Working Paper
Series

FINRISK
National Centre of Competence in Research
Financial Valuation and Risk Management

National Centre of Competence in Research
Financial Valuation and Risk Management

Working Paper No. 854

Fund Flows in Rational Markets

Francesco Franzoni

Martin C. Schmalz

First version: April 2013
Current version: July 2013

This research has been carried out within the NCCR FINRISK project on "Corporate Finance, Market Structure and the Theory of the Firm"

Die Nationalen Forschungsschwerpunkte (NFS) sind ein Förderinstrument des Schweizerischen Nationalfonds.
Les Pôles de recherche nationaux (PRN) sont un instrument d'encouragement du Fonds national suisse.
The National Centres of Competence in Research (NCCR) are a research instrument of the Swiss National Science Foundation.

Schweizerischer Nationalfonds
Fonds national suisse
Swiss National Science Foundation

Fund Flows in Rational Markets

Francesco Franzoni and Martin C. Schmalz*

July 29, 2013

PRELIMINARY. COMMENTS WELCOME.

Abstract

When investors are risk-averse, relative performance in downturns is more informative about managers' ability to generate utility for investors than relative performance in upturns. Rational investors thus reallocate more wealth between funds in response to downturn-performance than in response to upturn-performance, resulting in higher flow-performance sensitivities in downturns than in upturns. Combined with decreasing returns to scale, the model predicts that mutual fund returns are more persistent following market upturns than following market downturns. We identify the model by empirically testing its difference-in-differences predictions for the flow-performance sensitivity across fund types and market states.

*Franzoni: University of Lugano and Swiss Finance Institute, francesco.franzoni@usi.ch; Schmalz: University of Michigan Stephen M. Ross School of Business, schmalz@umich.edu. For helpful comments and discussions we would like to thank Yakov Amihud, Jonathan Berk, John Campbell, Roger Edelen, Stefano Giglio, Vincent Glode, Itay Goldstein, Denis Gromb, Rick Green, Alexander Guembel, Burton Hollifield, Marcin Kaczperczyk, Shimon Kogan, Robert Kosowski, Augustin Landier, Xuewen Liu, Massimo Massa, Roni Michaely, Lubos Pastor, Alberto Plazzi, Uday Rajan, Clemens Sialm, Luke Taylor, Jean Tirole, Russ Wermers, Youchang Wu, Josef Zechner, Hong Zhang, Sergey Zhuk, seminar participants at the University of Michigan and at IDC Herzliya. Shibao Liu provided excellent research assistance. All errors are our own.

1 Introduction

Economists have long been concerned with the question of whether mutual fund flows follow rational patterns, for at least three reasons. First, U.S. households invest 23% of their financial assets and more than 50% of their retirement savings through IRAs and 401(k)s in mutual funds (ICI, 2012). It would be a major concern from a societal welfare perspective if irrational forces were driving savings decisions of that magnitude. Second, mutual funds own a substantial share of financial assets in the US economy: the $13trn invested in U.S. mutual funds (end of 2011) correspond to 85% of U.S. stock market capitalization (ICI, 2012). Given the large share of institutional asset ownership, an efficient allocation of capital to these intermediaries is likely to be crucial for efficient pricing of these assets and, therefore, for the allocative efficiency in the economy. In other words, if informed fund managers persistently lack capital while uninformed managers are awash with liquidity, mispricing is likely to persist (Shleifer and Vishny, 1997). Understanding whether it is irrational investor decisions or other frictions that prevent an efficient allocation of capital across funds is key when evaluating normative prescriptions to improve allocative efficiency. Third, capital flows affect asset prices also because assets under management affect managers' compensation, incentives, and therefore their investment policy. In sum, understanding whether investors allocate capital across funds in a rational way is at the core of many questions relating to the functioning of the fund industry, and as a consequence for the economy more generally.

Yet, our theoretical understanding of the drivers of the relationship between performance and flows is limited. Given the potential welfare consequences outlined above, it seems worrisome that existing rational models cannot account for some of the most important factors that determine the flow-performance relation, as documented in the last decade of empirical research. As a consequence, authors have concluded that irrational forces or frictions other than parameter uncertainty are necessary to explain variation in the flow-performance relation. In particular, Spiegel and Zhang (2012) note that the existing literature has neglected substantial cross-sectional variation when estimating the average flow-performance sensitivity (FPS). Absent a rational model that predicts such variation, they attribute the variation at least partly to behavioral forces or other frictions that prevent an efficient reallocation of capital. Evidence of conditional performance persistence is another challenge to existing rational models. While much of the literature finds that performance persistence is scarce and short-lived (Carhart, 1997), Cremers and Petajisto (2009) and Petajisto (2013) find substantial performance persistence among particularly active funds and attribute their finding to manager skill. In the context of Berk and Green (2004)'s (BG) rational benchmark model[1], persistence in outperformance due to skill is prevented from the efficient reallocation of capital between funds. Hence, one has to amend BG by allowing for additional frictions if performance persistence has to survive in equilibrium. Similarly, Glode, Hollifield, Kacperczyk, and Kogan (2012) find persistence of mutual fund returns following market upturns but not following market downturns. Also taking BG as a benchmark, the authors conclude that investor irrationality prevents an efficient reallocation of capital to better funds in upturns, while investors make more rational decisions in downturns.[2] Lastly, the present paper documents that the state of the market is a first-order driver of the variation in the flow-performance sensitivity. Existing rational models cannot account for this variation either. In sum, the failure of existing rational models to explain key features of the flow-performance

[1] BG's model features rational Bayesian investors who allocate more capital to funds with high relative performance. As funds face decreasing returns to scale, investors reallocate just enough additional capital to overperforming funds to equalize expected returns across funds. These result is a lack of performance persistence despite the existence of fund manager skill.

[2] See also Sun, Wang, and Zheng (2013), who find outperformance of funds of funds that have outperformed in previous downturns. Choi, Kahraman, and Mukherjee (2012) find that performance in one fund of a given manager predicts subsequent performance in the manager's other fund, from which the authors conclude that investors respond insufficiently to a multi-fund manager's performance.

relation and fund returns has led researchers to rely on irrational investor behavior to explain the existing empirical evidence.

Our main contribution is to provide a new model of mutual fund flows that reconciles the above patterns with rational investor behavior. We allow for uncertainty not only with respect to fund manager's ability to generate average returns, but also with respect to the loading of fund returns on systematic risk. This generalization of existing models opens up a new dimension of theoretical thinking about fund flows: investors' rational response to news about fund performance should differ across upturns and downturns. We are thus able to address empirical puzzles relating to differences in the FPS as well as conditional performance persistence, which are so far unexplained within the domain of rational models. In particular, BG show that a lack of performance persistence is consistent with skilled managers and rational investors who equalize expected returns across funds. The present model has risk-averse investors who equalize expected marginal utilities across funds and makes explicit that the converse of BG's result is not true: (state-contingent) performance persistence does not imply that investors are irrational, or that frictions other than parameter uncertainty prevent an efficient allocation of capital between funds.

The intuition of the model is as follows. Investors realize that abnormal fund returns in upturns can stem either from managerial skill, or from abnormally high passive exposure of a given fund to the market benchmark. In contrast, positive abnormal returns in downturns cannot possibly be fueled by high passive exposure and therefore are more likely due to skill, which increases the likelihood that similar outperformance will repeat in future downturns. Risk-averse investors prefer funds endowed with higher skill and lower loading on systematic risk, as they care more for abnormal returns in downturns than in upturns. Therefore, investors react more strongly to outperformance in downturns than in upturns. In other

Table 1: Summary of the Main Empirical Result. Upturns and Downturns are defined, respectively, as the top and bottom 25% of periods according to the distribution of the CRSP value-weighted index since July 1926. The table reports slopes from Fama and MacBeth (1973) regressions of quarterly flows on prior-quarter mutual fund rank by style-adjusted performance. The last column reports the difference in coefficients between Upturns and Downturns. T-statistics are reported in parentheses.

	Upturns	Downturns	Up-Down
Flow-Performance Sensitivity	0.020***	0.045***	-0.042***
t-stat	(5.501)	(6.491)	(2.654)

words, abnormal returns in downturns are a stronger signal about a manager's ability to generate returns in future downturns than abnormal returns in upturns are. As risk-averse investors derive more marginal utility from consumption in downturns, they care more about abnormal fund returns in downturns than in upturns. Therefore, downturn returns are more informative about the utility that the manager generates for risk-averse investors. As investors maximize utility, and not expected returns, they rationally put higher weights on downturn-performance than on upturn-performance when making cross-sectional capital allocation decisions.[3]

A direct implication of this logic is that between-fund flows are larger following downturn-performance than following upturn-performance, leading to higher flow-performance sensitivity in downturns than in upturns. This key model prediction about the variation of the FPS across states of the economy finds robust support in the data. The FPS is almost twice as high in downturns than in upturns. Table 1 gives a summary of that finding.

[3]Risk-neutral investors only care about average returns but not their timing, and would therefore not react more strongly to downturn-performance than upturn-performance. That is, they key driver of our results is risk aversion, not an asymmetry in information processing across states of the market.

We verify our predictions about the variation of the FPS across market states in more detail. Specifically, our model correctly predicts the difference-in-differences for the FPS between upturns and downturns and across fund types, which helps eliminate endogeneity concerns that are present in estimations of the single-differences alone. In particular, the estimate of the upturn-FPS may be driven by the allocation of new capital that flows into the sector, rather than by the re-allocation of capital within the sector. The double-difference approach filters out that explanation under the (verifiable) assumption that different types of funds experience aggregate net inflows in upturns compared to downturns. More precisely, we first make cross-sectional predictions about the FPS similar to BG and other existing learning theories. For example, when investors have less precise prior beliefs about funds' performance parameters (e.g., because the funds are small, young, have a new manager, or follow a particularly active investment style), each observation leads to a more pronounced updating of prior beliefs, resulting in a higher FPS for these funds. Unique to our model is the further prediction that these cross-sectional differences vary across states of the economy: funds that have a steeper unconditional FPS (e.g. funds with a particularly active style and therefore a high dispersion of investors' ex ante beliefs) are predicted to have a higher FPS-difference between upturns and downturns.

Both the cross-sectional and difference-in-differences predictions for the FPS are borne out in the data as well. We empirically proxy for investors' dispersion in beliefs using measures of active share and tracking error that are drawn from Cremers and Petajisto (2009). In particular, we take "Concentrated funds," that is, those ranking high by both active share and tracking error, as examples of funds for which investors have less precise prior beliefs. Figure 1 illustrates the empirical results comparing the FPS of Concentrated funds to all other funds in different states of the economy. As predicted by the model, the flow-performance

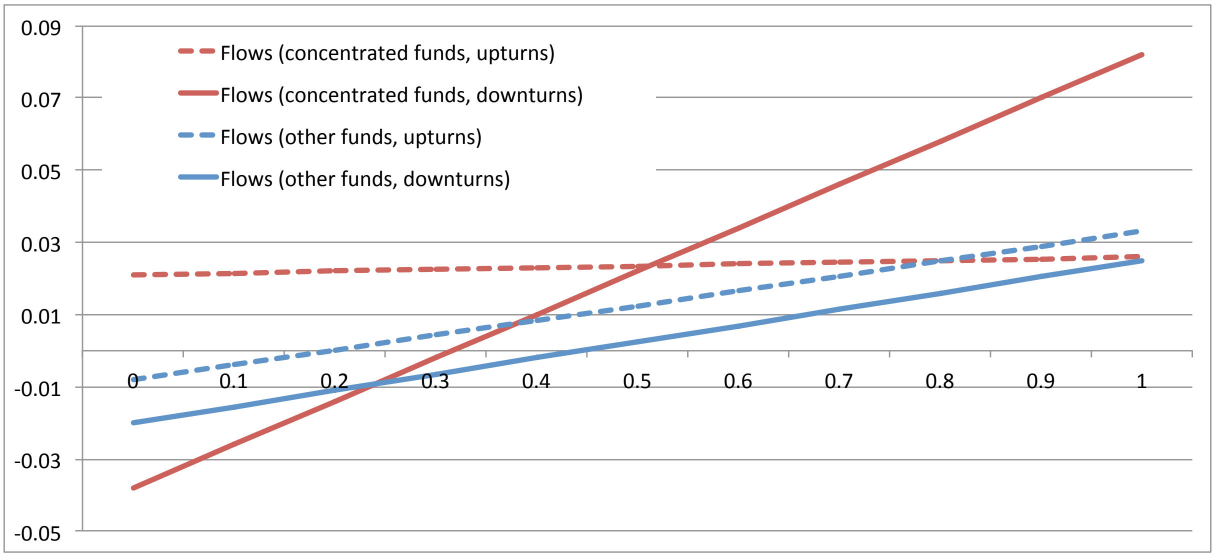

Figure 1: Flow-performance relation in upturns and downturns for "Concentrated funds" and all other funds (difference-in-differences results).

relation of Concentrated funds is steeper on average, and the difference of slopes across downturns and upturns is larger, compared to other funds. This difference-in-differences is highly significant.[4]

Both the model and empirics are robust to whether the flow-performance relation is convex or linear – a long-dating question (Chevalier and Ellison, 1997; Sirri and Tufano, 1998) recently reinvestigated by Spiegel and Zhang (2012): our theoretical model can be combined with participation costs, which generate convexity (Huang, Wei, and Yan, 2007), and our empirical results hold in both linear and convex specifications.[5]

While the focus of the paper is on the rationality of the flow-performance relation, the model also predicts the cross-sectional and time-series variation in performance persistence documented by the prior literature and confirmed in the present paper. First, trivially, funds

[4]We find similar results when Concentrated funds are replace with for younger, smaller, and more active funds, funds with managers of shorter tenure, and several interactions.

