Practice What You Preach:

Strategy Consistency and Mutual Fund Performance*

October 2022

C. Thomas Howard, Ph.D.
CEO and Director of Research
AthenaInvest Advisors LLC

Andrew L Detzel, Ph.D.
Associate Professor of Finance
Baylor University

*We thank Richard Evans for providing ticker assignment date data. All errors are our own. Contact: Andrew Detzel, Associate Professor of Finance, Baylor University, andrew_detzel@baylor.edu, +1(254)-710-6159; C. Thomas (Tom) Howard, CEO and Chief Investment Officer at AthenaInvest and Professor Emeritus of Finance, University of Denver, tom.howard@athenainvest.com, +1(877)430-5675x100.

1. Introduction

According to Morningstar, active U.S. equity mutual funds held $4.6 trillion in assets under management at the end of 2019. One of the most important and long-standing questions in financial markets is whether managers of any of these active funds possess sufficient skill to earn abnormal returns relative to their benchmarks. Predicting which funds will deliver superior performance ex ante is a critical input for evaluating market efficiency and allocating investor capital but is also notoriously challenging. Perhaps the most pervasive empirical fact on the performance of actively managed equity mutual funds is that, as a group, they significantly underperform their benchmarks, especially after expenses.[1] Moreover, Fama and French (2010) and Barras, Scaillet, and Wermers (2010) estimate that only about one or two percent of active funds have nontrivial positive abnormal returns (after costs), so identifying superior funds ex ante is analogous to finding a needle in a haystack. Worse yet, Jones and Mo (2021) and De Miguel, Gil-Bazo, Nogales, and Santos (2021) show that most of the variables that historically predicted mutual fund performance fail to do so in the last one to two decades.

In this paper, we propose a novel predictor of fund performance called strategy consistency based on a previously unexplored characteristic of fund holdings that we argue should be indicative of stock-picking skill. Strategy consistency ("consistency") is defined to be the degree to which a fund manager picks stocks in their portfolio that are most heavily invested in collectively by the group of managers with a similar self-declared principal investment strategy. Consistency should predict returns for at least three reasons. First, it reflects consensus among presumably skilled managers following a similar strategy. For example, if multiple skilled managers following a value strategy arrive at the same conclusion to purchase a given stock, it is more likely that stock was chosen wisely than if a single manager identifies it. This follows from the simple statistical fact that if multiple noisy signals convey the same message, it is simply more likely that the message is true as opposed to driven by noise. Second, over time, managers gain expertise in strategies they invest in, increasing the likelihood of future success in these strategies. If they naively extrapolate from this accumulated expertise and deviate into other strategies, they will not perform as well in expectation. This possibility is consistent with a large literature that explores the role of overconfidence in investing and explaining asset-pricing anomalies (e.g., Daniel, Hirshleifer, and Subrahmanyam, 1998, 2001; Grinblatt and Keloharju, 2001, 2009). Finally, achieving high consistency requires tilting weights from a manager's benchmark towards the stocks favored by a particular strategy thereby indicating high degrees of "activeness" and "conviction", both well-known harbingers of superior performance (see., e.g., Cremers and Petajisto, 2009; Amihud and Goyenko, 2013; Doshi, Elkhami, and Simutin, 2015; Cremers and Pareek, 2016; Cremers, 2017; and Antón, Cohen, and Polk, 2021).

We obtain a proprietary measure of active U.S. equity mutual fund strategy consistency, *Consistency*, from the asset management company AthenaInvest (hereafter Athena).[2] Though proprietary, this measure is based on literature-standard prospectus and holdings data from the SEC and Morningstar and is constructed in three intuitive steps as follows.[3] First, for each fund, Athena's algorithm examines the text of the Principal Investment Strategy in the prospectus and assigns that fund into a strategy group, such as valuation, future growth, etc.[4] Next, Athena assigns to each U.S. stock the strategy group that weights that stock most heavily. Finally, *Consistency* is an increasing function of the degree to which a given fund invests in stocks assigned to that fund's strategy group, which we refer to as own-

[1] See., e.g., Jensen (1968), Malkiel (1995), Carhart (1997), Fama and French (2010), and Ferreira, Keswani, Miguel, and Ramos (2013). Several studies find, however, that managers select stocks that outperform benchmarks before expenses. See., e.g., Grinblatt and Titman (1989, 1993), Daniel, Grinblatt, Titman, and Wermers (1997), Chen, Jegadeesh, and Wermers (2000), Wermers (2000), Alexander, Cici, and Gibson (2007), Berk and van Binsbergen (2015), and Antón, Cohen, and Polk, (2021).

[2] AthenaInvest.com

[3] Unlike many propriety measures in the literature, Athena's *consistency* measure, along with other strategy information, is available to researchers, at no cost to them, who are willing to sign a non-disclosure agreement.

[4] The principal investment strategy has been required by the SEC since 1998.

strategy stocks. Said differently, *Consistency* is a function of the degree to which a fund manager invests in own-strategy stocks.

We sort funds into five portfolios based on *Consistency* and find that high-*Consistency* funds earn significantly higher raw, benchmark-adjusted, and multifactor abnormal returns than low-*Consistency* funds by 1.9% to 3.6% per year. Before fees captured by the expense ratio (but after trading costs), high-*Consistency* funds earn significantly positive abnormal returns, indicating that consistency is evidence of skill (e.g., Berk and van Binsbergen, 2015). The performance of high-*Consistency* funds presents even though our sample period, 2007 through 2019, was especially bad for active funds as a group. For example, over this time, we find that the typical fund underperformed its benchmark even before costs and that highly active funds underperform funds with low levels of activeness, contrary to the result in the earlier sample period of Cremers and Petajisto (2009) and Amihud and Goyenko (2013). Sorting funds into portfolios based on both *Consistency* and past abnormal returns shows that high-*Consistency* funds that have performed well in the past continue to exhibit superior performance, with (net-of-costs) alphas over 4% per year with respect to the four-factor model of Cremers, Petajisto, and Zitzewitz (2013). Overall, our results are consistent with strategy consistency helping to identify the latent manager characteristic of skill.

Our study contributes to the growing literature that attempts to predict mutual fund performance. Recent studies in this vein largely focus on measures of managerial "activeness", i.e., the degree to which mutual fund portfolio weights deviate from those of their benchmark, and show that they predict fund performance (e.g., Kacperczyk, Sialm, and Zheng, 2005; Cremers and Petajisto, 2009; Amihud and Goyenko, 2013; Doshi, Elkhami, and Simutin, 2015; and Cremers and Pareek, 2016). Activeness is a necessary condition for superior mutual fund performance because managers cannot beat their benchmarks by copying them. As noted by Cremers (2017), however, activeness is not a sufficient condition since it does not directly measure the skill managers have to pick stocks, only the amount of stock-picking they actually do. De Miguel, Gil-Bazo, Nogales, and Santos (2021) use a machine-learning approach to combine many previously documented predictors of mutual fund performance, but find that, unlike *Consistency*, these predictors perform poorly during our sample period. Historically, a vast literature predicts returns on funds with past returns based on the premise that, if skill persists from one period to the next, then performance should too (see, e.g.., Bollen and Busse, 2005 for a recent survey). For example, Sirri and Tufano (2002), Del Guercio and Tkac (2008), and Berk and van Binsbergen (2016) show that past performance largely drives fund flows from investors. However, previous studies generally find that performance persistence is a short-lived phenomenon and past performance does not subsume other forward-looking information. For example, Armstrong, Genc, and Verbeek (2019) show that Morningstar ratings that incorporate analyst reports help predict funds with superior performance.

