Decomposing Fund Activeness

ANDREA M. BUFFA
University of Colorado Boulder

APOORVA JAVADEKAR
Indian School of Business

December 2020*

Abstract

We decompose the variance of mutual funds' active returns and define their *idiosyncratic activeness* (*IDA*) as the contribution of idiosyncratic shocks, relative to systematic shocks. This novel measure of the source of activeness is informative about the way mutual funds try to beat their benchmarks, and delivers new cross-sectional predictions of fund performance. We show that funds whose activeness is mostly idiosyncratic (high-*IDA*) significantly outperform those whose activeness is mostly systematic (low-*IDA*), and uncover that being more active only leads to better performance if fund activeness is sufficiently idiosyncratic. By further decomposing a fund's systematic activeness, we are able to identify funds that rely on smart-beta strategies and show that those that focus on multiple factors (within a quarter or over time) tend to underperform those that focus on a single factor.

*Contacts: `buffa@colorado.edu` and `apoorva_javadekar@isb.edu`. For helpful comments and discussions we thank Simona Abis, Vikas Agarwal, Philip Bond, Indraneel Chakraborty, Susan Christoffersen, Itay Goldstein, Amit Goyal, Omesh Kini, Anton Lines, Pedro Matos, Abhiroop Mukherjee, Clemens Sialm, Laura Starks, Shekhar Tomar, and Stijn Van Nieuwerburgh.

1 Introduction

Active mutual funds strive to beat their benchmarks by actively deviating from them. Recent contributions have provided different measures of the activeness of a fund, and shown that funds that are more active tend to outperform those that are less active (e.g., Kacperczyk, Sialm, and Zheng (2008); Cremers and Petajisto (2009); Amihud and Goyenko (2013)). Importantly, though, a fund can be active in different ways. It can deviate from its benchmark by over- or under-weighting specific stocks that are believed to be under- or over-valued, or it can be active by increasing or decreasing the exposure to risk factors when predicting upward or downward fluctuations of these factors in the near future. Surprisingly, little is known about the different possible risk profiles of the investment strategies that funds adopt to achieve activeness. So, the goal of this paper is to fill this gap by characterizing the ways in which funds are active—what we term as the *source of activeness*—and to provide new empirical evidence about cross-sectional differences in fund performance.

In the first part of the paper we introduce a novel measure of the source of activeness of mutual funds by identifying whether they deviate from their benchmarks by exposing their portfolios to idiosyncratic or systematic shocks. Specifically, our measure captures the relative contribution of idiosyncratic risk, vis-a-vis systematic risk, to the volatility of a fund's active returns— i.e., fund returns net of benchmark returns—and we refer to it as *idiosyncratic activeness* (*IDA*). *IDA* is always between 0 and 1, where the complement $(1-IDA)$ represents a fund's *systematic activeness*, that is the fraction of the total active return variation that is driven by the active exposure to systematic factors. The main idea is that two funds might have similar levels of activeness (similar average deviation from their benchmarks), but they could implement investment strategies with very different risk profiles. The *IDA* of the two funds would capture this difference, allowing us to further condition on this measure when investigating their performance. To our knowledge, ours is the first measure to decompose the variation in mutual funds' active returns to back-out their active risk exposures.

We obtain our measure *IDA* by means of two stages. In the first stage we estimate a fund's benchmark in a given quarter by identifying the weights of a long-only portfolio of index funds with no leverage that generates a performance that is "closest" to that of the fund in the previous four quarters. Taking the perspective of most investors, our benchmark construction is based on key "investability" criteria which guarantee that the constituents of fund benchmarks are easily tradable, reflect most common investors' portfolio constraints (leverage and shorting constraints), and are formed *ex-ante* without any look-ahead biases.[1] In the second stage, we apply a standard factor model (such as Fama and French (1993) and Carhart (1997)) to the fund's active returns obtained in the first stage to decompose the variance of active returns in its systematic and idiosyncratic components, and we compute a time series of *IDA* for each fund. The second stage is not about investability and hence we chose leading risk factors as established source of systematic risk to characterize the active risk profile of mutual funds. A non-zero loading on one of the risk factors in the second stage tells us that the fund manager has *actively* exposed the fund portfolio to that factor to generate active returns—returns in excess of its benchmark.[2] In our sample, *IDA* has a mean of 0.6, is persistent over time, and exhibits a substantial variation both within and across funds.

Our analysis delivers three sets of results. Our first set of results establishes that high-*IDA* funds—i.e., funds with high idiosyncratic activeness—have superior *contemporaneous* as well as *future* fund performance compared to low-*IDA* funds. This holds true for a variety of performance measures, including gross and net-of-fees active returns, traditional Fama-French alpha, value-added by Berk and van Binsbergen (2015), and stock picking measure proposed by Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014). Our result is robust to panel estimation that includes the typical set of controls used in the mutual find literature literature, as well, a strongest specification of fixed effects suggesting that the outperformance of high-*IDA* funds remains valid even within a given mandate-year. The difference in performance is economically important. A fund whose activeness is only idiosyncratic ($IDA = 1$) generates 3.13% higher annual active returns on average, or \$7.77M of additional *value-added*, relative to a fund whose activeness is only systematic ($IDA = 0$). We further show the relative importance of idiosyncratic active exposure, that defines a fund's *IDA*, is economically relevant over and above the role played by the level of fund risk (Jordan and Riley (2015)). In contrast with most of the existing literature which focuses on the predictive ability of measures of activeness, we also highlight the contemporaneous link between *IDA* and fund performance, which validates our measure as a proxy for skill.[3]

[1]Following Berk and van Binsbergen (2015) we use a set of Vanguard's passively managed index funds capturing the long and short legs of the well known size and value factors.

[2]In contrast, a non-zero loading in a regression of total fund returns on risk factors captures the fund's *total* exposure to that factor, which includes both passive exposure (to the extent the benchmark itself is exposed to that factor) and active exposure (exposure in excess of the benchmark's exposure).

[3]In a companion paper, Buffa and Javadekar (2020) provides a theoretical foundation for the outperformance of high-*IDA* funds. Their economic mechanism is based on the self-selection of fund managers with different skill to strategies with different exposure to idiosyncratic and systematic shocks. They show that the inherent difference in the nature of these shocks affects the ability of investors to learn about fund managers' skill. The equilibrium in their model is characterized by an endogenous allocation of talent in which high-skill managers adopt high-*IDA* strategies to reveal their skill, while low-skill managers adopt low-*IDA* strategies to hide theirs.

Our second set of results uncovers a novel link between the "level" and the "source" of activeness in the mutual fund industry. We first note that *IDA* is only moderately correlated with different established measures of the level of fund activeness (Active Shares by Cremers and Petajisto (2009), 1-R2 by Amihud and Goyenko (2013), Industry Concentration by Kacperczyk, Sialm, and Zheng (2008)), suggesting a wide cross-section of *IDA* within the group of active funds. In other words, the source of activeness varies substantially for any level of activeness. Consistent with the prior literature, we confirm these measures of activeness predict superior future performance unconditionally. However, when conditioning on the source of activeness, the level of activeness predicts superior performance *only* for high-*IDA* funds, i.e., funds whose active returns are mostly driven by idiosyncratic shocks. Moreover, within the group of low-*IDA* funds, whose activeness is mostly systematic, a higher level of activeness leads to underperformance. These startling results are derived using a standard double-sorting methodology, as well as a panel estimation, and are robust to different performance measures.

Higher *IDA*, instead, predicts better performance for any given level of fund activeness. In particular, we also show that the spread in performance between high- and low-*IDA* funds grows with the level of fund activeness, thus making the source of activeness especially important for the group of funds that are more active. Our results are robust to different benchmark portfolios used to identify funds' active returns in the first stage of our decomposition, and to different sets of factors used to determine the source of fund activeness in the second stage. Overall, our findings suggest that both the level *and* the source of activeness are important fund characteristics that help explaining the cross-sectional variation in fund performance. Therefore, we view our measure *IDA* as complementary, rather than as substitute, to the existing measures of activeness. Our contribution is to highlight that the way in which a fund deviates from its benchmark is as important as the extent of the deviation.

The last set of results focuses on the group of funds whose activeness is primarily systematic (i.e., low-*IDA* funds). In the last section of the paper, we further decompose mutual funds' systematic activeness ($1 - IDA$) to have a better picture of the type of strategies they employ. Two interesting facts emerge. First, the relative contribution of systematic factors in explaining fund's active strategies has increased over last 15 years. Moreover, the nature of systematic activeness has undergone significant changes with a declining importance of market exposure and growing importance of (other) factor exposures, primarily lead by SMB and HML factors. To be more precise, the market's contribution to funds' systematic activeness has dwindled from 60% in 2000 to merely 15% in 2020, while SMB and HML factors together currently account for 70% of the systematic activeness. Second, a typical low-*IDA* fund seems to focus on a single "dominant" factor within a given period, as well as over time. The dominant factor on average contributes

64% of the fund's systematic activeness within a quarter and 50% when averaged over the fund's sample life. Moreover, we find only few funds engaging in "multi-factor" or "factor-rotation" strategies—which we refer to as "smart-beta" funds—where the exposure to multiple factors is not consistently dominated by a single one, or changes over time.

In terms of performance, we show that funds whose systematic activeness is mainly driven by the exposure to the market tend to underperform those whose systematic activeness is driven by the other factors. We also find that funds focused on a single factor outperform smart-beta strategies. This finding is in line with those in Nanda, Wang, and Zheng (2004), which documents that fund families employing a large variety of strategies tend to underperform those that are more focused.

1.1 Related Literature

Our paper is tightly linked to the literature on fund activeness and holdings concentration measures. Cremers and Petajisto (2009) propose Active Share, a holdings-based measure of fund activeness, which is the sum of the absolute deviations of fund holdings from those of its benchmark. Although capturing the level of activeness of a fund, Active Share does not provide any information about the risk characteristics of the stocks in which the fund has deviated from its benchmark.[4] Our measure *IDA* takes this into account and, in fact, can be interpreted as a weighted version of Active Share where the weights reflect the contribution of stocks' idiosyncratic shocks to the volatility of fund active returns. Amihud and Goyenko (2013) propose 1-R2 as a measure of activeness, which is based on a variance decomposition of funds' total returns, where funds with higher R2 are considered as less active. Decomposing funds' total returns, as opposed to active returns, rules out the possibility that mutual funds can also be active by taking unpredictable exposure to risk factors. We, instead, allow for this and obtain a measure that reflects the source, rather than the level, of fund activeness. Overall, we contribute to this literature by characterizing benchmark deviations in terms of their risk exposures, thereby complementing existing measures of activeness with the additional information about the sources of risk that mutual funds' *active* portfolios are exposed to.

Our paper also contributes to the recent literature that aims to improve the way in which we infer fund benchmarks. Cremers, Petajisto, and Zitzewitz (2013) and Berk and van Binsbergen (2015) consider tradable versions of common risk factors as the underlying constituents of fund

[4]Doshi, Elkamhi, and Simutin (2015) consider a similar measure of fund activeness which considers the absolute deviation of the actual fund portfolio weights from the weights that the same fund portfolio would have if it was constructed as value-weighted.

benchmarks. However, although these factors are tradable, the resulting benchmark portfolios are not really investable since they are determined *ex-post* after having observed the fund performance. In order to be investable, the portfolio weights of fund benchmarks need to be determined "before the fact", as Sharpe (1992) puts it. Therefore, by pushing the "investability" requirements even further, we contribute to this literature by computing fund benchmarks as *ex-ante* tradable portfolios that have only positive weights on the underlying tradable constituents and employ no leverage. The restriction on the benchmark weights reflect the inability of most investors to have short and levered positions in their portfolios. Existing examples of *ex-ante* benchmarks in the literature include Ghysels and Jacquier (2006), Boguth, Carlson, Fisher, and Simutin (2011), and Frazzini, Kabiller, and Pedersen (2013). Ghysels and Jacquier (2006) and Boguth, Carlson, Fisher, and Simutin (2011) use lagged betas as instruments while estimating current factor betas. Frazzini, Kabiller, and Pedersen (2013) use lagged market beta to compute Buffet's CAPM active returns. However, to our knowledge, ours is the first paper that brings all aspects of investability together, and compute truly investable benchmark portfolios. Self-designated benchmarks may be an alternative to inferred benchmarks, but as documented by Sensoy (2009) they are very often mismatched, and hence not representative of the true fund strategy. Another possibility is to rely on peer-based benchmarks (Hoberg, Kumar, and Prabhala (2017)), which are very informative in some contexts, but do suffer from the lack of investability.

IDA measure can also be used as a proxy for picking strategies, and consequently $(1 - IDA)$ as a proxy for timing strategies. Intuitively, pickers deviate from their benchmarks by picking specific stock, and hence by betting against idiosyncratic shocks, while timers deviate from their benchmarks by increasing or decreasing their factor exposures, and hence by betting against systematic shocks. In Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014), for instance, pickers collect and use private information about idiosyncratic shocks, while timers learn about aggregate shocks. Relative to the traditional approach to identify pickers and timers by using the average picking and timing performance generated by a fund over a certain horizon, *IDA* has the advantage of being a performance-free measure of mutual fund strategies.[5] In Buffa and Javadekar (2020) timing (low-*IDA*) strategies reduce the ability of fund investor to learn about managerial skill, thus inducing low skilled managers to rely more heavily on those strategies.

[5]In unreported simulations, we confirm that picking and timing strategies are associated with high and low *IDA*, respectively.

2 Source of Activeness

In this section we introduce our novel measure to identify the *source of activeness* of mutual funds, i.e., whether funds deviate from their benchmark by exposing their portfolios to idiosyncratic or systematic shocks. Our measure captures the relative contribution of idiosyncratic risk, vis-a-vis systematic risk, to the volatility of a fund's active returns, and we refer to it as *idiosyncratic activeness.*

2.1 Idiosyncratic Activeness (*IDA*) Measure

While some funds generate active returns by betting on asset-specific news, others do so by betting on aggregate fluctuations. For the former funds, idiosyncratic shocks drive the majority of the variation in active returns; for the latter funds, it is systematic shocks. Although two funds may have similar average deviation from their benchmarks, i.e., similar level of activeness (e.g., similar active share, or similar tracking error), the way they deviate from them can be very different, thus generating a very different risk profile of their active returns.

Our goal is to distinguish between these two types of funds: those that deviate from their benchmark by taking active idiosyncratic risk, and those that increase or decrease, relative to their benchmark, the exposure to systematic risk. We do so by decomposing the volatility of a fund's active returns—i.e., the returns generated by the fund's active positions relative to its benchmark—into an idiosyncratic and a systematic component, and capture the source of activeness of that fund by measuring the relative contribution of the two components.