[5]See Lynch and Musto (2003) for an alternative explanation for why convexity obtains.

that are believed to have higher systematic risk exposure have to yield higher returns in equilibrium. Second, when managers face decreasing returns to scale, lower between-fund flows in upturns imply that performance is more persistent following upturns than following downturns: following upturns, funds that were too large (compared to the manager's ability to find profitable investments) remain too large and continue to underperform, while funds that are too small continue to overperform. The reason is that in upturns, investors – rationally suspecting that the returns are due to systematic risk exposure – do not allocate as much additional capital to outperforming managers as would be necessary for decreasing returns to scale to balance the fund manager's skill to generate high *average* returns. As a result, performance can be persistent following upturns but not downturns, while expected marginal utility is not persistently higher for investments in one fund than another at any point of time. Note that our paper does not make a claim whether there should be measureable (conditional) performance persistence or not. We merely posit that if performance persistence is found, it is consistent with a rational and largely frictionless model that performance is more persistent following upturns than following downturns. To reiterate, the motivation of the paper is to establish that investor decisions and fund flows are more consistent with rationality than previously thought. In this context, the focus on performance persistence serves as a convenient example because this issue has received considerable attention in the literature and has been closely linked to the question of whether investors' decisions are rational or not.

The paper proceeds as follows. Section 2 describes the relation of our model and empirical results to the existing literature. Section 3 presents the model. Section 4 describes the data, defines variables, and explains the identification strategy. Section 5 gives the empirical results. Section 6 concludes.

2 Related Literature

We first describe the relation of our model to the existing theoretical literature and, in particular, to the benchmark model of BG. Then, we describe the differences from the prior literature that concern both our theory and empirical results. Finally, we focus on the unique aspects of our empirical methodology and results.

Similar to BG, our model features: investors that provide capital to mutual funds in competitive ways; heterogeneity in the performance parameters of fund managers; decreasing returns to scale; and investors who rationally learn from past returns according to Bayes' law. A key difference from BG is that we allow for heterogeneous exposure of funds to time-varying benchmark returns. Moreover, we model investors who explicitly maximize a risk averse utility function rather than expected returns over a risk-adjusted benchmark.

Aside from risk aversion, the key new assumption of our model is that investors do not have perfect knowledge about the extent to which funds' cash flows load on the market return, i.e. systematic risk. This assumption may seem counterintuitive at first, as it is well-known that risk factor loadings of stock returns can be learned arbitrarily fast when they are continuously observed (Merton, 1980). However, estimating performance parameters of mutual funds is necessarily less precise because fund returns are not continuously observed and cannot be easily constructed as the portfolio is often rebalanced. Kacperczyk, Sialm, and Zheng (2008) show a large amount of portfolio rebalancing is concealed from investors.

In an effort to focus on the model's new predictions, we draw tighter boundaries than BG along a few dimensions. In particular, we do not derive the optimal compensation contract for the fund managers (Holmström, 1999), we do not endogenize the fee structure of the fund, and we do not explicitly discuss entry and exit from the mutual fund sector (Berk and Green, 2004).

While we allow for uncertainty with respect to funds' factor loadings, we are not the first to allow for uncertainty in more than one parameter. Pastor and Stambaugh (2012) allow for uncertainty with respect to the decreasing returns to scale parameter, which is assumed to be a known constant in our model. Their model explains the size of the actively managed fund industry, while ours focuses on cross-sectional and time-series differences in the sensitivity of investor flows to fund returns.

We model investor behavior in response to mutual fund performance in a given state of the economy, taking funds' performance parameters as given. Hence, the model is consistent with and does not take a stance on how assets are priced in the economy (in particular whether the CAPM holds or not), which asset pricing model fund managers use, whether managers are perfectly or only limitedly rational as in Kacperczyk, Nieuwerburgh, and Veldkamp (2012), and whether they generate abnormal performance by market timing or stock picking (Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2012), and whether the parameter distributions we assume are the result of strategic choice by managers or whether they are endowed with them.[6] Also, we do not model the fund-manager matching process (Gervais and Strobl, 2013), but we take the outcome as given.

Several authors have used the insight that risk-averse investors value mutual fund returns more in downturns than in upturns to study implications of the time-variation in the value of active management as a whole (e.g., Moskowitz (2000); Kallberg, Liu, and Trzcinka (2000); Kosowski (2006); Sun, Wang, and Zheng (2009); Glode (2011)). Our contribution is to study the implications of the same insight for the cross-sectional reallocation of capital within the

[6]A further distinction from Kacperczyk, Nieuwerburgh, and Veldkamp (2012) is that there is no asymmetric information in our model and that we do not assume parameter distributions or risk aversion to vary exogenously as the state of the economy changes. The latter element clarifies that asymmetric fund flows obtain in our model even if there is no asymmetry in the model parameters. Allowing for time-varying risk aversion or differences in parameter distributions as a function of the state of the economy would strengthen our conclusions.

mutual fund sector, which is reflected in the flow-performance relation.

Empirically, we complement the literature on the flow-performance relation (see Spiegel and Zhang (2012)) by documenting that the FPS reflecting within-sector flows across funds among retail mutual funds is almost twice as steep in downturns than in upturns. The paper also expands on existing empirical results on performance persistence. First, we confirm and extend the evidence on the time-variation of performance persistence: performance is persistent following market upturns, but not following market downturns (Glode, Hollifield, Kacperczyk, and Kogan, 2012). While these authors focus on persistence at horizons of three to twelve months, we focus on persistence at the monthly horizon following the model predictions. (For all specifications and subgroups, consistent with our model and the previous literature (Bollen and Busse, 2005), persistence is more pronounced at shorter horizons.) Second, we complement the existing evidence in Cremers and Petajisto (2009) and Petajisto (2013) by calculating cross-sectional differences in performance persistence. Consistent with the previous literature, we find that "Concentrated" funds tend to be associated with more persistence. Moreover, we document that this unconditional performance persistence of the Concentrated funds is driven by periods following market upturns.

3 Model

This section provides a rational model featuring Bayesian learning by mutual fund investors about managers' skill. We draw inspiration from Schmalz and Zhuk (2013) who model learning from cash flow news in an asset market. Aside from Bayesian learning, the model does not feature any frictions such as liquidity constraints or investors' limited rationality. Compared to BG, the crucial difference is that we explicitly allow for time-varying benchmark returns, heterogeneous exposure of funds' returns to that benchmark, and risk-averse investors who dislike such exposure to systematic risk, conditional on the average level of returns. Technically, this introduces a second parameter in investors' inference problem compared to BG.

3.1 Setup

There is a large number of funds, $i = 1, 2, ..., N$. The cash that fund i generates at time t from every dollar invested at time $t-1$ is denoted Y_t^i and can be described as

$$Y_t^i = 1 + a^i + b^i \cdot \xi_t - \frac{1}{\eta} S_t^i + \varepsilon_t^i \tag{1}$$

where ε_t^i are idiosyncratic shocks; S_t^i is the size of the fund; $\eta > 0$ is an efficiency parameter, such that $\frac{1}{\eta}$ indicates decreasing returns to scale; ξ_t is a market-wide shock with zero expected value that is normally distributed and *iid* over time, $\xi_t \sim \mathcal{N}(0, \sigma_\xi^2)$; and a^i and b^i are time-fixed fund-specific performance parameters that we assume to be distributed in an *iid* fashion. $a^i - \frac{1}{\eta} S_t^i$ is then fund i's expected return. b^i is the fund's exposure to the market-wide and time-dependent shock ξ_t.[7] Fund investors take a^i and b^i as exogenous.

Investors are uncertain about the precise value of the parameters a^i and b^i for a given firm, but they believe that the parameters are drawn from jointly Normal distributions with known mean, variance, and covariance. We omit the i-superscripts for simplicity for the

[7] We write a^i and b^i instead of α^i and β^i to avoid confusion with asset (as opposed to fund) characteristics, and to avoid an unwarranted association of the model with the Capital Asset Pricing Model. a^i and b^i can be understood as the cash-flow alpha and beta of the fund, respectively.

means and variance/covariance terms.

$$\begin{pmatrix} a^i \\ b^i \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix}, \begin{pmatrix} \bar{\sigma}_a^2 & \bar{\sigma}_{ab} \\ \bar{\sigma}_{ab} & \bar{\sigma}_b^2 \end{pmatrix}\right) \tag{2}$$

3.2 Discussion

Higher a^i, for given b^i, can be interpreted as either stock picking or market timing skill. The stock picking interpretation is straightforward. To gain intuition on market timing, one can assume for a moment that the b^i of manager i, is allowed to vary over time in a systematic fashion, b_t^i, while the manager has no stock picking skill (that is, $a^i = 0$). In particular, suppose b_t^i is always positive in upturns (when $\xi_t > 0$) and always negative in downturns (when $\xi_t < 0$) . This manager generates high average returns Y_t^i with zero market correlation. To an investor running the regression (1) that constrains market exposure to be time-fixed b^i, this manager appears to have high static a^i and zero static b^i. To the investor, this manager's performance is therefore observationally equivalent to the performance of a manager without timing ability (constant and relatively low b^i) but with stock picking skill (constant and relatively high a^i). In fact the investor is indifferent to how the manager generates returns. The investor merely evaluates cash flows which are appropriately weighted by her marginal utility in the state of the world in which the cash flow occurs. This example illustrates that the investor can be assumed to remain ignorant about the particular sources of skill of the manager. In sum, equation (1) is not necessarily the true cash flow process of funds, nor do investors need to believe that it is. It is merely one possible, and convenient, description of cash flows that is sufficient to describe the investors' inference problem.

While the model is generally compatible with any strategy managers may employ to generate returns, including stock picking and market timing, it is of course possible to construct cases in which the assumption of normality of the parameter distributions is violated. For example, if managers systematically have skill only in particular states of the economy, investors could not reasonably believe that the parameter distributions are normal. However, we deliberately assume symmetric parameter distributions to emphasize that asymmetric behavior across states of the economy obtains as an outcome of the model, even if parameters and their distributions do not change as a function of the state of the economy. For example, allowing for higher macro volatility in downturns than in upturns, i.e., a negatively skewed ξ_t, or increased risk aversion in downturns, would strengthen the predictions.

Moreover, we implicitly assume that investors do not directly invest in assets, but do so only through funds. As a result, they are not concerned about how assets are priced in this economy; investors evaluate only investments in mutual funds. In that sense, we present a partial equilibrium model. Similarly, which asset pricing model managers use is not important for our model. In particular, it is not assumed that the CAPM is used or holds.

Relatedly, investors are assumed to be unable to hedge their *aggregate* exposure to systematic risk through investments outside the mutual fund sector. In other words, non-trivial results obtain only if the investors' stochastic discount factor is allowed to vary with the state of the market. In contrast, we do not assume an inability to hedge systematic risk exposure of a particular fund. Rather, it is an equilibrium condition in the market for capital (which we model in general equilibrium) that no net hedging demand can exist.

3.3 Equilibrium

We model mutual fund investors as overlapping generations of a representative agent. Further, there is a risk-free asset that returns $R = 1 + r$. Without loss of generality, we set the risk-free rate to zero, $r = 0$. Equalizing the investors' marginal expected utility across funds yields the equilibrium fund size.