Another important finding of this paper is that, examining the correlations between the returns on all pairs of funds in the sample, we find funds' returns correlate more heavily with other funds following the same strategy than they do with funds in other strategies. No strategy classification scheme is perfect, but our correlation evidence vindicates the proprietary strategy classification used by Athena because it shows same-strategy managers pick economically related stocks. Ben-David, Li, Rossi, and Song (2020) find that investor fund flows "chase" Morningstar ratings, which are based on past performance. If these ratings are not adjusted for groupings, or "styles", of stocks in which there is a high degree of correlation, then ratings can cause a style momentum effect. The basic cause is that high within-style correlation inflates the perceived attractiveness of all funds in a successful style, thereby causing non-fundamental capital flows into these funds. Our correlation results suggest that investors and ratings providers should be careful to consider fund returns relative to other funds following similar strategies as well. Recognizing the importance of strategy-based fund categorization that is more comprehensive than commonly used alternatives like investment objective or the two-dimensional Morningstar Equity Style Box, Brown and Goetzman (1997) propose a statistical clustering-based measure of strategy categorization, though noting the limitation that the category boundaries lack economic motivation. We expand on this literature by providing a strategy classification based on economically motivated clustering and demonstrating the significant performance implications of maintaining consistency with respect to a given fund's strategy.[5]

The rest of this paper is organized as follows. Section 2 describes our data sources and variable construction. Section 3 presents our main results and Section 4 concludes.

2. Data and variable construction

2.1. *Data sources*

We collect fund-level return, turnover, expense ratios, and holdings data from Morningstar, along with index returns on the S&P500 and Russell indices from CRSP and Russell, respectively.[6] We choose Morningstar as the source of fund data to maintain consistency with Athena who use Morningstar holdings data to construct their *Consistency* measure. For funds with multiple share classes, we define the fund return to be the simple average of that of each share class. Before applying any data filters, the starting universe of funds includes all US active equity funds that existed in any month from January 1997 to December 2019. Returns on the Fama and French (1993) and Carhart (1997) market, size, value, and momentum factors, along with the one-month Treasury bill rate, come from the website of Kenneth French.[7]

2.2. Strategy and strategy consistency

2.2.1. Equity strategy identification

Investment strategy is the way a manager goes about analyzing, buying, and selling stocks. No two fund managers have identical strategies, even if they pick stocks based on economically similar variables. For example, two managers following a value strategy both try to find "undervalued" stocks, but they will not, in general, hold the same portfolio. However, despite these differences, it is critical to categorize funds into strategy peer groups based on objective empirical criterion for the purposes of performance evaluation. The performance of most fund managers is evaluated relative to a strategy-related benchmark and managers face a well-known moral hazard problem if they can choose their own benchmark. Moreover, if performance metrics are not adjusted for strategy, then investors can allocate capital to funds, regardless of skill, if a common factor drives these funds' returns and performs well (e.g., Ben-David et al., 2020). For example, if value stocks outperform growth stocks over a period, investors may naively invest in a value fund, even if that fund underperformed its value peers. Brown and Goetzman (1997) and Chan, Chen, and Lakonishok (2009) also show that common fund peer groups, such as the prospectus investment objective or Morningstar equity style grid are not necessarily well defined and are, in general, too course, leaving out important sources of common variation in the cross-section of mutual fund returns.

We argue that categorizing managers into economically motivated strategy groups based on their self-declared principal investment strategy is a natural choice. For example, managers claiming to follow a value strategy should have similar performance benchmarks. Athena strategy identifies active equity mutual funds by gathering the text of the "Principal Investment Strategies" section of each fund's prospectus and inputting this text into a proprietary strategy

[5] Chan, Dimmock, and Lakonishok (2009) also show commonly used size and value/growth-based styles are insufficient to capture cross-sectional variation in common fund strategies and therefore are inadequate to base benchmarks on. In an earlier paper, Howard (2010), using a cross-correlation analysis similar to the one described above, finds that self-declared strategy is much more effective than is the Morningstar style grid in forming fund clusters that pursue the same return factors. Even more surprising, he finds that random clustering outperforms style grid clustering in this regard. Among the three approaches, the style grid produces the worst clustering results.

[6] https://indexcalculator.ftserussell.com

[7] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

identification algorithm. This algorithm was developed using an iterative process involving manager interviews, gathering principal strategy information, eliminating key words that generated false signals, and creating a manageable number of strategies. The Principal Investment Strategies prospectus statement was first mandated by the SEC in 1998. After extensive testing, Athena finalized its algorithm by the start of 2007 and then backfilled their data to prior years. To avoid selection bias, we do not use observations prior to 2007, although untabulated tests show this choice does not impact our main inferences. To ensure accuracy, the strategy identifications assigned by the algorithm are subjected to a series of audits before being included in the fund data base. Athena updates the strategy classification whenever a fund issues a revised prospectus.[8] The strategy identification algorithm searches the principal investment strategy's text for key words and phrases that are matched with 40 "elements", which are specific items or concepts that managers use to pursue their strategy. Funds are then assigned to one of 10 equity strategies based on the combination of elements they use.

Panel A of Table 1 lists and describes the 10 equity fund strategies and Panel B lists the 40 strategy elements. Panel A shows, for example, that 'Competitive Position' managers focus on business principles, including quality of management, market power, product reputation, competitive advantage, sustainability of the business model, and history of adapting to market changes. On the other hand, 'Economic Conditions' managers take a top-down approach based on economic fundamentals which might include employment, productivity, inflation, and industrial output, then gauge where the overall economy is in the business cycle, the resulting supply and demand situations in various industries, and the best stocks to purchase as a result. While the exact algorithm used to map the principal investment strategy text onto the elements and then combine elements into strategies is proprietary, Panels A and B both shows that these elements and strategies are based on ubiquitous concepts used throughout the asset pricing and investments literature. We provide further empirical validation of these classifications below as well and, unlike other proprietary data used in the literature, Athena's measures are available at no cost to any researcher conditional on signing and non-disclosure agreement. Our strategy consistency variable can also be constructed using any alternative strategy classification.