We start by estimating the following factor model on the stock excess returns:

$$R_{j\tau} = \alpha_{jt+1} + \beta'_{jt+1} F_\tau + \epsilon_{j\tau}, \quad \text{for} \quad \tau \in (t, t+1], \tag{1}$$

with $\alpha_{jt+1} + \epsilon_{j\tau}$ capturing the stock-specific/idiosyncratic shocks and β_{jt+1} measuring the stock's exposure to the set of systematic factor F. If we let w_{ijt} denote the weight of stock j in fund i's portfolio between time t and $t+1$, and w^b_{ijt} denote the corresponding weight in the benchmark of fund i, the fund's active position in stock j is then given by $\Delta w_{ijt} = w_{ijt} - w^b_{ijt}$. It follows that the fund's *active return* at time $\tau \in (t, t+1]$ can be expressed as

$$R^A_{i\tau} \equiv \sum_{j=1}^{N} \Delta w_{ijt} \times R_{j\tau} = \underbrace{\sum_{j=1}^{N} \Delta w_{ijt} \times (\hat{\alpha}_{jt+1} + \hat{\epsilon}_{j\tau})}_{\substack{\text{idiosyncratic} \\ \text{active return}}} + \underbrace{\sum_{j=1}^{N} \Delta w_{ijt} \times \left(\hat{\beta}'_{jt+1} F_\tau\right)}_{\substack{\text{systematic} \\ \text{active return}}}. \tag{2}$$

The first term in (2) represents the fund return earned through its active exposure to idiosyncratic stock returns. The second term captures the return earned through the fund's active exposure to systematic factors. To identify the relative contribution of idiosyncratic (and systematic) shocks in determining active returns, we perform a variance decomposition of R^A based on the two orthogonal terms,

$$\mathbb{V}\mathrm{ar}_t(R^A_{i\tau}) = \mathbb{V}\mathrm{ar}_t\left(\sum_{j=1}^{N} \Delta w_{ijt} \times \hat{\epsilon}_{j\tau}\right) + \mathbb{V}\mathrm{ar}_t\left(\sum_{j=1}^{N} \Delta w_{ijt} \times \hat{\beta}'_{jt+1} F_\tau\right), \quad \text{for} \quad \tau \in (t, t+1], \tag{3}$$

and define the *idiosyncratic activeness* (*IDA*) of fund i between time t and $t+1$ (referred to as period t) as:

$$\mathit{IDA}_{it} \equiv \frac{\mathbb{V}\mathrm{ar}_t\left(\sum_{j=1}^{N} \Delta w_{ijt} \times \hat{\epsilon}_{j\tau}\right)}{\mathbb{V}\mathrm{ar}_t(R^A_{i\tau})}, \quad \text{for} \quad \tau \in (t, t+1]. \tag{4}$$

Since $\mathit{IDA} \in [0, 1]$, the fund's *systematic activeness* is given by $(1 - \mathit{IDA})$, thus representing the fraction of the total active return variation between t and $t+1$ that is driven by the active exposure to systematic factors. It is important to stress that our measure does not depend of the level of active returns but rather on their volatility, thus making it, at least a priori, independent of fund performance.

We note that if the idiosyncratic shocks $\epsilon_{j\tau}$ are orthogonal to each other (which is the case when the factor model in (1) is well specified), then *IDA* can be written as

$$\mathit{IDA}_{it} = \sum_{j=1}^{N} (\Delta w_{ijt})^2 \times \frac{\mathbb{V}\mathrm{ar}_t(\hat{\epsilon}_{j\tau})}{\mathbb{V}\mathrm{ar}_t(R^A_{i\tau})}, \quad \text{for} \quad \tau \in (t, t+1]. \tag{5}$$

According to (5), a fund's idiosyncratic activeness can be interpreted as the sum of the squared active positions of the fund (capturing the level of activeness), weighted by the contribution of each idiosyncratic shock to the volatility of active returns. Therefore, each deviation from the benchmark is judged by the type of active returns it generates.

In order to take into account that (i) idiosyncratic shocks might exhibit some pairwise comovement, and that (ii) a fund might change its portfolio weights in between time t and $t+1$, we compute *IDA* using fund returns data, without the need to rely on fund holdings and underlying stock returns data, as we discuss next.

2.2 Empirical Implementation

Given the conceptual foundation of our measure *IDA*, we now turn to the specifics of our empirical implementation, which relies on two stages. In the first stage, we estimate fund benchmarks to obtain mutual funds' active returns. In the second stage, we run a factor model to decompose the variance of active returns and compute a time series of *IDA* for each fund.

2.2.1 Stage 1: Estimating Fund Benchmarks

We estimate a fund's benchmark in a given quarter by forming a long-only portfolio of index funds with no leverage that generates a performance that is "closest" to that of the fund in the previous four quarters. In particular, we follow Berk and van Binsbergen (2015) and use a set of Vanguard's passively managed index funds that are easily tradable by most investors: Vanguard 500 Index fund (VFINX), Mid-Cap Index fund (VMISX), Small-Cap Index fund (NAESX), Value Index fund (VIVAX), and Growth Index fund (VIGRX). Denoting the index funds' excess returns by R^I, the daily benchmark excess return of fund i in quarter t is given by

$$R^B_{i\tau} = \hat{b}'_{it} R^I_\tau, \quad \text{for} \quad \tau \in (t, t+1]. \tag{6}$$

where the vector weights $\hat{b}_{it}$ are estimated by regressing fund i's daily excess returns on the excess returns of the index funds in the previous four quarters, without a constant and with the restrictions that each loading is non-negative and the sum of the loadings does not exceed one:

$$R_{i\tau} = b'_{it} R^I_\tau + e_{i\tau}, \quad \text{subject to} \quad \begin{cases} b_{it} \geqslant 0 \\ \mathbb{1}' b_{it} \leqslant 1 \end{cases}, \quad \text{for} \quad \tau \in (t-4, t]. \tag{7}$$

So, for each fund in a given quarter, we use daily excess returns over the past year to obtain a long-only portfolio of index funds without leverage that best resembles the total returns of that fund, and use that portfolio to determine the fund's active return in that quarter:[6]

$$R^A_{i\tau} = R_{i\tau} - R^B_{i\tau}, \quad \text{for} \quad \tau \in (t, t+1]. \tag{8}$$

2.2.2 Stage 2: Performing Variance Decomposition Using a Factor Model

Once we obtain the time series of active returns for each mutual fund, we consider a factor model to decompose the variation in active returns in its systematic and idiosyncratic components. Specifically, we estimate the factor model in (1), directly on the time series of a fund's active returns:

$$R^A_{i\tau} = \alpha^A_{it+1} + \beta^{A\prime}_{it+1} F_\tau + \epsilon^A_{i\tau}, \quad \text{for} \quad \tau \in (t, t+1], \tag{9}$$

where, as a baseline case, we use the three factors of Fama and French (1993), along with Carhart (1997)'s momentum as a fourth factor. Based on the estimated factor exposures $\hat{\beta}^A_{it+1}$, the variance decomposition in (3) becomes

$$\mathbb{V}\text{ar}_t(R^A_{i\tau}) = \mathbb{V}\text{ar}_t(\hat{\epsilon}^A_{i\tau}) + \mathbb{V}\text{ar}_t(\hat{\beta}^{A\prime}_{it+1} F_\tau), \quad \text{for} \quad \tau \in (t, t+1], \tag{10}$$

and we obtain a fund's *IDA* as

$$IDA_{it} = \frac{\mathbb{V}\text{ar}_t(\hat{\epsilon}^A_{i\tau})}{\mathbb{V}\text{ar}_t(R^A_{i\tau})}, \quad \text{for} \quad \tau \in (t, t+1]. \tag{11}$$

When stock-level idiosyncratic shocks are orthogonal to each other, and funds change their portfolio weights only at the beginning of each quarter, then *IDA* in (4) and (5) coincide with *IDA* in (11).[7] When, instead stock-level idiosyncratic shocks exhibit some comovement, and fund portfolio weights are adjusted also during a quarter, then *IDA* in (11) takes these into account, and hence is a more general measure of a fund's idiosyncratic activeness.

[6] In other words, we estimate benchmark weights in-sample and applied them out-of-sample. This is similar to the method used in Boguth, Carlson, Fisher, and Simutin (2011) to avoid over-conditioning.

[7] In this case, fund-level alpha, beta and idiosyncratic shocks are respectively equal to

$$\hat{\alpha}^A_{it+1} = \sum_{j=1}^{N} \Delta w_{ijt} \times \hat{\alpha}_{jt+1}, \quad \hat{\beta}^A_{it+1} = \sum_{j=1}^{N} \Delta w_{ijt} \times \hat{\beta}_{jt+1}, \quad \text{and} \quad \hat{\epsilon}^A_{i\tau} = \sum_{j=1}^{N} \Delta w_{ijt} \times \hat{\epsilon}_{j\tau}.$$

2.2.3 Discussion on the Two Stages

Our first stage is instrumental to estimate fund benchmarks that satisfy the following "investability" criteria:

(i) the underlying assets of a benchmark portfolio need to be easily available and tradable by most investors, as already emphasized by Cremers, Petajisto, and Zitzewitz (2013) and Berk and van Binsbergen (2015);

(ii) the weights of a benchmark portfolio need to be strictly positive and sum to less than one, reflecting the general inability of most investors to short and use leverage;

(iii) a benchmark portfolio needs to be *ex-ante*, and hence formed at the beginning of a quarter, as not to reflect any information about the active fund's performance in that quarter.

By taking the perspective of most investors, we view these three criteria as essential for any fund benchmark to be fair and meaningful. Using very liquid index funds as benchmark constituents, and estimating benchmark weights as in (7), we guarantee that our fund benchmarks satisfy all three criteria.[8] Our benchmark portfolios, therefore, represent the closest passive alternative to their corresponding active funds that can be easily traded at the beginning of each quarter.

By requiring these more stringent criteria, we depart from most of the extant literature, which more often relies on benchmark constituents that are not really tradable assets, and whose portfolio weights are obtained *ex-post* after observing the performance of the active fund for which the benchmark is computed. It is the *ex-post* nature of a benchmark that is most problematic. To see why this is the case, consider a common way to construct fund benchmarks, which entails running a simple factor model on the fund excess return,

$$R_{i\tau} = A_{it+1} + B'_{it+1} F_\tau + \eta_{i\tau}, \quad \text{for} \quad \tau \in (t, t+1], \tag{12}$$

and defining the fund benchmark as the predictable part of the returns, $\hat{B}'_{it+1} F_\tau$, essentially collapsing our two stages into one. Even if the factors F were to be tradable, by using the fund performance in quarter $t+1$ (i.e., between t and $t+1$) to obtain the benchmark loadings B_{t+1}, not only an investor would not be able to form the benchmark portfolio at the beginning of that quarter (and hence using it as a valid investment alternative), but the resulting active returns attributed to the fund, $\hat{A}_{it+1} + \hat{\eta}_{i\tau}$, would be, by construction, entirely idiosyncratic. This means

[8] Another advantage of using Vanguard's index funds is that their performance is net of transaction costs, thus reflecting the real cost of trading (Berk and van Binsbergen (2015)).

that any return that a fund generates by actively changing the exposures to different factors as part of its strategy would be considered entirely passive and completely attributed to its benchmark. This, for instance, would make any factor timing strategies passive. The part of B_{it+1} that is unpredictable at the start of quarter $t+1$ cannot be replicated by a fund investor, and for this reason should not be used to identify the fund benchmark. With the objective to isolate the predictable component of B_{it+1}, we estimate *ex-ante* weights in our first stage using data only up to quarter t.[9] Therefore, adopting two separate stages to identify a fund's benchmark and to characterize the sources of its active returns is key to meaningfully measure the source of the fund's activeness.

We also note that another advantage of the two stages is that the set of factors in the two stages do not necessarily need to coincide. This allows us to separate what is investible by most investors and what is a source of systematic variation in active returns. In our specification, in fact, the first stage estimation is carried out using tradable Vanguard Index funds while the decomposition of fund activeness in the second stage uses common risk factors to better capture the risk profile of the fund. For instance, momentum factor is used in the second stage to classify active returns as systematic, but it is not included as one of the constituents of a benchmark portfolio in the first stage, given that momentum Index fund/ETF did not exist for most part of the sample period. This implies that even a purely passive exposure to the momentum factor is treated as active by our empirical implementation, as it captures the contribution of the active fund in expanding the opportunity set of the investors. In other words, a fund is considered active if it provides an unpredictable exposure to factors that are available to the investors, or any exposure (predictable or otherwise) to factors that are not available to them. This approach is consistent with the recent evidence in Agarwal, Green, and Ren (2018) suggesting that hedge fund investors do value the exposure to the exotic factors, which are otherwise not easily investible.

3 Idiosyncratic Activeness and Fund Performance

3.1 Data and Measurement

We obtain mutual fund data on daily returns, expense ratios, total net assets (TNA) and other characteristics, such as fund objectives, from the CRSP Survivor Bias-Free Mutual Fund Database.

[9]Our method also differs from that in Daniel, Grinblatt, Titman, and Wermers (1997), who construct characteristic-based benchmarks. Although their benchmarks are based on a fund's asset allocation at the beginning of the period, they are not investable from the perspective of an investor, who observes the allocation only with lag.

Our sample covers the years 1999 to 2019. We also obtain data on Fama-French factors and momentum portfolios from Ken French's website.

We focus our analyses on U.S. domestic, actively managed equity-oriented mutual funds. We thus exclude index, balanced, international, sector, target-date, retirement, and accrual funds, funds with any exit-restrictions, and funds with more exotic investment strategies such as "Market Neutral" or "Dedicated Short Bias".[10] This leaves us with following CRSP codes – EDYG (Growth), EDYB/EDYI (Income, Value), EDCM (Mid-Cap), EDCS (Small-Cap), and EDCI (Micro-Cap).[11] Given that estimation of our measure requires a minimum sample of five quarters, we also exclude funds that have less than 8 quarters of data in our sample. After this restriction, the median in-sample life is 60 quarters. Following Elton, Gruber, and Blake (2001), we also eliminate smaller funds with assets less than $5M to overcome the upward bias in returns of smaller funds. We also eliminate the first year of data for every new fund to reduce the incubation bias documented by Evans (2010), and eliminate all funds with missing fund name, following Cremers and Petajisto (2009).

Finally, to correct for a potential misalignment, as documented by Kim, Shukla, and Tomas (2000); Kacperczyk, Sialm, and Zheng (2008), between a fund's portfolio and its stated objective, i.e. to be an equity fund, we require that a fund's average equity share in the sample is at least 80% and no greater than 100%. We also drop any fund-quarters where the equity share of a fund's portfolio falls below 60% or is higher than 120%.

Since mutual funds can have multiple share classes differing in their fee structure and clientele, but with identical portfolios and gross returns, we aggregate all the share classes of a fund following Huang, Wei, and Yan (2012),Amihud and Goyenko (2013), and Doshi, Elkamhi, and Simutin (2015). Consequently, fund assets are aggregate assets of all the share classes, fund age is the age of the oldest share class, and expense ratios, loads, turnover and fees are calculated as asset-weighted averages across all share classes for a given time period. Our final sample includes 3461 mutual funds with 126,315 fund-quarter observations.

[10] Sector funds are identified using crsp-obj-codes EDS, exotic strategies are identified by crsp-obj-codes EDYH and EDYS. Retirement funds/Target date funds are identified either from the fund name having those words and then additionally using Lipper Classification Codes such as MAT, MT to delete remaining retirement funds. Index funds are flagged in CRSP data. Additionally we look for key words such as "Index" in the fund name. Lipper Codes were used as an additional check to eliminate remaining index funds.

[11] Using CRSP codes is similar to Doshi, Elkamhi, and Simutin (2015). But we additionally apply appropriate Lipper codes to eliminate any misclassified funds using CRSP Objective codes. In particular, we restrict our attention to funds having the following Lipper Codes – LCCE, LCGE, LCVE, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, and SCVE.

3.1.1 Summary Statistics

Panel A of Table 1 presents the summary statistics for *IDA*. Mean *IDA* is 0.60 with median *IDA* of 0.62. The inter-quartile range is 0.45-0.77, which is symmetric around the mean. *IDA* variation is 0.21 in the pooled data, *within-fund* variation is 0.18 and across-fund variation is 0.13. Panel B provides a frequency distribution over five equal-length intervals between 0 and 1 – the range of *IDA*. The right tail is heavier with more than 20% of the observations falling in the $(0.80, 1]$ range, while only 4% lie in the $[0, 0.20]$ range.