Lemma 1. *Fund i's equilibirum size S_t^i, based on beliefs about performance parameters at time t is given by*

$$S_t^i = \eta \cdot (E_t[a^i - \phi b^i]) \tag{3}$$

where $\phi = -E_t[m_{t+1}\xi_{t+1}]$ and m_{t+1} is the stochastic discount factor. Moreover, when the investors' utility over cash flows is given by $u(Y) = -exp(-\gamma Y)$, where γ is the investors' effective risk aversion,

$$\phi = \gamma N\bar{b} \cdot \sigma_\xi^2. \tag{4}$$

The intuition of the lemma is straightforward. Investors determine allocations to funds so that the expected utility of a marginal dollar in each fund equals the outside option of zero. In doing that, the value of expected fund returns, $1 + a^i - \frac{1}{\eta}S_t^i$, is adjusted for the fund's sensitivity to the risk factor, b^i, multiplied by the factor ϕ, which represents how much investors care about the funds' risk exposure. With exponential utility, ϕ takes a simple tractable form. Specifically, how much investors care about systematic risk is given by the product of risk aversion and the total amount of risk in the market. The total amount of risk is the product of the size of the market, average sensitivity to the risk factor, and the riskiness of the market-wide factor. If investors are risk neutral, that is$\phi = 0$, they only care about expected returns but not about their timing. For intuition, BG's setup can be thought of as a special case of our economy, in which ϕ is set to zero.

3.4 Fund Flows

The main insight of the model is that the sensitivity of flows, F_t^i, to unexpected performance, $Y_t^i - E_{t-1}[Y_t^i]$, depends on the state of the market, ξ_t.

Lemma 2.

$$F_t^i := S_t^i - S_{t-1}^i = \eta \cdot \lambda(\xi_t) \cdot (Y_t^i - E_{t-1}[Y_t^i]) \tag{5}$$

where

$$\lambda(\xi_t) = \frac{\sigma_a^2 - \phi\sigma_{ab} - (\phi\sigma_b^2 - \sigma_{ab})\xi_t}{\sigma_a^2 + 2\sigma_{ab}\xi_t + \sigma_b^2\xi_t^2 + \sigma_\varepsilon^2} \tag{6}$$

The quantities σ_a^2, σ_{ab}, etc. denote prior beliefs about the variance and covariance of a and b, corresponding to equation (2). $\lambda(\xi_t)$ is the flow-performance sensitivity (FPS).[8]

The intution for expression (6) is best illustrated with the particular case of $\sigma_{ab} = 0$, which can be thought of the case in which the agent is symmetrically informed about fund behavior in upturns and downturns prior to the observation we study here. In this case,

$$\lambda(\xi_t; \sigma_{ab} = 0) = \frac{\sigma_a^2 - \phi\sigma_b^2\xi_t}{\sigma_a^2 + \sigma_b^2\xi_t^2 + \sigma_\varepsilon^2}. \tag{7}$$

First, notice that if fund i's parameter b^i is not uncertain, $\sigma_b^2 = 0$, then λ does not depend on the state of the economy, ξ_t, and there is no asymmetry in flows between upturns and downturns. Moreover, a familiar result is that the more dispersed cash-flow alphas are believed to be, i.e. the higher σ_a^2, the stronger the reaction to news, i.e., the steeper the FPS. Intuitively, if very high and low fund returns are deemed realistic and attributable to exceptionally high or low skill, rational investors are less prone to attribute abnormal returns to random noise. Conversely, if signals are less informative, i.e., if σ_ε^2 is higher relative to σ_a^2,

[8] For the difference in fund sizes $S_t^i - S_{t-1}^i$ to correspond to between-fund flows, it is implicitly assumed that each fund i distributes the net return $Y_{i,t} - 1$ at the end of period t.

then one-time abnormal performance of a given size triggers lower flows. Last, introducing uncertainty about b^i makes the FPS depend on the state of the economy, ξ_t. A positive σ_b^2 dampens the FPS in upturns and steepens it in downturns. This example illustrates that the state-dependence of the flow-performance relation explicitly depends on prior beliefs.

Schmalz and Zhuk (2013) formally analyze this dependence in more detail in the context of an asset market. As translated to the present context, they show that there exists a threshold prior belief $\bar{\sigma}_{ab}^*$, which is always positive and bounded by $\sigma_a\sigma_b$, $0 < \bar{\sigma}_{ab}^* < \sigma_a\sigma_b$, such that the expected change of fund size, and thus relative fund flows, are larger in downturns $\xi_t = -x$ than in upturns $\xi_t = +x$ of equal size x, for all prior beliefs below that threshold, $-\sigma_a\sigma_b \leq \sigma_{ab} < \bar{\sigma}_{ab}^*$. The opposite is true for $\bar{\sigma}_{ab}^* < \sigma_{ab} \leq \sigma_a\sigma_b$. Notice that the set of beliefs for which $-\sigma_a\sigma_b \leq \sigma_{ab} < \bar{\sigma}_{ab}^*$ is always larger than the set of beliefs for which $\bar{\sigma}_{ab}^* < \sigma_{ab} \leq \sigma_a\sigma_b$. Thus, the asymmetry emphasized in the present paper is shown to hold for a majority of prior beliefs.

We have no reason to believe that prior beliefs σ_{ab} are different from zero. We nevertheless present the general result in order to show that the qualitative predictions do not crucially depend on strong assumptions about prior beliefs. For the empirical predictions that follow, the disclaimer "for a majority of prior beliefs" is implicit, as is the notion that upturns and downturns of equal magnitude are compared. These conditions are made explicit in the proofs.

3.5 Empirical Predictions

We make two sets of empirical predictions. First, without further assumptions, the model predicts that FPS are higher in downturns than in upturns. We also show that performance is more persistent following upturns than following downturns. The second set of predictions concerns the cross-sectional variation and the difference-in-differences of the flow-performance sensitivity. The only additional assumption needed for these predictions is that investors have less precise ex ante beliefs about returns of funds that display higher ex-post dispersion of fund returns and return correlations with the market.

3.5.1 Upturn-Downturn Difference of Flow Performance Sensitivity

The first testable prediction is that the flow-performance sensitivity (FPS) is smaller in upturns (UT) than in downturns (DT).

Proposition 1. *Flow-performance sensitivities are larger in downturns than in upturns,*

$$\lambda(DT) - \lambda(UT) > 0.$$

The intuition for the asymmetry can best be seen as follows. Investors observe fund returns $Y_t^i = 1+a^i+b^i \cdot \xi_t - \frac{1}{\eta} S_t^i + \varepsilon_t^i$ (equation (1)), from which they can infer ξ_t. (Alternatively, one can think of ξ_t as being perfectly observed.) Note that the expected quantity that investors observe increases in a^i and increases in b^i when ξ_t takes positive values (upturns), but increases in a^i and decreases in b^i in downturns. In contrast, what investors try to infer – the equilibrium fund size – always increases in a^i and decreases in b^i: $S_t^i = \eta \cdot (E_t[a^i - \phi b^i])$ (equation (3)). Thus, what investors observe and what they are interested in tend to be more aligned in downturns than in upturns.

More formally, omitting variables without information content over and above prior beliefs such as fund size, investors make the following projection

$$E_t[a^i - \phi b^i | a^i + b^i \cdot \xi_t + \varepsilon_t^i] = E_t[a^i - \phi b^i | a^i - \phi b^i + b^i \cdot (\phi + \xi_t) + \varepsilon_t^i]. \tag{8}$$

Note that the righthand side of each expectation – the conditioning information: fund returns – consists of the quantity of interest on the lefthand side plus an error term. The error is smaller, in terms of its variance, when ξ_t takes a negative value than when it takes a positive value of equal magnitude, because ϕ is positive and $(\phi + \xi_t)$ is smaller in absolute value. It can be said that a component of the noise is "switched off" in downturns. Thus the asymmetry.

One can also think of equation (8) as a projection of a vector on a plane. In upturns, the vector stands more orthogonal to the plane; in downturns, it is more aligned with the plane. Thus, it is easier to infer the length of the vector in downturns than in upturns. To reiterate, no asymmetry between upturns and downturns obtains when investors are risk-neutral, and thus $\phi = 0$. In this case, the relative position of the vector the plane on which it is projected does not change. The formal proof is in the appendix.

3.5.2 Difference-in-Differences Prediction for FPS

When there are differences in the precision of the investors' ex ante beliefs across fund types, cross-sectional variation of the FPS across these fund types arises. In particular, for some positive constant $k > 0$, define

$$\bar{\sigma}^2_{a,Concentrated} = k \cdot \bar{\sigma}^2_{a,Other} \tag{9}$$

$$\bar{\sigma}^2_{b,Concentrated} = k \cdot \bar{\sigma}^2_{b,Other} \tag{10}$$

where *Concentrated* indexes funds that have both high active share and high tracking error as in Cremers and Petajisto (2009) and *Other* stands for all other funds, and furthermore the correlation of a and b remains the same, such that $\sigma_{ab,Concentrated} = k \cdot \sigma_{ab,Other}$. It then follows immediately from the $\lambda(\xi_t)$ expression in lemma 2 that investors react with higher flows to a given piece of news if it pertains to *Concentrated* funds. (A similar prediction obtains from established models of funds flows such as BG in comparing funds with higher parameter uncertaint to funds with lower uncertainty, such as young and old funds.)

Proposition 2. *Average flow-performance sensitivity is higher for Concentrated funds than for Other funds.*

$$\lambda_{Concentrated} - \lambda_{Other} > 0.$$

Note that the proposition relies on differences in the distributions of parameters introduced in equation (2), while it assumes that both funds have investors with similar preferences. Thus it should hold across types of funds that are held by similar investors, but it need not hold across funds with investors with distinct risk preferences. For example, we would not necessarily expect the proposition to hold across funds that are predominantly held by retail investors versus funds that are predominantly held by institutions, or across mutual funds and hedge funds.

We can now state the second main empirical prediction, the difference in FPS-differences across market states and fund types.

Proposition 3. *The difference in flow-performance sensitivites between downturns (DT) and upturns (UT) is larger for Concentrated funds than for Other funds.*

$$(\lambda_{DT} - \lambda_{UT})_{Concentrated} - (\lambda_{DT} - \lambda_{UT})_{Other} > 0$$

The intuition is that while learning about risk-adjusted performance is difficult in upturns for all types of funds, learning from downturn performance is particularly effective for funds with a high variance of prior beliefs about the fund's market correlation.

The difference-in-differences prediction in proposition 3 is unique to our model. If this prediction finds support in the data, we will say the model is "identified" in the sense of ruling out several alternative explanations for the upturn-downturn difference predicted in proposition (1). For example, one might suspect that "everybody is happy in upturns" and does not check fund performance, while investors scrutinize fund performance in downturns; therefore, the upturn-downturn difference in FPS ensues. However, such a behavioral theory would not easily explain why the upturn-downturn difference would differ across different types of funds. In section 5, we describe alternative theories in more detail.

3.5.3 Upturn-Dowturn Difference of Performance Persistence

Proposition 1 says that investors react with lower flows to a given performance in upturns than in downturns. As a result, the after-flow size of an outperforming fund is smaller following an upturn than it would have been following a similar performance in downturns. The manager facing decreasing returns to scale is therefore less burdened with an increased fund fize after upturns, and is more likely to outperform again. Specifically, we define performance persistence as a positive correlation of risk-adjusted performance in consecutive periods, $\rho(Y^i_{t+1} - \xi_{t+1}, Y^i_t - \xi_t)_{\xi_t}$, where the subscript indicates that the correlations will be calculated conditional on a set of market states $\xi_t \in \{UT, DT\}$. The following proposition states that risk-adjusted fund performance is more persistent following market upturns than following market downturns.[9]

Proposition 4. *Performance is more persistent following upturns than following downturns.*

$$\rho(Y^i_{t+1} - \xi_{t+1}, Y^i_t - \xi_t)_{UT} - \rho(Y^i_{t+1} - \xi_{t+1}, Y^i_t - \xi_t)_{DT} > 0$$

[9]Note that no claim is made about the absolute level of performance persistence, in particular whether there is persistence following downturns or whether persistence can be measured unconditioncally.

Note that the model allows the compensation for risk in steady state to be positive, zero, or negative, both for individual funds and for the industry. In other words, our predictions focus on the sign of the difference in performance persistence across market states, without making claims about the level of performance persistence that we should expect to measure.