Table 1: Active equity mutual fund strategies

AthenaInvest assigns strategies to mutual funds using an algorithm that analyzes the text of the 'principal investment strategies' section of the funds' prospectuses. The algorithm searches the principal investment strategy's text for key words and phrases that are matched with 40 "elements", which are specific items or concepts managers use to pursue their strategy. Funds are then assigned to one of 10 equity strategies based on the combination of elements they use. Panel A lists the ten mutual fund strategies. Panel B lists the 40 strategy "elements".

Panel A: Fund strategies
Competitive Position: Business principles, including quality of management, market power, product reputation, and competitive advantage. Considers the sustainability of the business model and history of adapting to market changes.
Economic Conditions: Top-down approach based on economic fundamentals; can include employment, productivity, inflation, and industrial output. Gauges where overall economy is in business cycle, the resulting supply and demand situations in various industries, and the best stocks to purchase as a result.
Future Growth: Companies poised to grow rapidly relative to others. The Future Growth and Valuation strategies are not mutually exclusive and can both be deemed important in the investment process.

[8] Their analysis used over 54,000 prospectuses and is available for the vast majority of U.S. equity mutual funds. About 7% of funds cannot be strategy identified because a prospectus cannot be located, or the prospectus did not provide useful strategy information (such a prospectus might simply state that the goal of the fund is to earn superior returns). While the Athena strategy is economically motivated, Abis and Lines (2022) use machine learning to make a statistical strategy identification also based on prospectus principal investment strategies and find that funds actions generally line up with their stated strategies.

Table 1: (continued)

Panel A: Fund strategies (continued)
Market Conditions: Consideration of stock's recent price and volume history relative to the market and similar stocks as well as the overall stock market conditions.
Opportunity: Unique opportunities that may exist for a small number of stocks or at different points in time. May involve combining stocks and derivatives and may involve use of considerable leverage. Many hedge fund managers follow this strategy, but a mutual fund manager may also be so classified.
Profitability: Company profitability, such as gross margin, operating margin, net margin and return on equity.
Quantitative: Mathematical and statistical inefficiencies in market and individual stock pricing. Involves mathematical and statistical modeling with little or no regard to company and market fundamentals.
Risk: Control overall risk, with increasing returns a secondary consideration. Risk measures considered may include beta, volatility, company financials, industry and sector exposures, country exposures, and economic and market risk factors.
Social Considerations: Company's ethical, environmental, and business practices as well as an evaluation of the company's business lines in light of the current social and political climate. A manager can look for these criteria or the lack of in selecting a stock.
Valuation: Stocks selling cheaply compared to peer stocks based on accounting ratios and valuation techniques. The Valuation and Future Growth strategies are not mutually exclusive and can both be deemed important in the investment process.

Panel B: Elements associated with each strategy

Strategy	Element	Strategy	Element
Competitive Position	Strong fundamentals	Profitability	Dividend yield
	Defensible market position		Strong financials
	Management quality		Return on equity
	Strong innovation		Return on invested capital
Economic Conditions	Economic output	Quantitative	Quantitative modeling
	Themes		Expected return modeling
	Interest rates		Stochastic modeling
	Inflation		Time sensitive anomalies
Future Growth	Overall company growth	Risk	Excess volatility
	Strong earnings growth		Country risk
	Sustainable growth		Downside risk
	Accelerated growth		Business risk
Market Conditions	Overall market conditions	Social Considerations	Political issues
	Momentum		Social Responsibility
	Technical analysis/charting		International issues
	Relative strength		Religious Issues
Opportunity	Behavioral considerations	Valuation	Intrinsic Valuation
	Arbitrage		Cash flow valuation
	Earnings surprise		Price ratios (e.g., P/E, P/S, P/B)
	Absolute Return		Contrarian

2.2.2. *Strategy consistency*

After the fund strategy identification defined above, Athena constructs its strategy consistency measure, Consistency, in two steps, described in detail in this section. First, they strategy identify stocks based on which strategy's funds invest in them most heavily. Second, they define *Consistency* for a mutual fund based on how much of its portfolio consists of own-strategy stocks.

For each stock, i, month, t, and strategy, S, Athena sums the weight, $w_{i,j,t}$, across all funds, j, with strategy S and AUM of $1 billion or less:

$$sum_{i,S,t} = \sum_{j \in S} w_{i,j,t} \,. \tag{1}$$

Using portfolio weights, rather than absolute amounts, eliminates the effect of fund size. Focusing on smaller funds (AUM $\leq$ $1 billion) helps avoid economically uninteresting variation in weights caused by the high degrees of diversification and benchmark-tracking common in large funds.[9] They then scale the sum given by Eq. (1) by dividing it by the average $sum_{i,S,t}$ for stocks in that strategy to temper the mechanical effect of the number of funds in the strategy, N_t^S:

$$\widetilde{sum}_{i,S,t} = \frac{sum_{i,S,t}}{\frac{1}{N_t^S}\sum_{k,S,t} sum_{k,S,t}} \tag{2}$$

Athena then defines the strategy profile for stock-strategy-month (i, S, t) by normalizing the $\widetilde{sum}_{i,S,t}$ to sum to 100% across the ten strategies:

$$SP_{i,S,t} = 100 * \frac{\widetilde{sum}_{i,S,t}}{\sum_{S'} \widetilde{sum}_{i,S',t}}. \tag{3}$$

Stock i's strategy is then defined to be the S with the largest $SP_{i,S,t}$ in month t.

Table 2 presents three Strategy Profile examples. Stock AAA's Social Considerations Strategy Profile weight is the largest at 27.8% and becomes its designated strategy for the month. Its weights sum to 100%, with no weighting for the Economic Conditions, Market Conditions, and Opportunity strategies. Similarly, BBB is designated a Future Growth Stock while CCC is an Opportunity stock. Untabulated results show that stocks remain in a particular strategy pool for an average of 14 months, with Market Conditions the least stable and Valuation the most stable pools. Thus, strategy stock pools move about the equity universe as the result of fund manager buy and sell decisions in response to an ever changing economic and market environment. The Competitive Position, Future Growth, and Valuation pools are the largest since the number of funds in the corresponding fund strategies are also the largest.

Table 2: Stock strategy profile examples

This table presents strategy profiles for three hypothetical stocks. Strategies are defined in Table 1 and this table uses the following abbreviations: CP=Competitive Conditions; EC=Economic Conditions; FG=Future Growth; MC=Market Conditions; Opp=Opportunity; Prof=Profitability; Quant=Quantitative; SC=Social Considerations; and Val=Valuation.

Stock	Stock Strategy	CP	EC	FG	MC	Opp	Prof	Quant	Risk	SC	Val
AAA	Social Considerations	8.2	0.0	0.7	0.0	0.0	27.3	6.9	6.8	27.8	22.5
BBB	Future Growth	26.4	1.6	44.5	2.8	5.8	3.1	1.0	0.0	9.5	5.2
CCC	Opportunity	1.6	1.9	21.1	0.0	61.6	3.3	2.2	0.0	0.7	7.7

[9] To be clear, they restrict the size of funds when assigning strategies to *stocks*, but we assign strategies and define strategy consistency for *all funds* regardless of size.

