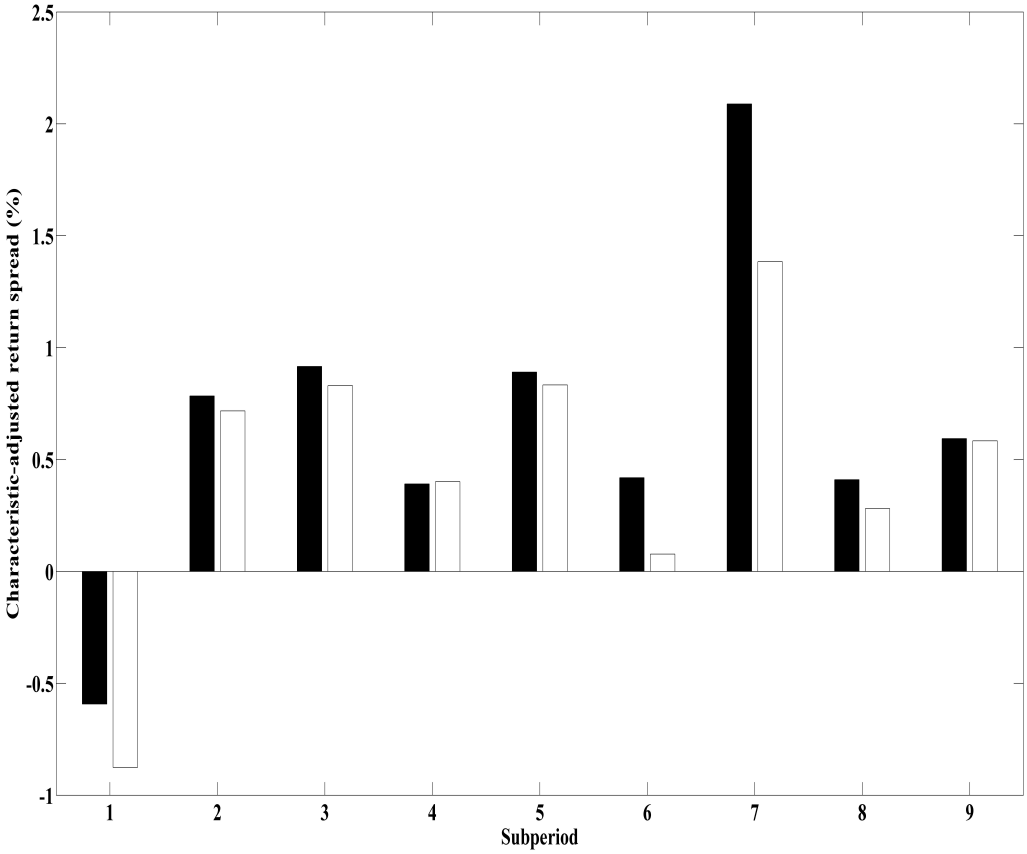


**Figure 1**

We plot the average quarterly characteristic-adjusted return spreads between the top and bottom decile portfolios formed on GIA and LGIA during nine subperiods. Each subperiod consists of three years, with the first subperiod being 1980-1982 and the last subperiod being 2004-2006. In each quarter we form equal-weighted decile portfolios on GIA and LGIA. The portfolios are held for four quarters and rebalanced quarterly. The four portfolios with the same decile rank but formed in the four different quarters are combined in equal-weights into a single portfolio (i.e., JT4). The performance within a subperiod is computed as the average quarterly top-bottom differences in characteristic-adjusted returns of the JT4 portfolios formed during the subperiod. For each period, the left-side bar represents the performance of the GIA strategy. The right-side bar represents the performance of the LGIA strategy.

# CFR working paper series

centre for financial research cologne

CFR working papers are available for download from www.cfr-cologne.de.

Hardcopies can be ordered from: centre for financial research (CFR), Albertus Magnus Platz, 50923 Koeln, Germany.

