North American Journal of Economics and Finance 40 (2017) 222–237

Contents lists available at ScienceDirect

North American Journal of Economics and Finance

journal homepage: www.elsevier.com/locate/ecofin

Precision about manager skill, mutual fund flows, and performance persistence ☆

Hyunglae Jeon [a], Jangkoo Kang [b], Changjun Lee [c,*]

[a] Samsung Asset Management, Seoul, Republic of Korea
[b] College of Business, Korea Advanced Institute of Science and Technology (KAIST), Seoul, Republic of Korea
[c] College of Business, Hankuk University of Foreign Studies, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 14 April 2016
Received in revised form 7 March 2017
Accepted 8 March 2017
Available online 24 March 2017

JEL classification:
G11
G12

Keywords:
Mutual fund flows
Flow-performance relationship
Precision about manager skill
Investor behavior

ABSTRACT

Constructing a measure for imprecision about manager skill, *IMP*, by standard deviation of multiple performance-based signals, we examine the role of *IMP* in flow-performance relationship. We find that future fund flows respond more (less) strongly to past performance measured with high (low) precision. Our finding is robust to the inclusion of control variables that are known to affect future fund flows. In addition, short-term performance persistence is observed among funds with lower *IMP*, implying that our *IMP* helps identify funds with performance persistence. Overall, our findings indicate that *IMP* plays a critical role in understanding the cross-section of future flows.

1. Introduction

Understanding the behavior of fund investors has long been a central research question. From the perspective of investors, one of the most important factors for investment decisions is the past performance of mutual funds. Since the investment decisions of mutual funds investors are reflected in fund flows, numerous researchers have investigated the relationship between fund performance and future fund flows. The literature has documented performance-chasing behavior in fund investors (Chevalier & Ellison, 1997; Goetzmann & Peles, 1997; Gruber, 1996; Ippolito, 1992; Sirri & Tufano, 1998); that is, funds with superior past performance attract disproportionately large money inflows.

In analyzing flow-performance relationship, however, previous studies have not paid attention to one important factor: the precision about past performance, although it should be included in the information set of mutual fund investors. Our motivation is related to the fact that mutual fund investors today receive multiple performance-based signals of uncertain precision.[1] Since past performance estimated from each measure is, at best, a noisy signal about managerial skill, for mutual

☆ We thank the editor and the two anonymous referees for helpful suggestion that have greatly improved the paper. This work was supported by Hankuk University of Foreign Studies Research Fund.

* Corresponding author.
E-mail addresses: hyunglaejeon@gmail.com (H. Jeon), jkkang@business.kaist.ac.kr (J. Kang), leechangjun@hufs.ac.kr (C. Lee).

[1] Indeed, numerous measures have emerged in the performance evaluation literature (Carhart, 1997; Daniel, Grinblatt, Titman, & Wermers, 1997; Elton, Gruber, Das, & Hlavka, 1993; Grinblatt & Titman, 1989a, 1989b; Ippolito, 1989; Jensen, 1968; Lehmann & Modest, 1987; Sharpe, 1966, 1992). In addition, different studies adopt different evaluation periods (Bollen & Busse, 2005; Carhart, 1997; Elton et al., 1993).

http://dx.doi.org/10.1016/j.najef.2017.03.002
1062-9408/

fund investors, it is unclear what particular measure is the most appropriate. Sirri and Tufano (1998, p. 1597) express their concern on the issue: "Even if performance and riskiness affect fund flows, it is unclear what particular measures and levels of performance (or risk) are most salient to consumers, or over what time period this measure should be calculated."

Therefore, a promising research direction is to investigate how investors take into account multiple signals of uncertain precision in investment decisions.

A key element of our work lies in constructing a proxy for imprecision about manager skill. Since the choices of the particular performance measure and the performance evaluation window affect past performance, we consider both these dimensions.[2] We use seven well-known fund performance measures proposed in both the academic and practice-oriented literature (performance relative to the market benchmark, Sharpe's ratio, Jensen's alpha, Treynor's ratio, information ratio, appraisal ratio, and Carhart's four-factor alpha) over multiple evaluation windows. We employ five past performance evaluation windows of one, two, three, four, and five years, respectively. Given 35 distinct past performance measures for each fund in month t, we define our measure of the imprecision of multiple performance-based signals, *imprecision* (*IMP*, hereafter), as the standard deviation of past performance rankings since the standard deviation has been the most widely used measure in estimating such imprecision.

Our paper provides new evidence on the mutual fund literature, focusing on the two following questions. First, we ask whether *IMP* affects the sensitivity of flow-performance relationship. Consider a Bayesian investor who receives multiple performance-based signals of uncertain precision about managerial skill. For this investor, it is unclear what particular measure is the most salient for investment decisions. Under the circumstances, the Bayesian investor is less likely to respond to signal measured with lower precision, regardless of the content of the signal. This means that when past performance is measured with lower precision, we expect to observe lower flow-performance sensitivity, i.e., less inflows for good past performance funds and less outflows for poor past performance funds. On the contrary, when past performance is measured with high precision, we expect to observe higher flow-performance sensitivity.

We find that future fund flows respond more (less) strongly to past performance measured with high (low) precision. For funds with the lowest *IMP*, the difference in the average future flow between funds with winners (top 20%) and losers (bottom 20%) is 3.84%, with a Newey–West t-statistic of 7.93. The difference monotonically decreases as *IMP* increases, and the difference in the average future flow between funds with winners and losers is 1.98% with a Newey–West t-statistic of 9.32 for funds with the highest *IMP*. Fama and MacBeth (1973) regressions confirm our finding. The interaction term of past performance and *IMP* is negative and statistically significant at the 1% level, implying that flow-performance relationship becomes weaker as imprecision of past performance increases. Our finding is robust to the inclusion of control variables such as fund size, total expense ratios, the logarithm of fund age, fund turnover ratios, and the flow of the previous month. We also confirm that our results are retained in the presence of convexity in flow-performance relationship.

Second, we investigate whether our proxy for imprecision about manager skill helps identify funds with strong performance persistence. This research question arises because *IMP*, by construction, has direct information about the skill of money managers. That is, funds with lower *IMP* have lower uncertainty about manager skill, whether the skill is good or bad, compared to funds with higher *IMP*. We find that performance persistence is observed among funds with lower *IMP*. This finding is intuitive because, among past losers, funds with lower *IMP* have a higher probability of being managed by money managers with low skills than funds with higher *IMP* do. If we apply the same reasoning, we should observe significant positive returns on a portfolio formed by the intersection of the best-performing group and the lowest *IMP* group. However, we fail to find this empirical evidence. While we observe that investors put disproportionately large money into funds with the highest past performance and the lowest *IMP*, our result indicates that these funds do not outperform a passive benchmark. Given that equity funds, on the average, do not outperform the benchmark, such a finding is not surprising. Overall, our empirical findings uncover a critical role for imprecision about manager skill in understanding the cross-section of mutual fund flows and performance persistence.

Our study is related with two recent works by Huang, Wei, and Yan (2012) and Li, Tiwari, and Tong (2016). Huang et al. (2012) show that flow sensitivity to past performance is decreasing with fund return volatility. Our paper differs from theirs in that, while we estimate imprecision about managerial skill by calculating the cross-sectional dispersion of 35 different performance measures, they obtain volatility by computing the time-series dispersion of monthly raw returns of mutual funds. In addition, Li et al. (2016) show that ambiguity-averse mutual fund investors tend to react more strongly to the worst-case scenario. Since the aforementioned studies contribute to the literature, we compare the explanatory power of our *IMP* with those suggested by recent works. We find that the inclusion of competing variables does not drive out the role of *IMP* in explaining flow-performance relationship. Our finding indicates that *IMP* conveys information relevant to explaining the cross-section of future fund flows.

This work contributes to our understanding of the investment decisions of mutual fund investors. An extensive literature studies mutual fund flows to examine the behavior of fund investors. Several empirical works suggest that investors make investment decisions based on past fund performance (e.g., Chevalier & Ellison, 1997; Goetzmann & Peles, 1997; Gruber, 1996; Ippolito, 1992; Sirri & Tufano, 1998) and it is now well-known that fund investors chase good past performance. In

[2] Previous studies either consider different performance measures or various time horizons in flow-performance relationship. However, for investors, it is hard to prioritize which dimension is more important, and what they understand is that both the two dimensions affect past performance. Since past performance across different performance measures and time horizons are readily available, we think that it is natural for investors to consider both dimensions.

addition, theoretical models show that investors learn about fund manager skill through past performance. Berk and Green (2004) explain performance chasing behavior combined with a lack of performance persistence. They propose a rational learning model that considers investor learning about managerial skill and diseconomies of scale in portfolio management. Developing this rational learning model, Lynch and Musto (2003) demonstrate convexity in the flow–performance relation by considering managers' incentives to abandon unsuccessful strategies. Huang, Wei, and Yan (2007) also explain the convexity in the flow–performance relation by incorporating investor participation costs. Pastor and Stambaugh (2012) model learning about a parameter governing the degree of decreasing returns to scale. However, none of these studies considers how investors take into account uncertain precision about past fund performance in their investment decisions. Our empirical results uncover previously unexplored implications of the behavior of mutual fund investors by documenting a critical role for precision about manager skill in mutual funds.

The remainder of this paper is organized as follows. Section 2 describes the data and sample selection criteria. Section 3 explains the construction of our proxy for imprecision about manager skill. Section 4 reports the empirical findings. Finally, Section 5 summarizes and presents our conclusions.