Panel C shows that *IDA* has a strong correlation structure with some important fund characteristics in the pooled data. Specifically, Panel C shows that *IDA* is not just driven by high idiosyncratic variance of active returns, but also by a lower variance of systematic returns. High-*IDA* fund-quarters (top 20%) have lower active volatility (5.15% relative to 7.47%) relative to low-*IDA* (bottom 20%) fund-quarters, due to both lower systematic active volatility (2.07% relative to 6.21%) and moderately higher idiosyncratic active volatility (4.50% relative to 3.94%). High- and low-*IDA* funds have similar age profile, prima-facie ruling out a life-cycle pattern of when fund prefer systematic or idiosyncratic strategies. Relative to low-*IDA* funds, high-*IDA* funds charge modestly higher fees (1.21% vs. 1.15%), are smaller (\$1114.16M vs. \$1283.66M), have lower turnover ratio (0.73 vs. 0.87), hold approximately one-third of the number of stocks in their portfolio (69 vs.188), and belong to smaller fund families (10 funds vs. 13 funds in a family). These findings are consistent with Pastor, Stambaugh, and Taylor (2020), who show that smaller funds trade less and hold more illiquid portfolios. Limited liquidity and higher trading costs for individual stocks relative to market/factor portfolios deter larger funds to operate strategies focusing on specific stocks or large idiosyncratic exposure.

In Panel D, we document the correlation between *IDA* and fund investment styles as categorized by the Morningstar style-matrix. Panel D(a) reports the average *IDA* for funds within each cell of the Morningstar matrix. Funds focused on Small-Cap stocks have significantly lower *IDA*, suggesting they have a relatively larger active exposure to systematic factors. In fact panel D(b) shows that only around 17% of Small-Cap focused funds have above-median *IDA*. Conversely, funds focusing on Mid-Cap stocks have higher active exposure to idiosyncratic sources. For funds with a small-cap mandate, these results seem to point towards the role of limited stock liquidity as a reason for actively engaging in factor exposure strategies. Meanwhile, funds with a large-cap mandate seem to engage in active strategies of both the systematic and idiosyncratic type.

3.1.2 Persistence of *IDA*

If *IDA* captures funds' investment strategies, we expect it to be persistent over time. Panel A of Table 2 documents *IDA*'s persistence in the data. To obtain these results, we first sort funds into five groups at the end of each quarter t based upon their time t *IDA*. We then compute the average standardized *IDA* for each group over the following 8 quarters. The "high-low" spread of *IDA* between Top 20% and Bottom 20% funds within the sorting quarter is on average 2.77 standard deviations. The spread remains large around 1.72% next quarter and then monotonically declines over the following quarters, but remaining statistically and economically significant even eight quarters out. Note that *IDA* is persistent for both high- and low-*IDA* funds with (high-) low-*IDA* funds on average remaining 0.46 (0.58) standard deviations above (below) the mean eight quarters out.

Panel B estimates the probability that a fund's *IDA* is above median for each of the 8 following quarters. Only 12% of low-*IDA* funds have above-median *IDA* in the next quarter, while 79.20% of the high-*IDA* funds have above-median *IDA* in the next quarter. In fact, only 21% of the low-*IDA* funds cross the median *IDA* after 8 quarters. Overall, the data shows that *IDA* is a persistent fund characteristic.

3.2 Results

3.2.1 *IDA* and Fund Performance

We next investigate whether and how *IDA* relates to fund performance. Table 3 reports our findings results for various measures of fund performance. We first establish a contemporaneous association between *IDA* and quarterly fund performance. To this end, we sort the funds in our sample at the end of each quarter into ten decile portfolios based on their *IDA*, and then compute the equally-weighted average performance for each portfolio for the sorting quarter. We repeat the exercise for each quarter.

Given that a skilled manager is more likely to outperform in any given period, documenting a contemporaneous association is particularly important if one thinks of high-*IDA* as a proxy for fund skill. Other measures of fund activeness that have been interpreted as a proxy for skill, such as $1-R2$ have been documented to predict higher future performance (Amihud and Goyenko (2013)), but not higher contemporaneous performance (Frazzini, Friedman, and Pomorski (2016)).

Raw and Active Returns. The first column of uses raw excess returns (above the risk-free rate) to measure fund performance. The portfolio of high-*IDA* funds in the top decile of the *IDA* distribution (D10) generates 1.21% higher returns annually, relative to the portfolio low-*IDA* funds in the bottom decile (D1).

The second column of Table 3 uses our preferred measure of performance, namely active returns capturing a fund's excess return over its ex-ante benchmark. The annual active return spread between high- and low-*IDA* funds is 1.67%. It is important to note that *IDA* is a variance decomposition of active returns and as such has no mechanical relation with the level of active returns itself. Intuitively, one can think of active returns as the size of the pie, while *IDA* as slicing the pie into components that contribute to it. What our results shows is that a particular split of the pie (with high-*IDA*) also implies larger pie size. Our results also suggest that high-*IDA* funds outperform low-*IDA* funds in spite of having substantially lower turnover, which Pastor, Stambaugh, and Taylor (2017) have documented to be positively related to risk-adjusted performance.

Value-Added. As a third performance measure, we use the *value-added* measure proposed by Berk and van Binsbergen (2015). Every quarter, it is obtained as a fund's active return for the quarter times the fund's assets under management q at the beginning of that quarter,

$$V_{it} = q_{it-1} R_{it}^{A}. \tag{13}$$

where R_{it}^{A} is fund's active return. *Value-added* V_{it} captures the dollar amount that fund i has generated in quarter t by deviating from its benchmark portfolio, and is a more appropriate measure of fund performance when funds face decreasing returns to scale. Given that high-*IDA* funds are smaller relative than low-*IDA* funds, it is important to rule out that the outperformance in terms of active returns is not solely due to the size differential. To handle the skewed distribution of assets under management, we take the natural logarithm of fund size to define the log value added as

$$\log V_{it} = \log\left(q_{it-1}\right) R_{it}^{A}. \tag{14}$$

Column 3 of Table 3 shows that a portfolio of high-*IDA* funds outperforms a portfolio of low-*IDA* funds even in terms of *log value-added* by generating incremental log value-added of \$0.12M. To interpret the coefficient in Dollars, we use the transformation $(high - low)^{\$} = (high - low) \times \left(\frac{q}{\log(q)}\right)$ and apply it to the fund with median fund size. The median fund size in the sample is \$214.20M implying that high-*IDA* funds generate \$4.82M excess Dollar value annually compared to low-*IDA* funds.[12] To put this result in context, Berk and van Binsbergen (2015) estimate that an average active mutual fund generates \$3.20M annually, suggesting that the value-added due high *IDA* is economically important. Our results also suggest that the active return spread of high-*IDA* funds is in excess of the spread implied merely by the fund size differential in the presence of fund-level decreasing returns to scale. Additionally, our results suggest that outperformance of high-*IDA* funds is robust to possible differences in the intensity of decreasing returns to scale faced funds with different types of active exposure.

Ex-post Alpha. In Column 4 of Table 3 we present findings using a more traditional *within-quarter* four-factor alpha of Fama and French (1993) and Carhart (1997). We compute ex-post alphas as

$$R_{i\tau} = \alpha_{it}^{ffc} + \beta_{it}' FFC_\tau + \epsilon_{i\tau} \quad \text{for} \quad \tau \in t \tag{15}$$

where FFC denotes Fama-French-Carhart four factors. A fund's estimated alpha differs from its active return in that any exposure to the systematic factors is absorbed by its *benchmark* – whether such exposure was predictable ex-ante (passive) or not (active). In other words, α_{it}^{ffc} can be thought of as active return with respect to an ex-post benchmark given by $R_{i\tau}^B = \beta_{it}' FFC_\tau$. Note, however, that this benchmark is known only ex-post and hence not investible at the start of quarter t (by construction). Hence, we label the alpha estimated from the above model as *ex-post alpha*. We find that high-*IDA* funds generate a 3.22% higher alpha compared to low-*IDA* funds.

Sources of Active Return. The last four columns of Table 3, Columns 5-8, measure performance by considering different sources of active returns. Specifically, in Columns 5-6, we decompose active returns into their idiosyncratic and systematic components. The systematic component is the part of the active return explained by risk factors, whereas the residual is the idiosyncratic component. Taken together, Column 5 and 6 show that outperformance of high-*IDA* funds is on account of superior idiosyncratic returns – high-*IDA* funds generate 3.53% in excess idiosyncratic returns which more than compensates the underperformance of 1.49% from the systematic returns. It is important to note that it is not necessary to have any link between *IDA* and sources of returns. In particular, a fund might be actively exposed to idiosyncratic shocks, with varying success over time, yielding overall zero idiosyncratic returns. An important motivation for our *IDA* measure is precisely that fund performance does not inform us about underlying strategies.

In Columns 7-8 of Table 3, we look at the picking and timing measures proposed by Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014). The picking measure captures a fund's portfolio devi-

[12] If one were to use mean fund size instead, the incremental value added by a fund of mean size is \$22.65M.

ations from the market portfolio that are associated with higher/lower future idiosyncratic stock returns. In particular, Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014) define "Picking" for fund i at time t as

$$Picking_{it} = \sum_{j=1}^{N} \left(w_{ijt-1} - w_{mjt-1} \right) \times \left(R_{jt} - \beta_{jt-1} R_t^m \right) \tag{16}$$

where $j = 1...N$ indexes stocks in the fund portfolio and w_{ij} and w_{mj} denote the portfolio and market weight on stock j. The timing measure "Timing" is defined as the covariance between deviations and future systematic returns on the stocks as follows

$$Timing_{it} = \sum_{j=1}^{N} \left(w_{ijt-1} - w_{mjt-1} \right) \times \left(\beta_{jt-1} R_t^m \right). \tag{17}$$

Column 7 shows that high-*IDA* funds generate higher a Picking value of 0.83% annually. On the other hand, Column 8 shows that there is no difference between the Timing value of high- and low-*IDA* funds. Both Columns 5 and 7 show a consistent pattern that attributes superior performance of high-*IDA* funds to Picking or idiosyncratic sources. At the same time, Column 8 highlights the need to have a measure that extracts fund strategies using an active risk decomposition instead of using fund returns. For example, it is common in the literature to define fund managers are "pickers" or "timers" if the managers generate higher or positive picking or timing value (See Zambrana and Zapatero (2018)). *IDA* on the other hand shows that extracting strategies from performance can be misleading. For example, even if the timing performance is similar across portfolios sorted on *IDA*, the underlying active risk profile is vastly different for these funds.

3.2.2 Panel Estimation

We expand upon the already presented evidence that *IDA* and fund performance are positively correlated by exploiting the panel nature of our data. Fund performance is related to multiple fund characteristics such as age (Chevalier and Ellison (1997)), size (Berk and Green (2004); Chen, Hong, Huang, and Kubik (2004)), investment style, risk (Jordan and Riley (2015)), turnover (Chen, Jegadeesh, and Wermers (2000); Dahlquist, Engström, and Söderlind (2000); Pastor, Stambaugh, and Taylor (2017)), expense ratios (Gil-Bazo and Ruiz-Verdu (2009)). Moreover, as summarized in Table 1, *IDA* is also related to some of these fund characteristics: high-*IDA* funds are smaller, have lower turnover, hold a lower number of stocks, and exhibit lower active risk. Furthermore, there is a strong association between *IDA* and investment style – funds focused on Mid-Cap stocks feature higher *IDA* on average and funds focused on Small-Cap stocks have low *IDA*. When examining the relation between *IDA* and fund performance using a panel estimation, it is therefore important that we systematically control for these possible confounding effects.

As covariates in our estimation, we include log age, log assets, square of the log assets, expense ratio, turnover, lagged fund flows, log of number of stocks in the portfolio, and lagged active returns to control for performance persistence. Controlling for lagged flows is important in our context. Past Factor-Related-Returns (FRR's) generate higher fund flows and depresses future fund performance as documented by Barber, Huang, and Odean (2016); Song (2020). Given that current factor exposures and hence *IDA* can be correlated with past FRR's, it is important to control for this effect. We also address the issue that *IDA* – the decomposition of variation in active returns, makes economic sense only when fund has enough variation in active returns. Hence we drop all the fund-quarters where active volatility is below the threshold of 1%.[13]

Considering the same set of performance measures as Table 3, Table 4 present the results of our panel estimations. As a reference point, Column 1 regresses (gross) active returns on fund's *IDA* without any controls or fixed effects to obtain an OLS estimate of 1.145%. This estimate implies that a fund with *IDA* of 1 generates 1.15% more active returns annually relative to a fund with *IDA* of 0. In Column 2, implement our baseline panel estimation, and include covariates in our regression. We also include Style×Quarter fixed effects, thereby comparing the impact of *IDA* on a fund's active return across funds within a given style for a given quarter. This addresses the potential concern that the superior performance of high-*IDA* funds is solely explained by the underlying investment style (Frazzini, Friedman, and Pomorski (2016)).[14] We cluster the standard errors at the fund-level to account for serial correlation of errors within-fund.[15] This is the baseline panel structure we use unless otherwise mentioned. Using this panel structure, our estimated coefficient on *IDA* jumps from 1.14% to 3.13% after correctly controlling for the effects of covariates and comparing funds within a given style in the same quarter. As a robustness, we absorb style fixed effects and cluster the errors by fund and time both in Column 3 to generate a spread of 6.11% annually.

While Columns 1-3 consider gross active returns, before netting of any fees charged by fund, Column 4 instead considers active returns *net-of-fees*, which captures the net value created for the investor. Given that expense ratios are not very different across high- and low-*IDA* funds, it is not surprising that coefficient on *IDA* remains 3.13%, or approximately the same as in Column 2.

[13]Our results are robust to a lower or higher threshold.

[14]We must note however that benchmark already nets-off any predictable exposure to well-known factors/styles.

[15]Note that Style×Quarter fixed effects eliminate the correlation across the funds if such correlation is driven by a common component for each style-quarter combination.

We next test alternate performance measures. In Columns 5 and 6, we use the log *value-added* measure of Berk and van Binsbergen (2015) and again confirm the positive association between a fund's *IDA* and gross value-added. Moving from Column 5 to Column 6, the coefficient on *IDA* jumps from 0.10 to 0.17, a jump that is similar to the one we observed when comparing Columns 1 an 2. The result is robust to using traditional Fama-French-Carhart alpha with a positive coefficient of 4.57% on *IDA* (Column 7). Turning last to the sources of active returns as in Table 3, we again find that *IDA* is positively associated with a fund's idiosyncratic component (4.83%, Column 8) and negatively correlated with its systematic component (-1.91%, Column 9). However, after controlling for covariates we find that *IDA* is positively associated with both "picking" and "timing" value proposed by Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014). That is, funds with active systematic exposure actually destroy systematic value.

In Table 5, we confirm that the outperformance of *IDA* is a more general result that holds for funds belonging to smaller or larger fund families (Column 1), funds of varying sizes (Column 4), and is not explained by the number of stocks held by the fund (Column 2), its turnover (Column 3) or investment style (Columns 5-8). To this end, we estimate a non-linear model allowing the effect of *IDA* on fund performance to vary across terciles of the fund characteristics listed in Columns 1-4.[16]

One reason to consider these characteristics is that high-*IDA* funds belong to smaller fund families, are smaller in size, hold substantially lower number of stocks and trade less. It is important therefore to validate that *IDA* is associated with performance on a broad range of values for these characteristics. But there are important economic reasons for doing so as well. For example, it could be that only relatively large fund families are able to exploit stock-specific/idiosyncratic opportunities given bigger research departments that can track more number of stocks. On the contrary, smaller families might allow more flexibility to the fund manager to deviate and undertake riskier stock-specific bets. Pastor, Stambaugh, and Taylor (2017) document that funds that trade more outperform. Hence, it could be that the outperformance is driven only by handful of funds with large turnover. It must be noted though, that average turnover for high-*IDA* funds is lower. There is also a large literature on portfolio concentration and fund performance (Kacperczyk, Sialm, and Zheng (2008); Sapp and Yan (2008); Cohen, Polk, and Silli (2020)). For example, Cohen, Polk, and Silli (2020) document that the largest holdings of fund managers outperform by significant margins. Turning to fund size, given that average fund size is small for high-*IDA* funds,

[16]Bottom tercile group serves as the base group. The standalone coefficient on *IDA* indicates if *IDA* is associated with performance for bottom tercile of the characteristic, while the interaction of *IDA* with middle and top terciles indicates if the association differs as we move up the distribution of characteristic.

it could be that outperformance is concentrated within smaller funds only. At the same time, one would expect that exploiting stock-specific opportunities could be difficult for larger funds.