While we omit a derivation of the model predictions when investors care about exposure of fund returns to more than one factor, of course, our empirical pendant of risk-adjusted permance also controls for the other Carhart (1997) factors besides market returns.

3.5.4 Cross-sectional Prediction for Performance Persistence

While not a key prediction of the paper, the model also predicts cross-sectional differences in performance persistence that obtain empirically. It follows immediately from equation (**??**) and proposition (4) that more persistence can be found within subsets of funds with higher dispersion of prior beliefs σ_b^2 .

Proposition 5. *Performance is more persistent for Concentrated funds than for Other funds on average.*

$$\rho(Y_{t+1}^i - \xi_{t+1}, Y_t^i - \xi_t)_{Concentrated} - \rho(Y_{t+1}^i - \xi_{t+1}, Y_t^i - \xi_t)_{Other} > 0$$

3.6 Model limitations

The first limitation is that the model is essentially static. Taking the model as is seems to imply that after sufficiently many observations, investors learn parameter values well enough for the FPS first to fall, and then for all flows to disappear. A more realistic model would assume that funds periodically disappear (for exogenous reasons or because they perform below a threshold) and get replaced with new ones, about which little is known. Doing so, we could derive cross-sectional predictions about the FPS as a function of the age of funds, or fund manager tenure. However, these predictions have been derived by Berk and Green (2004) and many others before and would therefore not add new insights. Also, Schmalz and Zhuk (2013) discuss such an extension in a learning framework similar to the one in the present paper. We omit such a derivation here to focus on the key predictions, which are about variations across the state of the economy. That key prediction would become substantially more involved and less intuitive in a more elaborate, truly dynamic model. A closely related limitation results because our model of cross-sectional learning is tractable only because there is no learning about the stochastic discount factor at the same time. To our knowledge, no existing model is able to track both the cross-sectional and time-series dimension at the same time.

One observation about dynamics is in place however. When the model is modified in a way that parameters get periodically re-assigned for a fraction of funds, the average precision of beliefs about fund value is highest following downturns, compared to following downturns. As a result, fund sizes most closely match a first-best allocation without parameter uncertainty following downturns. We view this as reminiscent of Schumpeter's assertion that recessions have a cleansing effect on the economy, as applied to the funds sector.

The second main limitation is that the parameters a^i and b^i are exogenous in the model. There are several reasons for this modeling choice. First is that the key predictions would be very difficult to obtain if the parameters were endogenous and a result of fund managers' choice. Second, including the managers' choice would come at the expense of having to make assumptions about their preferences and incentives, which would obscure which part of our results comes from assumptions about investor preferences, and which part would obtain from assumptions about managers' incentives. We want to make clear that the present model is about investor behavior alone. It is not precluded in the model, however, that the parameter distributions are already the outcome of an optimization on behalf of the fund managers. Studying the interaction of investor and manager behavior when a^i is exogenously distributed and known to the manager but uncertain to investors and b^i a strategic choice of the manager and likewise uncertain to the investor may be an interesting subject for future research.

A third observation is that any cross-sectional predictions, and namely the difference-in-difference prediction on the FPS, rely on a homogenous set of investors across different types of funds – recall the key role the risk aversion parameter γ plays in the composition of ϕ, which, among others determines the extent of upturn-downturn differences of the FPS. The model predictions should therefore hold true within a set of funds with a reasonably homogenous investor base.

4 Description of the Data

The primary data source for this study is the CRSP Survivorship Bias Free Mutual Fund Database. These data contain fund returns, total net assets (TNA), investment objectives, and other fund characteristics. Following the prior literature, we select Domestic Equity open-end mutual funds and exclude sector funds using the CRSP objective code (which maps Strategic Insights, Wiesenberger, and Lipper objective codes). Because the reported objectives do not always indicate whether the fund is balanced, we exclude funds that on average hold less than 80% of their assets in stocks. Given that the focus of this study is on actively managed mutual funds, we also exclude index funds.

To address the potential bias resulting from the fact that the fund incubation period is also reported, we exclude observations for which the data is prior to the reported starting date of the fund, similar to Kacperczyk, Van Nieuwerburgh, and Veldkamp (2012). As incubated funds tend to be smaller, we exclude observations with reported assets under management smaller than $5 million in the prior quarter.

Mutual funds in CRSP include both retail and institutional share classes. The predictions of our model are based on a homogenous set of investors. So, pooling two classes of investors would blur the empirical tests of the model predictions. Besides, institutional funds are subject to a number of constraints in terms of minimum investment size, long term investment agreements, and limited choice set whenever they are offered to individuals through a 401(k) plan. These arguments prompt us to restrict our empirical analysis to mutual funds that are sold to retail investors and exclude institutional funds. The retail-fund indicator is available in CRSP starting in December 1999. For the prior years, we backward impute the retail indicator whenever available and we use the names of share classes to identify institutional funds. The funds for which no information can be gathered on whether they are retail or institutional are excluded from the sample.

The sample spans the years from 1980 to 2012, when complete information on investment objectives is available. Since CRSP does not report monthly TNA until 1990, we follow the existing literature and use quarterly data for the flow-performance sensitivity analysis (Huang, Wei, and Yan, 2007).

Using the quarterly net asset values and returns from CRSP, we compute net flows according to the literature standard as

$$Flows_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}\left(1 + R_{i,t}\right)}{TNA_{i,t-1}} \tag{11}$$

where $TNA_{i,t}$ is total net assets in quarter t for fund i and $R_{i,t}$ is the quarterly return, which are obtained from cumulating monthly returns. Elton, Gruber, and Blake (2001) point out a number of errors in the CRSP mutual fund database that could lead to extreme values of returns and flows. For this reason, following Huang, Wei, and Yan (2007), we filter out the top and bottom 2.5% tails of the the returns and net flows distributions.

Between 1980:Q1 and 2012:Q4, we have 144,382 mutual fund-quarter observations with valid information on returns and TNA in quarter t and quarter $t + 1$, corresponding to 5,763 funds.[10] The other variables that are used in the analysis, and for which we require availability for sample inclusion, are the expense ratio, the portfolio turnover ratio, and return volatility, which is computed over the prior twelve months. These variables are winsorized at the 1^{st} and 99^{th} percentiles. We compute fund age as the time (in quarters) since the first appearance of the fund in the overall CRSP sample. Table 2 reports summary statistics for these variables. From Panel A, we notice that the average (median) fund has a size of $678 million ($82 million). The maximum fund size is about $109 billion. Fund age ranges from five to 151 quarters. Our sample is comparable to other studies in terms of return volatility, asset turnover, and expense ratio (see Huang, Wei, and Yan (2007)). The performance persistence analysis is run on a monthly version of the above-described sample.

Part of our study makes use of data on active share and tracking error, which are defined as in Cremers and Petajisto (2009) and Petajisto (2013).[11] These variables are constructed using information on portfolio composition of mutual funds as well as their benchmark indexes. The stock holdings of mutual funds come from the CDA/Spectrum database provided by Thomson Financial. The authors currently make their data available between 1980 and

[10]Starting in the 1990's, some funds offer multiple share classes that represent claims to the same portfolio. Some authors aggregate different share classes at the portfolio level (see, e.g., Glode, Hollifield, Kacperczyk, and Kogan (2012)). Following Huang, Wei, and Yan (2007), we abstain from this aggregation as our purpose is to study fund flows, which differ at the share class level. Nevertheless, our results are not materially impacted by this choice.

[11]We are grateful to Antti Petajisto for making the data available on his website: www.petajisto.net

2009.

To define upturns and downturns, we procede as in Glode, Hollifield, Kacperczyk, and Kogan (2012) and use the distribution of the excess return on the market up to quarter t. A quarter is denoted as an upturn if the excess return on the CRSP value weighted index for that quarter lies in the top 25% of the distribution of the quarterly excess market returns up to quarter t. Symmetrically, a quarter is a downturn if the realization of the market in that quarter is in the bottom 25% of the distribution. In computing the distribution of the market excess return, we use the history going back to the third quarter of 1926. As a result, out of the 131 quarters in our sample, 32 are upturns and 30 are downturns. When using monthly data, we proceed similarly in defining upturns and downturns. Panels B and C of Table 2 have summary statistics on the relevant variables at the quarterly frequency in the subsamples of upturns and downturns. As expected, returns and flows are on average larger in upturns, while other variables are similar in magnitude across market states.

5 Empirical Methodology and Results

Next, we turn to the empirical analysis. In Section 5.1, we describe the variation of the flow-performance sensitivity across states of the economy and fund types, as well as the difference-in-differences of the FPS, which identifies the model in the sense of ruling out a variety of alternative stories. In Section 5.2, we extend the existing evidence on performance persistence by showing that, consistent with the model, (i) persistence is present only in periods following upturns, also for subgroups of funds, and that (ii) performance persistence, while driven mostly by upturns, is particularly strong for concentrated funds, as defined by Cremers and Petajisto (2009).

5.1 Flow-Performance Sensitivity Results

The predictions of the model are expressed in terms of the difference in slope of the relation between flows and performance (flow-performance sensitivity, FPS) between upturns and downturns. The most direct way to test these predictions is through a linear a specification. Then, our focus is on the slope b in the regression:

$$Flows_{i,t+1} = a + b \cdot frank\_style_{i,t} + \varepsilon_{i,t} \tag{12}$$

where $frank\_style_{it}$ is the fractional rank of fund i in period t with respect to funds in the same style. For mutual funds, the style is defined by the CRSP Objective variable. We estimate the regression in equation (12) using the Fama and MacBeth (1973) methodology.

A large body of literatur, (starting with Ippolito (1992), Gruber (1996), Chevalier and Ellison (1997) and Sirri and Tufano (1998), identifies a convex flow-performance relation. More recently, other authors (Spiegel and Zhang, 2012) argue that convexity originates from a mispecified empirical model, and that the relation between flows and performance is truly linear. This paper does not intend to contribute to this debate, given that our predictions on the state-dependency of the flow-performance relation are insensitive to the shape of this relation. Still, to assess the robustness of our predictions to alternative empirical specifications of the shape of the flow-performance relation, we also estimate a piecewise linear relation

$$Flows_{i,t+1} = a + b_1 \cdot trank\_style1_{i,t} + b_2 \cdot trank\_style2_{i,t} + b_3 \cdot trank\_style3_{i,t} + \varepsilon_{i,t} \tag{13}$$

where $trank\_style1_{i,t} = \min(\frac{1}{3}, frank\_style_{i,t})$, $trank\_style2_{i,t} = \min(\frac{1}{3}, frank\_style_{i,t} - trank\_style1_{i,t})$, and $trank\_style3_{i,t} = \min(\frac{1}{3}, frank\_style_{i,t} - trank\_style1_{i,t} - trank\_style2_{i,t})$

Table 3 presents the main results for the linear specification. This analysis provides a test of proposition 1 which states that the flow-performance relation is steeper in downturns than in upturns. The table describes the variation of FPS across states of the economy. The first three columns give results that do not include additional controls, as in equation (12). Column (1) reports a significant FSP of 0.043, without conditioning on the state of the market. The FPS is more than twice as large in downturns (0.051) compared to upturns (0.021) for the average fund, emphasizing the economic significance of the result. The difference is also highly statistically significant (bottom of column (3)). Columns (4)-(6) include all controls suggested by Spiegel and Zhang (2012). These are the aggregate flows in quarter $t+1$ into the funds that have the same objective as fund i, the total expense ratio, the logarithm of TNA, the portfolio turnover ratio, the return volatility over the prior twelve months, and the logarithm of the fund's age. Given that flows display some persistence, we also include the fund's flows in quarter t. After adding these controls, the magnitudes as well as the statistical significance of the upturn/downturn difference (bottom of column (6)) are preserved. Overall, this evidence provides an empirical validation of proposition 1.

Table 4 has the estimates for the piecewise linear specification in equation (13). Consistent with the prior literature, we find evidence of convexity of the flow-performance relation (columns (1) and (4)). More relevant for our purposes, the evidence strongly supports the predictions of the model. In each interval of the domain of the piecewise linear specification, the FPS is larger in downturns than in upturns (columns (2) and (3)). This result holds also when we include the controls (columns (5)-(6)). At the bottom of columns (3) and (6), we report p-values from a chi-squared test for the equality of the three slopes b_1, b_2, and b_3 between upturns and downturns. The test rejects the null hypothesis. Given the consistency of the conclusions between Tables 3 and 4, we feel legitimized to proceed with the linear specification, which more easily allows us to test the difference-in-difference predictions of the model.