2. Data and sample selection

We obtain monthly data on mutual funds from the Center for Research in Security Prices (CRSP) Survivorship Bias Free Mutual Fund Database. Our data set spans from January 1980 to December 2012. The CRSP mutual fund database has fund returns, total net assets (TNA), different types of fees, investment objectives, and other fund characteristics. The CRSP returns are net after fees, expenses, and brokerage commissions but before any front-end or back-end loads.

Since we study actively managed U.S. equity mutual funds, we construct our sample of funds as follows. First, based on the Wiesenberger and Lipper objective codes, we select funds with the following Wiesenberger objectives: G, SCG, LTG, MCG, GCI, or IEQ or the following Lipper objectives: EI, CA, G, GI, MC, MR, or SG. When both the Wiesenberger and Lipper codes are unclassified or missing, we include funds with a strategic insight objective code of AGC, GMC, GRI, GRO, ING, or SCG into our sample. If the fund style cannot be identified, it is excluded from our sample. Second, we exclude index funds by deleting those whose name includes the word *index* or the term *ind*, *S&P*, *DOW*, *Wilshire*, or *Russell*. Third, for consistency with prior studies, we exclude sector funds, international funds, bond funds, balanced funds, and exchange-traded funds. Finally, we require funds to have at least 80% of their assets in common stocks.

Reviewing the related literature, we use an additional screening process. First, following Elton, Gruber, and Blake (1996), we require funds to have TNA of at least $15 million since the inclusion of smaller funds could cause a survivorship bias problem. Second, we exclude observations before the fund's starting year reported by the CRSP to address Evans's (2010) comment on incubation bias. Third, following Cremers and Petajisto (2009), we eliminate funds with missing fund names in the CRSP from our analysis. Finally, when a fund has multiple share classes in the CRSP database, we aggregate all the observations pertaining to different share classes into one observation. For the qualitative attributes of funds (e.g., name, objective, year of origination), we retain the observation of the oldest fund. For TNA, we sum the TNAs of the different share classes. For the other quantitative fund attributes (e.g., return, expense, turnover ratio), we compute the weighted average of the attributes of the individual share classes, the weights being the lagged TNA of the share class in question.

Our final sample includes 3,881 distinct funds and 349,192 fund–month observations. Panel A of Table 1 reports summary statistics for fund characteristics. The average net return is 0.64% per month and the average monthly net flow into funds is 0.24%. The average values of the other fund characteristics are very similar to those in related studies, giving us confidence that our sample selection is not biased. For example, using 3,416 actively managed U.S. equity mutual funds from April 1980 to June 2012, Busse, Chordia, Jiang, and Tang (2015) report that the average TNA, fund age, expense ratio, and turnover are $838 million, 16.8 years, 1.16%, and 84.3%, respectively, in their sample.

3. Construction of the imprecision measure

3.1. Measures of flow and performance

We examine whether information contained in multiple performance-based measures affect the flow-performance relationship. To this end, we measure the monthly net flow into funds and past performance. Consistent with prior studies, we define monthly net flow into a fund as

$$Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t})}{TNA_{i,t-1}} \quad (1)$$

where $R_{i,t}$ denotes fund *i*'s return net of expenses during month *t* and $TNA_{i,t}$ is the fund's TNA at the end of month *t*. To mitigate the impact of outliers, we eliminate the largest and smallest 1% of flow observations from our sample. Our definition of flows reflects the percentage growth of the fund's assets after adjusting for the appreciation of the mutual fund's assets, assuming that all cash flows are invested at the end of the period.

Table 1
Summary statistics on actively managed U.S. equity funds. Panel A shows summary statistics on actively managed U.S. equity funds. The net return (%) is the monthly return after fees, expenses, and brokerage commissions, TNA is the fund's TNA (in millions of dollars), fund age is the number of years since the fund was first offered, expenses is the annual expense ratio (%), turnover is the minimum of aggregated sales or aggregated purchases of securities divided by the average twelve-month TNA of the fund, and flow is defined as $Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1+R_{i,t})}{TNA_{i,t-1}}$. Panel B reports the average fund performance estimated from the seven measures (Sharpe ratio, appraisal ratio, Treynor ratio, information ratio, performance relative to the market benchmark, Jensen's alpha, and Carhart's four-factor alpha) based on five evaluation windows (one, two, three, four, and five years). For each estimation, funds that have return data for at least five-sixths of the observations in each estimation period are included in the sample. The excess return, Jensen's alpha, and the four-factor alpha are annualized returns in percentages. The sample period is from January 1980 to December 2012. Our sample includes 3,881 distinct funds and 349,192 fund–month observations.

Panel A: Fund characteristics

	MEAN	MIN	MAX	STD
Net return (%)	0.64	−47.96	53.53	5.49
TNA ($ millions)	891.8	15.0	105,938.5	2,814.4
Fund age (years)	13.4	0.1	88.5	13.1
Expenses (%)	1.23	0.01	6.42	0.43
Turnover (%)	85.45	0.00	5,388.00	91.41
Flow (%)	0.24	−59.93	758.65	5.65

Panel B: Fund performance

Evaluation period	Sharpe ratio	Appraisal ratio	Treynor ratio	Information ratio	Excess return	Jensen's alpha	Four-factor alpha
1 yr	0.1647	−0.1340	0.0047	−0.0753	−0.73	−0.91	−0.62
2 yr	0.1409	−0.1266	0.0047	−0.0652	−0.68	−1.14	−0.68
3 yr	0.1280	−0.1172	0.0047	−0.0615	−0.69	−1.18	−0.81
4 yr	0.1172	−0.1014	0.0043	−0.0611	−0.73	−1.13	−0.87
5 yr	0.1131	−0.0899	0.0044	−0.0596	−0.71	−1.05	−0.85

While it is relatively simple to measure the net flows of funds, the literature proposes various ways to measure performance in mutual funds.[3] Two features are worth mentioning regarding our study. First, previous studies suggest different metrics to measure performance (Carhart, 1997; Ferson & Schadt, 1996; Jensen, 1969; Sharpe, 1966; Treynor, 1965). Second, the literature shows that the choice of evaluation period affects fund performance and performance persistence (Bollen & Busse, 2005; Carhart, 1997; Elton et al., 1993). From the perspective of investors, they make investment decisions based on multiple performance-based signals. To take this aspect into account, we use seven well-known fund performance measures proposed in both the academic and practice-oriented literature—performance relative to the market benchmark, Sharpe's ratio, Jensen's alpha, Treynor's ratio, information ratio, appraisal ratio, Carhart's four-factor alpha—over multiple evaluation windows. We employ past performance evaluation windows of one, two, three, four, and five years.

First, we measure performance relative to the market benchmark. We choose the Standard & Poor's (S&P) 500 Index as a benchmark portfolio. This measure is not a risk-adjusted performance measure but just excess return relative to passive investment. We use this measure because (1) investors can easily calculate this metric and (2) Del Guercio and Tkac (2002) find that mutual fund investors pay more attention to simple measures than to complex measures.

We use the Sharpe (1966) ratio, the simplest risk-adjusted performance measure. The Sharpe ratio for fund i is defined as

$$Sharpe\ ratio_i = \frac{E[R_i - R_f]}{\sigma(R_i - R_f)} \tag{2}$$

where $R_i - R_f$ is the excess return of fund i and $\sigma(R_i - R_f)$ is the standard deviation of the excess returns.

Jensen's alpha is one of the most well-known measures of performance evaluation. It is defined as an intercept α_i in the following time series regression:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i MKT_t + \varepsilon_{i,t} \tag{3}$$

where the market factor, MKT, represents the monthly excess return over the riskless asset.

Treynor (1965) proposes a measure that penalizes the portfolio in proportion to the amount of leverage employed, as follows:

$$Treynor\ ratio_i = \frac{E[R_i - R_f]}{\beta_i} \tag{4}$$

Similar to the Sharpe ratio, a higher value of the Treynor ratio implies superior performance. Unlike the Sharpe ratio, for the Treynor ratio, the excess return is normalized relative to systematic risk, the capital asset pricing model (CAPM) beta, and not total risk or volatility.

[3] We employ a return-based approach to measure performance and performance persistence, because (1) holdings-based analysis requires the stock composition of each fund, which is not generally available, and (2) information about holdings is disclosed on a quarterly basis in most funds and is sometimes inaccurate.

Black and Treynor (1973) show that the optimal deviations from the benchmark holdings for each security depend on the appraisal ratio:

$$Appraisal\ ratio_i = \frac{\alpha_i}{\sigma(\varepsilon_i)} \tag{5}$$

where $\sigma(\varepsilon_i)$ is the standard deviation of the residual for security i in the CAPM. The authors argue that the appraisal ratio is a measure of a portfolio manager's skill in gathering and using information specific to individual securities.

The information ratio is a measure of a portfolio's risk-adjusted return. It is defined as the expected active return divided by a tracking error, as follows:

$$Information\ ratio_i = \frac{E[R_i - R_b]}{\sigma(R_i - R_b)} \tag{6}$$

where the active return is the difference between the return of fund i, R_i, and the return of a selected benchmark index R_b. The tracking error is the standard deviation of the active return. We use the S&P 500 Index as a benchmark portfolio.

Finally, we employ a risk-adjusted performance measure proposed by Carhart (1997).[4] Carhart's four-factor alpha is the intercept α_i in the following time series regression:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i^{MKT} MKT_t + \beta_i^{SMB} SMB_t + \beta_i^{HML} HML_t + \beta_i^{UMD} UMD_t + \varepsilon_{i,t} \tag{7}$$

where the market factor, MKT, represents the monthly market excess return. The factors SMB, HML, and UMD represent the monthly returns on size, value, and momentum factor-mimicking portfolios, respectively.