Columns 1-3 confirm that the outperformance of high-*IDA* is robust to fund family size, portfolio concentration, and fund turnover, given that the standalone coefficient on *IDA* is significant and economically large and none of the interaction is significant (either statistically or economically). Surprisingly, Column 4 shows that even though high-*IDA* smaller funds outperform, larger funds with higher *IDA* outperform with an even larger margin. The result suggests that larger funds are better able to exploit stock-specific strategies.

In Columns 5-8, we estimate the relationship between *IDA* and performance separately for prominent investment styles using the Morningstar style matrix (or Lipper classification when data on a fund's Morningstar style is not available). All the styles (Small-Cap, Large-Cap, Value, and Growth) exhibit higher active returns for funds with higher *IDA*. The coefficient on *IDA* is 1.57% and 1.68% for funds following Large-Cap and Value styles. However, the coefficients jump to 5.37% and 3.92% for Small-Cap and Growth styles. This is consistent with the notion that Small-Cap and Growth stocks are more likely to present stock-specific opportunities due to mispricing given asymmetric information in the market, compared to Large-Cap or bigger Value stocks which are heavily researched by large number of analysts.

3.2.3 *IDA* and Future Performance

Next we examine the performance predictability of high-*IDA* funds over low-*IDA* funds. Combining the facts that *IDA* is positively linked with contemporaneous fund performance, and that *IDA* is persistent over long horizons, suggests that *IDA* could be predictive of future fund performance. Moreover, if *IDA* reflects fund skill, we would expect fund performance to persist even if a fund changes its strategy in the short-run.

In order to test performance predictability, we sort funds by *IDA* into five portfolios at the end of each quarter and track the average performance of each portfolio over next 8 quarters. Table 6 reports the average results of this exercise when repeated over the entire sample. Column 1 reports an active return spread in the sorting period between high- and low-*IDA* funds ("high-low") of 1.32%. This spread jumps to 2.49% in the next quarter (t+1) and then gradually diminishes over time, while still remaining economically and statistically significant even after eight quarters. These findings provide strong evidence that funds with higher *IDA* are likely to outperform in the coming quarters.

4 Level vs. Source of Activeness

A more recent empirical finding is that more active funds outperform less active funds, and multiple measures have been proposed to estimate fund activeness such as Active Share of Cremers and Petajisto (2009), $1 - R^2$ of Amihud and Goyenko (2013), Active Weight of Doshi, Elkamhi, and Simutin (2015). Intuitively, these measures of fund activeness quantify how much a fund deviates from its benchmark. For example, Active Share computes a fund's total deviation from the benchmark portfolio by aggregating stock-by-stock deviations – measured as the absolute difference between stock's weight in fund's portfolio and the benchmark. In contrast, our measure Idiosyncratic Activeness or *IDA* aims to characterize the source of a fund's activeness. Particularly, *IDA* tells us whether a fund's active exposure is idiosyncratic or systematic. In other words, current measures capture the *level* of funds' activeness while *IDA* captures the *source* of funds' activeness.

In this section we present compelling evidence that the source of a fund's activeness matters. Even though, the level of a fund's activeness predicts superior performance on average, we document that this relation is entirely concentrated within funds with high *IDA*. In other words, activeness predicts superior performance only if such activeness is mostly idiosyncratic. On the other hand, we find that funds that are engaged in active market/factor bets or timing strategies in fact destroy value if they are more active. It is important to stress that our measure *IDA* is not in "competition" with the existing measures of activeness, but rather complements them by providing information on how funds try to be active. Level and source of activeness, therefore, should be viewed as complementary to each other. We present our evidence using the following measures of fund activeness: Active Share (AS) by Cremers and Petajisto (2009), One minus R-square (1-R^2) by Amihud and Goyenko (2013), Industry Concentration (IC) by Kacperczyk, Sialm, and Zheng (2008), and Active Volatility (i.e., the volatility of active returns)

Table 7 (Panel A) presents summary statistics for all of the considered measures of fund activeness. Mean active share, $1 - R^2$, Industry Concentration, and Active Volatility are 0.82, 5.3%, 4.2% and 4.63%, respectively. These statistic are in line with the estimates from the original papers.[17] Panel B gives the average of the cross-sectional correlations across various measures of activeness and *IDA*. A first observation is that the four measures of activeness are highly correlated with each other. The correlation of Active Share with $1 - R^2$ and Industry Concentration (IC) is

[17] Although Cremers and Petajisto (2009) do not mention a point estimate of the mean active share, the time series plot (Figure 4 in their paper) indicate that the mean range of active share over time is 0.75-0.80. The mean 1-R2 is 8.9%, slightly above our estimates as we compute $1 - R2$ using Vanguard factors instead of Fama-French-Carhart factors. Mean estimate of IC in the Kacperczyk, Sialm, and Zheng (2008) is 5.98%.

0.48 and 0.30, while that of $1 - R^2$ with IC is 0.42. Active volatility is also strongly correlated with Active Share, $1 - R^2$, and IC, with correlations of 0.56 and 0.54, and 0.43, respectively.

IDA, instead, is not very correlated with any of these measures of activeness. For instance, the correlation of *IDA* with Active Share and Industry Concentration is merely 0.11 and 0.03, respectively. This suggests that active funds are active in diverse ways - some by exposing themselves to idiosyncratic shocks while others to systematic shocks. The fact that *IDA* and industry concentration are almost uncorrelated goes against the natural prior that funds having larger industry concentration are heavily exposed to systematic factors. In the data, some high-IC funds seem to be specializing in stocks within industries, while other high-IC funds overweight industry portfolios to eliminate idiosyncratic risk. The correlation of *IDA* with $1 - R^2$ and Active Volatility are slightly higher (in absolute value), with coefficients of 0.34 and -0.27, respectively.

4.1 Double Sorts

To understand how the level and the source of fund activeness interact with each other and affect fund performance, we first use portfolios of funds double-sorted on the source and level of activeness (as is common in the literature). To this end, at the end of each quarter, we first sort the funds based upon their average *IDA* over the past four quarters (including the sorting quarter) into five portfolios and further sort each of these portfolios into five portfolios based upon a fund's average activeness over the same time window, producing 25 portfolios of funds exhibiting varying levels and sources of activeness. Then, for each of the portfolios, we compute future active returns over the following quarters. Table 8 presents the average future performance for each of these portfolios.

Active Share (AS). Panel A presents the evidence for Active Share. Consistent with Cremers and Petajisto (2009), Column 1 (All) confirms that higher active share predicts higher future fund performance with active return spread of 81 basis points in the next quarter and 56 basis points over next four quarters (measured in annual terms) between high-AS and low-AS portfolios. [18] However, conditional on funds having low-*IDA* in the past (Column 2), higher active share predicts lower, not higher future performance – the high-AS portfolio underperforms the low-AS portfolio by 41 basis points over the next quarter (though not statistically significant) and even statistically by 45 basis points over next four quarters. In other words, systematic activeness

[18] Lower spread between high-low portfolios sorted on active share compared to the original paper is consistent with Jones and Mo (2020) who document that out-of-sample spreads for many mutual fund predictors drop by at least half.

hurts fund performance. This underperformance persists as we move from the bottom to second quintile of *IDA* (Column 3). Active share predicts future performance positively only as we move to higher quintiles of past *IDA*. For high-*IDA* funds, the high-AS portfolio outperforms the low-AS portfolio by 74 and 62 basis points over the next quarter and year, respectively. This evidence indicates that the predictive ability of active share is solely driven by high-*IDA* funds. Economically, *any* activeness is not guaranteed to predict better future performance as highlighted by Frazzini, Friedman, and Pomorski (2016). Our evidence shows that only funds having high idiosyncratic activeness predict better future performance. Comparing the extreme portfolios, Panel A shows that a "high-*IDA*, high-AS" portfolio generates 1.24% excess active return annually compared to a "low-*IDA*, low-AS" portfolio.

The last column of Panel A (high-low *IDA*) estimates the active return spread between high- and low-*IDA* portfolios for a given level of active share over the next quarter. *IDA* predicts statistically and economically higher future performance for each group of active share. Importantly, the "high-low" spread that *IDA* generates increases with active share from 1.78% for the portfolio of the least active funds to 2.94% for the portfolio of most active funds. Again, it is important to stress that this evidence is not a horse-race between the two measures. In fact, *IDA* captures the decomposition of a fund's activeness, and as such loses its economic meaning when a fund is not sufficiently active. What our evidence shows is that for funds which are most active, those having larger active exposure to idiosyncratic shocks perform significantly better.

1-$\mathbf{R}^2$ and Industry Concentration (IC). Panel B and C consider $1 - R^2$ and Industry Concentration (IC), respectively, and show that the evidence presented for active share generalizes to these two measures as well. In spite of having positive correlation between *IDA* and $1 - R^2$, Panel B confirms that the two measures are not substitutes for each other. In fact, the average or unconditional "high-low" spread of 1.55% between high- and low- portfolios of $1 - R2$, is concentrated within the fourth and fifth quintile of *IDA*, while high $1 - R^2$ generates underperformance of 50 basis points if *IDA* is low. The same results carry over to Industry Concentration (Panel C), where the active return spread between high- and low- portfolios of IC is almost four times for high-*IDA* portfolio (1.27%) compared to the average spread of 45 basis points. The evidence shows that funds which overweight selective industries in their portfolio by relying on specific stocks within industries tend to perform better compared to those funds who hold diversified industry portfolios. In each of the Panels B and C, the last column confirms that higher *IDA* translates into superior future performance for each group of $1 - R^2$ or IC as the case may be. In fact, for the portfolios of high-activeness in these two panels, the "high-low" spread produced by *IDA* is in excess of 3% – twice that of the spread for group of low-activeness funds.

Active Volatility (AV). Panel D considers the active volatility (AV) measure. Recall that *IDA* is obtained by decomposing AV into its idiosyncratic and systematic components. In this sense, active volatility can be considered as a measure of a fund's activeness. Column "All" shows that higher AV in the past predicts poor fund performance in the future with a "high-low" spread of -2.12% in the next quarter and -1.72% over next year. This is consistent with Jordan and Riley (2015) who show that fund return volatility is associated with poor fund performance. The correlation between AV and return volatility is 0.67, explaining the negative link between active returns and AV. But as we move from low-*IDA* to high-*IDA* groups, the negative impact of AV on fund performance diminishes. For high-*IDA* groups, higher AV does not predict poor performance over the next quarter.

The last column of Panel D again confirms that *IDA* generates higher active return for each of the groups of AV. This is of particular importance given that AV is negatively correlated with both *IDA* and fund performance. Hence it is important to rule out that the *IDA*-induced active return spread is not concentrated only within low-AV groups. Evidence shows that in fact, *IDA* induced spread is significant for all AV groups. Moreover, the spread increases from 1.16% for the low-AV group to 2.72% for the high-AV group. This confirms that the link between *IDA* and active returns does not merely proxy the link between volatility and active returns.

Robustness With Four-Factor α. The results we have presented so far focus on performance in terms of active returns. However, the majority of the measures have been shown to predict more traditional performance measures, such as Fama-French-Carhart four factor alpha. In row (c) of each panel, we report the spread in α over the next quarter between high- and low-activeness portfolios. All the panels confirm again that the average outperformance in terms of α is driven only by high-*IDA* portfolios.

Taken together, Panels A-D provide consistent evidence that the level of fund's activeness on its own does not predict performance. Our contribution is in isolating a common component of activeness that consistently predicts better future performance across all measures of activeness.

4.2 Panel Estimation

To further verify the conclusions from our double-sorting exercise, we estimate the impact of the level and the source of activeness on fund performance jointly, exploiting the panel structure of our data. We use the same controls as before and apply Style×Quarter fixed effects with fund clustering. At the end of each quarter, we sort the funds into five groups based on average *IDA* over the past four quarters (as we did for the double sorting exercise described above). For convenience, we club the middle three groups. We interact these *IDA* groups with average fund activeness computed over the past four quarters. Table 9 presents our findings.

Consider column 1 where activeness is measured using Active Share. First, relative to the bottom *IDA* group, middle and top *IDA* groups earn 1.99% and 2.05% higher active returns, respectively, over the next quarter (measured in annual terms). This confirms that *IDA* matters for fund performance controlling for fund activeness. Second, the coefficient on activeness is negative indicating that higher activeness does not produce higher active returns when *IDA* is low. Third, and most importantly, the estimated coefficient on the interaction between activeness and the top *IDA* group is positive and economically large (0.54%), indicating that for the subset of high-*IDA* funds, active share produces higher future active returns. We obtain consistent patterns across Columns 2-4 where we consider other measures of activeness. In fact, Column 3 indicates that industry concentration which loads more systematic active exposure hurt fund performance with an estimated negative coefficient of -0.69%.

5 Robustness

We perform two sets of robustness analyses for the computation of *IDA*. First, we modify the systematic factors used to decompose a fund's active returns, and second we also modify the benchmark used to compute a fund's active return.

Robustness to Systematic Factors. Fund managers are likely to have active systematic exposure to a wider set of factors beyond traditional market, size, value, and momentum factors, especially given the emergence of ETFs. If omitted factors are present in the decomposition, it implies that the true systematic active exposure gets categorized as idiosyncratic, thereby inflating our estimate of the *IDA* measure. With this motivation, we employ two alternative factor models. First, we test a "Six-Factor model" which includes five-factors of Fama and French (2015) and the momentum factor of Carhart (1997). This model adds "profitability" and "investment" factors to market, size and value factors of Fama and French (1993).[19] Second, we test a conditional performance model of Ferson and Schadt (1996) to account for a fund's systematic exposure that is correlated with lagged public information. Following Kacperczyk, Sialm, and Zheng (2006); Doshi, Elkamhi, and Simutin (2015), we use the dividend yield on the S&P 500 index, the default

[19] Recent research shows that profitable firms generate higher returns (Novy-Marx (2013); Fama and French (2015)) and that firms exhibiting faster asset growth underperform.

spread given by AAA-Baa yield spread, the term spread given by yield difference on 10 year and 2 year treasury bond, and the rate on the 3 month T-bill, and estimate the following model of decomposition

$$R_{i\tau}^{A} = \alpha_{it}^{A} + \beta_{it}^{A} F_{\tau} + \gamma_{it}^{A} \left(F_{\tau}^{mkt} . Z_{\tau-1} \right) + \epsilon_{i\tau} \quad \text{for} \quad \tau \in (t-1, t], \tag{18}$$

where $Z_{\tau-1}$ denotes the daily lags of four Ferson-Schadt instruments, and F^{mkt} denotes the market factor. [20] The Ferson-Schadt argument to include lagged public information is not to credit the fund manager for any return that is earned using such public information. We use the Ferson-Schadt model to correctly classify a fund's active return as systematic if such active returns were earned using lagged public information, say by adjusting market exposure accordingly. With our "typical investor's benchmark", such intelligent use of public information is still considered as active.

Robustness Relating to Benchmarks. Our ex-ante "investible" benchmark derived from tradable low-cost Vanguard Index funds represents a typical investor's point of view. However, more sophisticated investors might want to benchmark their fund managers against factors "known" to outperform, even if they are not readily investible. To this end, we modify the benchmark model from Vanguard index funds to the Four-Factor model of Fama and French (1993); Carhart (1997). While doing so, we maintain the ex-ante feature of the benchmark by considering historical factor loadings of funds on these four-factors as we did with Vanguard index funds. In the second stage, we again use the Four-Factor model as in our baseline specification, so as to understand the role of the benchmark clearly.