2012

| No. | Author(s) | Title |
|---|---|---|
| 12-12 | Y. Gündüz, J. Nasev, M. Trapp | The Price Impact of CDS Trading |
| 12-11 | Y. Wu, R. Wermers, J. Zechner | Governance and Shareholder Value in Delegated Portfolio Management: The Case of Closed-End Funds |
| 12-10 | M. Trapp, C. Wewel | Transatlantic Systemic Risk |
| 12-09 | G. Cici, A. Kempf, C. Sorhage | Are Financial Advisors Useful? Evidence from Tax-Motivated Mutual Fund Flows |
| 12-08 | S. Jank | Changes in the composition of publicly traded firms: Implications for the dividend-price ratio and return predictability |
| 12-07 | G. Cici, C. Rosenfeld | The Investment Abilities of Mutual Fund Buy-Side Analysts |
| 12-06 | A. Kempf, A. Pütz, F. Sonnenburg | Fund Manager Duality: Impact on Performance and Investment Behavior |
| 12-05 | R. Wermers | Runs on Money Market Mutual Funds |
| 12-04 | R. Wermers | A matter of style: The causes and consequences of style drift in institutional portfolios |
| 12-03 | C. Andres, A. Betzer, I. van den Bongard, C. Haesner, E. Theissen | Dividend Announcements Reconsidered: Dividend Changes versus Dividend Surprises |
| 12-02 | C. Andres, E. Fernau, E. Theissen | Is It Better To Say Goodbye? When Former Executives Set Executive Pay |
| 12-01 | L. Andreu, A. Pütz | Are Two Business Degrees Better Than One? Evidence from Mutual Fund Managers' Education |

2011

| No. | Author(s) | Title |
|---|---|---|
| 11-16 | V. Agarwal, J.-P. Gómez, R. Priestley | Management Compensation and Market Timing under Portfolio Constraints |
| 11-15 | T. Dimpfl, S. Jank | Can Internet Search Queries Help to Predict Stock Market Volatility? |

| | | |
|---|---|---|
| 11-14 | P. Gomber,<br>U. Schweickert,<br>E. Theissen | Liquidity Dynamics in an Electronic Open Limit Order Book:<br>An Event Study Approach |
| 11-13 | D. Hess, S. Orbe | Irrationality or Efficiency of Macroeconomic Survey Forecasts?<br>Implications from the Anchoring Bias Test |
| 11-12 | D. Hess, P. Immenkötter | Optimal Leverage, its Benefits, and the Business Cycle |
| 11-11 | N. Heinrichs, D. Hess,<br>C. Homburg, M. Lorenz,<br>S. Sievers | Extended Dividend, Cash Flow and Residual Income Valuation<br>Models – Accounting for Deviations from Ideal Conditions |
| 11-10 | A. Kempf, O. Korn,<br>S. Saßning | Portfolio Optimization using Forward - Looking Information |
| 11-09 | V. Agarwal, S. Ray | Determinants and Implications of Fee Changes in the Hedge Fund Industry |
| 11-08 | G. Cici, L.-F. Palacios | On the Use of Options by Mutual Funds: Do They Know What They Are Doing? |
| 11-07 | V. Agarwal, G. D. Gay,<br>L. Ling | Performance inconsistency in mutual funds: An investigation of window-dressing behavior |
| 11-06 | N. Hautsch, D. Hess,<br>D. Veredas | The Impact of Macroeconomic News on Quote Adjustments,<br>Noise, and Informational Volatility |
| 11-05 | G. Cici | The Prevalence of the Disposition Effect in Mutual Funds' Trades |
| 11-04 | S. Jank | Mutual Fund Flows, Expected Returns and the Real Economy |
| 11-03 | G.Fellner, E.Theissen | Short Sale Constraints, Divergence of Opinion and Asset Value: Evidence from the Laboratory |
| 11-02 | S.Jank | Are There Disadvantaged Clienteles in Mutual Funds? |
| 11-01 | V. Agarwal, C. Meneghetti | The Role of Hedge Funds as Primary Lenders |

2010

| No. | Author(s) | Title |
|---|---|---|
| 10-20 | G. Cici, S. Gibson,<br>J.J. Merrick Jr. | Missing the Marks? Dispersion in Corporate Bond Valuations Across Mutual Funds |
| 10-19 | J. Hengelbrock,<br>E. Theissen, C. Westheide | Market Response to Investor Sentiment |
| 10-18 | G. Cici, S. Gibson | The Performance of Corporate-Bond Mutual Funds:<br>Evidence Based on Security-Level Holdings |
| 10-17 | D. Hess, D. Kreutzmann,<br>O. Pucker | Projected Earnings Accuracy and the Profitability of Stock Recommendations |
| 10-16 | S. Jank, M. Wedow | Sturm und Drang in Money Market Funds: When Money Market Funds Cease to Be Narrow |
| 10-15 | G. Cici, A. Kempf, A. Puetz | The Valuation of Hedge Funds' Equity Positions |
| 10-14 | J. Grammig, S. Jank | Creative Destruction and Asset Prices |
| 10-13 | S. Jank, M. Wedow | Purchase and Redemption Decisions of Mutual Fund Investors and the Role of Fund Families |