Panel B of Table 1 shows the average fund performance estimated from the seven measures based on the five evaluation windows. For each estimation, funds that have return data for at least five-sixths of the observations in each estimation period are included in the sample. For example, if we use 24-month evaluation periods, we include a fund that has at least 20 return observations in the evaluation window. The excess return, Jensen's alpha, and the four-factor alpha are annualized returns in percentages. The results indicate that actively managed equity funds, on average, do not outperform passive benchmarks, regardless of performance measures.

3.2. Correlation analysis across various performance measures

We use seven different performance measures and five different evaluation periods. Thus, for each fund i in month t, investors have 35 distinct performance-based signals. For each month t, we rank all funds using each of the 35 performance measures. For each measure, the worst fund has a percentile rank of zero and the best fund obtains a percentile rank of one. In Table 2, we present the Spearman rank correlation coefficients across distinct performance measures pooled over funds and over months. We fix the evaluation period in calculating the correlation coefficients in Table 2 and each panel represents the rank correlations for the specific estimation period.

The rank correlations are high among the performance measures. Correlations between the Sharpe ratio, Jensen's alpha, the appraisal ratio, and the Treynor ratio are very high (0.87–0.98). In addition, correlations between the information ratio and the S&P 500 benchmark excess return are very high (0.95–0.96). The rank correlation of the four-factor alpha in relation to the other performance measures is lower than the others (0.49–0.69). One possible explanation for this result is the choice of benchmark portfolio when investors evaluate past performance. Sharpe's ratio, Jensen's alpha, the appraisal ratio, and the Treynor ratio are performance measures based on the CAPM or a mean–variance-efficient portfolio. On the other hand, the information ratio and S&P 500 benchmark excess return are performance measures based on the benchmark portfolio's excess return. The four-factor alpha is a performance measure based on a multi-factor benchmark portfolio. Since fund performance measures incorporate benchmark portfolio's information, the choice of benchmark portfolio could affect performance evaluation. While acknowledging that all performance measures capture common features of past performance, our analysis reveals that different performance measures produce different information about manager skill. Thus, to the extent that it is unclear what particular measures are most appropriate for mutual fund investors, investors who receive multiple performance-based signals of uncertain precision about manager skill are likely to assess the accuracy of the signals and make decisions based on the level of information precision.

Table 3 presents the Spearman rank correlation coefficients across different estimation windows when we use the same performance measure. The overall pattern shows that the correlation coefficient rises as the overlapping period increases. For example, in Panel A, we calculate a correlation coefficient of 0.69 between the performance ranks based on one year and on two years. On the other hand, the correlation coefficient between performance ranks based on one year and five years is only 0.39. Our result indicates that information contained in the same performance measure is different if we employ different evaluation periods. Since mutual funds typically provide past performance based on different evaluation windows, investors are likely to combine multiple performance-based signals to make their investment decisions.

[4] We use the Carhart four-factor model (1997) because it is well-known that mutual funds follow momentum (Grinblatt, Jostova, Petrasek, & Philipov, 2016; Grinblatt, Titman, & Wermers, 1995). Under this circumstance, we should include the momentum factor in measuring fund performance since one wants to know whether a manager outperforms a mechanical strategy.

Table 2
Rank correlations among different performance measures. This table shows the Spearman rank correlation coefficients across performance measures for a fixed evaluation period. We use evaluation periods from one year to five years and each panel reports the rank correlation coefficients for different evaluation periods. The term SR means the Sharpe ratio and α refers to Jensen's alpha, AR is the appraisal ratio, TR is the Treynor ratio, IR represents the information ratio, and EXRET means the excess return relative to the S&P 500 index. Finally, 4Fα denotes the Carhart four-factor alpha. For each calculation, funds that have return data for at least five-sixths of the observations in each evaluation period are included in the sample. The sample period is from January 1980 to December 2012.

Panel A: Evaluation period = 1 year

	SR	α	AR	TR	IR	EXRET	4Fα
SR	1.00						
α	0.90	1.00					
AR	0.87	0.95	1.00				
TR	0.98	0.92	0.89	1.00			
IR	0.78	0.81	0.86	0.79	1.00		
EXRET	0.80	0.84	0.82	0.81	0.95	1.00	
4Fα	0.51	0.56	0.56	0.52	0.50	0.49	1.00

Panel B: Evaluation period = 2 years

	SR	α	AR	TR	IR	EXRET	4Fα
SR	1.00						
α	0.89	1.00					
AR	0.87	0.96	1.00				
TR	0.98	0.91	0.88	1.00			
IR	0.77	0.82	0.88	0.79	1.00		
EXRET	0.79	0.85	0.84	0.81	0.96	1.00	
4Fα	0.56	0.63	0.63	0.58	0.60	0.60	1.00

Panel C: Evaluation period = 3 years

	SR	α	AR	TR	IR	EXRET	4Fα
SR	1.00						
α	0.89	1.00					
AR	0.86	0.96	1.00				
TR	0.98	0.90	0.88	1.00			
IR	0.78	0.84	0.89	0.80	1.00		
EXRET	0.80	0.86	0.86	0.82	0.96	1.00	
4Fα	0.60	0.67	0.67	0.62	0.63	0.64	1.00

Panel D: Evaluation period = 4 years

	SR	α	AR	TR	IR	EXRET	4Fα
SR	1.00						
α	0.89	1.00					
AR	0.87	0.96	1.00				
TR	0.98	0.90	0.88	1.00			
IR	0.81	0.86	0.91	0.82	1.00		
EXRET	0.83	0.89	0.88	0.85	0.96	1.00	
4Fα	0.62	0.69	0.69	0.63	0.64	0.65	1.00

Panel E: Evaluation period = 5 years

	SR	α	AR	TR	IR	EXRET	4Fα
SR	1.00						
α	0.87	1.00					
AR	0.86	0.97	1.00				
TR	0.98	0.89	0.88	1.00			
IR	0.80	0.87	0.91	0.82	1.00		
EXRET	0.83	0.91	0.89	0.84	0.97	1.00	
4Fα	0.63	0.71	0.71	0.64	0.65	0.66	1.00

In sum, we have shown that the past performance of mutual funds differs depending on the choice of performance measure and evaluation period. Therefore, in a world of multiple performance-based signals of uncertain precision, mutual fund investors combine large amount of information in making investment decisions. In the next section, we study how investors take into account multiple signals in investment decisions.

3.3. Constructing the measure of imprecision about managerial skill

Since the past performance estimated from each measure is at best a noisy signal about managerial skill, investors assess the precision of signals when they make investment decisions. We construct our measure for imprecision about managerial skill using multiple past performance measures. Given 35 distinct past performance measures for each fund in month t, a natural way to estimate the imprecision of multiple performance-based signals is to calculate the standard deviation of past performance rankings.[5]

For this purpose, define the performance ranking of fund i in month t as $\mathrm{Perf}_{i,\,j}[t-12k,t]$,$j=1,\ldots$J, $k=1,\ldots,$K, where the index j indicates a particular performance measure and the index k denotes a particular past performance evaluation period, J = 7 and K = 5 in our case. The variable $\mathrm{Perf}_{i,j}[t-12k,t]$ represents fractional ranks ranging from zero to one based on fund i's performance in month t when we evaluate past performance using measure j during the past 12k-month evaluation window. Then, our measure of imprecision of the signal, *IMP*, is defined as

$$IMP_{i,t}=\sqrt{\frac{\sum_j\sum_k(Perf_{i,j}[t-12k,t]-avgrank_{i,t})^2}{JK-1}} \tag{8}$$

[5] Recently, there are some papers investigating relative importance of performance measures in investment decision of mutual funds. For example, Barber, Huang, and Odean (2015) study the sensitivity of fund flows to alphas estimated using market-adjusted returns, the CAPM, the Fama-French three-factor model, and Carhart four-factor model, and find that the CAPM-based alpha better describes fund flows. By developing a new method of testing asset pricing models, Berk and van Binsbergen (2016) also find that the CAPM better explains fund flows. In this paper, however, we do not study the role of individual performance measure since our focus is to examine the role of imprecision of signal, estimated from multiple performance-based signals, in flow-performance relationship.

Table 3
Rank correlation across different performance evaluation periods. This table shows the Spearman rank correlation coefficients across performance evaluation periods for a fixed performance measure. We use seven performance measures (Sharpe ratio, appraisal ratio, Treynor ratio, information ratio, performance relative to the market benchmark, Jensen's alpha, and Carhart's four-factor alpha) and each panel reports the rank correlation coefficients for each different performance measure. For each calculation, funds that have return data for at least five-sixths of the observations in each evaluation period are included in the sample. The sample period is from January 1980 to December 2012.