5.1 Results

Summary. Table 10 presents summary statistics for these alternative models. Panel A shows that adding more factors in the decomposition stage (six-factors or Ferson-Schadt instruments) reduces a fund's *IDA* only modestly. Compared to the median of 0.62 under our baseline VI+Four-Factor model, median *IDA* drops to 0.55 under the VI+Six-Factor model, and to 0.56 under the VI+Ferson-Schadt model. Panel B shows that the correlation of *IDA* computed under these three models is very high (in excess of 90%). Jointly, these findings suggest that an average fund's

[20] As we use daily factor model over a quarter, we consider daily lags of the instrument. Alternatively, one can consider start of the month values as a proxy for lags while estimating a quarterly model.

active idiosyncratic exposure cannot be explained by lagged public information or some of the newer asset pricing factors.

Panel A also reveals that changing the constituents of fund benchmarks from Vanguard Indexes to Fama-French-Cahart four-factors pushes-up funds' *IDA* significantly with the median *IDA* increasing from 0.62, under our baseline case, to 0.81.[21] Benchmarking the manager using the same set of factors adopted in the second stage makes what was a significant portion of systematic activeness in our baseline case passive. This reflects the fact that most investors cannot easily get exposure to the long-short positions that characterize canonical risk factors. Panel B also shows that *IDA* computed under four-factor benchmark has only modest correlation with *IDA* computed using Vanguard benchmark.

Performance. Table 11 documents that high-*IDA* funds produce higher raw returns, active returns, value-added, and traditional four-factor α under these extended models. Panel A uses a univariate sort based on funds' *IDA* to form ten portfolios, and reports the spread in performance between funds with top 10% and bottom 10% *IDA*. High-*IDA* funds generate 1.83% (1.85%) higher active returns, and 3.26% (3.34%) higher four-factor α under the VI+Six-Factor (VI+Ferson-Schadt) model.

Given that *IDA* under baseline VI+Four-Factor model and under VI+Six-Factor and VI+Ferson-Schadt models is strongly correlated, it is not surprising to see high-*IDA* funds outperform under these models. Panel A shows, that even after computing the benchmark using Four-Factor model, high-*IDA* funds subject to that benchmark outperform low-*IDA* funds with high-low spread of active return of 0.94% and 1.88% for α.

Panel B shows that our findings regarding the relation between *IDA* and fund outperformance under various models is also robust to panel estimation, where we control for typical covariates and absorb Style$\times$Quarter fixed effects, and cluster the standard errors at the fund level.

Double-Sort and Performance of Activeness Measures. We next test the robustness of our findings that outperformance predicted by fund activeness is concentrated within funds with high *IDA*, using a Four-Factor benchmark instead of a benchmark computed using Vangaurd Index funds. To this end, we conduct exactly the same double-sort exercise as before; sorting the funds upon their average *IDA* of the prior year into five buckets, and then further sorting each bucket into five buckets of average activeness over the prior year. For each of the resulting 25

[21] Note that in both the models, we are using four-factor model to decompose the active returns. Hence, differences in *IDA* are only driven by differences in active returns.

buckets, we then compute the performance over next year. Table 12 reports the results of this exercise. For brevity, we report only the performance spread between the portfolio of high- and low-activeness within each portfolio of *IDA*. Using either active returns or Fama-French-Carhart α, we confirm that outperformance associated with fund activeness increases if such activeness is idiosyncratic. For example, for the lowest quintile of *IDA*, the spread of active returns between high- and low-Active Share portfolios is merely 36 basis points. It jumps to 82 basis points for the highest quintile of *IDA* funds. Similarly, the spread more than doubles when we consider $1 - R^2$. We find similar improvement in performance when activeness is measured using active volatility.

6 Fund Strategies

In this section we focus our attention on different types of strategies that are associated with a funds' systematic activeness. In particular, we shed light on the sources of systematic risk to which mutual funds actively expose their portfolios, and compare the performance of these various strategies

6.1 Sources of Systematic Activeness

Which systematic factors contribute, on average, most to funds' systematic activeness ($SA \equiv 1 - IDA$)? Has the emergence of ETFs and newer factor strategies (including smart-beta funds) changed funds' systematic risk profile? To answer these questions, we decompose funds' *SA* into components due to each factor, the same way we decompose funds' active risk into idiosyncratic and systematic components. In particular, we define the contribution of factor ℓ to a fund's *SA* as

$$FA^{\ell}_{it} = \frac{\mathbb{V}\mathrm{ar}_t(\hat{\beta}^A_{i\ell t+1} F^{\ell}_{\tau}) + \sum_{k \neq \ell} \mathbb{C}\mathrm{ov}_t(\hat{\beta}^A_{i\ell t+1} F^{\ell}_{\tau}, \hat{\beta}^A_{ikt+1} F^{k}_{\tau})}{\mathbb{V}\mathrm{ar}_t(\hat{\beta}^{A\prime}_{it+1} F_{\tau})}, \quad \text{for} \quad \tau \in (t, t+1]. \tag{19}$$

where the denominator $\mathbb{V}\mathrm{ar}_t(\hat{\beta}^{A\prime}_{it+1} F_{\tau})$ gives the total active volatility of systematic component or simply the fund's *SA*.[22] The first term of the numerator $\mathbb{V}\mathrm{ar}_t(\hat{\beta}^A_{i\ell t+1} F^{\ell}_{\tau})$ computes the variance of the part of the active return that is directly attributable to factor ℓ. Meanwhile, the second term in the numerator splits the pairwise covariance terms equally across two factors; that is we assign half of the covariance between factor ℓ and k to factor ℓ against each of the other factor k.[23] Note

[22] F^{ℓ}_{τ} gives the factor ℓ daily returns and $\hat{\beta}^A_{i\ell t+1}$ gives the beta of fund's active return on factor ℓ during time $t+1$.
[23] Intuitively, $\mathbb{V}\mathrm{ar}(A+B) = \mathbb{V}\mathrm{ar}(A) + \mathbb{V}\mathrm{ar}(B) + 2C(A,B)$, so the contribution of A to the variance of $(A+B)$ is given by $(\mathbb{V}\mathrm{ar}(A) + C(A,B))/\mathbb{V}\mathrm{ar}(A+B)$, while the contribution of B is given by $(\mathbb{V}\mathrm{ar}(B) + C(A,B))/\mathbb{V}\mathrm{ar}(A+B)$.

that we scale the contribution by funds' *SA*, therefore the individual factor contributions add up to 1.

Focusing on fund-quarters with above median *SA*, Panel A of Table 13 provides summary statistics for the sources of systematic activeness. In Column 1, we report the decomposition for our baseline Four-Factor model. The mean *IDA* for the resulting group of funds is 45.3%, significantly below the unconditional mean of 60.60%. This implies that on average, systematic factors explain 54.70% of the variance of active returns for this group of funds. The key observation is that the broad market is not the most important factor explaining a fund's systematic active exposure. Rather, we find that "SMB" is the largest contributor to a fund's *SA*, explaining 37% of it. "SMB" is followed by the market factor, which explains another 25%. The remaining *SA* is split equally between HML and Momentum factors. A traditional market timing strategy is among the oldest factor strategies known in academia and industry since Treynor and Mazuy (1966) or Henriksson and Merton (1981), so perhaps the non-dominance of the market factor in funds' *SA* is a surprising finding. Later we show how market activeness lost its importance within the industry over time.

In Columns 2 and 3, we use the extended Six-Factor and Ferson-Schadt model to decompose funds' active returns. "SMB" remains the largest contributor explaining 37% and 38% of funds' *SA* under these two models, respectively. Under the Six-Factor model, on average, the Profitability and Investment factors explain 11% and 9.7% of a fund's *SA*. Because factor contributions add-up to 100%, some of the other factors have to make room for newly added factors compared to a Four-Factor model. We find that the market factor loses its importance significantly, explaining on average only 9.6% of a fund's *SA*. In fact, among all factors, it is the market factor that contributes the least to a fund's *SA*.[24] In the Ferson-Schadt model in Column 3, the interaction of interest rates with the market factor turns out to be single largest contributor out of the four macroeconomic variables we consider. It explains 21.4% of a fund's *SA*. Dividend yield explains another 7.5% of fund's *SA*. These two macro variables reduce the importance of the standalone market factor even further. These findings suggest that managers actively adjust market-beta in response to lagged macroeconomic variables which are known to predict future market returns such as interest rates or dividend yields (REFERENCE). Accounting for the direct and indirect contributions it has through the macro-variables, the market factor still only contributes roughly 20% of a fund's *SA* – much less than the "SMB" factor.

Note that given a weak correlation across returns on Fama-French-Carhart factors, the covariance terms are almost negligible.

[24]Note that we have orthogonalized the all the other factors with respect to market factor. Hence, the loss of significance of market factor is not due to covariance assignment.

6.1.1 Sources of Systematic Activeness Over Time

Next, we study the evolution of factor contributions to systematic exposure over time for funds having below-median *IDA*. We refer to these funds as "systematically active". Panel A of Figure 1 plots de-trended (using Hodrick-Prescott filter) equal-weighted average of the broad components of a fund's activeness, namely *IDA* and systematic activeness (*SA*) split into market activeness (*MA*), which is the market's contribution to *SA* and activeness due to rest of the factors which we denote as factor activeness (*FA*). First, Panel A confirms that average fund within the group of "systematically active" funds has become even more systematically active over time. Specifically, we find that *IDA* falls from more than 50% in the early part of the sample to less than 40% as we approach 2020. The second important observation is that the composition of systematic activeness (*SA*) has shifted dramatically over last two decades. While the market factor explained systematic activeness for the "systematically active" funds almost entirely in early 2000's, this is no longer true. Over time, and most dramatically in the first decade of 2000, *SA* explained by "market" dropped to 15% by the end of our sample. This suggests that mutual funds rely more and more on other factors (size, value, and momentum) in constructing their strategies. Unreported results confirm that this drop is obtained even "within-fund" when we absorb fund fixed effects, suggesting that the diminished importance of the market factor is not explained, at least entirely by newer fund mandates.

Panel B plots the importance of each of the four factors over time, where the contributions add-up to 1. The drop in importance of the market factor is compensated by the "SMB" factor; it explained less than 10% of fund's *SA* in 2000, and now accounts for close to 40% of it. The "HML" factor has also shown a steady rise in its importance, with its contribution rising from less than 10% to around 30% in 2020. Momentum on the other hand has lost its share from more than 25%, to around 15% in 2015, rising again slightly in the last 3-4 years. One potential reason why funds are relying less and less on the market factor for their systematic strategies is the emergence of market ETFs, that have allowed even unsophisticated investors to easily obtain exposure to the market. This implies that to create value for investors and gather more assets, fund managers need to employ strategies involving more sophisticated factors. With the advent of factor-based ETFs (e.g. low-volatility ETFs), it will be interesting to study how mutual funds will change their systematic strategies in response.

6.2 Are Funds "Smart-Beta"?

Having documented the average patterns of factor exposures over time, we now dig deeper to understand the nature of factor-strategies employed by the mutual funds. For example, funds could be specialized in targeting one of the factors (say "Momentum") every quarter. In this case, we see "momentum" explaining the majority of a fund's *SA* every quarter, as well as over the longer period. Alternatively, funds can employ factor-rotation strategies – a type of "smart-beta" strategy, where funds aim to target factor "seasons" moving in and out of factors regularly. In this case, the dominant factor would rotate from period to period. To this end, we compute the contribution of the largest factor which we refer to as "dominant factor" to a fund's *SA* and report the summary statistics of this variable in Panel B of Table 13. We report the numbers only for "systematically active" funds as before (identified as funds having below-median *IDA*).

The mean contribution of the "dominant" factor is 66.1%, and the 25^{th} percentile is 52%. On the other hand, the "least important factor" contributes close to 0% on average. Both of these numbers suggests that in a given quarter, funds derive their systematic exposure largely from a single "dominant factor" and that they are able to diversify away the risk of unwanted factors. In Row 3, we compute the average contribution of each factor over a fund's life and report the contribution of a "life-time dominant factor" – the factor that dominates a fund's active systematic risk profile over its life-time. The median contribution of the "life-time dominant factor"is 48.5% with a mean of 50.2%, suggesting that funds persists with the same factor over time while employing factor strategies. Had funds employed factor-rotation strategies, we would expect the contribution to be more equally spread across.

6.3 Factor Strategies and Fund Performance

Does a particular profile of systematic activeness or factor strategies matter for fund performance? This is the question we next explore. First, we distinguish between market and factor activeness (*MA* and *FA*). Columns 1-2 of Table 14 report the results. We again focus on the group of systematically active funds (funds having below-median average *IDA*) to yield economically meaningful results. Column 1 regresses a fund's active return on that fund's *MA* during the quarter, and shows a negative coefficient of -5.57% per year. We control for the level of a fund's *IDA* to isolate the variation due to a fund's *MA*. Our results indicate that within the "systematically active" funds, those having larger exposure to non-market factors (size, value, or momentum) perform much better compared to the funds employing active market-factor strategies. The result is robust as we include all the covariates as well as a strong set of fixed effects. In Column 2, instead of using the market's contribution (*MA*) as a continuous variable, we use a dummy identifying funds within top-quartile of *MA*, and we show that those funds underperform the rest by 2.64% annually.

These results, however, raise the question whether the outperformance of high-*IDA* funds (relative to low-*IDA* funds) is concentrated within funds having high-*MA* only. To understand whether this is the case, in Column 3, we split the funds into three exclusive groups. First, a group of high-*IDA* funds having the highest 20% of *IDA*. Next, we split the group of low-*IDA* funds (bottom 20% of *IDA*) into those having below-median market activeness (low-*IDA* and low-*MA*) and those having above-median market activeness (low-*IDA* and high-*MA*). Column 3 confirms that the group of high-*IDA* funds outperforms both groups by 0.96% and 1.91%, respectively.

Next, we compare the performance of "dominant factor" strategies (focusing on a single factor) against "multi-factor" strategies (targeting multiple factors at a time). To this end, Column 4 regresses a fund's active return on the share of a fund's *SA* explained by a "dominant factor" during the quarter. We obtain a positive coefficient of 1.15% on the share of the dominant factor, suggesting that systematically active funds focusing on a single factor during the quarter outperform funds with "multi-factor" strategies. This is consistent with Nanda, Wang, and Zheng (2004), who document that fund-families exhibiting a larger variation of investment strategies across its funds tend to underperform on average. We provide a complementary result at the fund-level, focusing on investment factors. Last, we also study the performance of "factor rotation" strategies (rotating of factors over time) compared to "persistent" strategies, which focus on a single factor over time. We again confirm the result that a persistent factor strategy outperforms a "factor rotation" strategy by 1.85% in terms of active returns.

7 Concluding Remarks

In which way is a mutual fund active? We answer this question by decomposing the variance of funds' active returns into its idiosyncratic and systematic components. We then use the relative importance of the former component to define the *idiosyncratic activeness* (*IDA*) of that fund. High-*IDA* funds deviate from their benchmarks by primarily taking idiosyncratic risk and exposing their portfolios to asset-specific shocks, while low-*IDA* funds deviate by primarily taking systematic risk and exposing their portfolios to aggregate shocks (in ways that are not predictable by the investors). To our knowledge, ours is the first measure that uses mutual funds' active risk exposures, instead of the level of their performance, to infer the way that they intend to beat their benchmarks.

Our analysis provides new empirical evidence on cross-sectional differences in performance in the mutual fund industry. We uncover that: (i) high-*IDA* funds outperform low-*IDA* funds by an economically significant margin; (ii) the level of fund activeness is associated with outperformance only for high-*IDA* funds—it is indeed associated with underperformance for low-*IDA* funds; (iii) the nature of systematic activeness has changed significantly in the last twenty years, with a drastic switch from market exposure to factor exposure; (iv) smart-beta funds, defined as those that have exposures to multiple factors (and identified by further decomposing the volatility of systematic active returns) tend to underperform those that focus on a single factor.