| | | |
|---|---|---|
| 10-12 | S. Artmann, P. Finter, A. Kempf, S. Koch, E. Theissen | The Cross-Section of German Stock Returns: New Data and New Evidence |
| 10-11 | M. Chesney, A. Kempf | The Value of Tradeability |
| 10-10 | S. Frey, P. Herbst | The Influence of Buy-side Analysts on Mutual Fund Trading |
| 10-09 | V. Agarwal, W. Jiang, Y. Tang, B. Yang | Uncovering Hedge Fund Skill from the Portfolio Holdings They Hide |
| 10-08 | V. Agarwal, V. Fos, W. Jiang | Inferring Reporting Biases in Hedge Fund Databases from Hedge Fund Equity Holdings |
| 10-07 | V. Agarwal, G. Bakshi, J. Huij | Do Higher-Moment Equity Risks Explain Hedge Fund Returns? |
| 10-06 | J. Grammig, F. J. Peter | Tell-Tale Tails |
| 10-05 | K. Drachter, A. Kempf | Höhe, Struktur und Determinanten der Managervergütung-Eine Analyse der Fondsbranche in Deutschland |
| 10-04 | J. Fang, A. Kempf, M. Trapp | Fund Manager Allocation |
| 10-03 | P. Finter, A. Niessen-Ruenzi, S. Ruenzi | The Impact of Investor Sentiment on the German Stock Market |
| 10-02 | D. Hunter, E. Kandel, S. Kandel, R. Wermers | Endogenous Benchmarks |
| 10-01 | S. Artmann, P. Finter, A. Kempf | Determinants of Expected Stock Returns: Large Sample Evidence from the German Market |

2009

| No. | Author(s) | Title |
|---|---|---|
| 09-17 | E. Theissen | Price Discovery in Spot and Futures Markets: A Reconsideration |
| 09-16 | M. Trapp | Trading the Bond-CDS Basis – The Role of Credit Risk and Liquidity |
| 09-15 | A. Betzer, J. Gider, D.Metzger, E. Theissen | Strategic Trading and Trade Reporting by Corporate Insiders |
| 09-14 | A. Kempf, O. Korn, M. Uhrig-Homburg | The Term Structure of Illiquidity Premia |
| 09-13 | W. Bühler, M. Trapp | Time-Varying Credit Risk and Liquidity Premia in Bond and CDS Markets |
| 09-12 | W. Bühler, M. Trapp | Explaining the Bond-CDS Basis – The Role of Credit Risk and Liquidity |
| 09-11 | S. J. Taylor, P. K. Yadav, Y. Zhang | Cross-sectional analysis of risk-neutral skewness |
| 09-10 | A. Kempf, C. Merkle, A. Niessen-Ruenzi | Low Risk and High Return – Affective Attitudes and Stock Market Expectations |
| 09-09 | V. Fotak, V. Raman, P. K. Yadav | Naked Short Selling: The Emperor`s New Clothes? |
| 09-08 | F. Bardong, S.M. Bartram, P.K. Yadav | Informed Trading, Information Asymmetry and Pricing of Information Risk: Empirical Evidence from the NYSE |

| | | |
|---|---|---|
| 09-07 | S. J. Taylor , P. K. Yadav, Y. Zhang | The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks |
| 09-06 | S. Frey, P. Sandas | The Impact of Iceberg Orders in Limit Order Books |
| 09-05 | H. Beltran-Lopez, P. Giot, J. Grammig | Commonalities in the Order Book |
| 09-04 | J. Fang, S. Ruenzi | Rapid Trading bei deutschen Aktienfonds: Evidenz aus einer großen deutschen Fondsgesellschaft |
| 09-03 | A. Banegas, B. Gillen, A. Timmermann, R. Wermers | The Cross-Section of Conditional Mutual Fund Performance in European Stock Markets |
| 09-02 | J. Grammig, A. Schrimpf, M. Schuppli | Long-Horizon Consumption Risk and the Cross-Section of Returns: New Tests and International Evidence |
| 09-01 | O. Korn, P. Koziol | The Term Structure of Currency Hedge Ratios |