Panel A: Jensen's alpha

	1 yr	2 yr	3 yr	4 yr	5 yr
1 yr	1.00				
2 yr	0.69	1.00			
3 yr	0.56	0.81	1.00		
4 yr	0.45	0.68	0.84	1.00	
5 yr	0.39	0.58	0.72	0.87	1.00

Panel B: Sharpe ratio

	1 yr	2 yr	3 yr	4 yr	5 yr
1 yr	1.00				
2 yr	0.62	1.00			
3 yr	0.49	0.72	1.00		
4 yr	0.40	0.61	0.77	1.00	
5 yr	0.35	0.52	0.66	0.81	1.00

Panel C: Appraisal alpha

	1 yr	2 yr	3 yr	4 yr	5 yr
1 yr	1.00				
2 yr	0.71	1.00			
3 yr	0.58	0.82	1.00		
4 yr	0.49	0.71	0.86	1.00	
5 yr	0.43	0.62	0.76	0.89	1.00

Panel D: Treynor ratio

	1 yr	2 yr	3 yr	4 yr	5 yr
1 yr	1.00				
2 yr	0.65	1.00			
3 yr	0.52	0.76	1.00		
4 yr	0.44	0.65	0.80	1.00	
5 yr	0.38	0.56	0.69	0.83	1.00

Panel E: Information ratio

	1 yr	2 yr	3 yr	4 yr	5 yr
1 yr	1.00				
2 yr	0.71	1.00			
3 yr	0.58	0.82	1.00		
4 yr	0.50	0.71	0.86	1.00	
5 yr	0.46	0.63	0.77	0.89	1.00

Panel F: Excess return relative to the S&P 500 Index

	1 yr	2 yr	3 yr	4 yr	5 yr
1 yr	1.00				
2 yr	0.68	1.00			
3 yr	0.55	0.80	1.00		
4 yr	0.47	0.69	0.85	1.00	
5 yr	0.43	0.61	0.75	0.88	1.00

Panel G: Carhart four-factor alpha

	1 yr	2 yr	3 yr	4 yr	5 yr
1 yr	1.00				
2 yr	0.58	1.00			
3 yr	0.43	0.75	1.00		
4 yr	0.36	0.61	0.81	1.00	
5 yr	0.30	0.52	0.69	0.86	1.00

where $avgrank_{i,t}$ is the average of various performance rankings of fund i in month t, that is, $avgrank_{i,t} = \frac{\sum_j \sum_k (Perf_{i,j}[t-12k,t])}{JK}$. Our measure of *avgrank* can be interpreted as an investor's expectations about managerial skill and *IMP* can be interpreted as the investor's uncertainty about managerial skill. Using our full sample of data, the average of *IMP* is estimated as 0.168. The minimum and maximum of *IMP* are 0.033 and 0.326, respectively. In addition, the standard deviation of *IMP* is 0.056. Therefore, considerable dispersion exists along with our *IMP*, implying that our measure of imprecision about managerial skill helps explain the behavior of mutual fund investors.

Panel A of Table 4 presents the fund characteristics of decile portfolios classified by *IMP*. For each month, all eligible funds are ranked by *IMP*. The characteristics of each portfolio are calculated as the mean values at the time of portfolio formation and the time-series averages of portfolio characteristics are presented. Future flows tend to decrease as *IMP* increases, and especially, funds with the lowest *IMP* have disproportionately large flows implying that the imprecision of signal plays an important role in explaining the cross-section of mutual fund flows in subsequent analysis. In addition, young funds and funds with high expenses and a high turnover ratio have high values of imprecision about managerial skill.

In Panel B of Table 4, we replicate Panel A using decile portfolios based on the *avgrank*. We observe the well-known performance chasing behavior of mutual fund investors documented by Chevalier and Ellison (1997) and Sirri and Tufano (1998), among others. That is, future fund flows tend to increase with past performance measured by *avgrank*. In addition, *IMP* has an inverted U-shaped relation with average past performance ranking. *IMP* peaks at decile 6 and monotonically decreases to decile 1 or decile 10. This pattern shows that *IMP* and past performance are not perfectly correlated, suggesting that, in addition to past performance, *IMP* has additional information on the cross-section of mutual fund flows.

Table 4
Fund characteristics sorted by *IMP* or *avgrank*. The table reports the characteristics of equally weighted decile portfolios classified by *IMP* or *avgrank*. Panel A presents fund characteristics sorted by *IMP* and Panel B reports the results sorted by *avgrank*. For each month *t*, all available funds are sorted based on *IMP* or *avgrank*. The variable *IMP* is calculated as the standard deviation of 35 fund performance ranking variables and *avgrank* is calculated as their mean. The characteristics of each portfolio are computed as the mean characteristic values at the time of portfolio formation and the time-series averages of the portfolio characteristics are reported. The sample period is from January 1980 to December 2012.

Panel A: Sort based on *IMP*										
	Low	2	3	4	5	6	7	8	9	High
IMP	0.074	0.107	0.126	0.142	0.156	0.169	0.185	0.202	0.224	0.263
avgrank	0.512	0.490	0.495	0.497	0.504	0.501	0.498	0.497	0.496	0.499
Future flow (%)	0.51	0.20	−0.04	−0.28	−0.22	−0.26	−0.25	−0.28	−0.20	−0.19
Future flow ($ millions)	4.69	1.10	0.30	−0.95	−0.49	−0.70	−1.14	−2.15	−0.62	−2.25
Four-factor alpha (%)	−0.30	−0.69	−0.59	−0.53	−0.44	−0.56	−0.58	−0.60	−0.66	−0.37
Jensen's alpha (%)	−0.70	−1.13	−1.09	−0.98	−0.71	−0.88	−0.90	−0.86	−1.09	−1.27
Excess return (%)	−0.19	−0.64	−0.74	−0.56	−0.36	−0.59	−0.67	−0.64	−0.73	−0.73
TNA ($ millions)	706.1	768.4	795.5	784.6	828.1	806.7	786.0	774.0	815.3	778.6
Expenses (%)	1.19	1.13	1.10	1.11	1.10	1.11	1.13	1.15	1.17	1.21
Fund age (year)	19.2	21.1	21.8	21.2	21.3	20.9	20.5	20.0	19.0	18.1
Turnover (%)	79.11	76.35	75.60	76.07	75.91	76.24	78.87	81.26	85.32	89.31
Flow (%)	0.38	−0.25	−0.22	−0.40	−0.29	−0.37	−0.41	−0.41	−0.38	−0.39
Panel B: Sort based on *avgrank*										
	Low	2	3	4	5	6	7	8	9	High
avgrank	0.129	0.237	0.319	0.390	0.458	0.528	0.601	0.678	0.765	0.873
IMP	0.110	0.155	0.176	0.187	0.192	0.193	0.188	0.176	0.156	0.109
Future flow (%)	−1.34	−1.19	−1.06	−0.71	−0.45	−0.20	−0.01	0.41	1.11	2.40
Future flow ($ millions)	−7.54	−6.93	−5.75	−4.60	−3.01	−2.31	−0.25	2.44	7.70	18.46
Four-factor alpha (%)	−6.54	−3.84	−2.65	−1.67	−0.93	−0.22	0.57	1.42	2.68	5.62
Jensen's alpha (%)	−10.00	−6.31	−4.35	−3.04	−1.57	−0.39	0.95	2.39	4.41	8.00
Excess return (%)	−9.28	−5.77	−3.95	−2.65	−1.24	−0.11	1.30	2.72	4.68	8.19
TNA ($ millions)	450.3	547.5	622.8	724.3	792.8	835.4	912.3	941.8	997.8	1000.1
Expenses (%)	1.26	1.20	1.17	1.14	1.10	1.11	1.09	1.09	1.12	1.15
Fund age (year)	19.8	20.4	21.2	21.2	21.1	20.6	20.7	19.9	19.5	18.3
Turnover (%)	87.56	86.13	84.69	83.08	79.62	79.74	74.16	72.26	73.39	73.42
Flow (%)	−1.60	−1.22	−1.04	−0.79	−0.61	−0.42	−0.12	0.19	0.81	2.02

4. Empirical results

4.1. Imprecision about managerial skill and future fund flows

We now examine the influence of imprecision of manager skill on flow–performance relationship. To this end, we use two approaches commonly used in the literature: (1) sorts of future flows based on *IMP* and past performance, and (2) the Fama and MacBeth (1973) cross-sectional regressions.

To obtain a big picture of how future fund flows vary across levels of imprecision about manager skill after controlling for past performance, we rank all available funds based on *avgrank* and construct three portfolios (Low, Mid, and High) using breakpoints of 30% and 70%. Then, for each portfolio, we sort funds into quintile portfolios based on *IMP*. For each of the 15 portfolios, we calculate the fund flows of the following month and obtain the time-series averages of future fund flows. Panel A of Table 5 reports the results.

We find that the role of *IMP* is especially pronounced with winners, where the difference in future flows between funds with the lowest *IMP* and those with the highest is estimated as 1.14%, with a Newey–West *t*-statistic of 4.82. This result occurs because funds with the highest past performance and the lowest *IMP* have disproportionately huge future flows of 2.13%, with a Newey–West *t*-statistic of 5.64. Our finding is intuitive in that investors are likely to put their money into that fund if it is classified as a past winner and its superior past performance is estimated with high precision. Our finding, however, implies that investors hesitate to put their money into funds with low precision of signal even though the past performance is superior. On the other hand, we find that precision about manager skill is less important among losers. Our finding indicates that, when investors recognize that past performance is low, on average, they do not consider the precision of signal in their investment decisions. Instead, they place greater weight on the level of the signal. Inferior performance itself is a sufficient reason for investors to withdraw money from funds.

Now, to examine our main research question, we construct Panel B of Table 5. To investigate the influence of imprecision of manager skill on flow–performance relationship, we first rank all available funds based on *IMP* and construct three portfolios (Low, Mid, and High) using breakpoints of 30% and 70%. We then sort funds into quintile portfolios based on *avgrank*. For each of the 15 portfolios, we calculate the fund flows of the following month and obtain the time-series averages of future fund flows.

Table 5
Future fund flows sorted by *IMP* and *avgrank*. This table reports average future fund flows based on *IMP* and *avgrank*. In Panel A (B), for each month, we first rank all available funds based on *avgrank* (*IMP*) and construct three portfolios (Low, Mid, and High) using breakpoints of 30% and 70%. Then, for each portfolio, we sort funds into quintile portfolios based on *IMP* (*avgrank*). For each of the 15 portfolios, we calculate the fund flows the following month and the time series averages of the future fund flows are reported. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 1980 to December 2012.