The next step is to identify the model by testing its difference-in-difference prediction across fund types and market states given in Proposition 3. Econometrically, the double-difference result rules out a number of alternative theories that could drive the variation in the FPS across market states. For example, suppose that new capital that flows into the mutual fund sector gets primarily allocated with medium performers, possibly attenuating our upturn-FPS estimate, while outflows from the sector primarily hit underperformers, which might steepen the FPS-estimate. Taking the difference across fund types of the upturn-downturn difference would eliminate such an effect.

The first step to construct the difference-in-differences test is to define what constitutes the cross-sectional variation. To proxy for the heterogeneity in degree of ex ante uncertainty about a particular fund's parameters (captured by the model parameters $\sigma_a$and σ_b) we use the variables constructed by Cremers and Petajisto (2009) and Petajisto (2013). In detail, we conjecture that, for the funds that these authors label 'Concentrated,' investors have higher uncertainty about risk loadings and skill. According to the authors, Concentrated funds are those that rank highest by both active share and tracking error. In our application, a Concentrated fund is one that appears in the top half of the distribution of these two variables. Our intutition is that the extent of active management that characterizes these funds, both in terms of stock picking and sector rotation, makes it more difficult for investors to precisely know the underlying parameters of the distribution of these funds' returns. Similar results obtain when other proxies for less certain distributions are chosen, such as younger funds, smaller funds, or funds with a manager with shorter tenure. The second step is to calculate the upturn-downturn differences for each of these types of funds, and then taking the difference in differences.

Table 5 summarizes the variation of the FPS across funds and differences of that variation across market states. The first column shows that the flow-performance relationship is almost 50% steeper for Concentrated funds (coefficient on the interaction $frank\_style \times concentrated$), as predicted by proposition 2. The difference is statistically significant (t-stat=2.015). Columns 2 and 3 show that the cross-sectional difference is entirely driven by downturns: the FPS for Concentrated funds in downturns is 0.12 compared to the FPS for Other funds of 0.045, whereas the FPS in upturns is not significantly different across fund types. The difference between the FPS of Concentrated and Other funds is therefore much higher in downturns than in upturns. In other words, the difference-in-difference (between downturns and upturns and between Concentrated and Other funds) is 0.111 and is highly significant (p-value<0.01, see the test at the bottom of columns 2 and 3). This result confirms the prediction made in proposition 3 and therefore identifies the model. Columns 4-6 report qualitatively and quantitatively similar results after the introduction of controls.

5.2 Performance Persistence Main Results

In this section, we look for evidence that performance persistence is stronger following market upturns than following market downturns. We thereby test the empirical relevance of proposition 4.

We look for evidence of persistence by testing whether fund ranking in month t predicts abnormal performance in month $t + 1$. Each month, we rank funds within each investment objective by their monthly return. We focus on the top and bottom 20% of funds by this ranking (Good and Bad performers). Then, we form two equally-weighted portfolios of Good and Bad performers and hold them for a month. In so doing, we obtain two time series of monthly portfolio returns.

We estimate alphas from four-factor-model regressions (Carhart, 1997) for Bad and Good performers and for the long-short portfolio of Good minus Bad performers. The estimates are reported in Table 6. We estimate the model for all months (columns 1-3) and we also condition the estimation on whether the ranking month was an upturn (columns 4-6) or a downturn (columns 7-9).

When considering all months, we find that fund performance is persistent and the predictability mostly originates from Bad performers. This finding resonates with Carhart (1997). The long-short portfolio of Good and Bad performers earns an abnormal return of 0.7% on average in the month after the ranking period. However, consistent with proposition 4, there is a distinct upturn-downturn asymmetry. Following upturns, the evidence of persistence is much stronger and it originates mostly from Good performers. The persistence in performance of top funds resonates with the results in Kosowski, Timmermann, Wermers, and White (2006). The long-short portfolio earns a sizeable 1.2% abnormal return. Also predicted by our model, no significant outperformance is detected following market downturns. This holds for the Good as well as the Bad performers.

Overall, the results in Table 6 validate proposition 4. Also, this evidence corroborates the finding in Glode, Hollifield, Kacperczyk, and Kogan (2012) that performance persistence is stronger following upturns. We extend their results by presenting novel evidence that predictability in upturns originates mostly from Good performers.

5.3 Extension of Performance Persistence Results

This section extends existing evidence on performance persistence both in the cross-sectional dimension and across states of the economy. In addition to the above-described performance rank, each month we also group funds by a proxy for parameter uncertainty. We use the Concentrated vs. Other funds classification from Section 5.1. Then, four portfolios are formed combining the performance sorts and the grouping by parameter uncertainty.

Table 7 compares the average performance persistence of Concentrated and Other funds over the entire period. A long-short portfolio of Good minus Bad performers earns positive abnormal returns for both classes of funds. However, the return of the long-short portfolio for Concentrated funds earns 0.8% per month, which is twice the return on the long-short portfolio for Other funds. The differences are statistically significant (column 7). This evidence is an empirical validation of the model predicition that is summarized in proposition 5.

Table 8 repeats the analysis separately for periods following upturns and for periods following downturns. Consistent with the main results in Table 6, the performance persistence of all portfolios and all fund types is driven by periods following upturns.[12]

5.4 Robustness Checks

The online appendix provides several robustness checks. In particular, it reports similar results on FPS and performance persistence as those given in the paper for different subsets of funds including small vs. large, young vs. old, etc. Furthermore, it shows robustness with respect to a different specification of the FPS analysis in which we replace fund flows with market shares, as suggested by Spiegel and Zhang (2012).

[12]Note that a higher flow-performance sensitivity of Concentrated funds does not imply that they should have lower performance persistence. The FPS is driven by variation in the a^i parameter, while performance persistence is driven by variation in b^i. A fund type with higher variance both along the a^i and b^i dimension can have higher FPS and higher performance persistence in a given state of the economy. Difference-in-differences predictions for performance persistence therefore require more restrictive assumptions about the relative variance of the parameter distributions. For clarity of exposition, we omit these predictions.

6 Conclusion

We provide a rational and largely frictionless model that captures key regularities left unexplained by existing rational models. In particular, our model predicts that rational investors re-allocate less capital between funds following upturns than following downturns, leading to lower flow-performance sensitivity and higher performance persistence following upturns, compared to downturns. We show that the model predictions about the flow-performance sensitivity in different states of the economy, across types of funds, and about the difference-in-differences are strongly confirmed empirically. Specifically, the flow-performance relation is more than twice as steep following downturns than following upturns. Finally, we extend existing empirical evidence on performance persistence both across states of the economy and fund types.

Appendix

It is useful to derive an additional lemma before proving Lemma 1.

Lemma 3. *Based on current beliefs, investors value each dollar invested in the fund according to*

$$p_t^i = E_t[1 + a^i - \phi b^i - \frac{1}{\eta} S_t^i] \quad (14)$$

where

$$\phi = \gamma N \bar{b} \cdot \sigma_\xi^2 \quad (15)$$

Proof of Lemma 3 (Fund Value)

First, we derive the stochastic discount factor in the economy. Then we use this result to derive the value of a fund, conditional on beliefs and its equilibrium size. Assuming a set of overlapping generations mutual fund investors with homogeneous preferences allows us to obtain analytic solutions for cross-sectional learning, because we do not need to keep track of learning about the aggregate economy. (See Schmalz and Zhuk for a more detailed explanation.)

Assume an agent consuming Y_{t+1} at $t+1$ is buying x units of an asset (here: a fund) that pays Z_{t+1} at $t+1$ and cost p_z at t. Then, the expected utility of the agent

$$U(x) = E_t[u(Y_{t+1} + x(Z_{t+1} - Rp_z))].$$

The agents' utility must be maximized when $x = 0$

$$0 = U'(x)|_{x=0} = E_t[u'(Y_{t+1})(Z_{t+1} - Rp_z)]$$

$$0 = E_t[u'(Y_{t+1})Z_{t+1}] - E_t[u'(Y_{t+1})]Rp_z$$

$$p_z = \frac{1}{R} \cdot \frac{E_t[u'(Y_{t+1})Z_{t+1}]}{E_t[u'(Y_{t+1})]}.$$

Therefore,

$$p_z = \frac{1}{R} E_t \left[\frac{u'(Y_{t+1})}{E_t[u'(Y_{t+1})]} Z_{t+1} \right]$$

and the stochastic discount factors (SDF) is equal to

$$m_{t+1} = \frac{u^{'}(Y_{t+1})}{E_t[u^{'}(Y_{t+1})]}.$$

Next, note that Y is normally distributed. We can therefore write

$$E[e^{\gamma Y}] = e^{\gamma E[Y] + \frac{\gamma^2}{2} V[Y]}$$

$$E[Y \cdot e^{\gamma Y}] = (E[Y] + \gamma V[Y]) \cdot e^{\gamma E[Y] + \frac{\gamma^2}{2} V[Y]} = (E[Y] + \gamma V[Y]) \cdot E[e^{\gamma Y}].$$

Therefore, for exponential utility with $u'(Y_{t+1}) = \gamma e^{-\gamma Y_{t+1}}$

$$\phi = E_t[m_{t+1}\xi_{t+1}] = \frac{1}{E_t[e^{-\gamma Y_{t+1}}]} E_t \left[e^{-\gamma Y_{t+1}} \xi_{t+1} \right].$$

With the aggregate consumption equal to $Y_{t+1} = \Sigma_{i=1}^{N} Y_{t+1}^{i} = N(1 + \bar{a} + \bar{b} \cdot \xi_{t+1} - \frac{1}{\eta} S_{t+1}^{i})$,

$$\begin{aligned} E_t[e^{-\gamma Y_{t+1}} \xi_{t+1}] &= E_t[e^{-\gamma N(1+\bar{a}+\bar{b}\cdot\xi_{t+1} - \frac{1}{\eta} S_{t+1}^{i})} \xi_{t+1}] \\ &= -\gamma N \bar{b} \sigma_{\xi}^2 \cdot E_t[e^{-\gamma Y_{t+1}}]. \end{aligned}$$

Thus,

$$\phi = -E_t[m_{t+1}\xi_{t+1}] = \gamma N \bar{b} \sigma_{\xi}^2$$

and

$$p_t^i = \frac{1}{R} E_t[m_{t+1} Y_{t+1}^i] = E_t[1 + a^i - \phi b^i - \frac{1}{\eta} S_t^i].$$

□

Proof of Lemma 1 (Fund Size)

The equilibrium condition is that the marginal utility from the last dollar invested in each fund must be equal to the marginal utilty invested in the risk-free asset. As the risk-free rate is normalized to zero, it must be that the value of a dollar invested in each fund i is one dollar. Combining this equilibrium condition with lemma 3, $p_t^i = E_t[1 + a^i - \phi b^i - \frac{1}{\eta} S_t^i] = 1$, immediately yields the result. □

Proof of Lemma 2 (Fund Flows)

The key to the proof is that beliefs about fund returns conditional on the market shock ξ_t are normally distributed. As a result, the standard formulas for Bayesian updating of beliefs apply (see Schmalz and Zhuk for details). It is useful to write the fund size in vector notation:

$$S_t^i = \eta \cdot E_t[a^i - \phi b^i] = \eta(1, -\phi) \cdot \mu_t^i$$

where

$$\mu_t^i = \begin{pmatrix} E_t[a^i] \\ E_t[b^i] \end{pmatrix}$$

With

$$var[Y_t^i] = \sigma_a^2 + 2\sigma_{ab}\xi_t + \sigma_b^2\xi_t^2 + \sigma_\epsilon^2,$$

$$cov\left[\mu_t^i, Y_t^i\right] = \begin{pmatrix} \sigma_a^2 + \sigma_{ab}\xi_t \\ \sigma_{ab} + \sigma_b^2\xi_t \end{pmatrix},$$

and

$$\mu_t = \mu_{t-1} + cov\left[\mu_t^i, Y_t^i\right] \frac{(Y_t^i - E[Y_t^i])}{var[Y_t^i]},$$

we have

$$\begin{aligned} S_t^i - S_{t-1}^i &= \eta(1, -\phi) \cdot (\mu_t^i - \mu_{t-1}^i) \\ &= \eta(1, -\phi) \cdot cov\left[\mu_t^i, Y_t^i\right] \frac{(Y_t^i - E[Y_t^i])}{var[Y_t^i]} \end{aligned}$$

which yields the desired expression for $\lambda(\xi_t)$. □

Proof of Proposition 1 (Schmalz and Zhuk (2013))