2008

| No. | Author(s) | Title |
|---|---|---|
| 08-12 | U. Bonenkamp, C. Homburg, A. Kempf | Fundamental Information in Technical Trading Strategies |
| 08-11 | O. Korn | Risk Management with Default-risky Forwards |
| 08-10 | J. Grammig, F.J. Peter | International Price Discovery in the Presence of Market Microstructure Effects |
| 08-09 | C. M. Kuhnen, A. Niessen | Public Opinion and Executive Compensation |
| 08-08 | A. Pütz, S. Ruenzi | Overconfidence among Professional Investors: Evidence from Mutual Fund Managers |
| 08-07 | P. Osthoff | What matters to SRI investors? |
| 08-06 | A. Betzer, E. Theissen | Sooner Or Later: Delays in Trade Reporting by Corporate Insiders |
| 08-05 | P. Linge, E. Theissen | Determinanten der Aktionärspräsenz auf Hauptversammlungen deutscher Aktiengesellschaften |
| 08-04 | N. Hautsch, D. Hess, C. Müller | Price Adjustment to News with Uncertain Precision |
| 08-03 | D. Hess, H. Huang, A. Niessen | How Do Commodity Futures Respond to Macroeconomic News? |
| 08-02 | R. Chakrabarti, W. Megginson, P. Yadav | Corporate Governance in India |
| 08-01 | C. Andres, E. Theissen | Setting a Fox to Keep the Geese - Does the Comply-or-Explain Principle Work? |

2007

| No. | Author(s) | Title |
|---|---|---|
| 07-16 | M. Bär, A. Niessen, S. Ruenzi | The Impact of Work Group Diversity on Performance: Large Sample Evidence from the Mutual Fund Industry |

| | | |
|---|---|---|
| 07-15 | A. Niessen, S. Ruenzi | Political Connectedness and Firm Performance: Evidence From Germany |
| 07-14 | O. Korn | Hedging Price Risk when Payment Dates are Uncertain |
| 07-13 | A. Kempf, P. Osthoff | SRI Funds: Nomen est Omen |
| 07-12 | J. Grammig, E. Theissen, O. Wuensche | Time and Price Impact of a Trade: A Structural Approach |
| 07-11 | V. Agarwal, J. R. Kale | On the Relative Performance of Multi-Strategy and Funds of Hedge Funds |
| 07-10 | M. Kasch-Haroutounian, E. Theissen | Competition Between Exchanges: Euronext versus Xetra |
| 07-09 | V. Agarwal, N. D. Daniel, N. Y. Naik | Do hedge funds manage their reported returns? |
| 07-08 | N. C. Brown, K. D. Wei, R. Wermers | Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices |
| 07-07 | A. Betzer, E. Theissen | Insider Trading and Corporate Governance: The Case of Germany |
| 07-06 | V. Agarwal, L. Wang | Transaction Costs and Value Premium |
| 07-05 | J. Grammig, A. Schrimpf | Asset Pricing with a Reference Level of Consumption: New Evidence from the Cross-Section of Stock Returns |
| 07-04 | V. Agarwal, N.M. Boyson, N.Y. Naik | Hedge Funds for retail investors? An examination of hedged mutual funds |
| 07-03 | D. Hess, A. Niessen | The Early News Catches the Attention: On the Relative Price Impact of Similar Economic Indicators |
| 07-02 | A. Kempf, S. Ruenzi, T. Thiele | Employment Risk, Compensation Incentives and Managerial Risk Taking - Evidence from the Mutual Fund Industry - |
| 07-01 | M. Hagemeister, A. Kempf | CAPM und erwartete Renditen: Eine Untersuchung auf Basis der Erwartung von Marktteilnehmern |

2006

| No. | Author(s) | Title |
|---|---|---|
| 06-13 | S. Čeljo-Hörhager, A. Niessen | How do Self-fulfilling Prophecies affect Financial Ratings? - An experimental study |
| 06-12 | R. Wermers, Y. Wu, J. Zechner | Portfolio Performance, Discount Dynamics, and the Turnover of Closed-End Fund Managers |
| 06-11 | U. v. Lilienfeld-Toal, S. Ruenzi | Why Managers Hold Shares of Their Firm: An Empirical Analysis |
| 06-10 | A. Kempf, P. Osthoff | The Effect of Socially Responsible Investing on Portfolio Performance |
| 06-09 | R. Wermers, T. Yao, J. Zhao | Extracting Stock Selection Information from Mutual Fund holdings: An Efficient Aggregation Approach |
| 06-08 | M. Hoffmann, B. Kempa | The Poole Analysis in the New Open Economy Macroeconomic Framework |
| 06-07 | K. Drachter, A. Kempf, M. Wagner | Decision Processes in German Mutual Fund Companies: Evidence from a Telephone Survey |
| 06-06 | J.P. Krahnen, F.A. Schmid, E. Theissen | Investment Performance and Market Share: A Study of the German Mutual Fund Industry |
| 06-05 | S. Ber, S. Ruenzi | On the Usability of Synthetic Measures of Mutual Fund Net-Flows |

| | | |
|---|---|---|
| 06-04 | A. Kempf, D. Mayston | Liquidity Commonality Beyond Best Prices |
| 06-03 | O. Korn, C. Koziol | Bond Portfolio Optimization: A Risk-Return Approach |
| 06-02 | O. Scaillet, L. Barras, R. Wermers | False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas |
| 06-01 | A. Niessen, S. Ruenzi | Sex Matters: Gender Differences in a Professional Setting |