Panel A: Percent flows of funds sorted by *IMP* after controlling for *avgrank*

	IMP(t − 1)					
Avgrank(t − 1)	Low	2	3	4	High	Low − High
Loser	−1.07**	−1.26***	−1.14***	−1.19***	−1.15***	0.08
	(−2.27)	(−4.44)	(−4.06)	(−4.25)	(−4.04)	(0.20)
Mid	−0.44	−0.30	−0.44	−0.20	−0.35	−0.10
	(−1.65)	(−1.06)	(−1.59)	(−0.57)	(−1.02)	(−0.60)
Winner	2.13***	1.35***	0.84***	0.54*	0.99**	1.14***
	(5.64)	(3.60)	(2.67)	(1.85)	(2.16)	(4.82)

Panel B: Percent flows of funds sorted by *avgrank* after controlling for *IMP*

	Avgrank(t − 1)					
IMP(t − 1)	Loser	2	3	4	Winner	Winner − Loser
Low	−1.33***	−1.07	−0.32	0.99***	2.51***	3.84***
	(−2.81)	(−3.72)	(−1.07)	(2.88)	(6.31)	(7.93)
Mid	−1.29***	−0.88***	−0.44*	0.11	1.18**	2.47***
	(−4.65)	(−3.11)	(−1.65)	(0.38)	(2.55)	(7.68)
High	−1.14***	−0.71***	−0.20	0.03	0.84**	1.98***
	(−3.75)	(−2.64)	(−0.56)	(0.10)	(2.47)	(9.32)

A couple of features deserve to be highlighted. First, we observe performance chasing behavior of mutual fund investors regardless the level of imprecision of the signal, *IMP*. For any level of *IMP*, the difference in future fund flows between winners and losers is positively significant. Second, and more importantly, the difference in future fund flows between winners and losers becomes larger as the imprecision of the signal becomes lower. That is, we find that when past performance is precise (imprecise), stronger (weaker) flow-performance relationship is present. For funds with the lowest *IMP*, the difference in the average future flow between funds with winners and losers is 3.84%, with a Newey–West *t*-statistic of 7.93. The difference monotonically decreases as the imprecision of the signal increases, and the difference in the average future flow between funds with winners and losers is 1.98% with a Newey–West *t*-statistic of 9.32 for funds with the highest *IMP*. Third, the last columns in Panels A and B indicate that the difference in future fund flows is higher across past performance than across *IMP*. Therefore, past performance is still a very important determinant of the cross-section of future fund flows. However, the fact that the difference in future fund flows between winners and losers varies substantially across *IMP* means that imprecision about manager skill helps explain flow-performance relationship.

Why is the sensitivity of flow-performance relationship higher (lower) for funds with lower (higher) *IMP*? In fact, our *IMP* represents the imprecision of multiple performance-based signals. Consider a Bayesian investor who receives multiple performance-based signals of uncertain precision about managerial skill. From the perspective of this investor, it is unclear what particular measure is the most salient for investment decisions. Under the circumstances, the Bayesian investor is less likely to respond to signal with lower precision, regardless of the content of the signal. This means that when past performance is measured with lower precision, we observe lower flow-performance sensitivity, i.e., less inflows for good past performance funds and less outflows for poor past performance funds. On the contrary, when past performance is measured with high precision, we observe higher flow-performance sensitivity. Therefore, the imprecision of multiple performance-based signals affects sensitivity of flow-performance relationship.

Next, as an alternative method, we perform Fama and MacBeth (1973) regressions where the dependent variable is future fund flows and the independent variables are *avgrank*, *IMP*, interaction term of $IMP \times avgrank$, and a number of control variables known to affect future fund flows. We control for fund return volatility computed by the standard deviation of monthly raw returns during the past three years, fund size measured by the logarithm of fund TNA, total expenses ratios, the logarithm of fund age, fund turnover ratios measured by the minimum of aggregated sales or aggregated purchases of securities divided by the fund's average 12-month TNA, and the flow of the previous month. We also include the growth of the fund objective category in period *t*-1 to control for aggregate fund market conditions. Since we examine the influence of imprecision of manager skill on flow–performance relationship, the main variable of interest is the interaction term of $IMP \times avgrank$, and we expect the coefficient term to be negative and statistically significant.

Table 6 presents the Fama–MacBeth regression results, and each column represents one regression result. For each model, we report the time-series averages of the estimated coefficients and Newey–West *t*-statistics are reported in parentheses. We run the following cross-sectional regressions to estimate the Model 1 and Model 2:

$$\begin{aligned} flow_{i,t} = {} & a + b_1 avgrank_{i,t-1} + b_2 (IMP \times avgrank)_{i,t-1} + b_3 IMP_{i,t-1} + b_4 vol_{i,t-1} + b_5 \log(TNA)_{i,t-1} + b_6 expenses_{i,t-1} \\ & + b_7 \log(AGE)_{i,t-1} + b_8 Turnover_{i,t-1} + b_9 flow_{i,t-1} + b_{10} Categoryflow_{i,t-1} + \varepsilon_{i,t} \end{aligned} \tag{9}$$

Table 6
Effect of imprecision about manager skill on fund flow. This table reports estimation results from the Fama–MacBeth cross-sectional regressions of future fund flows on *avgrank* and interaction term of *IMP* × *avgrank* in addition to control variables. Flow is measured as the percentage growth of the fund's assets. The variable *IMP* is calculated from the standard deviation of 35 fund performance ranking variables and *avgrank* is calculated from the variables' mean. In Models 3 and 4, the fractional rank in the bottom performance quintile (*LOW*) is Min(0.2, *avgrank*), in the three median quintiles (*MID*) is Min(0.6, *avgrank–Low*), and in the top quintile (*HIGH*) is *avgrank–LOW–MID*. In Models 5 and 6, the fractional rank in the bottom 30% (*LOW*) is Min(0.3, *avgrank*), in the median 40% (*MID*) is Min(0.4, *avgrank–Low*), and in the top 30% (*HIGH*) is *avgrank–LOW–MID*. The control variables are lagged fund characteristics, including *vol* (volatility of monthly raw returns during the prior three years), fund size (log of TNA), the expense ratio, the log of fund age, the turnover ratio, lagged flow, and the fund objective category flows. The time-series averages of the coefficients and Newey–West *t*-statistics (in parentheses) are reported. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 1980 to December 2012.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	−0.016***	−0.017**	−0.012**	−0.007	−0.012***	−0.008
	(−8.16)	(−2.16)	(−2.36)	(−0.50)	(−3.03)	(−0.65)
avgrank	0.045***	0.037***				
	(13.68)	(5.54)				
HIGH			0.120***	0.092***	0.085***	0.067***
			(8.23)	(7.82)	(8.86)	(8.46)
MID			0.036***	0.033***	0.034***	0.031***
			(11.45)	(6.39)	(8.42)	(5.79)
LOW			0.026	0.021	0.028**	0.023*
			(1.03)	(1.12)	(2.05)	(1.78)
IMP × *avgrank*	−0.096***	−0.098***	−0.086***	−0.098***	−0.086***	−0.094***
	(−6.55)	(−2.88)	(−6.55)	(−3.36)	(−6.36)	(−2.93)
IMP		0.026		0.031		0.028
		(1.18)		(1.62)		(1.32)
vol		−0.662		−0.935		−0.925
		(−1.11)		(−1.53)		(−1.54)
log(TNA)		0.000		0.000		0.000
		(−0.44)		(−0.66)		(−0.71)
expenses		0.319		0.215		0.212
		(1.42)		(0.96)		(0.96)
log(AGE)		−0.001**		−0.001		−0.001
		(−2.01)		(−1.46)		(−1.44)
turnover		0.003**		0.003**		0.003**
		(2.24)		(2.43)		(2.37)
flow		0.212***		0.212***		0.212***
		(3.79)		(3.84)		(3.83)
category flow		−0.011		−0.012		−0.010
		(−0.93)		(−1.02)		(−0.95)

The first column of Table 6 shows the result of Model 1 when the control variables are not included. In this model, the coefficient of *IMP* × *avgrank* is −0.096 and is statistically significant at the 1% significance level. Therefore, consistent with the results in Table 5, we find that flow-performance relationship becomes weaker as imprecision of past performance increases. In Model 2, the significance of interaction term of *IMP* × *avgrank* remains in the presence of the control variables which are known to affect future fund flows. In sum, the regression results confirm our conjecture that the sensitivity of flow-performance relationship is higher (lower) for funds with lower (higher) *IMP*.

Now, to examine whether our results are retained in the presence of convexity in flow-performance relationship, we control for the convexity in the regressions. To this end, following Sirri and Tufano (1998), we use a piecewise-linear specification to allow flow-performance sensitivities to differ for the lowest quintile (Low), the three middle quintiles (Mid), and the top quintile (High) of the *avgrank*. Specifically, we define the following performance ranking variables:

$$\begin{aligned} Low_{i,t-1} &= \min(0.2, avgrank_{i,t-1}) \\ Mid_{i,t-1} &= \min(0.6, avgrank_{i,t-1} - Low_{i,t-1}) \\ High_{i,t-1} &= avgrank_{i,t-1} - (Low_{i,t-1} + Mid_{i,t-1}) \end{aligned} \tag{10}$$

Model 3 represents the regression result without control variables. We find clear convexity in the flow-performance relationship for our sample. The slope on the lowest quintile is 0.026, and it is 0.120 on the highest quintile. The level of convexity is economically large in that improvement in performance ranking from the 90th percentile to the 100th percentile is associated with an increase in fund flows of 1.2%. More importantly, when we allow flow-performance sensitivities to differ across past performance, the slope on our main variable, *IMP* × *avgrank*, is still negatively significant. When we add control variables in Model 4, fund flows again respond less strongly to past performance estimated with low precision. In addition, for robustness check, we allow flow-performance sensitivities to differ for the lowest 30% (Low), the middle 40% (Mid), and the top 30% (High) of the *avgrank* in Models 5 and 6. The results show that our main findings remain the same. In sum, the results in Tables 5 and 6 implies that imprecision of the signal about manager skill helps understand flow-performance relationship.[6]

[6] We replicate Tables 5 and 6 when we use dollar flow instead of percentage flow. The empirical findings are qualitatively similar.