Overall, our findings allow to better understand the source of risks mutual fund mangers take by actively managing their portfolios, and how they generate value for their investors. Importantly, if the source of activeness of a mutual fund is a proxy for its skill (as proposed in the theoretical framework of Buffa and Javadekar (2020) and validated by their empirical evidence), fund investors should care more about the way funds deviate from their benchmarks rather than the extent of their deviations.

Table 1: Summary Statistics

This table reports the summary statistics for Idiosyncratic Activeness (*IDA*) in Panel A and B, its association with other fund characteristics in Panel C and with fund styles in Panel D. Panel B reports the number of fund-quarters and the sample fraction that lies within each of the five equal intervals of *IDA* from $[0, .20]$ to $(0.80, 1]$. In Panel C, $\rho(IDA, x)$ denotes the average of cross-sectional correlation between *IDA* and characteristic x. Panel C also reports the mean of each characteristic in the sample, and mean conditional on low-*IDA* (bottom 20%) and high-*IDA* (Top 20%) fund-quarters. It also reports the "High-Low" difference and associated standard errors. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively. Panel D (a) reports average *IDA* for each of the Morningstar style, while Panel D(b) reports the average fraction of funds with *IDA* above median *IDA* each period.

Panel A: *IDA* Distribution

Mean	P25	P50	P75	SD	Within SD	Between SD
0.606	0.455	0.627	0.778	0.216	0.181	0.132

Panel B: *IDA* and Fund-Quarter Frequency

Bins of *IDA*	[0-0.20]	(0.20-0.40]	(0.40-0.60]	(0.60-0.80]	(0.80-1]
Fund-Quarters	5,355	18,082	34,363	41,392	27,123
Frequency (%)	4	14	27.2	33	21.47

Panel C: *IDA* and Fund Characteristics

		Mean				
	$\rho(IDA, x)$	Sample	Low-*IDA*	High-*IDA*	High-Low	(SE)
Active Volatility (AV) %	-0.278***	5.911	7.476	5.150	-2.325***	(0.043)
Idiosyncratic AV %	0.124***	4.133	3.943	4.506	0.563***	(0.022)
Systematic AV %	-0.645***	3.861	6.217	2.079	-4.137***	(0.038)
Lagged Fund Age	0.010*	56.713	53.635	54.886	1.251***	(0.461)
Lagged Assets	-0.015***	1350.274	1283.662	1114.168	-169.495***	(44.080)
Expense Ratio	0.051***	1.184	1.156	1.216	0.059***	(0.004)
Turnover	-0.074***	0.790	0.875	0.734	-0.141***	(0.016)
No. of Funds in Family	-0.090***	11.861	13.271	10.180	-3.091***	(0.115)
No. of Stocks	-0.245***	105.848	188.995	69.935	-119.060***	(2.447)

Panel D: *IDA* and Morningstar Style Matrix

	Panel D (a): Average IDA			Panel D(b): Fraction of Funds with Above Median IDA		
	Large-Cap	Mid-Cap	Small-Cap	Large-Cap	Mid-Cap	Small-Cap
Blend	0.649	0.682	0.490	0.638	0.791	0.175
Growth	0.629	0.642	0.509	0.558	0.619	0.175
Value	0.591	0.646	0.473	0.427	0.704	0.177

Table 2: Persistence of *IDA*

This table documents the persistence of Idiosyncratic Activeness (*IDA*). Funds are sorted on quarterly *IDA* at the end of quarter t into five bins and average *IDA* for each bin is computed for each of the next 8 quarters and average is reported in Panel A using standardized units of *IDA*, where standardization is performed within each quarter. "High-Low" row reports the spread of standardized *IDA* between top 20% and bottom 20% of the bins formed during the sorting period. Panel B reports the probability that a fund belonging to bottom 20% (Lo) and top 20% (Hi) groups formed at time t has above-median *IDA* for each of the next 8 quarters. Standard errors are in parenthesis.

Panel A: Standardized *IDA*

Q of *IDA*	t	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$
Low-*IDA*	-1.438***	-0.961***	-0.806***	-0.715***	-0.672***	-0.630***	-0.607***	-0.597***	-0.585***
	(0.002)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Q2	-0.563***	-0.297***	-0.249***	-0.224***	-0.204***	-0.185***	-0.176***	-0.172***	-0.181***
	(0.002)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Q3	0.044***	0.084***	0.069***	0.061***	0.045***	0.057***	0.056***	0.059***	0.060***
	(0.002)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Q4	0.620***	0.412***	0.342***	0.299***	0.285***	0.256***	0.252***	0.243***	0.240***
	(0.002)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
High-*IDA*	1.341***	0.766***	0.648***	0.584***	0.551***	0.507***	0.479***	0.467***	0.466***
	(0.002)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
High-Low	2.779***	1.727***	1.454***	1.299***	1.223***	1.136***	1.086***	1.064***	1.052***
	(0.003)	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)

Panel B: Probability of Above-Median *IDA*

Q of *IDA*	t	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$
Low-*IDA*	0.000	0.123***	0.177***	0.202***	0.210***	0.219***	0.220***	0.214***	0.210***
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
High-*IDA*	1.000***	0.792***	0.719***	0.670***	0.636***	0.597***	0.574***	0.553***	0.535***
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
High-Low	1.000***	0.669***	0.542***	0.468***	0.426***	0.378***	0.354***	0.339***	0.325***
	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)

Table 3: *IDA* and Fund Performance

This table reports the association between fund's Idiosyncratic activeness (*IDA*) and fund performance computed using various measures. We sort the funds using *IDA* each quarter into deciles, where *IDA* is computed using Vanguard Index + Four-Factor (VI+FF4) model and report the average fund Raw Return — returns in excess of risk-free rates (Column 1), Active return (Column 2), Log *Value-Added* (Column 3), and Four-Factor α of Fama and French (1993)-Carhart (1997) (Column 4). In columns 5-6, we report the sources of active returns namely idiosyncratic and systematic components. In column 7-8, we compute Picking and Timing measures of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014). For each column, we report the "D10-D1" performance spread between Top and Bottom decile of *IDA*. Standard errors are in parenthesis. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

	Raw Return (1)	Active Ret (AR) (2)	Value-Added (3)	Four-Factor Alpha (4)	Idiosyncratic AR (5)	Systematic AR (6)	Picking (7)	Timing (8)
Bottom 10%	5.501***	-1.647***	-0.115***	-2.540***	-3.594***	1.650***	2.494***	6.429***
IDA (D1)	(0.331)	(0.122)	(0.007)	(0.129)	(0.116)	(0.040)	(0.131)	(0.363)
D2	6.316***	-0.990***	-0.069***	-1.075***	-2.075***	1.047***	2.566***	7.042***
	(0.331)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.130)	(0.361)
D3	6.812***	-0.187	-0.020***	-0.013	-0.849***	0.705***	2.810***	7.219***
	(0.331)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.131)	(0.362)
D4	7.141***	0.171	0.002	0.485***	-0.294**	0.582***	2.774***	7.279***
	(0.331)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.131)	(0.362)
D5	7.079***	0.143	-0.000	0.499***	-0.217*	0.487***	3.165***	7.244***
	(0.331)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.132)	(0.365)
D6	7.190***	0.295**	0.011*	0.643***	-0.057	0.453***	3.029***	7.267***
	(0.331)	(0.122)	(0.007)	(0.129)	(0.116)	(0.040)	(0.132)	(0.365)
D7	7.024***	0.132	0.001	0.568***	-0.149	0.373***	3.121***	6.962***
	(0.331)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.133)	(0.367)
D8	6.903***	0.110	0.009	0.613***	-0.121	0.285***	3.198***	6.711***
	(0.331)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.133)	(0.368)
D9	6.716***	-0.006	0.006	0.559***	-0.143	0.213***	3.074***	6.989***
	(0.331)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.134)	(0.372)
Top 10%	6.715***	0.026	0.006	0.685***	-0.055	0.157***	3.329***	6.598***
IDA (D10)	(0.332)	(0.123)	(0.007)	(0.129)	(0.116)	(0.040)	(0.136)	(0.376)
D10-D1	1.214***	1.674***	0.121***	3.224***	3.538***	-1.493***	0.836***	0.169
	(0.468)	(0.173)	(0.009)	(0.183)	(0.164)	(0.057)	(0.189)	(0.523)

Table 4: *IDA* and Fund Performance: Panel Estimation

This table reports the association between fund's idiosyncratic activeness (*IDA*) and various fund performance measures with panel estimation. We consider active returns (AR) - both gross (Columns 1-3) and net-of-fees (Column 4), gross Value-Added (Column 5-6), and Four-Factor α using Carhart (1997) (Column 7). Lastly we consider idiosyncratic and systematic components of active returns (Column 8-9) and Picking and Timing measures of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014) (Column 10-11). Fixed effects and standard cluster specifications are mentioned at the bottom. *IDA* and performance are measured at the same time. All the other covariates are lagged one period. Style Fixed effects indicate CRSP Objective Code. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

	Active Return (AR) (AR)			Net (AR)	Value Added		Four-Factor α	Idiosy AR	Syste AR	Picking	Timing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
IDA	1.145***	3.131***	6.118***	3.136***	0.106***	0.178***	4.576***	4.837***	-1.191***	1.681***	0.945***
	(0.180)	(0.289)	(2.262)	(0.289)	(0.010)	(0.017)	(0.311)	(0.296)	(0.109)	(0.269)	(0.269)
Log Age		0.094	0.476**	0.091		0.004	0.142*	0.032	0.063***	0.095	0.256***
		(0.071)	(0.225)	(0.071)		(0.004)	(0.073)	(0.068)	(0.022)	(0.073)	(0.069)
Log Assets		-0.446***	-0.667***	-0.442***		-0.016*	-0.347***	-0.367***	-0.063	-0.358***	-0.090
		(0.129)	(0.221)	(0.129)		(0.009)	(0.130)	(0.125)	(0.039)	(0.128)	(0.103)
Expense Ratio		-0.205	-0.543	-1.167***		-0.016**	-0.121	-0.241*	0.040	0.282*	0.420***
		(0.134)	(0.416)	(0.134)		(0.007)	(0.142)	(0.135)	(0.039)	(0.146)	(0.120)
Turnover		-0.541***	-0.678*	-0.541***		-0.030***	-0.485***	-0.687***	0.150***	-0.584***	0.140*
		(0.086)	(0.344)	(0.086)		(0.004)	(0.089)	(0.091)	(0.030)	(0.082)	(0.074)
Flows		0.001	-0.004	0.001		0.000	0.009**	0.009**	-0.006***	0.003	-0.004
		(0.004)	(0.015)	(0.004)		(0.000)	(0.004)	(0.004)	(0.001)	(0.004)	(0.003)
Lagged Active Return		0.060***	0.035	0.060***		0.003***	0.060***	0.062***	0.007***	0.022***	0.043***
		(0.005)	(0.055)	(0.005)		(0.000)	(0.005)	(0.005)	(0.002)	(0.005)	(0.004)
Log Assets Sq		0.023**	0.027*	0.023**		0.000	0.020*	0.022**	0.000	0.021**	0.000
		(0.011)	(0.015)	(0.011)		(0.001)	(0.011)	(0.010)	(0.003)	(0.010)	(0.008)
Log Number of Stocks		0.388***	0.843***	0.391***		0.021***	0.442***	0.394***	0.046*	0.005	-0.159*
		(0.067)	(0.285)	(0.067)		(0.004)	(0.071)	(0.067)	(0.024)	(0.072)	(0.085)
Style FE	N		Y		N						
Style×Time FE	N	Y		Y	N	Y	Y	Y	Y	Y	Y
Fund Cluster	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Time Cluster	N		Y		N						
Fund-Quarters	126241	83577	83586	83577	122381	83577	83577	83577	83577	74074	74074
Adj R-Sq	0.000	0.184	0.013	0.186	0.001	0.177	0.256	0.171	0.418	0.197	0.918

Table 5: *IDA* and Fund Performance: Mechanisms

This table reports the association between fund's Idiosyncratic Activeness (*IDA*) and fund performance under various conditions. Columns 1-4 estimates a non-linear model allowing the relationship between *IDA* and performance to vary over the lagged terciles of family size (measured using number of funds within a family in a given quarter), fund size, and number of stocks in the portfolio, and fund's turnover. Bottom tercile serves as the base-group. Hence the reported coefficients on the terciles and the interaction of terciles with *IDA* are interpreted as incremental performance over and above the base-group. In columns 4-7, we estimate the model for sub-sample of funds belonging to specific Morningstar/Lipper styles. Sub-sample of "Small" (Column 4) includes all funds with small-cap label (whether blend, growth or value). Other sub-samples are created similarly. *IDA* and performance are measured at the same time. All the other covariates are lagged one period. All the columns include Style× Quarter fixed effects and standard errors are clustered at fund level. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

	Dependent Variable: Active Return							
	Terciles of				Styles			
	Family Size	Stocks	Turnover	Fund Size	Small Cap	Large Cap	Growth	Value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IDA	2.935***	3.518***	2.948***	2.337***	5.374***	1.579***	3.928***	1.687***
	(0.440)	(0.520)	(0.418)	(0.456)	(0.757)	(0.352)	(0.391)	(0.365)
Middle Tercile	-0.337	0.581	-0.378	-0.623				
	(0.359)	(0.458)	(0.338)	(0.386)				
Top Tercile	0.143	0.415	-0.823**	-0.547				
	(0.346)	(0.446)	(0.358)	(0.419)				
Middle Tercile×*IDA*	0.754	-0.918	0.276	1.213**				
	(0.533)	(0.647)	(0.492)	(0.558)				
Top Tercile×*IDA*	-0.144	-0.222	0.087	1.070**				
	(0.517)	(0.587)	(0.532)	(0.540)				
Log Age	0.097	0.098	0.094	0.096	0.162	0.126	0.080	-0.175*
	(0.072)	(0.071)	(0.071)	(0.071)	(0.184)	(0.081)	(0.101)	(0.105)
Log Assets	-0.468***	-0.435***	-0.409***	-0.546***	-0.638*	-0.335**	-0.383**	-0.339*
	(0.130)	(0.129)	(0.129)	(0.178)	(0.352)	(0.147)	(0.180)	(0.193)
Expense Ratio	-0.203	-0.205	-0.187	-0.205	-0.072	0.002	-0.076	0.071
	(0.134)	(0.134)	(0.134)	(0.134)	(0.331)	(0.158)	(0.181)	(0.177)
Turnover	-0.547***	-0.545***	-0.274**	-0.543***	-0.988***	-0.545***	-0.532***	-0.264**
	(0.087)	(0.086)	(0.110)	(0.086)	(0.198)	(0.111)	(0.100)	(0.110)
Flows	0.001	0.001	0.001	0.001	0.005	-0.000	0.004	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)	(0.005)	(0.006)	(0.006)
Lagged Active Return	0.060***	0.060***	0.059***	0.060***	0.087***	0.042***	0.026***	0.045***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.008)	(0.007)	(0.008)
Log Assets Sq	0.024**	0.022**	0.019*	0.030**	0.012	0.019	0.020	0.016
	(0.011)	(0.011)	(0.011)	(0.013)	(0.031)	(0.012)	(0.015)	(0.016)
Log Number of Stocks	0.386***	0.285***	0.425***	0.394***	0.964***	0.175**	0.485***	0.359***
	(0.067)	(0.098)	(0.068)	(0.067)	(0.163)	(0.088)	(0.115)	(0.101)
Style×Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Fund-Quarter Obs	83577	83577	83577	83577	23533	41043	41511	35226
Adj R-Sq	0.184	0.184	0.185	0.184	0.260	0.131	0.255	0.141