2005

| No. | Author(s) | Title |
|---|---|---|
| 05-16 | E. Theissen | An Analysis of Private Investors´ Stock Market Return Forecasts |
| 05-15 | T. Foucault, S. Moinas, E. Theissen | Does Anonymity Matter in Electronic Limit Order Markets |
| 05-14 | R. Kosowski, A. Timmermann, R. Wermers, H. White | Can Mutual Fund „Stars“ Really Pick Stocks? New Evidence from a Bootstrap Analysis |
| 05-13 | D. Avramov, R. Wermers | Investing in Mutual Funds when Returns are Predictable |
| 05-12 | K. Griese, A. Kempf | Liquiditätsdynamik am deutschen Aktienmarkt |
| 05-11 | S. Ber, A. Kempf, S. Ruenzi | Determinanten der Mittelzuflüsse bei deutschen Aktienfonds |
| 05-10 | M. Bär, A. Kempf, S. Ruenzi | Is a Team Different From the Sum of Its Parts? Evidence from Mutual Fund Managers |
| 05-09 | M. Hoffmann | Saving, Investment and the Net Foreign Asset Position |
| 05-08 | S. Ruenzi | Mutual Fund Growth in Standard and Specialist Market Segments |
| 05-07 | A. Kempf, S. Ruenzi | Status Quo Bias and the Number of Alternatives - An Empirical Illustration from the Mutual Fund Industry |
| 05-06 | J. Grammig, E. Theissen | Is Best Really Better? Internalization of Orders in an Open Limit Order Book |
| 05-05 | H. Beltran-Lopez, J. Grammig, A.J. Menkveld | Limit order books and trade informativeness |
| 05-04 | M. Hoffmann | Compensating Wages under different Exchange rate Regimes |
| 05-03 | M. Hoffmann | Fixed versus Flexible Exchange Rates: Evidence from Developing Countries |
| 05-02 | A. Kempf, C. Memmel | Estimating the Global Minimum Variance Portfolio |
| 05-01 | S. Frey, J. Grammig | Liquidity supply and adverse selection in a pure limit order book market |

2004

| No. | Author(s) | Title |
|---|---|---|
| 04-10 | N. Hautsch, D. Hess | Bayesian Learning in Financial Markets – Testing for the Relevance of Information Precision in Price Discovery |
| 04-09 | A. Kempf, K. Kreuzberg | Portfolio Disclosure, Portfolio Selection and Mutual Fund Performance Evaluation |
| 04-08 | N.F. Carline, S.C. Linn, P.K. Yadav | Operating performance changes associated with corporate mergers and the role of corporate governance |

| | | |
|---|---|---|
| 04-07 | J.J. Merrick, Jr., N.Y. Naik, P.K. Yadav | Strategic Trading Behaviour and Price Distortion in a Manipulated Market: Anatomy of a Squeeze |
| 04-06 | N.Y. Naik, P.K. Yadav | Trading Costs of Public Investors with Obligatory and Voluntary Market-Making: Evidence from Market Reforms |
| 04-05 | A. Kempf, S. Ruenzi | Family Matters: Rankings Within Fund Families and Fund Inflows |
| 04-04 | V. Agarwal, N.D. Daniel, N.Y. Naik | Role of Managerial Incentives and Discretion in Hedge Fund Performance |
| 04-03 | V. Agarwal, W.H. Fung, J.C. Loon, N.Y. Naik | Risk and Return in Convertible Arbitrage: Evidence from the Convertible Bond Market |
| 04-02 | A. Kempf, S. Ruenzi | Tournaments in Mutual Fund Families |
| 04-01 | I. Chowdhury, M. Hoffmann, A. Schabert | Inflation Dynamics and the Cost Channel of Monetary Transmission |

centre for financial research
cfr/university of cologne
Albertus-Magnus-Platz
D-50923 cologne
Fon +49(0)221-470-6995
Fax +49(0)221-470-3992
kempf@cfr-cologne.de
www.cfr-cologne.de