The precise statement we wish to prove is that there exists a threshold prior belief $\sigma_{ab}^* > 0$, such that for all $\sigma_{ab} < \sigma_{ab}^*$

$$\lambda(\xi_t = -x) > \lambda(\xi_t = +x) \tag{16}$$

for any x, where $\lambda(\xi_t) = \frac{\sigma_a^2 - \phi\sigma_{ab} - (\phi\sigma_b^2 - \sigma_{ab})\xi_t}{\sigma_a^2 + 2\sigma_{ab}\xi_t + \sigma_b^2\xi_t^2 + \sigma_\varepsilon^2}$. (Because the threshold is positive and feasible beliefs are symmetric with respect to zero, we will say that condition (16) holds "for a majority of prior beliefs.") We have

$$\lambda(\xi_t) = \frac{\sigma_a^2 - \phi\sigma_{ab} - (\phi\sigma_b^2 - \sigma_{ab})\xi_t}{\sigma_a^2 + 2\sigma_{ab}\xi_t + \sigma_b^2\xi_t^2 + \sigma_\varepsilon^2} =$$

$$\frac{A + B\xi_t}{C + D\xi_t}$$

where

$$A = \sigma_a^2 - \phi\sigma_{ab} \qquad B = \sigma_{ab} - \phi\sigma_b^2$$

$$C = \sigma_a^2 + \sigma_b^2\xi_t^2 + \sigma_\varepsilon^2 \qquad D = 2\sigma_{ab}.$$

Therefore,

$$\lambda_{\xi=+x} > \lambda_{\xi=-x} \quad \Leftrightarrow \quad A \cdot D < B \cdot C$$

which is equivalent to

$$(\sigma_a^2 - \phi\sigma_{ab})2\sigma_{ab} < (\sigma_{ab} - \phi\sigma_b^2)(\sigma_a^2 + \sigma_b^2x^2 + \sigma_\varepsilon^2) \quad \Leftrightarrow$$

$$\phi\sigma_b^2(\sigma_a^2 + \sigma_b^2x^2 + \sigma_\varepsilon^2) + \sigma_{ab}(\sigma_a^2 - \sigma_b^2x^2 - \sigma_\varepsilon^2) - 2\phi\sigma_{ab}^2 < 0.$$

The left hand side is a quadratic function of σ_{ab}

$$f(\sigma_{ab}) = \phi\sigma_b^2(\sigma_a^2 + \sigma_b^2x^2 + \sigma_\varepsilon^2) + \sigma_{ab}(\sigma_a^2 - \sigma_b^2x^2 - \sigma_\varepsilon^2) - 2\phi\sigma_{ab}^2.$$

The statement of the proposition follows from the following two facts

1. $f(\sigma_{ab} = 0) = \phi\sigma_b^2(\sigma_a^2 + \sigma_b^2x^2 + \sigma_\varepsilon^2) > 0.$

2. $f(\sigma_{ab} = -\sigma_a\sigma_b) + f(\sigma_{ab} = \sigma_a\sigma_b) = 2\phi(\sigma_b^2x^2 + \sigma_\varepsilon^2) > 0 \Rightarrow$ either $f(\sigma_{ab} = -\sigma_a\sigma_b) > 0$, or $f(\sigma_{ab} = \sigma_a\sigma_b) > 0$, or both conditions hold.

$\square$

Proof of Proposition 2

We wish to show that $\lambda_{Concentrated} > \lambda_{Other}$ when

$$\bar{\sigma}^2_{a,Concentrated} = k \cdot \bar{\sigma}^2_{a,Other}$$

and

$$\bar{\sigma}^2_{a,Concentrated} = k \cdot \bar{\sigma}^2_{a,Other}$$

with $k > 1$. Plugging in equations (9) and (10) into expression (6) yields:

$$\begin{aligned}
\lambda_{Concentrated} &= \frac{\sigma^2_{a,Concentrated} - \phi\sigma_{ab,Concentrated} - (\phi\sigma^2_{b,Concentrated} - \sigma_{ab,Concentrated})\xi_t}{\sigma^2_{a,Concentrated} + 2\sigma_{ab,Concentrated}\xi_t + \sigma^2_{b,Concentrated}\xi_t^2 + \sigma^2_\varepsilon} \\
&= \frac{k \cdot (\sigma^2_{a,Other} - \phi\sigma_{ab,Other}) - k \cdot (\phi\sigma^2_{b,Other} - \sigma_{ab,Other})\xi_t}{k \cdot (\sigma^2_{a,Other} + 2\sigma_{ab,Other}\xi_t + \sigma^2_{b,Other}\xi_t^2) + \sigma^2_\varepsilon} \\
&= \frac{(\sigma^2_{a,Other} - \phi\sigma_{ab,Other}) - (\phi\sigma^2_{b,Other} - \sigma_{ab,Other})\xi_t}{(\sigma^2_{a,Other} + 2\sigma_{ab,Other}\xi_t + \sigma^2_{b,Other}\xi_t^2) + \frac{\sigma^2_\varepsilon}{k}} \\
&> \frac{(\sigma^2_{a,Other} - \phi\sigma_{ab,Other}) - (\phi\sigma^2_{b,Other} - \sigma_{ab,Other})\xi_t}{(\sigma^2_{a,Other} + 2\sigma_{ab,Other}\xi_t + \sigma^2_{b,Other}\xi_t^2) + \sigma^2_\varepsilon} \\
&= \lambda_{Other}.
\end{aligned}$$

$\square$

Proof of Proposition 3

Using the notation from the proof of Proposition (1), we have

$$
\begin{aligned}
\lambda_{DT} - \lambda_{UT} &= \left(\frac{A + B\xi_t}{C + D\xi_t}\right)_{DT} - \left(\frac{A + B\xi_t}{C + D\xi_t}\right)_{UT} \\
&= \frac{A - Bx}{C - Dx} - \frac{A + Bx}{C + Dx} \\
&= \frac{(A - Bx)(C + Dx) - (A + Bx)(C - Dx)}{C^2 - D^2x^2} \\
&= \frac{2(AD - BC)x}{C^2 - D^2x^2} \\
&= \frac{2\left(2(\sigma_a^2 - \phi\sigma_{ab})\sigma_{ab} - (\sigma_{ab} - \phi\sigma_b^2)(\sigma_a^2 + \sigma_b^2\xi_t^2 + \sigma_\varepsilon^2)\right)x}{(\sigma_a^2 + \sigma_b^2\xi_t^2 + \sigma_\varepsilon^2)^2 - 4\sigma_{ab}^2x^2}.
\end{aligned}
$$

When $k > 1$, we have

$$
\begin{aligned}
(\lambda_{DT} - \lambda_{UT})_{Concentrated} &= \frac{2\left(2(\sigma^2_{a,Concentrated} - \phi\sigma_{ab,Concentrated})\sigma_{ab,Concentrated} - (\sigma_{ab,Concentrated} - \phi\sigma^2_{b,Concent}\right.}{(\sigma^2_{a,Concentrated} + \sigma^2_{b,Concentrated}\xi_t^2 + \sigma_\epsilon^2)^2 - 4\sigma^2_{ab,C}} \\
&= \frac{4(\sigma^2_{a,Concentrated} - \phi\sigma_{ab,Concentrated})\sigma_{ab,Concentrated}x - (\sigma_{ab,Concentrated} - \phi\sigma^2_{b,Concentra}}{(\sigma^2_{a,Concentrated} + \sigma^2_{b,Concentrated}\xi_t^2 + \sigma_\varepsilon^2)^2 - 4\sigma^2_{ab,Co}} \\
&= \frac{4k^2 \cdot (\sigma^2_{a,Other} - \phi\sigma_{ab,Other})\sigma_{ab,Other}x - k^2 \cdot (\sigma_{ab,Other} - \phi\sigma^2_{b,Other})(\sigma^2_{a,Other} + \sigma^2_{b,Othe}}{k^2 \cdot (\sigma^2_{a,Other} + \sigma^2_{b,Other}\xi_t^2 + \frac{\sigma_\varepsilon^2}{k})^2 - 4k^2 \cdot \sigma^2_{ab,Other}x^2} \\
&= \frac{4(\sigma^2_{a,Other} - \phi\sigma_{ab,Other})\sigma_{ab,Other}x - (\sigma_{ab,Other} - \phi\sigma^2_{b,Other})(\sigma^2_{a,Other} + \sigma^2_{b,Other}\xi_t^2 + \frac{\sigma_\varepsilon^2}{k}}{(\sigma^2_{a,Other} + \sigma^2_{b,Other}\xi_t^2 + \frac{\sigma_\varepsilon^2}{k})^2 - 4\sigma^2_{ab,Other}x^2} \\
&> \frac{4(\sigma^2_{a,Other} - \phi\sigma_{ab,Other})\sigma_{ab,Other}x - (\sigma_{ab,Other} - \phi\sigma^2_{b,Other})(\sigma^2_{a,Other} + \sigma^2_{b,Other}\xi_t^2 + \sigma_\varepsilon^2}{(\sigma^2_{a,Other} + \sigma^2_{b,Other}\xi_t^2 + \frac{\sigma_\varepsilon^2}{\kappa})^2 - 4\sigma^2_{ab,Other}x^2} \\
&> \frac{4(\sigma^2_{a,Other} - \phi\sigma_{ab,Other})\sigma_{ab,Other}x - (\sigma_{ab,Other} - \phi\sigma^2_{b,Other})(\sigma^2_{a,Other} + \sigma^2_{b,Other}\xi_t^2 + \sigma_\varepsilon^2}{(\sigma^2_{a,Other} + \sigma^2_{b,Other}\xi_t^2 + \sigma_\varepsilon^2)^2 - 4\sigma^2_{ab,Other}x^2} \\
&= (\lambda_{DT} - \lambda_{UT})_{Other}.
\end{aligned}
$$

□

Proof of Proposition 4

We wish to show that

$$\rho(Y_{t+1}^i - \xi_{t+1}, Y_t^i - \xi_t)_{UT} - \rho(Y_{t+1}^i - \xi_{t+1}, Y_t^i - \xi_t)_{DT} > 0.$$

SHOULD WE REWRITE THE PROPOSITION AS Y - E[BETA]* XI? WOULDN'T DO ANYTHING AS PHI IS STILL THERE. As the model variances are the same for upturns and downturns, we can focus on comparing the numerators of the time-series correlation following upturns, compared to downturns, and show:

$$Cov(Y_{t+1}^i - \xi_{t+1}, Y_t^i - \xi_t)_{UT} - Cov(Y_{t+1}^i - \xi_{t+1}, Y_t^i - \xi_t)_{DT} > 0.$$

Using equation (1) and the equilibrium condition (3), we have BUT THE EXPECTATION HERE SHOULD BE WITH A SMALLER t.

$$Y_t^i - \xi_t = 1 + a^i + (b^i - 1)\xi_t - E_t[a^i - \phi b^i] + \varepsilon_t^i,$$

and NOT SURE ABOUT THAT

$$\begin{aligned} E[Y_t^i - \xi_t] &= 1 + E[a^i] + \xi_t E[b^i - 1] - E_t[a^i - \phi b^i] \\ &= 1 - \xi_t + (\phi + \xi_t)E[b^i]. \end{aligned}$$

Then,

$$Y_t^i - \xi_t - E[Y_t^i - \xi_t] = (a^i - E_t[a^i]) + (b^i - \bar{b})\xi_t + \varepsilon_t^i.$$

The time-series covariance of mutual fund returns is then calculated as

$$
\begin{aligned}
&E\left[\left(Y^i_{t+1} - \xi_{t+1} - E[Y^i_{t+1} - \xi_{t+1}]\right)\left(Y^i_t - \xi_t - E[Y^i_t - \xi_t]\right)\right] \\
&= E[(a^i - \bar{a})^2] + E[(a^i - \bar{a})(b^i - \bar{b})\xi_{t+1}] + E[(a^i - \bar{a})(b^i - \bar{b})\xi_t] + E[(b^i - \bar{b})^2\xi_t\xi_{t+1}] \\
&= \sigma_a^2 + \sigma_{ab}\xi_t.
\end{aligned}
$$