A fund's *Consistency* in a particular month is based on the percent, by count, of own-strategy stocks held by the fund. For example, if a fund holds fifty different stocks, 10 of which belong to its strategy, the fund's percent is 20%. This percent is converted into a standard normal deviate, based on the distribution of the percent of own strategy stocks held by each fund in the strategy, truncated at three standard deviations above and below the mean, rescaled to range between 0 to 10, and then rounded to the nearest whole number (0 through 10). *Consistency* is based on the number of distinct stocks, ignoring weights of stock, to empirically distinguish the notion of strategy consistency with that of conviction, which is described in the next section and based on the degree to which managers weight their top positions.

2.2.3. Activeness and conviction

In the spirit of Berk and van Binsbergen (2015), we estimate best-fit benchmarks for each fund in our sample from the nine domestic Vanguard index funds formed on size and style, which they argue proxy for the passive investment opportunity set of mutual fund investors.[10] These index funds are listed in Table 3. For each fund, i, and month, t, we assign a best-fit benchmark as the Vanguard fund with the highest R^2 statistic from the nine regressions of fund i's excess returns over month $t-36$ through $t-1$ (with a minimum of twelve observations) on those of each one of the Vanguard funds. The midcap value and growth funds are not available at the beginning of the sample. We include them as soon as they exist for at least 12 months.

Table 3: Vanguard Funds used as Fund Benchmarks

This table lists the name, ticker symbol, and inception date of the nine Vanguard index funds used as benchmarks in this paper, along with their size and style dimensions, which are listed in the row and column headings, respectively.

	Value	Blend	Growth
Large-Cap	Value Index Fund VIVAX 11/02/92	S&P 500 Index Fund VFINX 8/31/76	Growth Index Fund VIGRX 11/2/92
Mid-Cap	Mid-Cap Value Index Fund VMVIX 8/24/06	Extended Market Index Fund VEXMX 12/21/87	Mid-Cap Growth Index Fund VMGIX 8/24/06
Small-Cap	Small-Cap Value Index Fund VISVX 5/21/98	Small-Cap Index Fund NAESX 5/15/84	Small-Cap Growth Index Fund VISGX 5/21/98

Recent studies that predict mutual fund performance focus on measures of mutual fund "activeness", the degree to which active funds deviate from their benchmarks instead of "closet indexing". Said differently, activeness measures the quantity of stock picking a manager does, though Cremers (2017) notes that it does not directly measure stock-selection skill. The seminal study of Cremers and Petajisto (2009) measures fund activeness using the absolute deviation of fund portfolio weights from those of the corresponding benchmark. Amihud and Goyenko (2013) argue that activeness is more easily measured (inversely) using a fund's R^2 from regressions of that fund's returns on those of benchmark factors. The intuition is that funds that essentially mimic their benchmark, or "closet index", will exhibit near perfect correlations between their returns and those of the benchmark, thereby exhibiting an R^2 close to one. Following this intuition, we measure a fund's activeness as:

[10] Berk and van Binsbergen (2015) also include two international Vanguard funds, but we do not since we only consider the performance of U.S. equity funds.

$$Activeness = 1 - R^2 \tag{4}$$

where R^2 for a given fund-month comes from the regression that is used to assign that fund's best-fit Vanguard benchmark.[11]

A closely related notion to activeness is "conviction", that is, the willingness for managers to take and maintain significant positions in stocks they think are most likely to outperform. We define our proxy of this notion, *Conviction*, at the beginning of each month using the sum of the portfolio weights a manager places in their top ten stocks ranked by relative weight, $rw_{i,j,t}$, defined as:

$$rw_{i,j,t} = \left(w_{i,j,t}^{act} - w_{i,j,t}^{cap}\right), \tag{5}$$

where $w_{i,j,t}^{act}$ denotes the actual portfolio weight of stock i in fund j at the end of month t, and $w_{i,j,t}^{cap}$ denotes the corresponding market-capitalization-based portfolio weight.[12] Doshi et al. (2015) argue that the cap weights used in Eq. (5) are a reliable proxy for fund's benchmark returns for the purposes of capturing the notion of "relative weight". Antón et al. (2020), among others, show that the few stocks that receive the greatest weight in each fund tend to outperform stocks that get less weight, the implication being that mangers with high conviction invest most heavily in their "best ideas". Motivated by this finding, we use *Conviction* as a potential predictor of fund returns.

2.3. Summary statistics and strategy validation

Table 4 presents time-series means of the number of funds within each strategy and strategy-level month-by-month (equal-weighted) averages of fund-level statistics: assets under management (*AUM*), the number of different stocks held by a given fund *(#Stocks*), annualized return (*Ret*), expense ratio (*Exp ratio*), strategy consistency (*Consistency*), the R^2 statistic from a regression of the fund returns on those of its best-fit benchmark over the previous 36 months (used in Eq. (4)), and the sum of the portfolio weights in the fund's top-10 stocks by relative weight at the end of the previous month (*Conviction*). The bottom row, labeled 'All', presents corresponding statistics for all funds in our sample, those for which we can calculate *Consistency*. The sample period is January 2007 through December 2019. The 'All' row shows that, on average, there are 1,773 funds in our sample per month. The rows above show that, not surprisingly, the most common strategies are those based on valuation (555 funds on average), future growth (304 funds), and competitive position (563 funds), which has perhaps the broadest definition. Table 4 further shows that while there is variation across strategies in most characteristics, all strategies have average *Consistency* of approximately five.

Table 4: Summary statistics by strategy

This table presents time-series means of the number of funds within each strategy and strategy-level month-by-month averages of fund-level statistics: assets under management (*AUM*), the number of different stocks held by a given fund *(#Stocks*), annualized return (*Ret*), expense ratio (*Exp ratio*), strategy consistency (*Consistency*), the R^2 statistic from a regression of the fund returns on those of its best-fit benchmark over the previous 36 months (with a minimum of 12 observations), and the sum of the portfolio weights in the fund's top-10 stocks by relative weight at the end of the previous month (*Conviction*. See section 2.2.3 for details). The bottom row, 'All', presents corresponding statistics for all funds in our sample, those for which we can calculate *Consistency*. The sample period is January 2007 through December 2019.

[11] Amihud and Goyenko use the R^2 from the regression of fund returns on the Fama-French-Carhart four-factor model although the intuition is based on how closely a fund tracks its benchmark. We form R^2 relative to the best-fit benchmark following the intuition, but untabulated results show that results in this paper are robust to using the R^2 based on the Fama-French-Carhart model.