Table 7
Comparison with other measures. This table reports estimation results from the Fama–MacBeth cross-sectional regressions of future fund flows on *avgrank* and interaction term of $IMP \times avgrank$ in the presence of competing explanatory variables. Flow is measured as the percentage growth of the fund's assets. The variable *IMP* is calculated from the standard deviation of 35 fund performance ranking variables and *avgrank* is calculated from the variables' mean. To control for convexity, we use fractional performance ranks of four-factor alphas over the low, medium, and high performance ranges for the past one, three, and five years. For example, *Low_1* yr is Min(0.2, *avgrank*), *Mid_1* yr is Min(0.6, *avgrank–Low_1* yr), and *High_1* yr is *avgrank–Low_1* yr–*Mid_1* yr where *avgrank* is computed during the past one year. The variable min is computed as the minimum performance rank of four-factor alpha for the past one, three, and five years. The variable *vol* is computed as the volatility of monthly raw returns during the prior 3 years. The control variables are lagged fund characteristics, including fund size (log of TNA), the expense ratio, the log of fund age, the turnover ratio, lagged flow, and fund objective category flows. Time series averages of the coefficients and Newey–West *t*-statistics (in parentheses) are reported. To save the space, we do not displays slopes on control variables. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 1980 to December 2012.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	−0.019**	−0.005	−0.007	−0.015**	−0.019**	−0.022***
	(−1.97)	(−0.50)	(−0.51)	(−1.97)	(−2.10)	(−2.90)
LOW_1 yr	−0.001	0.008	0.005	−0.003	0.000	−0.003
	(−0.07)	(0.76)	(0.58)	(−0.19)	(−0.01)	(−0.19)
MID_1 yr	0.006**	0.006***	0.005**	0.006**	0.007**	0.006*
	(2.05)	(2.72)	(2.20)	(1.94)	(1.97)	(1.71)
HIGH_1 yr	0.007	0.018*	0.012	0.007	0.001	0.001
	(0.59)	(1.78)	(1.24)	(0.56)	(0.11)	(0.06)
LOW_3 yr	0.028*	0.024	0.035**	0.013	0.025*	0.025*
	(1.88)	(1.61)	(2.43)	(1.01)	(1.66)	(1.73)
MID_3 yr	0.027***	0.020***	0.033***	0.010***	0.026***	0.025***
	(4.65)	(2.65)	(3.45)	(3.51)	(3.98)	(3.21)
HIGH_3 yr	0.052***	0.064***	0.067***	0.045***	0.054***	0.052***
	(3.81)	(3.61)	(3.46)	(3.47)	(3.61)	(3.18)
LOW_5 yr	−0.009	−0.008	−0.009	−0.006	−0.007	−0.005
	(−0.59)	(−0.44)	(−0.43)	(−0.36)	(−0.40)	(−0.28)
MID_5 yr	0.003	0.003	0.002	0.001	0.003	0.003
	(1.23)	(1.17)	(0.66)	(0.36)	(1.17)	(1.11)
HIGH_5 yr	0.037***	0.029***	0.031***	0.026**	0.035***	0.035***
	(3.13)	(2.46)	(2.61)	(2.24)	(3.21)	(3.11)
$IMP \times avgrank$	−0.089***		−0.093***		−0.083***	−0.083**
	(−3.17)		(−3.87)		(−2.77)	(−2.10)
IMP	0.025		0.027**		0.022	0.029*
	(1.53)		(2.22)		(1.27)	(1.83)
$IMP \times min$						−0.015
						(−0.29)
min				0.007**	0.001	0.005
				(2.26)	(0.19)	(0.50)
$vol \times avgrank$		−2.489	−1.286			
		(−1.03)	(−0.53)			
vol		0.003	−0.049			
		(0.00)	(−0.03)			
Controls	YES	YES	YES	YES	YES	YES

4.2. Comparison with other measures

The present study is related to two recent papers. Huang et al. (2012) show that flow sensitivity to past performance decreases with fund return volatility. While we estimate imprecision about managerial skill by calculating the cross-sectional dispersion of 35 different performance measures, the authors obtain volatility by computing the time-series dispersion of the monthly raw returns of mutual funds. They argue that a fund with higher return volatility has a less precise signal of skill. Li et al. (2016) show that ambiguity-averse mutual fund investors tend to react more strongly to the worst-case scenario. In other words, according to the model of Li et al. (2016), investors place greater weight on the worst signal when they receive multiple signals of uncertain quality. Since the aforementioned studies contribute to the literature, we compare the explanatory power of our main variable with those suggested by recent works.

Following Huang et al. (2012), the fund return volatility measure, *vol*, is computed with the volatility of monthly raw returns during the past three years. In addition, the worst performance measure, *min*, of Li et al. (2016) is defined as the minimum performance rank of the Carhart four-factor alpha in the past one, three, and five years. To control for convexity, we use variables in Table 7 of Li et al. (2016). For example, *Low_1yr*, *Mid_1yr*, and *High_1yr* are defined as Eq. (10), but in this case, the rankings are computed based on the Carhart four-factor alphas measured over one year. Similarly, we adopt fractional ranks for 3 and 5-year performance. Table 7 presents the estimation results and each model represents one regression model. For each model, the time-series averages of the estimated coefficients and Newey–West *t*-statistics are reported in parentheses. In addition, for each model, we include control variables used in Table 6, but we do not report coefficients on control variables to conserve the space.

Table 8
Fund portfolio returns based on univariate sorting on lagged *IMP* or lagged *avgrank.* In Panel A (B), for each month, we rank all available funds based on our *avgrank* (*IMP*) measure and construct decile portfolios. We then hold 10 portfolios for one month and calculate the equally weighted portfolio returns, which yields the complete time series returns of the 10 portfolios. We estimate the annualized excess returns and the four-factor alphas for each of the 10 portfolios. Newey–West *t*-statistics are reported in parentheses. *IMP* is calculated from the standard deviation of 35 fund performance ranking variables and *avgrank* is calculated from the variables' mean. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 1980 to December 2012.

Panel A: *avgrank* decile portfolios			Panel B: *IMP* decile portfolios		
avgrank(t − 1)	Excess returns	Four-factor alphas	*imprecision*(t − 1)	Excess returns	Four-factor alphas
Low	3.18	−2.56**	Low	5.21	−1.02
	(0.80)	(−2.16)		(1.37)	(−1.30)
2	4.06	−1.93**	2	5.02	−1.36*
	(1.06)	(−2.03)		(1.33)	(−1.92)
3	5.68	−0.42	3	5.35	−0.72
	(1.50)	(−0.48)		(1.45)	(−1.08)
4	4.51	−1.68**	4	4.81	−1.40**
	(1.20)	(−2.21)		(1.29)	(−1.98)
5	5.34	−0.80	5	5.38	−0.89
	(1.44)	(−1.26)		(1.43)	(−1.46)
6	5.86	−0.43	6	5.60	−0.60
	(1.57)	(−0.67)		(1.51)	(−0.99)
7	6.05	−0.26	7	5.74	−0.58
	(1.61)	(−0.40)		(1.53)	(−0.92)
8	6.17	−0.30	8	6.20	−0.06
	(1.64)	(−0.42)		(1.63)	(−0.09)
9	7.18*	0.39	9	6.25	−0.10
	(1.85)	(0.52)		(1.64)	(−0.13)
High	7.44*	0.47	High	5.47	−1.21*
	(1.86)	(0.53)		(1.36)	(−1.65)
High − Low	4.26**	3.03*	High − Low	0.27	−0.19
	(2.28)	(1.86)		(0.31)	(−0.25)

Model 1 shows our model, and the slope on *IMP* × *avgrank* is negatively significant when we control for convexity following Li et al. (2016). In Model 2, we include *vol* and the interaction term of *vol* × *avgrank* to replicate the work of Huang et al. (2012). Consistent with the result of Huang et al. (2012), slope on *vol* × *avgrank* is negative, but the coefficient is not statistically significant.[7] In Model 3, we find that our main variable, *IMP* × *avgrank*, is negatively significant in the presence of the measure used in Huang et al. (2012). Looking at the Model 4, we find that the worst performance measure, *min*, is positively significant, which is consistent with the work of Li et al. (2016). However, the explanatory power of the worst performance measure disappears when the interaction term, *IMP* × *avgrank*, is added in the regressions as shown in Model 5. More importantly, our *IMP* × *avgrank* is negatively significant, and the estimated coefficient on *IMP* × *avgrank* in the presence of *min* is very similar to the slope in Model 1.

Li et al. (2016) argue that ambiguity-averse investors react more strongly to the worst-case scenario. If our *IMP* is a good proxy for the fund's ambiguity level, we expect to observe greater flow sensitivity to the minimum performance for funds with higher *IMP*. That is, the interaction term, *IMP* × *min* is positive since an increase in *IMP* increases the fund's ambiguity level. Model 6 displays the result and the interaction term, *IMP* × *min*, is negative and statistically insignificant, indicating that our *IMP* is not a good proxy for the fund's ambiguity level. Again, our interaction term, *IMP* × *avgrank*, is negatively significant. Overall, the results in Table 7 show that the inclusion of competing variables does not drive out the role of *IMP* in explaining flow-performance relationship.