The difference of conditional covariances, using $\xi_t = +x$ for a downturn and $\xi_t = -x$ for a downturn, is then given by

$$
\begin{aligned}
&Cov(Y^i_{t+1} - \xi_{t+1}, Y^i_t - \xi_t)_{UT} - Cov(Y^i_{t+1} - \xi_{t+1}, Y^i_t - \xi_t)_{DT} \\
&= \sigma_a^2 + \sigma_{ab}x - (\sigma_a^2 - \sigma_{ab}x) = 2\sigma_{ab}x > 0.
\end{aligned}
$$

□

Proof of Proposition 5

Similar to the proof of Proposition (4), we have

$$
\begin{aligned}
&Cov(Y^i_{t+1} - \xi_{t+1}, Y^i_t - \xi_t)_{Concentrated} - Cov(Y^i_{t+1} - \xi_{t+1}, Y^i_t - \xi_t)_{Other} \\
&= blubb > 0.
\end{aligned}
$$

□

References

(2012): "2012 Investment Company Factbook," Investment Company Institute. 1

Berk, J. B., and R. C. Green (2004): "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, pp. 1269–1295. 1, 2, 3.6

Bollen, N. P., and J. A. Busse (2005): "Short-term persistence in mutual fund performance," *Review of Financial Studies*, 18(2), 569–597. 2

Carhart, M. M. (1997): "On Persistence in Mutual Fund Performance," *The Journal of Finance*, 52(1), 57–82. 1, 3.5.3, 5.2, 6, 7, 8

Chevalier, J. A., and G. D. Ellison (1997): "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 105, 1167–1200. 1, 5.1

Choi, D., B. Kahraman, and A. Mukherjee (2012): "Understanding Performance-Chasing Behavior in Mutual Funds: Evidence from Multi-Fund Managers," *Available at SSRN 2023580*. 2

Cremers, K. M., and A. Petajisto (2009): "How active is your fund manager? A new measure that predicts performance," *Review of Financial Studies*, 22(9), 3329–3365. 1, 1, 2, 3.5.2, 4, 5, 5.1

Elton, E. J., M. J. Gruber, and C. R. Blake (2001): "A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases," *Journal of Finance*, 56(6), 2415–2430. 4

Fama, E. F., and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *The Journal of Political Economy*, pp. 607–636. 1, 5.1, 3, 4, 5

Gervais, S., and G. Strobl (2013): "The Industrial Organization of Money Management," Discussion paper, Available at SSRN: http://ssrn.com/abstract=2023884 or http://dx.doi.org/10.2139/ssrn.2023884. 2

Glode, V. (2011): "Why mutual funds "underperform"," *Journal of Financial Economics*, 99(3), 546–559. 2

Glode, V., B. Hollifield, M. Kacperczyk, and S. Kogan (2012): "Time-varying Predictability in Mutual Fund Returns," *Available at SSRN 1364351*. 1, 2, 10, 4, 5.2

Gruber, M. J. (1996): "Another puzzle: The growth in actively managed mutual funds," *The journal of finance*, 51(3), 783–810. 5.1

Holmström, B. (1999): "Managerial incentive problems: A dynamic perspective," *The Review of Economic Studies*, 66(1), 169–182. 2

Huang, J., K. D. Wei, and H. Yan (2007): "Participation Costs and the Sensitivity of Fund Flows to Past Performance," *Journal of Finance*, 62(3), 1273–1311. 1, 4, 4, 10

Ippolito, R. A. (1992): "Consumer reaction to measures of poor quality: Evidence from the mutual fund industry," *Journal of law and Economics*, 35(1), 45–70. 5.1

Kacperczyk, M., S. V. Nieuwerburgh, and L. Veldkamp (2012): "Rational Attention Allocation Over the Business Cycle," NYU Working Paper No. FIN-11-050. 2, 6

Kacperczyk, M., C. Sialm, and L. Zheng (2008): "Unobserved actions of mutual funds," *Review of Financial Studies*, 21(6), 2379–2416. 2

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2012): "Time-varying fund manager skill," *Journal of Finance*, Forthcoming. 2, 4

Kallberg, J. G., C. L. Liu, and C. Trzcinka (2000): "The Value Added from Investment Managers: An Examination of Funds of REITs," *Journal of Financial and Quantitative Analysis*, 35(3), 387–408. 2

Kosowski, R. (2006): "Do Mutual Funds Perform When it Matters Most to Investors? US Mutual Fund Performance and Risk in Recessions and Expansions," *US Mutual Fund Performance and Risk in Recessions and Expansions (August 1, 2006)*. 2

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*, 61(6), 2551–2595. 5.2

Lynch, A. W., and D. K. Musto (2003): "How Investors Interpret Past Fund Returns," *The Journal of Finance*, 58(5), 2033–2058. 5

Merton, R. C. (1980): "On estimating the expected return on the market: An exploratory investigation," *Journal of Financial Economics*, 8(4), 323–361. 2

Moskowitz, T. J. (2000): "Discussion," *The Journal of Finance*, 55(4), 1695–1703. 2

Pastor, L., and R. F. Stambaugh (2012): "On the Size of the Active Management Industry," *Journal of Political Economy*, 120(4), 740 – 781. 2

Petajisto, A. (2013): "Active Share and Mutual Fund Performance," *Financial Analysts Journal*, Forthcoming. 1, 2, 4, 5.1

Schmalz, M. C., and S. Zhuk (2013): "Revealing Downturns," Discussion paper, University of Michigan. 3, 3.4, 3.6, 6, 6, 6

Shleifer, A., and R. W. Vishny (1997): "The limits of arbitrage," *The Journal of Finance*, 52(1), 35–55. 1

Sirri, E. R., and P. Tufano (1998): "Costly Search and Mutual Fund Flows," *The Journal of Finance*, 53(5), 1589–1622. 1, 5.1

Spiegel, M., and H. Zhang (2012): "Mutual Fund Risk and Market Share-adjusted Fund Flows," *Journal of Financial Economics*. 1, 1, 2, 5.1, 5.1, 5.4

Sun, Z., A. Wang, and L. Zheng (2009): "Do Active Funds Perform Better in Down Markets? New Evidence from Cross-Sectional Study," *New Evidence from Cross-Sectional Study (September 1, 2009)*. 2

——— (2013): "Only When the Tide Goes Out: Downside Returns and Hedge Fund Performance," *Available at SSRN*. 2

Tables

Table 2: Summary Statistics. The table reports summary statistics for the variables that are used in the analysis: the fund's monthly (net) return, assets under management (TNA), the total expense ratio, turnover, return volatility over the prior twelve months, fund age computed as the number of quarters since the first appearance in CRSP, quarterly flows. The sample ranges from 1980:Q1 to 2012:Q4. Upturns and Downturns are defined, respectively, as the top and bottom 25% of periods according to the distribution of the CRSP value-weighted index since July 1926.

Panel A: WHOLE SAMPLE

	N	Mean	SD	Min	Median	Max	CORRELATIONS	Ret	TNA	Expense	Turnover	Volatility	Age
Ret	144,382	0.015	0.092	-0.230	0.025	0.205	Ret	1.00					
TNA	144,382	678	3134	5	82	109073	TNA	0.01	1.00				
Expense ratio	144,382	0.015	0.005	0.000	0.015	0.090	Expense ratio	-0.01	-0.22	1.00			
Turnover	144,382	0.8	0.8	0.0	0.6	34.5	Turnover	-0.02	-0.07	0.17	1.00		
Volatility	144,382	0.048	0.022	0.000	0.0457	0.338	Volatility	0.03	-0.03	0.06	0.17	1.00	
Age (quarters)	144,382	25.600	21.600	1.000	20.000	171.000	Age (quarters)	0.05	0.12	-0.17	-0.04	0.01	1.00
Flows	144,382	-0.002	0.085	-0.171	-0.018	0.415	Flows	0.07	0.00	-0.10	-0.05	-0.02	-0.19

Panel B: UPTURNS

	N	Mean	SD	Min	Median	Max	CORRELATIONS	Ret	TNA	Expense	Turnover	Volatility	Age
Ret	30,850	0.120	0.042	-0.187	0.121	0.205	Ret	1.00					
TNA	30,850	642	2899	5	80	105939	TNA	0.00	1.00				
Expense ratio	30,850	0.015	0.005	0.000	0.015	0.089	Expense ratio	-0.02	-0.22	1.00			
Turnover	30,850	0.824	0.790	0.000	0.630	24.000	Turnover	-0.03	-0.07	0.18	1.00		
Volatility	30,850	0.1	0.0	0.0	0.1	0.3	Volatility	0.49	-0.03	0.04	0.11	1.00	
Age (quarters)	30,850	27.400	22.500	1.000	21	169	Age (quarters)	0.11	0.12	-0.17	-0.05	0.11	1.00
Flows	30,850	-0.003	0.086	-0.171	-0.020	0.415	Flows	0.04	0.00	-0.09	-0.08	-0.11	-0.18

Panel C: DOWNTURNS

	N	Mean	SD	Min	Median	Max	CORRELATIONS	Ret	TNA	Expense	Turnover	Volatility	Age
Ret	35,758	-0.100	0.068	-0.230	-0.103	0.177	Ret	1.00					
TNA	35,758	654	3086	5	75	104718	TNA	0.03	1.00				
Expense ratio	35,758	0.015	0.005	0.000	0.015	0.090	Expense ratio	-0.02	-0.22	1.00			
Turnover	35,758	0.849	0.810	0.000	0.660	34.500	Turnover	-0.07	-0.07	0.17	1.00		
Volatility	35,758	0.1	0.0	0.0	0.0	0.3	Volatility	-0.38	-0.03	0.10	0.22	1.00	
Age (quarters)	35,758	24.200	21.300	1.000	18	167	Age (quarters)	-0.03	0.11	-0.16	-0.05	-0.10	1.00
Flows	35,758	-0.006	0.079	-0.171	-0.020	0.414	Flows	0.12	0.00	-0.08	-0.02	0.02	-0.21

Table 3: Flow-Performance Sensitivity Main Results. The table reports slopes from Fama and MacBeth (1973) regressions of quarterly flows on prior-quarter mutual fund rank by style-adjusted performance (frank_style) and controls. T-statistics are reported in parentheses. At the bottom of the table, we report the test statistic and p-value (assuming normality) for the test of the null hypothesis that the difference between dowturns and upturns in the slopes on frank_style is zero. The sample ranges from 1980:Q1 to 2012:Q4. Upturns and Downturns are defined, respectively, as the top and bottom 25% of periods according to the distribution of the CRSP value-weighted index since July 1926.

Flows (t+1)	All quarters	Upturns	Downturns	All quarters	Upturns	Downturns
frank_style	0.043***	0.021***	0.051***	0.034***	0.020***	0.045***
	(12.188)	(2.771)	(6.178)	(11.498)	(3.489)	(5.551)
flows_style				0.222*	-0.281	0.295***
				(1.780)	(-0.596)	(3.116)
fee				-0.248	0.116	-0.155
				(-1.070)	(0.191)	(-0.332)
logsize				-0.001*	-0.001	0.001
				(-1.917)	(-0.942)	(0.765)
turn_ratio				-0.003**	-0.009***	-0.000
				(-1.992)	(-2.881)	(-0.087)
vol				0.028	-0.326*	0.410*
				(0.310)	(-1.801)	(1.704)
logage				-0.011***	-0.015***	-0.024***
				(-3.717)	(-2.819)	(-2.766)
flows				0.500***	0.530***	0.464***
				(24.744)	(12.816)	(8.853)
Constant	-0.016***	-0.002	-0.023***	0.032***	0.079***	0.055*
	(-6.385)	(-0.452)	(-4.742)	(2.893)	(3.447)	(1.778)
Observations	144,382	30,850	35,758	144,382	30,850	35,758
R-squared	0.044	0.027	0.058	0.439	0.467	0.397
Number of groups	131	32	30	131	32	30
z-stat		2.623			2.564	
p-val		0.00872			0.0103	

Table 4: Flow-Performance Sensitivity Main Results (Piecewise Linear Specification). The table reports slopes from Fama and MacBeth (1973) regressions of quarterly flows on prior-quarter mutual fund rank by style-adjusted performance and controls. The rank variable is defined to separately capture performance between 0 and 1/3 (trank_style1), between 1/3 and 2/3 (trank_style2), and between 2/3 and the top (trank_style3) of the distribution. T-statistics are reported in parentheses. At the bottom of the table, we report the p-value (assuming normality) for the test of the hypothesis that the differences between dowturns and upturns in the slopes on trank_style1, trank_style2, and trank_style3 are jointly zero. The sample ranges from 1980:Q1 to 2012:Q4. Upturns and Downturns are defined, respectively, as the top and bottom 25% of periods according to the distribution of the CRSP value-weighted index since July 1926.