[12] For calculating the cap relative weight, we employ the Antón et. Al. (2020) approach of dividing a stock's market capitalization by the sum of the market capitalization of all stocks held in the fund's portfolio.

Strategy	#Funds	AUM	#Stocks	Ret (%)	Exp ratio (%)	Consistency	R^2	Conviction (%)
Competitive Position	563	2,601.8	113	8.11	1.29	5.01	0.87	17.58
Economic Conditions	58	826.7	118	6.86	1.41	4.96	0.81	16.30
Future Growth	304	1,772.9	122	9.03	1.30	5.00	0.91	15.72
Market Conditions	14	307.6	252	7.06	1.43	4.91	0.89	8.54
Opportunity	68	751.8	135	4.32	1.60	5.13	0.77	18.79
Profitability	40	1,544.9	105	7.14	1.24	4.97	0.87	18.45
Quantitative	83	736.3	202	6.91	1.20	5.13	0.90	12.53
Risk	35	684.9	153	4.99	1.51	5.04	0.83	13.42
Social Considerations	53	526.2	168	7.99	1.27	5.02	0.89	18.76
Valuation	555	1,491.8	130	6.61	1.27	5.06	0.89	16.63
All	1773	1,756.4	128	7.44	1.30	5.03	0.88	16.61

As noted by Brown and Goetzmann (1997), no strategy categorization scheme is perfect. However, they can be useful if they result in high within-strategy correlations, which in turn indicates that a given scheme defines useful fund peer groups for performance evaluation. Thus, if our strategy definitions are well defined, then we should expect to see relatively strong correlations between the returns on funds within each strategy rather than across strategies, at least controlling for common exposure to the market.[13] Table 5 presents average pairwise correlations of market-adjusted fund returns within and across strategies. Specifically, for each fund, i, and month, t, we estimate rolling 36-month market adjusted returns, $\hat{\epsilon}_{i,\tau,t}, \tau = t - 35, \dots, t$, as the residuals from a regression of the form:

$$rx_{i\tau} = \alpha_{\mathrm{it}} + \beta_{it} MKT_\tau + \epsilon_{i,\tau,t}, \qquad \tau = t - 35, \dots, t \tag{6}$$

where $rx_{i\tau}$ denotes i's return excess return over the one-month Treasury bill in month τ and MKT_τ denotes the corresponding excess return on the CRSP value-weighted index. For every pair of funds i, j in month t with at least 12 observations, we estimate the correlation coefficient between the returns, $\hat{\rho}_{i,j,t} = \mathrm{corr}(\hat{\epsilon}_{i,\tau,t}, \hat{\epsilon}_{j,\tau,t}),\ \tau = t - 35, \dots, t$. For each pair of fund strategies, we then form equal-weighted averages of the correlations within the strategy pair each month t. The array consisting of the first ten rows of Table 5 present the time-series average of the resulting equal-weighted average correlations for each strategy pair, defined by the row and column headings, over all months in the sample. The untabulated average of all of cells is 0.055. The row beneath the array of average correlations contains the average of the off-diagonal entries from the column above and the bottom row contains the ratio of the diagonal entry to the corresponding off-diagonal average.

Table 5: Average 36-month rolling correlations of market-adjusted mutual fund returns within and across strategy groups

For each fund i and month t, we estimate rolling 36-month market adjusted returns, $\hat{\epsilon}_{i,\tau,t}, \tau = t - 35, \dots, t$, as the residuals from a regression of the form: $rx_{i\tau} = \alpha_{\mathrm{it}} + \beta_{it} MKT_\tau + \epsilon_{i,\tau,t}$ $(\tau = t - 35, \dots, t)$, where $rx_{i\tau}$ denotes i's return in excess of the one-month Treasury bill in month τ and MKT_τ denotes the corresponding excess return on the CRSP value-weighted index. For every pair of funds i, j in month t with at least 12 observations, we estimate the correlation coefficient between the returns, $\hat{\rho}_{i,j,t} = \mathrm{corr}(\hat{\epsilon}_{i,\tau,t}, \hat{\epsilon}_{j,\tau,t}), \tau = t - 35, \dots, t$. For each pair of fund strategies, we then form equal-weighted averages of the correlations within the strategy pair each month t. The array consisting of the first ten rows of this table present the time-series average of the resulting average correlations for

[13] Two managers following very different strategies could have high correlations purely because they have very similar market betas, and vice versa. Hence, we remove the common variation driven by the market to try to isolate excess correlation driven by strategy. However, it is important to note that managers picking similar stocks will also have similar betas so correlations in market-adjusted returns will understate own-strategy correlation.

each strategy pair, defined by the row and column headings, over all months in the sample. The untabulated average of all these cells is 0.055. The row beneath the array of average correlations contains the average of the off-diagonal entries from the column above and the bottom row contains the ratio of the diagonal entry to the corresponding off-diagonal average.

	Competitive Position	Economic Conditions	Future Growth	Market Conditions	Opportunity	Profitability	Quantitative	Risk	Social Considerations	Valuation
Competitive Position	0.083	0.068	0.126	0.089	0.042	0.000	0.052	0.004	0.073	0.048
Economic Conditions	0.068	0.071	0.098	0.069	0.043	0.023	0.057	0.020	0.064	0.044
Future Growth	0.126	0.098	0.249	0.138	0.021	-0.049	0.067	-0.003	0.094	-0.003
Market Conditions	0.089	0.069	0.138	0.161	0.051	0.003	0.106	0.023	0.074	0.060
Opportunity	0.042	0.043	0.021	0.051	0.070	0.024	0.054	0.003	0.047	0.085
Profitability	0.000	0.023	-0.049	0.003	0.024	0.111	0.038	0.061	0.045	0.074
Quantitative	0.052	0.057	0.067	0.106	0.054	0.038	0.092	0.037	0.058	0.061
Risk	0.004	0.020	-0.003	0.023	0.003	0.061	0.037	0.060	0.022	0.015
Social Considerations	0.073	0.064	0.094	0.074	0.047	0.045	0.058	0.022	0.096	0.081
Valuation	0.048	0.044	-0.003	0.060	0.085	0.074	0.061	0.015	0.081	0.142
Average off diagonal	0.056	0.054	0.054	0.068	0.041	0.024	0.059	0.020	0.062	0.052
On/off diagonal	1.482	1.317	4.598	2.370	1.695	4.576	1.563	2.992	1.548	2.750

Inspection of the bottom row of Table 5 shows that, on average, the correlations between market-adjusted returns of a given fund are higher with those of other funds in the same strategy than they are with those of funds in other strategies by about 32% (Economic Conditions) to 360% (Future Growth). This finding helps validate the strategy classification scheme and, like any significant source of commonality in returns, also suggests that investors would likely benefit from evaluating fund performance relative to other same-strategy funds.