4.3. Imprecision about managerial skill and performance persistence

Up to now, we have studied the role of *IMP* in flow-performance relationship. Our construction of *IMP* enables us to examine another important issue in mutual fund literature, performance persistence. This research question arises because our *IMP*, by construction, has direct information about the skill of money managers. That is, funds with lower *IMP* have lower uncertainty about whether managerial skill is high or low compared to funds with higher *IMP*. Therefore, we conjecture that strong performance persistence is observed among funds with lower *IMP*. In this section, we investigate whether our *IMP* helps identify funds with strong performance persistence.

Before examining performance persistence conditional on *IMP*, as a preliminary analysis we conduct two univariate sorts and study whether past performance or our *IMP* can predict future performance. First, we examine whether performance

[7] We believe that our findings are not consistent with those of Huang et al. (2012) due to different data frequencies and sample selection criteria. First, while they use quarterly flow data, we calculate fund flows on a monthly basis. Second, when a fund has multiple share classes, we aggregate all the observations pertaining to different share classes into one observation. Huang et al., on the other hand, perform their analysis at the fund share level. Therefore, while our sample includes 3,881 distinct funds, their sample covers 9,791 distinct fund shares.

Table 9
Fund portfolio returns based on a double sort by lagged *IMP* and lagged *avgrank*. Panel A (B) presents the annualized excess returns (Carhart four-factor alphas) of fund portfolios based on a double sort by lagged *IMP* and lagged *avgrank*. For each month t, all available funds are sorted into quintile portfolios based on $avgrank_{t-1}$. Independently, all available funds are sorted into quintile portfolios based on IMP_{t-1}. We construct 25 portfolios from the intersection of the two criteria and calculate the returns on each portfolio for the following month. Then, for the following month test period (t), we calculate the monthly average excess returns and the four-factor alpha for each portfolio. Panel A presents the annualized excess returns of the fund portfolios and Panel B reports the annualized four-factor alphas of the fund portfolios. *IMP* is calculated from standard deviations of 35 fund performance ranking variables and *avgrank* is calculated from the variables' mean. Newey–West *t*-statistics are presented in parentheses. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. The sample period is from January 1980 to December 2012.

Panel A: Results using excess returns

	avgrank(t − 1)					
IMP(t − 1)	Loser	2	3	4	Winner	Winner − Loser
Low	3.22	2.08	3.04	7.39**	7.68*	4.41**
	(0.83)	(0.57)	(0.70)	(1.99)	(1.92)	(2.56)
2	3.90	4.38	5.65	3.80	6.62*	2.71*
	(0.99)	(1.16)	(1.53)	(1.02)	(1.71)	(1.71)
3	5.49	4.51	3.98	6.10	6.92*	1.43
	(1.38)	(1.21)	(1.08)	(1.62)	(1.71)	(0.66)
4	4.90	4.06	6.06	7.00*	6.60	1.71
	(1.17)	(1.07)	(1.56)	(1.83)	(1.54)	(0.91)
High	5.08	4.99	5.61	5.89	4.46	−0.61
	(1.13)	(1.24)	(1.48)	(1.45)	(0.99)	(−0.23)

Panel B: Results using the Carhart four-factor alpha

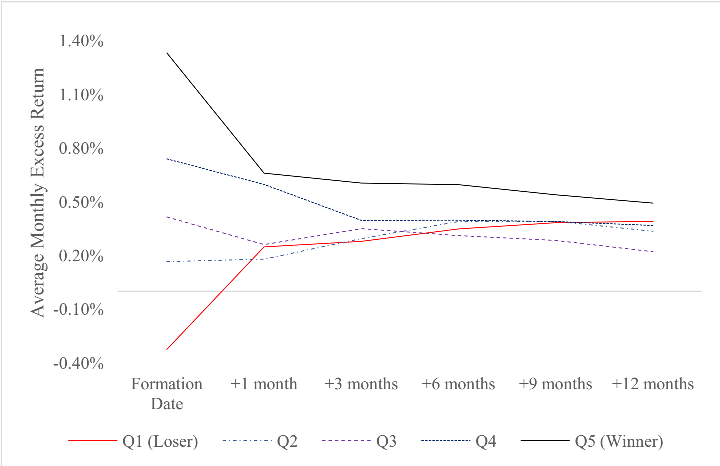
	avgrank(t − 1)					
IMP(t − 1)	Loser	2	3	4	Winner	Winner − Loser
Low	−2.50**	−2.27**	−2.12**	0.59	0.40	2.93**
	(−2.20)	(−2.40)	(−2.11)	(0.62)	(0.47)	(2.02)
2	−2.34*	−1.58*	−2.02***	0.01	0.21	2.44*
	(−1.91)	(−1.69)	(−2.74)	(0.01)	(0.25)	(1.66)
3	−0.54	−1.60*	−1.10	0.07	−0.01	0.60
	(−0.45)	(−1.74)	(−1.56)	(0.10)	(−0.01)	(0.41)
4	−1.31	−1.25	−0.58	0.66	0.32	1.48
	(−0.98)	(−1.48)	(−0.70)	(0.68)	(0.28)	(0.94)
High	−1.08	−1.44	−0.67	−0.49	−0.56	0.74
	(−0.63)	(−1.39)	(−0.81)	(−0.51)	(−0.40)	(0.51)

Table 10
Performance persistence with different holding periods. This table presents the Carhart four-factor alphas of fund portfolios based on a double sort by lagged *IMP* and lagged *avgrank*. For each month t, all available funds are sorted into quintile portfolios based on $avgrank_{t-1}$. Independently, all available funds are sorted into quintile portfolios based on IMP_{t-1}. We construct 25 portfolios from the intersection of the two criteria and calculate the returns on each portfolio up to twelve months. For the lowest and highest *IMP* quintiles, we display the annualized four-factor alphas of the winners, losers, and the winners minus the losers. Newey–West *t*-statistics are presented in parentheses. *IMP* is calculated from the standard deviations of 35 fund performance ranking variables and *avgrank* is calculated from the variables' mean. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is from January 1980 to December 2012.

Sorting variables		Holding period				
IMP(t − 1)	*avgrank*(t − 1)	one month	three months	six months	nine months	twelve months
Low	Winner	0.40	−0.20	−0.17	−0.31	−0.49
		(0.47)	(−0.24)	(−0.22)	(−0.38)	(−0.61)
	Loser	−2.50**	−2.06*	−1.33	−0.97	−0.90
		(−2.20)	(−1.91)	(−1.31)	(−1.02)	(−1.00)
	Winner − Loser	2.93**	1.86	1.16	0.66	0.41
		(2.02)	(1.35)	(0.91)	(0.53)	(0.35)
High	Winner	−0.56	−1.69	−0.74	0.25	0.10
		(−0.40)	(−1.15)	(−0.63)	(0.23)	(0.09)
	Loser	−1.08	−0.23	0.37	−2.15	−1.39
		(−0.63)	(−0.09)	(0.15)	(−1.44)	(−1.05)
	Winner − Loser	0.74	−1.46	−1.12	2.39	1.49
		(−0.51)	(−0.57)	(−0.43)	(1.20)	(0.85)

persistence is present in our sample. To this end, for each month t, all available funds are sorted into decile portfolios based on $avgrank_{t-1}$. We then calculate the returns for each decile for the following month. The average annualized excess returns and the Carhart four-factor alphas are presented in Panel A of Table 8. Newey–West *t*-statistics are presented in parentheses. When performance is measured in terms of excess returns over the T-bill rate, the annualized difference between the

Panel A: Funds with the lowest *IMP*



Panel B: Funds with the highest *IMP*

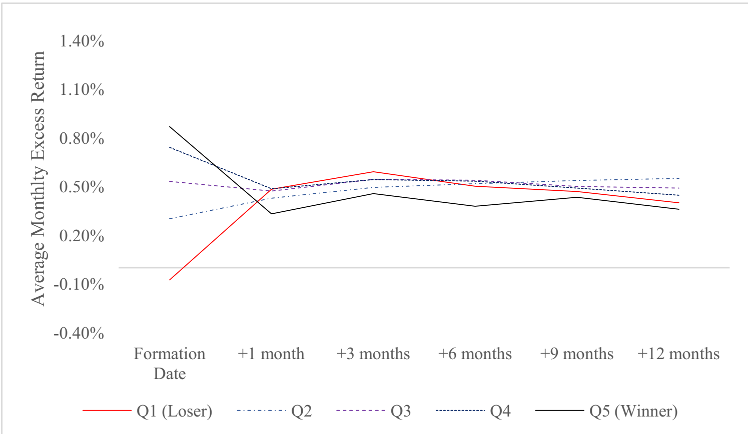


Fig. 1. Post-formation returns on portfolios of mutual funds sorted on past performance. For each month t, all available funds are sorted into quintile portfolios based on our *IMP*. Within each quintile, we construct quintile portfolios based on *avgrank*. Then, for funds with the lowest *IMP* and highest *IMP* quintile, we calculate post-formation excess returns on *avgrank* quintile portfolios in the month subsequent to the initial ranking (formation date) and in each of the next three months after formation. Panel A shows the results for the lowest *IMP* funds and Panel B displays the results for the highest *IMP* funds, with Q5 denoting funds with the highest *avgrank* and Q1 funds with the lowest *avgrank*.

best- and worst-performing funds is 4.26%, with a t-statistic of 2.28. When measured using the Carhart's (1997) four-factor model, performance persistence is observed among losers. In addition, the annualized four-factor alpha of the zero-cost portfolio that purchases the best-performing funds and sells the worst-performing funds is 3.03%, significant at the 10% level. Our results suggest that short-term performance persistence exists and persistence is especially pronounced among losers.