	Panel A: No controls			Panel B: With controls		
Flows (t+1)	All quarters	Upturns	Downturns	All quarters	Upturns	Downturns
trank_style1	0.042***	0.005	0.040**	0.030***	0.020	0.033
	(4.282)	(0.189)	(2.486)	(2.972)	(0.748)	(1.219)
trank_style2	0.032***	0.037	0.049**	0.029***	0.025	0.047**
	(3.699)	(1.481)	(2.591)	(3.167)	(0.876)	(2.315)
trank_style3	0.062***	0.011	0.066***	0.047***	0.015	0.053**
	(5.929)	(0.406)	(3.034)	(5.215)	(0.698)	(2.749)
flows_style				0.225*	-0.268	0.297***
				(1.805)	(-0.568)	(3.145)
fee				-0.208	0.242	-0.132
				(-0.839)	(0.462)	(-0.201)
logsize				-0.001*	-0.001	0.001
				(-1.829)	(-1.348)	(1.103)
turn_ratio				-0.002*	-0.009***	0.000
				(-1.947)	(-3.037)	(0.049)
vol				-0.005	-0.375*	0.367
				(-0.050)	(-1.832)	(1.461)
logage				-0.012***	-0.021***	-0.023**
				(-4.009)	(-3.424)	(-2.711)
flows				0.499***	0.522***	0.468***
				(24.039)	(11.869)	(8.753)
Constant	-0.015***	0.000	-0.021***	0.039***	0.103***	0.052
	(-4.768)	(0.060)	(-3.629)	(3.477)	(4.963)	(1.663)
Observations	144,382	30,850	35,758	144,382	30,850	35,758
R-squared	0.072	0.068	0.086	0.460	0.501	0.424
Number of groups	131	32	30	131	32	30
p-val(χ^2)		0.0348			0.0456	

Table 5: Flow-Performance Sensitivity Double-Difference Results. The table reports slopes from Fama and MacBeth (1973) regressions of quarterly flows on prior-quarter mutual fund rank by style-adjusted performance (frank_style) and controls. The rank variable is interacted with a dummy variable denoting "Concentrated" funds, which are the funds with above-median levels of active share and tracking error. T-statistics are reported in parentheses. At the bottom of the table, we report the test statistic and p-value (assuming normality) for the test of the null hypothesis that the difference between dowturns and upturns in the slope on frank_style×concentrated is zero. The sample ranges from 1980:Q1 to 2012:Q4. Upturns and Downturns are defined, respectively, as the top and bottom 25% of periods according to the distribution of the CRSP value-weighted index since July 1926.

	Panel A: No controls			Panel B: With controls		
Flows (t+1)	All quarters	UT	DT	All quarters	UT	DT
frank_style × concentr.	0.022**	-0.036	0.075***	0.015	-0.023	0.063**
	(2.015)	(-1.449)	(2.833)	(1.334)	(-0.872)	(2.321)
frank_style	0.050***	0.041***	0.045***	0.039***	0.032*	0.034***
	(9.173)	(2.974)	(3.632)	(7.330)	(2.028)	(3.249)
concentrated	0.001	0.029	-0.018*	-0.001	0.009	-0.016
	(0.098)	(1.308)	(-1.837)	(-0.160)	(0.308)	(-1.615)
flows_style				0.248***	0.209	0.250**
				(3.933)	(1.703)	(2.716)
fee				0.602	2.449**	0.092
				(1.351)	(2.260)	(0.087)
logsize				-0.001	-0.000	-0.000
				(-1.185)	(-0.187)	(-0.026)
turn_ratio				-0.003	-0.017	0.006
				(-0.909)	(-1.488)	(1.306)
vol				0.087	0.183	0.268
				(0.304)	(0.157)	(0.848)
logage				-0.023	-0.071	-0.009
				(-1.606)	(-1.214)	(-0.587)
flows				0.539***	0.523***	0.529***
				(16.734)	(9.058)	(5.912)
Constant	-0.018***	-0.008	-0.020***	0.076	0.272	0.006
	(-6.065)	(-0.965)	(-3.303)	(1.453)	(1.301)	(0.076)
Observations	19,577	3,540	4,881	19,577	3,540	4,881
R-squared	0.119	0.111	0.138	0.430	0.473	0.410
Number of groups	117	27	27	117	27	27
p-val		0.00219			0.0232	
z-stat		3.063			2.270	

Table 6: Performance Persistence Main Results. Each month, funds are ranked according to their monthly style-adjusted return. Equally-weighted portfolios are formed of the Good Performers (top 20%) and Bad Performers (bottom 20%). The table reports alphas and factor loadings from time-series regressions of the Bad, Good, and Good minus Bad monthly portfolio returns on the Carhart (1997) four factors. T-statistics are reported in parentheses. The sample ranges from Juanuary 1980 to December 2012. Upturns and Downturns are defined, respectively, as the top and bottom 25% of periods according to the distribution of the CRSP value-weighted index since July 1926. Upturns and Downturns are defined, respectively, as the top and bottom 25% of periods according to the distribution of the CRSP value-weighted index since July 1926.

	All funds								
	All months			Upturns			Downturns		
E[Return]	Bad Performer	Good Perfomer	Good - Bad	Bad Performer	Good Perfomer	Good - Bad	Bad Performer	Good Perfomer	Good - Bad
alpha	-0.005***	0.002*	0.007***	-0.002	0.010***	0.012***	-0.004*	-0.000	0.003
	(-4.615)	(1.804)	(3.860)	(-0.710)	(3.524)	(2.830)	(-1.710)	(-0.177)	(0.793)
mktrf	1.046***	0.909***	-0.137***	0.761***	0.756***	-0.005	1.143***	0.896***	-0.246***
	(41.580)	(36.040)	(-3.275)	(9.455)	(10.006)	(-0.042)	(31.754)	(21.236)	(-3.617)
hml	-0.028	-0.012	0.016	0.031	-0.314***	-0.345**	-0.215***	0.267***	0.481***
	(-0.735)	(-0.309)	(0.256)	(0.311)	(-3.386)	(-2.457)	(-3.275)	(3.466)	(3.875)
smb	0.274***	0.339***	0.065	0.293***	0.264***	-0.029	0.339***	0.290***	-0.048
	(7.573)	(9.350)	(1.082)	(3.407)	(3.266)	(-0.240)	(5.883)	(4.301)	(-0.443)
mom	-0.025	0.029	0.054	0.026	-0.173***	-0.199***	-0.107**	0.101*	0.208**
	(-1.076)	(1.237)	(1.391)	(0.539)	(-3.842)	(-2.919)	(-2.386)	(1.930)	(2.454)
Observations	395	395	395	88	88	88	98	98	98
R-squared	0.859	0.826	0.040	0.627	0.762	0.152	0.944	0.860	0.370

Table 7: Performance Persistence Cross-sectional Results. Each month, funds are ranked according to their monthly style-adjusted return as well as, independently, according to the Concentrated vs. Other funds dummy. “Concentrated” funds are those with above-median levels of active share and tracking error. Good (Bad) performers rank in the top (bottom) 20% of funds by style-adjusted return. Four equally-weighted portfolios are from the intersection of the two dimensions. The table reports alphas and factor loadings from time-series regressions of monthly portfolio returns on the Carhart (1997) four factors. T-statistics are reported in parentheses. The sample ranges from Juanuary 1980 to December 2012.

	All months						
	Other Funds			Concentrated			Concentrated - Others
E[Return]	Bad Performer	Good Perfomer	Good - Bad	Bad Performer	Good Perfomer	Good - Bad	Good - Bad
alpha	-0.003***	0.001	0.004***	-0.004**	0.004**	0.008***	0.004***
	(-3.892)	(1.387)	(3.236)	(-2.525)	(2.581)	(3.032)	(3.090)
mktrf	1.049***	0.952***	-0.098***	1.080***	0.928***	-0.201***	-0.108***
	(58.925)	(50.781)	(-3.323)	(31.031)	(27.605)	(-3.122)	(-3.229)
hml	-0.004	-0.009	-0.005	0.007	0.021	-0.005	-0.003
	(-0.147)	(-0.319)	(-0.114)	(0.136)	(0.469)	(-0.064)	(-0.079)
smb	0.092***	0.131***	0.039	0.393***	0.485***	0.097	0.052
	(3.781)	(5.114)	(0.967)	(8.660)	(11.893)	(1.258)	(1.299)
mom	-0.025	0.029*	0.054**	-0.030	0.004	0.003	-0.018
	(-1.578)	(1.737)	(2.059)	(-1.039)	(0.147)	(0.062)	(-0.696)
Observations	354	354	354	247	244	226	226
R-squared	0.929	0.907	0.056	0.855	0.845	0.052	0.051

Table 8: Performance Persistence Across Fund Types and Market States. Each month, funds are ranked according to their monthly style-adjusted return as well as, independently, according to the Concentrated vs. Other funds dummy. Concentrated funds are those with above-median levels of active share and tracking error. Good (Bad) performers rank in the top (bottom) 20% of funds by style-adjusted return. Four equally-weighted portfolios are from the intersection of the two dimensions. The table reports alphas and factor loadings from time-series regressions of monthly portfolio returns on the Carhart (1997) four factors. In the last column, the dependent variable is the double difference of returns between Good and Bad performers and Concentrated vs. Other funds. T-statistics are reported in parentheses. The sample ranges from Juanuary 1980 to December 2012.

PANEL A: UPTURNS

	Other Funds			Concentrated			Concentrated - Others
E[Return]	Bad Performer	Good Perfomer	Good - Bad	Bad Performer	Good Perfomer	Good - Bad	Good - Bad
alpha	-0.001	0.005*	0.006*	-0.004	0.014***	0.018**	0.009**
	(-0.394)	(1.900)	(1.699)	(-1.240)	(2.790)	(2.650)	(2.316)
mktrf	0.888***	0.893***	0.005	0.944***	0.750***	-0.248	-0.170*
	(20.926)	(13.539)	(0.061)	(12.585)	(6.195)	(-1.443)	(-1.800)
hml	0.081	-0.227***	-0.308***	0.240**	-0.389***	-0.667***	-0.331***
	(1.494)	(-2.712)	(-2.898)	(2.654)	(-2.739)	(-3.312)	(-2.987)
smb	0.085*	0.071	-0.014	0.482***	0.318***	-0.161	-0.025
	(1.777)	(0.956)	(-0.148)	(6.455)	(2.772)	(-0.985)	(-0.277)
mom	-0.008	-0.077**	-0.068	0.043	-0.227***	-0.292***	-0.158***
	(-0.347)	(-2.062)	(-1.451)	(1.101)	(-3.718)	(-3.377)	(-3.325)
Observations	76	76	76	52	44	44	44
R-squared	0.907	0.852	0.180	0.860	0.845	0.321	0.290

PANEL B: DOWNTURNS

	Other Funds			Concentrated			Concentrated - Others
E[Return]	Bad Performer	Good Perfomer	Good - Bad	Bad Performer	Good Perfomer	Good - Bad	Good - Bad
alpha	-0.002	0.001	0.003	-0.002	0.002	0.008	0.003
	(-1.251)	(0.760)	(1.217)	(-0.530)	(0.557)	(1.099)	(0.822)
mktrf	1.081***	0.938***	-0.143***	1.117***	0.878***	-0.337**	-0.159**
	(36.833)	(33.179)	(-2.989)	(15.151)	(15.010)	(-2.583)	(-2.225)
hml	-0.135**	0.238***	0.373***	-0.303**	0.280***	0.530**	0.204*
	(-2.582)	(4.710)	(4.370)	(-2.346)	(3.082)	(2.655)	(1.865)
smb	0.115**	0.135***	0.020	0.421***	0.428***	0.011	-0.026
	(2.604)	(3.172)	(0.277)	(3.856)	(5.596)	(0.065)	(-0.287)
mom	-0.065*	0.060*	0.125**	-0.177**	0.084	0.169	0.071
	(-1.869)	(1.806)	(2.215)	(-2.083)	(1.351)	(1.251)	(0.951)
Observations	89	89	89	58	65	55	55
R-squared	0.959	0.940	0.388	0.886	0.847	0.343	0.256