Table 6 summarizes fund characteristics by Consistency portfolio. Panel A presents time-series means of the number of funds *(#Funds*) within each portfolio and strategy-level month-by-month averages fund statistics: assets under management (*AUM*), the number of distinct stocks held by the fund *(#Stocks*), the expense ratio (*Exp ratio*), the R^2 statistic from a regression of the fund returns on those of its best-fit benchmark over the previous 36 months (minimum of 12), and the sum of the portfolio weights in the fund's top-10 stocks by relative weight at the end of the previous month (*Conviction*). Panel B of this table presents time-series averages of slopes and R^2 statistics from monthly cross-sectional regressions of Consistency on the fund-level characteristics defined by the column headings. $Log(AUM)$ denotes the natural logarithm of *AUM*. On average, both panels of Table 6 show that high-consistency funds tend to be smaller in terms of assets under management than low-consistency funds. They also tend to have relatively high expense ratios, activeness, portfolio concentration, and *Conviction*. However, panel B shows these characteristics combined explain only a small portion, 6.3% on average, of the cross-sectional variation in consistency. The extreme consistency groups consist of a small number of funds; just over 3% of funds are in the high-consistency (5) portfolio on average. Importantly, this paucity is a prerequisite for predicting the miniscule percentage of funds that are expected to earn superior returns, which Barras, Scaillet, and Wermers (2010) and Fama and French (2010) estimate to be about 1% or 2% of funds.

Table 6: Correlates of mutual fund strategy consistency

Panel A of this table presents time-series means of the number of funds (*#Funds*) within each *Consistency* portfolio and strategy-level month-by-month averages fund statistics: assets under management (*AUM*), the number of distinct stocks held by the fund (*#Stocks*), the expense ratio (*Exp ratio)*, the R^2 statistic from a regression of the fund returns on those of its best-fit benchmark over the previous 36 months (minimum of 12), and the sum of the portfolio weights in the fund's top-10 stocks by relative weight at the end of the previous month (*Conviction*. see section 2.2.3 for details). Panel B of this table presents time-series averages of slopes and R^2 statistics from monthly cross-sectional regressions of *Consistency* on the fund-level characteristics defined by the column headings. $Log(AUM)$ denotes the natural logarithm of *AUM*. Time-series *t* statistics are below the corresponding slopes in parentheses. The sample period is January 2007 through December 2019.

Panel A: Summary statistics by *Consistency* portfolio						
Portfolio	*#Funds*	*AUM* ($millions)	*#Stocks*	*Exp ratio* (%)	R^2	*Conviction (%)*
1	74	2,338.4	131	1.33	0.82	18.0
2	629	2,479.2	175	1.27	0.88	15.2
3	769	1,537.8	114	1.28	0.89	16.2
4	247	688.0	69	1.38	0.87	19.3
5	54	287.6	53	1.54	0.78	24.5

Panel B: Average slopes from cross-sectional regressions of *Consistency* on fund characteristics						
$Log(AUM)$	$Log(AUM)^2$	$\#Stocks$	$Expense\ ratio$	$Activeness$	$Conviction$	R^2
0.078	-0.017	-0.001	0.098	-0.401	0.014	0.063
(8.88)	(-22.01)	(-46.43)	(11.24)	(-10.45)	(28.83)	

3. Main results

3.1 Consistency and fund performance

We examine whether *Consistency* predicts fund performance relative to two benchmarks. The first benchmark is the best-fit Vanguard fund defined in Section 2. The second benchmark is that produced by the four-factor model of Cremers, Petajisto, and Zitzewitz (2013), denoted CPZ:

$$r_{it} = \alpha_i + \beta_i(S5_t) + s_i(R2_t - S5_t) + v_i(R3V_t - R3G_t) + m_i MOM_t + \epsilon_{it}, \tag{7}$$

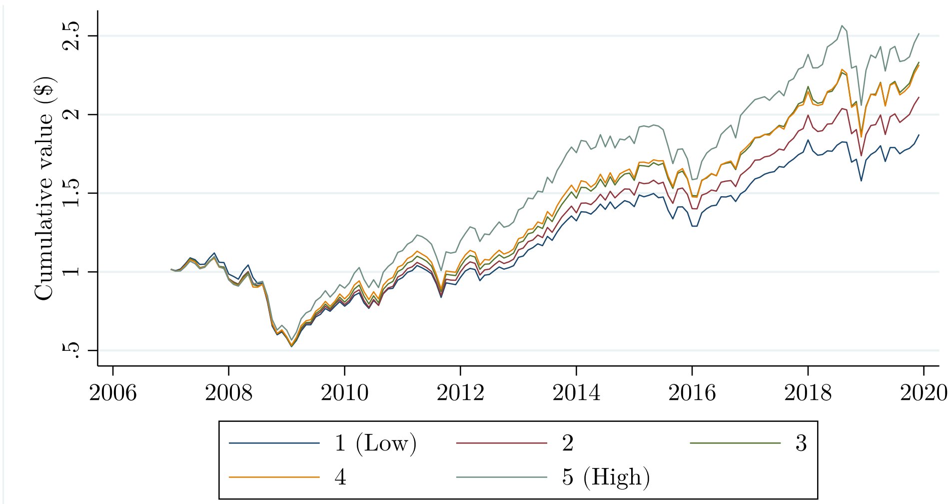
where r_{it}, is the return on fund i in excess of the one-month risk-free rate; $S5_t$ is the excess return on the S&P 500 index; $R2_t - S5_t$ is the return on the Russell 2000 index minus that of the S&P 500; $R3V_t - R3G_t$ is the return on the Russell 3000 Value index minus that of the Russell 3000 Growth index; and MOM_t denotes the return on the Fama and French (1993) and Carhart (1997) momentum factor.[14] These factors are chosen to be based on similar characteristics as the Fama-French-Carhart model while eliminating spurious alphas earned by passive benchmark indices relative to the latter model whose factors are not easily investable.

[14] Cremers et al (2013) also propose a seven-factor model with mid-cap factors and separate value-minus-growth factors for the three size groups but find the four-factor model has lower tracking-error volatility when using monthly data as opposed to daily data. Untabulated results show that all inferences in this paper are unchanged using the Fama-French-Carhart model instead of the Cremers et al. (2013) model. Huij and Verbeek (2009) also show that using factors that ignore implementation costs, such as those in the Fama-French-Carhart model, biases fund performance results.