Table 6: *IDA* and Performance Predictability

This table reports the predictability of performance for funds with higher Idiosyncratic Activeness (*IDA*). In each quarter, funds are sorted into 5 bins and for each bin, average active return is computed over next 8 quarters. Table reports the averaged-out values over entire sample period. "High-Low" indicates the spread in active returns between top 20% and bottom 20% of funds sorted on *IDA* during quarter *t*. Parenthesis reports the robust standard errors. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

	Future Active Return								
	t	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8
IDA Quintile(t)									
Low-*IDA* (Q1)	-1.319***	-2.287***	-1.404***	-1.647***	-1.211***	-0.811***	-0.898***	-0.773***	-0.794***
	(0.087)	(0.084)	(0.084)	(0.085)	(0.086)	(0.084)	(0.085)	(0.086)	(0.086)
Q2	-0.008	-0.548***	-0.498***	-0.732***	-0.572***	-0.673***	-0.424***	-0.413***	-0.626***
	(0.087)	(0.084)	(0.084)	(0.085)	(0.086)	(0.084)	(0.085)	(0.086)	(0.086)
Q3	0.219**	-0.071	-0.168**	-0.312***	-0.107	-0.545***	-0.353***	-0.377***	-0.578***
	(0.087)	(0.084)	(0.084)	(0.085)	(0.086)	(0.084)	(0.085)	(0.086)	(0.086)
Q4	0.121	0.053	0.048	-0.063	0.281***	-0.279***	-0.117	-0.145*	-0.313***
	(0.087)	(0.085)	(0.084)	(0.085)	(0.086)	(0.084)	(0.085)	(0.086)	(0.086)
High-*IDA* (Q5)	0.010	0.208**	0.151*	0.297***	0.632***	0.137	-0.133	0.211**	0.101
	(0.087)	(0.085)	(0.085)	(0.086)	(0.086)	(0.085)	(0.085)	(0.086)	(0.086)
High-Low	1.329***	2.495***	1.555***	1.945***	1.843***	0.948***	0.765***	0.983***	0.895***
(se)	(0.123)	(0.120)	(0.119)	(0.121)	(0.122)	(0.119)	(0.120)	(0.121)	(0.122)

Table 7: Correlation With Measures of Activeness

The table reports the summary statistics for various measures of activeness (Panel A). Active Share is computed as theminium active share for a fund with respect to six Vanguard Index funds (Large, Mid, Small, Market, Value, and Growth) and the details are in appendix. $1-R2$ indicates the Amihud and Goyenko (2013) measure computed as the $1-R^2$ of the regression of fund returns on the Four-Factor model of Carhart (1997) each quarter using daily data. Industry Concentration measure is computed as in Kacperczyk, Sialm, and Zheng (2008). Active Volatility indicates the volatility of active returns each quarter. Panel B reports the average of cross-sectional correlation between *IDA* and various measures of activeness. The standard errors are in parenthesis. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

Panel A: Summary Statistics

	P25	P50	P75	Mean	σ
Active Share (AS)	0.690	0.825	0.925	0.796	0.149
1-R2	0.029	0.053	0.094	0.076	0.082
Industry Concentration (IC)	0.022	0.042	0.072	0.056	0.056
Active Volatility % (AV)	3.287	4.564	6.644	6.088	8.814

Panel B: Correlation Matrix

	AS	1-R2	IC	AV
IDA	0.111***	0.349***	0.033***	-0.276***
	(0.001)	(0.000)	(0.000)	(0.000)
Active Share		0.484***	0.297***	0.562***
		(0.000)	(0.000)	(0.000)
$1-R^2$			0.428***	0.541***
			(0.000)	(0.001)
Industry Concentration				0.438***
				(0.000)

Table 8: Activeness and Performance Conditional on *IDA*

The table reports the fund's future performance for each of the 5×5 portfolio, double sorted at the end of $t-1$ first on fund's average *IDA* between quarters $t-4$ and $t-1$ and further sorted on fund's average activeness over the same period. We consider Active Share (Panel A), $1-R2$ (Panel B), Industry Concentration (Panel C), and Active Volatility (Panel D) as four measures of fund's activeness. For each of the 25 portfolios, we report the average active returns during next quarter t, as well as "High-Low" spread within each *IDA* quintile. We also report "High-Low" spread for active returns between quarter t to $t+3$ (labeled as b) and "High-Low" spread of Four-Factor α of Carhart (1997) (labeled as c). The last column of each panel reports active return spread between high-*IDA* and low-*IDA* denoted by "High-Low" *IDA* for each of the quintile of Activeness. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively. Standard errors are in parenthesis.

Panel A: Active Return Spread With Active Share

	$IDA(t-4,t-1)$						
Active Share $(t-4,t-1)$	All	Low	Q2	Q3	Q4	High	High-Low *IDA*
Low (Bottom 20%)	-0.486***	-1.715***	-0.439**	-0.312	-0.009	0.072	1.787***
	(0.096)	(0.254)	(0.210)	(0.200)	(0.198)	(0.203)	(0.207)
Q2	-0.496***	-2.136***	-1.029***	-0.050	-0.134	0.041	2.177***
	(0.096)	(0.255)	(0.211)	(0.201)	(0.198)	(0.204)	(0.252)
Q3	-0.282***	-1.950***	-0.762***	-0.123	0.041	0.843***	2.793***
	(0.096)	(0.256)	(0.211)	(0.201)	(0.198)	(0.204)	(0.297)
Q4	-0.213**	-1.842***	-0.423**	0.584***	0.960***	1.380***	3.222***
	(0.096)	(0.255)	(0.211)	(0.201)	(0.199)	(0.204)	(0.335)
High (Top 20%)	0.333***	-2.126***	-0.783***	0.125	0.882***	0.819***	2.945***
	(0.097)	(0.257)	(0.212)	(0.202)	(0.200)	(0.205)	(0.409)
(a) High-Low Active Return (t)	0.819***	-0.411	-0.344	0.437	0.891***	0.748***	
	(0.136)	(0.361)	(0.299)	(0.284)	(0.281)	(0.289)	
(b) High-Low Active Return $(t,t+3)$	0.563***	-0.458**	-0.566***	-0.238	0.069	0.621***	
	(0.076)	(0.198)	(0.172)	(0.163)	(0.163)	(0.170)	
(c) High-Low Four-Factor α (t)	1.179***	-0.181	0.062	1.189***	1.534***	1.278***	
	(0.148)	(0.353)	(0.324)	(0.318)	(0.320)	(0.329)	

Panel B: Active Return Spread With 1-R2

	$IDA(t-4,t-1)$						
1-R2 $(t-4,t-1)$	All	Low	Q2	Q3	Q4	High	High-Low *IDA*
Low (Bottom 20%)	-0.991***	-2.093***	-0.636***	-0.325*	-0.386**	-0.185	1.908***
	(0.086)	(0.227)	(0.190)	(0.179)	(0.174)	(0.176)	(0.176)
Q2	-0.864***	-1.973***	-0.657***	-0.215	0.278	0.223	2.196***
	(0.086)	(0.229)	(0.191)	(0.180)	(0.175)	(0.176)	(0.221)
Q3	-0.566***	-2.220***	-0.618***	0.177	0.431**	0.842***	3.062***
	(0.086)	(0.229)	(0.190)	(0.180)	(0.175)	(0.176)	(0.257)
Q4	-0.078	-2.180***	-0.779***	0.155	0.770***	1.062***	3.241***
	(0.086)	(0.229)	(0.191)	(0.180)	(0.175)	(0.176)	(0.291)
High (Top 20%)	0.568***	-2.593***	-1.016***	0.120	0.738***	0.830***	3.423***
	(0.086)	(0.230)	(0.192)	(0.181)	(0.176)	(0.179)	(0.371)
(a) High-Low Active Return (t)	1.558***	-0.500	-0.380	0.445*	1.124***	1.015***	
	(0.121)	(0.323)	(0.270)	(0.254)	(0.248)	(0.251)	
(b) High-Low Active Return $(t,t+3)$	1.353***	-0.020	-0.250	0.073	0.713***	0.707***	
	(0.075)	(0.187)	(0.165)	(0.157)	(0.159)	(0.161)	
(c) High-Low Four-Factor α (t)	2.607***	-0.234	-0.302	0.443	1.255***	0.719**	
	(0.133)	(0.309)	(0.292)	(0.290)	(0.292)	(0.298)	

Table 8 (Continued): Activeness and Performance Conditional on *IDA*

Panel C: Active Return Spread With Industry Concentration (IC)

	$IDA(t-4,t-1)$						
IC $(t-4,t-1)$	All	Low	Q2	Q3	Q4	High	High-Low *IDA*
Low (Bottom 20%)	-0.569***	-1.520***	-0.870***	-0.482**	-0.271	0.024	1.545***
	(0.096)	(0.252)	(0.208)	(0.199)	(0.196)	(0.202)	(0.216)
Q2	-0.312***	-1.944***	-0.960***	-0.409**	0.057	0.445**	2.388***
	(0.096)	(0.253)	(0.209)	(0.199)	(0.197)	(0.203)	(0.269)
Q3	-0.289***	-2.447***	-0.783***	-0.050	0.416**	0.578***	3.025***
	(0.096)	(0.253)	(0.209)	(0.199)	(0.197)	(0.202)	(0.304)
Q4	0.050	-2.011***	-0.314	0.711***	0.745***	1.009***	3.020***
	(0.096)	(0.253)	(0.209)	(0.200)	(0.197)	(0.203)	(0.324)
High (Top 20%)	0.015	-2.050***	-0.489**	0.723***	1.207***	1.299***	3.349***
	(0.096)	(0.255)	(0.210)	(0.201)	(0.198)	(0.204)	(0.387)
(a) High-Low Active Return (t)	0.584***	-0.530	0.381	1.205***	1.478***	1.274***	
	(0.136)	(0.359)	(0.296)	(0.282)	(0.279)	(0.287)	
(b) High-Low Active Return $(t,t+3)$	0.453***	-0.232	0.612***	0.678***	0.824***	0.887***	
	(0.076)	(0.194)	(0.171)	(0.162)	(0.162)	(0.171)	
(c) High-Low Four-Factor α (t)	1.728***	-0.205	0.915***	2.470***	2.100***	1.594***	
	(0.147)	(0.346)	(0.321)	(0.319)	(0.322)	(0.330)	

Panel D: Active Return Spread With Active Volatility

	$IDA(t-4,t-1)$						
Active Volatility $(t-4,t-1)$	All	Low	Q2	Q3	Q4	High	High-Low *IDA*
Low (Bottom 20%)	0.299***	-0.931***	0.198	0.045	0.430**	0.230	1.161***
	(0.085)	(0.227)	(0.189)	(0.179)	(0.174)	(0.177)	(0.154)
Q2	0.190**	-2.175***	0.104	0.177	0.165	0.502***	2.677***
	(0.086)	(0.228)	(0.190)	(0.180)	(0.175)	(0.176)	(0.203)
Q3	0.147*	-2.830***	-0.864***	0.372**	0.895***	1.012***	3.841***
	(0.086)	(0.228)	(0.188)	(0.180)	(0.175)	(0.176)	(0.250)
Q4	-0.923***	-2.637***	-1.481***	-0.222	0.324*	0.790***	3.426***
	(0.085)	(0.227)	(0.191)	(0.178)	(0.173)	(0.177)	(0.300)
High (Top 20%)	-1.822***	-2.499***	-1.705***	-0.477***	-0.010	0.222	2.721***
	(0.087)	(0.232)	(0.194)	(0.183)	(0.178)	(0.179)	(0.391)
(a) High-Low Active Return (t)	-2.121***	-1.568***	-1.903***	-0.522**	-0.440*	-0.008	
	(0.122)	(0.324)	(0.271)	(0.256)	(0.249)	(0.251)	
(b) High-Low Active Return $(t,t+3)$	-1.725***	-0.741***	-1.324***	-0.430***	-0.857***	-0.720***	
	(0.075)	(0.187)	(0.166)	(0.158)	(0.160)	(0.161)	
(c) High-Low Four-Factor α (t)	-2.178***	-1.912***	-2.350***	-0.237	-0.045	1.136***	
	(0.134)	(0.310)	(0.293)	(0.292)	(0.294)	(0.297)	

Table 9: Activeness and Performance Conditional on *IDA*: Panel Estimation

The table documents the association between fund active returns and fund's Idiosyncratic Activeness (*IDA*) conditional on fund activeness employing panel estimation. Columns 1-4 uses Active share, 1-R2, Industry Concentration and Active Volatility respectively as the four measures of fund activeness. For each quarter t, "Lagged Activeness" measures the average activeness between quarters $t-4$ and $t-1$. Similarly, we measure "Lagged *IDA*" for each fund over last four quarters and assign each fund to one of the three buckets – bottom 20%, middle 60% and top 20%, where bottom group serves as the control group. We use lagged covariates and Style × Quarter fixed effects in all the columns and cluster the standard errors at fund level. Parenthesis reports the robust standard errors. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

	Active Return			
Activeness Measure	Active Share (1)	1-R2 (2)	Industry Concentration (3)	Active Volatility (4)
Lagged Activeness	-0.117	0.038	-0.693***	-0.939***
	(0.098)	(0.235)	(0.145)	(0.130)
Lagged Activeness×$\mathbb{1}$(Mid 60% of Lagged *IDA*)	0.798***	0.717***	1.134***	0.310**
	(0.103)	(0.257)	(0.161)	(0.158)
Lagged Activeness×$\mathbb{1}$(Top 20% of Lagged *IDA*)	0.540***	0.675**	1.114***	0.420**
	(0.124)	(0.266)	(0.181)	(0.181)
$\mathbb{1}$(Mid 60% of Lagged *IDA*)	1.990***	2.155***	2.149***	1.487***
	(0.150)	(0.199)	(0.144)	(0.131)
$\mathbb{1}$(Top 20% of Lagged *IDA*)	2.050***	2.239***	2.119***	1.484***
	(0.165)	(0.211)	(0.156)	(0.147)
Log Age	0.128*	0.113	0.098	0.107
	(0.072)	(0.070)	(0.071)	(0.072)
Log Assets	-0.410***	-0.363***	-0.418***	-0.522***
	(0.133)	(0.130)	(0.130)	(0.131)
Expense Ratio	-0.317**	-0.392***	-0.218	0.025
	(0.139)	(0.134)	(0.133)	(0.134)
Turnover	-0.560***	-0.598***	-0.557***	-0.431***
	(0.083)	(0.082)	(0.079)	(0.084)
Lagged Flows	-0.003	-0.001	-0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)
Lagged Active Return	0.050***	0.055***	0.055***	0.055***
	(0.006)	(0.005)	(0.005)	(0.005)
Log Assets Square	0.019*	0.014	0.019*	0.028***
	(0.011)	(0.011)	(0.011)	(0.011)
Log Number of Stocks	0.708***	0.824***	0.571***	0.182**
	(0.081)	(0.079)	(0.072)	(0.073)
Style×Time FE	Y	Y	Y	Y
Fund-Quarters	81034	83673	82980	83673
Adj R-Sq	0.185	0.188	0.187	0.188

Table 10: *IDA* Summary Statistics Various Models

This table reports the summary statistics for Idiosyncratic Activeness (*IDA*) computed under various models. VI+Four Factor is the baseline model considered so far. For each of the VI+Four Factor, VI+Six Factor, and VI+FS models, we use Vanguard Index funds to compute the ex-ante benchmark and the active returns. In the second stage to decompose the active returns, we use four-factors of Carhart (1997) for VI+Four-Factor model, six-factor model including Profit and Momentum factors of Fama and French (2015), and Ferson-Schadt instruments as in Ferson and Schadt (1996) for VI+FS model. The four instruments used are dividend yield, 3 Month treasury rate, default spread (measured as the spread between Baa and AAA bonds of 10 year maturity), and term-spread (measured as the difference between AAA bonds of 10 years vs. 2 years). Four-Factor model instead uses Carhart (1997) four-factors in both the stages. Panel B reports the average cross-sectional correlations between *IDA* computed under various measures. Standard errors are in parenthesis and ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