Second, we investigate whether *IMP* has forecasting power for future performance. For this purpose, for each month t, all available funds are sorted into decile portfolios based on IMP_{t-1} and we then calculate the returns for each decile for the following month. Panel B in Table 8 shows the annualized excess returns and the four-factor alphas. Focusing on the four-factor alphas, we do not find any noticeable patterns. Therefore, it seems that future performance is not forecasted by imprecision about managerial skill.

Now, we investigate whether performance persistence is especially pronounced among funds with lower *IMP*. For each month t, all available funds are sorted into quintile portfolios based on $avgrank_{t-1}$. Independently, all available funds are sorted into quintile portfolios based on IMP_{t-1}. We construct 25 portfolios from the intersection of the two criteria and calculate the returns on each portfolio for the following month. The average excess returns are displayed in Panel A of Table 9 and the four-factor alphas are presented in Panel B. Newey–West t-statistics are presented in parentheses.

Several features of the empirical findings are worth highlighting. First, looking at the future returns on each of the 25 portfolios, we observe performance persistence among funds with lower *IMP*. Looking at Panel B of Table 9, in the lowest *IMP* quintile, we find the three worst-performing portfolios have statistically significant four-factor alphas. In particular, the portfolio from the intersection of the worst-performing group and the lowest *IMP* group has the worst future return, −2.50% per year (*t*-statistic of −2.20), of the 25 portfolios. This finding is intuitive because, among past losers, funds with lower *IMP* have a higher probability of being managed by money managers with low skill than funds with higher *IMP* do. If we apply the same reasoning, we should observe significant positive returns on the portfolio formed by the intersection of the best-performing group and the lowest *IMP* group. Indeed, the four-factor alpha is 0.40% per year, but it is economically small and statistically insignificant. In Table 5, we observed that investors put disproportionately large amounts of money into funds with the highest past performance and the lowest *IMP*. However, the results in Table 9 indicate that these funds do not significantly outperform the passive benchmark. Given that equity funds, on the average, do not outperform the benchmark, such a finding is not surprising.

Second, returns on the zero-cost portfolios that purchase the best-performing funds and sell the worst-performing funds also indicate performance persistence among funds with lower *IMP*. Looking at the Panel B in Table 9, the zero-cost strategy creates annual returns of 2.93% (*t*-statistic of 2.02) in the lowest *IMP* group and 2.44% (*t*-statistic of 1.66) in the second lowest *IMP* group. On the other hand, for other quintiles, the zero-cost strategies do not generate significant abnormal returns. This finding is also observed when we measure performance using excess returns in Panel A. Overall, the results in Table 9 indicate that the negative performance persistence observed in Table 8 arises due to strong performance persistence among funds with lower *IMP*.

To examine the performance persistence of different horizons, we estimate the Carhart four-factor alphas of holding periods up to twelve months subsequent to the fund's initial ranking (formation date). Table 10 reports the results. For the lowest *IMP* quintile, we find that negative performance persistence is present for up to three months: Losers have a four-factor alpha of −2.50% (*t*-statistic of −2.20) for a holding period of one month and a four-factor alpha of −2.06% (*t*-statistic of −1.91) for a holding period of three months. In addition, the difference in abnormal returns between winners and losers is statistically significant for a holding period of one month. For a holding period greater than three months, however, we fail to find performance persistence. Our finding of short-term performance persistence is consistent with the empirical evidence of Bollen and Busse (2005). On the other hand, for the highest *IMP* quintile, we find no evidence regarding performance persistence. In sum, the empirical findings in Table 10 indicate that performance persistence is observed among funds with the lowest *IMP* and performance persistence is short-lived.

To visualize performance persistence for different holding periods, we replicate Carhart's (1997) Figure 2 (see Fig. 1). Specifically, for funds with the lowest and highest *IMP* quintiles, we calculate post-formation excess returns in each quintile portfolio sorted by *avgrank* in the month subsequent to the initial ranking (formation date) and in each of the next three months after formation. Panel A shows the results for funds with the lowest *IMP* and Panel B displays the results for funds with the highest *IMP*. Consistent with Table 10, we find that while some performance persistence is observed up to three months among funds with the lowest *IMP*, none of the funds with the highest *IMP* exhibit performance persistence. In sum, our finding indicates that our *IMP* helps identify funds with performance persistence. Therefore, conditioning on *IMP* helps understand performance persistence.

5. Conclusion

In examining the flow-performance relationship, previous studies have not taken into account precision about past performance, although it should be included in the information set of mutual fund investors. To fill this gap, we construct a measure for imprecision about manager skill, *IMP*, using the standard deviations of the rankings of various performance measures over multiple performance evaluation periods, to study the role of our *IMP* in explaining the flow-performance relationship.

We find that future fund flows respond more (less) strongly to past performance measured with high (low) precision. Our finding is robust to the inclusion of control variables such as fund size, total expense ratios, the logarithm of fund age, fund turnover ratios, and the flow of the previous month. We also confirm that our results are retained in the presence of convexity in flow-performance relationship. Finally, short-term performance persistence is observed among funds with lower *IMP*. Therefore, our proxy for imprecision about manager skill helps identify funds with strong performance persistence. Overall, our empirical evidence indicates that our *IMP* plays an important role in explaining the cross-section of future fund flows. Therefore, imprecision about manager skill should be considered an important factor when analyzing the behavior of mutual fund investors.

References

Barber, B. M., Huang, X., & Odean, T. (2015). *Which factors matter to investors? Evidence from mutual fund flows* Working Paper. Berkeley: University of California.
Berk, J. B., & van Binsbergen, H. (2016). Assessing asset pricing models using revealed preference. *Journal of Financial Economics, 119*(1), 1–23.
Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy, 112*(6), 1269–1295.
Black, F., & Treynor, J. L. (1973). How to use security analysis to improve portfolio selection. *Journal of Business, 46*(1), 66–86.
Bollen, N. P., & Busse, J. A. (2005). Short-term persistence in mutual fund performance. *Review of Financial Studies, 18*(2), 569–597.
Busse, J. A., Chordia, T., Jiang, L., & Tang, Y. (2015). *Mutual fund trading costs and diseconomies of scale* Working Paper. Emory University.
Carhart, M. M. (1997). On the persistence in mutual fund performance. *Journal of Finance, 52*(1), 57–82.
Chevalier, J., & Ellison, G. (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy, 105*(6), 1167–1200.

Cremers, K. M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies, 22*(9), 3329–3365.
Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance, 52*(3), 1035–1058.
Del Guercio, D. D., & Tkac, P. A. (2002). The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds. *Journal of Financial and Quantitative Analysis, 37*(4), 523–557.
Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivorship bias and mutual fund performance. *Review of Financial Studies, 9*(4), 1097–1120.
Elton, E. J., Gruber, M. J., Das, S., & Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from manager portfolios. *Review of Financial Studies, 6*(1), 1–22.
Evans, R. B. (2010). Mutual fund incubation. *Journal of Finance, 65*(4), 1581–1611.
Fama, E. F., & MacBeth, J. D. (1973). Risk, return and equilibrium: Empirical tests. *Journal of Political Economy, 81*(3), 607–636.
Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance, 51*(2), 425–461.
Goetzmann, W. M., & Peles, N. (1997). Cognitive dissonance and mutual fund investors. *Journal of Financial Research, 20*(2), 145–158.
Grinblatt, M., & Titman, S. (1989a). Mutual fund performance: An analysis of quarterly portfolio holdings. *Journal of Business, 62*(3), 393–416.
Grinblatt, M., & Titman, S. (1989b). Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies, 2*(3), 393–421.
Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *American Economic Review, 85*(5), 1088–1105.
Grinblatt, M., Jostova, G., Petrasek, L., & Philipov, A. (2016). *Style and skill: Hedge funds, mutual funds, and momentum* Working Paper. Yale University.
Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *Journal of Finance, 51*(3), 783–810.
Huang, J. C., Wei, K. D., & Yan, H. (2007). Participation costs and the sensitivity of fund flows to past performance. *Journal of Finance, 62*(3), 1273–1311.
Huang, J. C., Wei, K. D., & Yan, H. (2012). *Investor learning and mutual fund flows* Working Paper. University of Texas.
Ippolito, R. A. (1989). Efficiency with costly information: A study of mutual fund performance. *Quarterly Journal of Economics, 104*(1), 1–23.
Ippolito, R. A. (1992). Consumer reactions to measure of poor quality: Evidence from the mutual fund industry. *Journal of Law and Economics, 35*(1), 45–70.
Jensen, M. C. (1968). The performance of mutual funds in period 1945–1964. *Journal of Finance, 23*(2), 389–416.
Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business, 42*(2), 167–247.
Lehmann, B. N., & Modest, D. M. (1987). Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance, 42*(2), 233–265.
Li, C. W., Tiwari, A., & Tong, L. (2016). Investment decisions under ambiguity: Evidence from mutual fund investor behavior. *Management Science*. . forthcoming.
Lynch, A. W., & Musto, D. K. (2003). How investors interpret past fund returns. *Journal of Finance, 58*(5), 2033–2058.
Pastor, L., & Stambaugh, R. F. (2012). On the size of the active management industry. *Journal of Political Economy, 120*(4), 740–781.
Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business, 39*(1), 119–138.
Sharpe, W. F. (1992). Asset allocation management style and performance measurement. *Journal of Portfolio Management, 18*(2), 7–19.
Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *Journal of Finance, 53*(5), 1589–1622.
Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*(1), 63–75.