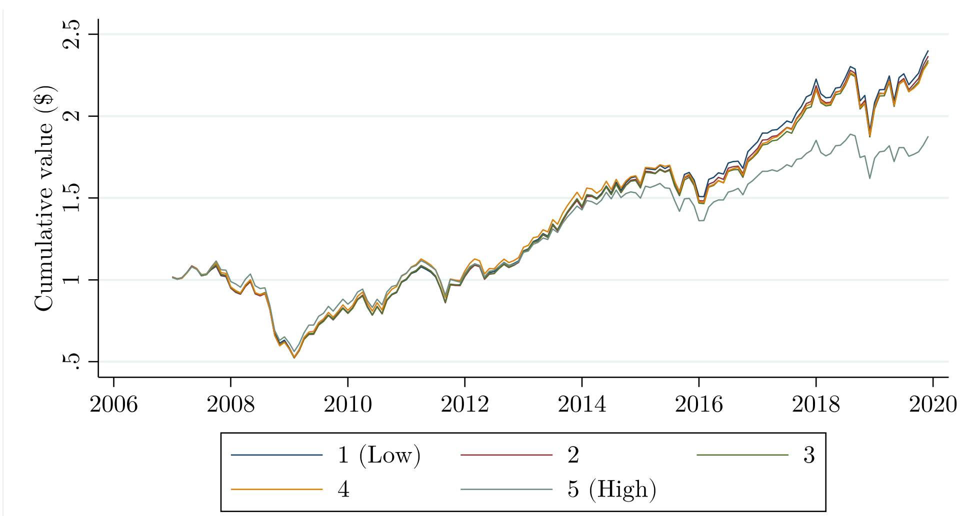
Panel A of Figure 1 depicts cumulative returns on the five equal-weighted portfolios of funds by consistency group. This panel shows that high-*Consistency* funds outperform low-*Consistency* funds in terms of cumulative returns by about 40% over our 13-year sample period (2007 to 2019). Conversely, Panels B and C, which show similar returns for portfolios formed on *Activeness* and *Conviction*, show that high-*Activeness* and high-*Conviction* funds underperform low-*Activeness* and low-*Conviction* funds over this period, which reflects a common practitioner view that active mutual funds categorically underperformed the market over the 2010s.

Figure 1: Cumulative returns of $1 invested in quintile portfolios based on *Consistency* and *Activeness*

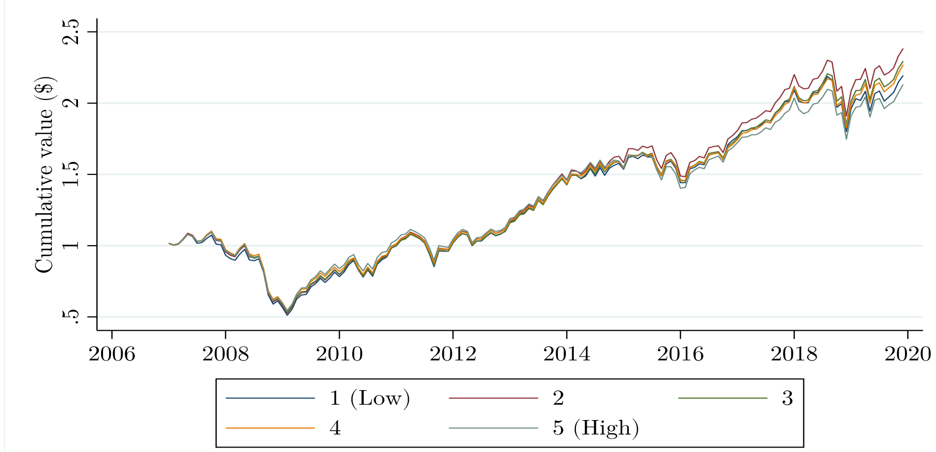
Panel A shows the cumulative value of a one-dollar initial investment in each of the five equal-weighted portfolios of mutual funds based on *Consistency*. Panels B and C are similar, but use portfolios based on *Activeness* or *Conviction*, respectively, instead of *Consistency*. The sample period for both panels is January 2007 through December 2019.



Panel A: *Consistency* portfolios



Panel B: *Activeness* portfolios



Panel C: *Conviction* Portfolios

Table 7 presents time-series averages of slopes and R^2 statistics from month-by-month cross-sectional regressions of benchmark-adjusted returns (Panel A) and CPZ alphas (Panel B) on *Consistency* and the other variables from Table 2. For each fund, i, and month, T, the estimated benchmark-adjusted return used as the dependent variable, $\hat{\alpha}_{i,T}$, is given by:

$$\hat{\alpha}_{i,T} = r_{it} - \left(\hat{\beta}_i(S5_t) + \hat{\mathrm{s}}_{\mathrm{i}}(\mathrm{R2_t} - \mathrm{S5_t}) + \hat{\mathrm{v}}_i(\mathrm{R3V_t} - R3G_t) + \hat{\mathrm{m}}_{\mathrm{i}} MOM_t\right), \tag{8}$$

where the slopes are estimated via Eq. (7) over the same preceding 36 months ($T - 36, \ldots, T - 1$) as those used to identify the best-fit Vanguard benchmarks and rolling correlations in Section 2.

Table 7: Fama-Macbeth regressions of risk-adjusted returns on fund-characteristics

This table presents time-series averages of slopes and R^2 statistics from monthly cross-sectional regressions of benchmark-adjusted fund returns (Panel A) and alphas (Panel B) on the fund-level characteristics specified by the row headings and defined in Table 1. Time-series t statistics are below the corresponding slopes in parentheses. The sample period is January 2007 through December 2019. On average, there are 1,632 funds per month.

Panel A: Benchmark-adjusted alpha							
$Consistency$	1.814				1.781	1.871	1.904
	(2.34)				(2.26)	(2.41)	(2.45)
$Activeness$		-63.157			-64.578	-63.545	-56.143
		(-2.19)			(-2.10)	(-2.07)	(-1.84)
$Conviction$			0.044		0.082	0.081	0.088
			(0.32)		(0.50)	(0.48)	(0.52)
$Security\ Count$				-0.003		-0.002	-0.004
				(-0.95)		(-0.67)	(-1.25)
$Log(AUM)$						1.136	0.007
						(0.87)	(0.01)
$Log(AUM)^2$						-0.031	0.003
						(-0.29)	(0.03)
$Expense\ Ratio$							-9.329
							(-5.02)
R^2	0.007	0.050	0.006	0.002	0.064	0.070	0.073

Panel B: Cremers-Petajisto-Zitzewitz four-factor alpha							
$Consistency$	1.709				1.845	1.893	1.920
	(2.47)				(2.62)	(2.78)	(2.79)
$Activeness$		0.553			5.190	5.275	12.413
		(0.03)			(0.28)	(0.28)	(0.68)
$Conviction$			-0.125		-0.212	-0.192	-0.190
			(-1.33)		(-1.83)	(-1.66)	(-1.65)
$Security\ Count$				0.004		0.004	0.002
				(1.59)		(1.40)	(0.83)
$Log(AUM)$						1.493	0.512
						(1.33)	(0.44)
$Log(AUM)^2$						-0.127	-0.098
						(-1.39)	(-1.07)
$Expense\ Ratio$							-8.448
							(-4.05)
R^2	0.005	0.022	0.004	0.001	0.031	0.036	0.040