Panel A: Summary Statistics of *IDA*

Model	VI+Six Factor	VI+FS	Four-Factor
Stage 1 (Benchmark)	Vanguard	Vanguard	Four-Factor
Stage 2 (Decomposition)	Six-Factor	Ferson-Schadt	Four-Factor
Mean	0.546	0.553	0.773
P25	0.398	0.412	0.691
P50	0.557	0.568	0.815
P75	0.707	0.709	0.897
SD	0.207	0.201	0.165

Panel B: Correlation of *IDA* Under Various Models

	VI+Four Factor	VI+Six Factor	VI+FS	Four-Factor
VI+Four Factor		0.950*** (0.000)	0.971*** (0.000)	0.257*** (0.000)
VI+Six Factor			0.924*** (0.000)	0.252*** (0.000)
VI+FS				0.258*** (0.000)

Table 11: *IDA* Under Various Models and Performance

This table reports the association between fund performance and Idiosyncratic Activeness (*IDA*) under various models and for various performance measures. In Panel A, for each of the model and for each performance measure, we report the average "High-Low" performance spread between funds with Top 10% and Bottom 10% *IDA* each quarter. In Panel B, we report the coefficient on *IDA* in the regression of active returns on *IDA* and other covariates. Each of the models include controls, Style×Quarter fixed effects and standard errors are clustered at fund level. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

Panel A: Performance Spread (Top 10% - Bottom 10% *IDA*)

Models	VI+Six Factor	VI+FS	Four-Factor
Raw Returns	1.459***	1.432***	1.140**
	(0.468)	(0.468)	(0.471)
Active Return	1.831***	1.854***	0.945***
	(0.173)	(0.173)	(0.142)
Log Dollar Value	0.132***	0.129***	0.047***
	(0.009)	(0.009)	(0.008)
Four-Factor Alpha	3.265***	3.341***	1.881***
	(0.183)	(0.183)	(0.182)

Panel B: Performance With Panel Estimation (Coefficient on *IDA*)

Models	VI+Six Factor	VI+FS	Four-Factor
Active Returns	3.448***	3.659***	0.684**
	(0.304)	(0.304)	(0.322)
Style×Time FE	Y	Y	Y
Controls	Y	Y	Y
Clustering	Fund	Fund	Fund

Table 12: Outperformance of Measures of Activeness With FFC Model

The table reports the fund's future performance for each of the 5×5 portfolio, double sorted at the end of $t-1$ first on fund's average *IDA* between quarters $t-4$ and $t-1$ computed using Four-Factor Model and further sorted on fund's average activeness over the same period. We consider Active Share (Panel A), $1-R2$ (Panel B), Industry Concentration (Panel C), and Active Volatility (Panel D) as four measures of fund's activeness as in table 8. For each of the 25 portfolios, we report the average "High-Low" active return spread between High (Top 20%) and Low (Bottom 20%) activeness within each group of *IDA*, where active returns are measured between quarter t and $t+3$. We also report "High-Low" spread of Four-Factor α (between quarters t and $t+3$) computed using Carhart (1997). ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

	$IDA(t-4,t-1)$					
	All	Low	Q2	Q3	Q4	Q5
Panel A: Active Share						
High-Low Active Returns $(t,t+3)$	0.718***	0.360**	0.600***	0.578***	0.763***	0.825***
	(0.062)	(0.140)	(0.138)	(0.134)	(0.135)	(0.136)
High-Low Four-Factor α $(t,t+3)$	0.897***	0.520***	0.711***	0.496***	0.820***	1.036***
	(0.082)	(0.179)	(0.183)	(0.177)	(0.174)	(0.177)
Panel B: 1-R2						
High-Low Active Returns $(t,t+3)$	0.987***	0.402***	1.038***	0.885***	0.979***	0.959***
	(0.058)	(0.125)	(0.122)	(0.120)	(0.120)	(0.121)
High-Low Four-Factor α $(t,t+3)$	2.241***	1.126***	1.930***	1.861***	2.098***	2.189***
	(0.076)	(0.160)	(0.162)	(0.159)	(0.155)	(0.155)
Panel C: Industry Concentration						
High-Low Active Returns $(t,t+3)$	0.229***	0.135	0.049	0.213	0.125	0.051
	(0.062)	(0.139)	(0.136)	(0.133)	(0.134)	(0.135)
High-Low Four-Factor α $(t,t+3)$	1.122***	0.685***	0.617***	0.744***	0.917***	0.790***
	(0.082)	(0.178)	(0.182)	(0.176)	(0.173)	(0.176)
Panel D: Active Volatility						
High-Low Active Returns $(t,t+3)$	-0.344***	-0.700***	-0.279**	-0.234*	-0.121	-0.218*
	(0.054)	(0.125)	(0.122)	(0.120)	(0.120)	(0.121)
High-Low Four-Factor α $(t,t+3)$	-0.788***	-1.470***	-0.999***	-0.721***	-0.449***	-0.185
	(0.071)	(0.160)	(0.163)	(0.159)	(0.156)	(0.156)

Table 13: Sources of Systematic Activeness

Panel A of the Table reports the decomposition of fund's systematic activeness (*SA*) which is given by 1 minus fund's idiosyncratic Activeness (*IDA*) for each of the VI+Four-Factor, VI+Six-Factor and VI+FS (Ferson-Schadt) models. The components are scaled so that they add upto 1. Panel B reports the mean and the other distributional quantities of the "largest contributing factor" to the fund *SA*.

Models	VI+ Four-Factor (1)	VI+ Six-Factor (2)	VI+ Ferson-Schadt (3)
IDA	0.453	0.396	0.409
1-*IDA*	0.547	0.604	0.591
Sources of 1-IDA			
Market	0.246	0.096	-0.075
SMB	0.369	0.366	0.379
HML	0.194	0.170	0.198
Mom	0.190	0.160	0.222
Profitability		0.111	
Investment		0.097	
Dividend Yield			0.070
Default Spread			-0.009
Term Spread			0.001
Interest Rate			0.214

Panel B: Share of Fund's Systematic Activeness (*SA*)

	P25	P50	P75	Mean
SA Share of "Dominant Factor"	0.520	0.641	0.790	0.661
SA Share of Least Contributing Factor	-0.011	0.003	0.026	0.007
SA Share of "Life-Time Dominant Factor"	0.388	0.485	0.609	0.502

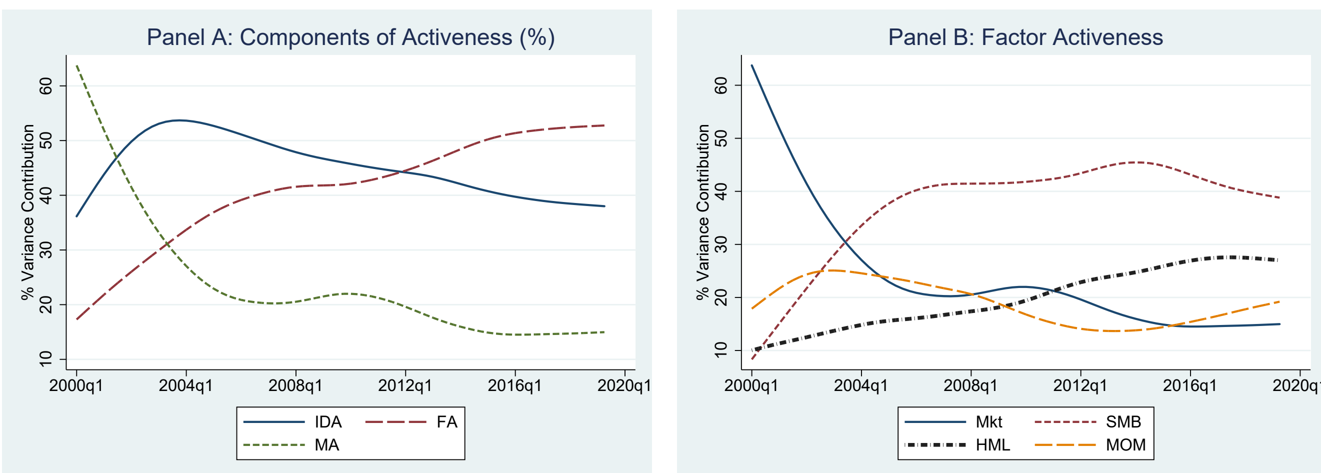
Table 14: Sources of Systematic Activeness and Performance

This table reports fund performance as a function of sources of systematic activeness (*SA*). To this end, in each column we consider only those funds that have below-median *IDA* each quarter. Market Activeness (*MA*) is the contribution of market factor to fund's *SA*. $\mathbb{1}$(Top 25% *MA*) dummy indicates if a fund has *MA* that falls in the top quartile of the distribution of *MA* that quarter. In Column 3, we construct three exclusive groups of funds each quarter – base group of funds having Top 20% *IDA* that quarter. Second group of "Low-*IDA* + Low-*MA*" indicates funds with bottom 20% *IDA* that quarter and additionally having below-median *MA* within that group of low-*IDA*. The last group of "Low-*IDA* + High-*MA*" indicates the remaining funds in the Low-*IDA* group – namely funds with above-median *MA*. In columns 4-5, we test the importance of factor rotation strategies. In Column 4, we measure the share of the largest contributing factor to fund's *SA* each quarter denoted by "Var Share of Largest Factor Within Quarter". In column 5 instead, we average out the variance contribution shares for each factor over fund's life and then consider the share of largest contributing factor. Each column is estimated using Panel data and controlling for typical covariates used in the paper as well as absorbing Style × Quarter fixed effects and clustering the errors at the fund level. Parenthesis reports the robust standard errors. ***, **, * indicates significance at less than 1%, 5%, and 10% respectively.

	Active Return (Sample: Average *IDA* < Median)				
	Market vs. Non-Market Factors			"Multi-Factor"	"Factor Rotation"
	(1)	(2)	(3)	(4)	(5)
Market Activeness (*MA*)	-5.579***				
	(0.368)				
IDA	4.946***	5.365***		6.749***	4.459***
	(0.663)	(0.669)		(0.683)	(0.460)
$\mathbb{1}$(Top 25% *MA*)		-2.642***			
		(0.177)			
Factor Concentration (Within-Quarter)				1.158***	
				(0.383)	
Factor Concentration (Over-Time)					1.851***
					(0.621)
High-*IDA*			Base Group		
Low-*IDA* + Low-*MA*			-0.964***		
			(0.208)		
Low-*IDA* + High-*MA*			-1.914***		
			(0.206)		
Controls	Y	Y	Y	Y	Y
Style×Time FE	Y	Y	Y	Y	Y
Fund Cluster	Y	Y	Y	Y	Y
Fund-Quarter Obs	42627	42627	33151	42627	42808
Adj R-Sq	0.268	0.267	0.232	0.263	0.249

Figure 1: Sources of Fund Activeness

Panel A plots the average Idiosyncratic, Market and Factor activeness (*IDA*, *MA* and *FA* respectively) for funds in our sample. Panel B plots the break-up of *FA* over important factors using Four-Factor model. "Mkt", "SMB", "HML", and "Mom" denotes Market-rf, Small-Big, Value-Growth, and Momentum factors as in Carhart (1997). For each quarter, the average denotes the equal-weighted average of all the funds in the sample.



References

Agarwal, Vikas, Clifton T. Green, and Honglin Ren, 2018, Alpha or beta in the eye of the beholder: What drives hedge fund flows?, *Journal of Financial Economics* 127, 417 – 434.

Amihud, Yakov, and Ruslan Goyenko, 2013, Mutual fund's R^2 as predictor of performance, *Review of Financial Studies* 26, 667–694.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which Factors Matter to Investors? Evidence from Mutual Fund Flows, *The Review of Financial Studies* 29, 2600–2642.

Berk, Jonathan, and Richard Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan, and Jules H. van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 18, 1–20.

Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, 2011, Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas, *Journal of Financial Economics* 102, 363–389.

Buffa, Andrea M., and Apoorva Javadekar, 2020, The allocation of talent across mutual fund strategies, working paper.

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

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, 2000, The value of active mutual fund management: An examination of the stockholdings and trades of fund managers, *Journal of Financial and Quantitative Analysis* 35, 343–368.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276–1302.

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

Cohen, Randolph B., Christopher Polk, and Bernhard Silli, 2020, Best ideas, *IESE Working Paper*.

Cremers, K. J. Martijn, and Antti Petajisto, 2009, How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22, 3329–3365.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2013, Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review* 2, 1–48.

Dahlquist, Magnus, Stefan Engström, and Paul Söderlind, 2000, Performance and characteristics of swedish mutual funds, *Journal of Financial and Quantitative Analysis* 35, 409–423.

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

Doshi, Hitesh, Redouane Elkamhi, and Mikhail Simutin, 2015, Managerial Activeness and Mutual Fund Performance, *Review of Asset Pricing Studies* 5, 156–184.

Elton, Edwin J., Martin J. Gruber, and Christopher 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, 2415–2430.

Evans, Richard, 2010, Mutual fund incubation, *Journal of Finance* 65, 1581–1611.

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

——— , 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1 – 22.

Ferson, Wayne, and Rudi Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Frazzini, Andrea, Jacques Friedman, and Lukasz Pomorski, 2016, Deactivating active share, *Financial Analysts Journal* 72, 14–21.

Frazzini, Andrea, David Kabiller, and Lasse H. Pedersen, 2013, Buffett's Alpha, NBER Working Papers 19681 National Bureau of Economic Research, Inc.

Ghysels, Eric, and Eric Jacquier, 2006, Market beta dynamics and portfolio efficiency, *University of North Carolina at Chapel Hill*.

Gil-Bazo, Javier, and Pablo Ruiz-Verdu, 2009, The relation between price and performance in the mutual fund industry, *Journal of Finance* 64, 2153–2183.

Henriksson, Roy, and Robert C. Merton, 1981, On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, *Journal of Business* 54, 513–533.

Hoberg, Gerard, Nitin Kumar, and Nagpurnanand Prabhala, 2017, Mutual Fund Competition, Managerial Skill, and Alpha Persistence, *The Review of Financial Studies* 31, 1896–1929.

Huang, Jennifer, Kelsey Wei, and Hong Yan, 2012, Investor learning and mutual fund flows, working paper.

Jones, Christopher S, and Haitao Mo, 2020, Out-of-Sample Performance of Mutual Fund Predictors, *The Review of Financial Studies*.

Jordan, Bradford, and Timothy Riley, 2015, Volatility and mutual fund manager skill, *Journal of Financial Economics* 118, 289–298.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2006, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379–2416.

——— , 2008, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2011.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2014, Time-varying fund manager skill, *Journal of Finance* 69, 1455–1484.

Kim, Moon, Ravi Shukla, and Michael Tomas, 2000, Mutual fund objective misclassification, *Journal of Economics and Business* 52, 309–323.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004, Family Values and the Star Phenomenon: Strategies of Mutual Fund Families, *The Review of Financial Studies* 17, 667–698.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1 – 28.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2017, Do funds make more when they trade more?, *Journal of Finance* 72, 1483–1528.

——— , 2020, Fund tradeoffs, *Journal of Financial Economics*.

Sapp, Travis, and Xuemin (Sterling) Yan, 2008, Security concentration and active fund management: Do focused funds offer superior performance?, *Financial Review* 43, 27–49.

Sensoy, Berk A., 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25 – 39.

Sharpe, William F., 1992, Asset allocation, *Journal of Portfolio Management* 18, 7–19.

Song, Yang, 2020, The mismatch between mutual fund scale and skill, *Journal of Finance* 75, 2555–2589.

Treynor, Jack, and Kay Mazuy, 1966, Can mutual funds outguess the market, *Harvard Business Review* 44, 131–136.

Zambrana, Rafael, and Fernando Zapatero, 2018, A tale of two types: Generalists vs. specialists in asset management, working paper.