Panels A and B both show that *Consistency* significantly and positively predicts mutual fund returns in the cross-section by itself and controlling for other predictors. However, in contrast to the prior literature on fund activeness whose sample period is largely prior to our own, Table 7 shows *Activeness* and *Conviction* do not reliably predict returns over this sample, and they do so negatively if anything. This result has a few possible interpretations. First, our post-2006 sample period is largely after those used to identify the importance of activeness (e.g., the sample period of Cremers et al. 2009 was 1980 to 2003; that of Amihud and Goyenko 2013 was 1988 to 2010). Thus, it is possible that *Activeness* was a spurious predictor that will not work out-of-sample, much like many predictors of the cross-section of stock returns are thought by some to be artifacts of data snooping (e.g., Harvey et al, 2016). This possibility is also consistent with Jones and Mo (2021) who find that most mutual fund predictors fail out-ot-sample. A second related interpretation is that prior findings were not data snooping, but investors directed capital to funds with high *Activeness* after these measures were published until they eliminated the ability of highly active managers to continue delivering superiors returns. This interpretation is analogous to post-publication decline in returns on anomalies in the stock market documented by McLean and Pontiff (2016) and like the Berk and Green (2004) effect in which competition among investors' capital flows eliminates any expected abnormal returns after managerial fees. A third interpretation is that our sample is unique in the sense that it was an unusually bad time for active mutual fund managers, and especially active fund managers were more adversely affected as a result. Untabulated results show that, during this time, not only were average benchmark-adjusted returns of active funds negative, consistent with historical data, but so was the aggregate dollar-value added by fund managers, in contrast to the longer sample of Berk and van Binsbergen (2015).

Table 8 expands on the analysis in Table 7 by examining the performance of portfolios of funds sorted on *Consistency* and the other fund predictors. Specifically, each month, we sort funds into five equal-weighted portfolios, denoted 1 (Low) to 5 (High), based on the previous month's $Consistency$ (Panel A), $Activeness$ (Panel B), or $Conviction$ (Panel C).[15] For each of these portfolios and the hypothetical long-short portfolio whose return is that of portfolio 5 minus that of portfolio 1, Table 8 presents average excess returns, Sharpe ratios, benchmark-adjusted

[15] $Consistency$ takes discrete values 1 through 10, so we define consistency portfolio 1 (Low) consists of funds with consistency of 1 or 2 and portfolio 5 (High) consists of funds with consistency of 9 or 10. Portfolios 2 through 4 are defined similarly. $Activeness$ and $Conviction$ are continuous, so for these variables, we define portfolio 1 (Low) through 5 (High) by simple quintiles.

returns and CPZ alphas estimated using Eq. (7) over the full sample.[16] In addition to reporting CPZ alphas using the standard net-of-expenses returns, we also report CPZ alphas using before-expenses returns (given by the fund returns plus the expense ratio).

Table 8: Performance of fund portfolios based on *Consistency*, *Activeness*, and *Conviction*

Each month, we sort mutual funds into five equal-weighted portfolios, denoted 1 (Low) to 5 (High), based on *Consistency* (Panel A), *Activeness* (Panel B), or *Conviction* (Panel C). Each column presents performance statistics for one of these portfolios or the hypothetical long-short portfolio that is long 5 and short 1. The performance statistics are the average excess return over the one-month treasury bill (Excess return), Sharpe ratio, average return in excess of the best-fit benchmark return (Benchmark alpha), and the alpha from the four-factor model of Cremers, Petajisto, and Zitzewitz (2013), 'CPZ alpha', and the CPZ alpha using fund returns before fund expenses are removed, 'CPZ alpha (gross)'. The sample is January 2007 through December 2019.

Panel A: *Consistency* portfolios						
	1 (Low)	2	3	4	5 (High)	5-1
Excess return	5.06	6.03	6.94	6.92	7.54	2.47
	(1.24)	(1.45)	(1.57)	(1.54)	(1.69)	(1.98)
Sharpe ratio	0.44	0.49	0.52	0.51	0.57	0.59
Benchmark alpha	-2.75	-1.81	-1.10	-1.14	-0.83	1.92
	(-3.89)	(-4.62)	(-3.42)	(-2.54)	(-1.01)	(2.29)
CPZ alpha	-3.54	-2.48	-1.76	-1.50	-0.18	3.36
	(-3.28)	(-3.88)	(-3.68)	(-3.31)	(-0.25)	(3.45)
CPZ alpha (gross)	-2.08	-1.19	-0.48	-0.09	1.52	3.59
	(-1.93)	(-1.87)	(-1.00)	(-0.20)	(2.12)	(3.69)
Panel B: *Activeness* portfolios						
	1 (Low)	2	3	4	5 (High)	5-1
Excess return	7.14	7.06	6.96	6.94	5.02	-2.12
	(1.63)	(1.59)	(1.58)	(1.56)	(1.26)	(-2.70)
Sharpe ratio	0.53	0.52	0.52	0.52	0.46	-0.62
Benchmark alpha	-0.92	-1.01	-1.18	-1.10	-2.89	-1.97
	(-3.98)	(-3.73)	(-3.36)	(-2.66)	(-3.30)	(-2.31)
CPZ alpha	-1.43	-1.67	-1.83	-1.95	-2.88	-1.44
	(-3.66)	(-3.45)	(-3.59)	(-3.23)	(-3.17)	(-2.09)
CPZ alpha (gross)	-0.26	-0.44	-0.53	-0.58	-1.34	-1.08
	(-0.66)	(-0.92)	(-1.05)	(-0.96)	(-1.47)	(-1.55)
Panel C: *Conviction* portfolios						
	1 (Low)	2	3	4	5 (High)	5-1
Excess return	6.49	7.11	6.74	6.61	6.15	-0.34
	(1.46)	(1.61)	(1.57)	(1.57)	(1.44)	(-0.57)
Sharpe ratio	0.48	0.53	0.52	0.52	0.50	-0.05
Benchmark alpha	-1.48	-0.91	-1.29	-1.28	-1.76	-0.27
	(-4.58)	(-2.80)	(-3.49)	(-3.20)	(-3.31)	(-0.59)
CPZ alpha	-1.77	-1.69	-1.98	-2.04	-2.36	-0.59
	(-3.36)	(-3.36)	(-3.57)	(-3.75)	(-3.76)	(-1.56)
CPZ alpha (gross)	-0.50	-0.42	-0.73	-0.72	-0.85	-0.35
	(-0.94)	(-0.84)	(-1.31)	(-1.31)	(-1.36)	(-0.98)

[16] Results are similar if instead of estimating Eq. (7) over the whole sample for the five portfolios, we instead use the time-series average of the monthly portfolio-level average fund alphas estimated via Eq. (8).

The results in Table 8 are consistent with those in Table 7. Panel A shows that high-*consistency* funds significantly outperform low-*consistency* funds by 1.9% to 3.6% per annum using both raw, and abnormal returns relative to the Vanguard and CPZ benchmarks. However, high-*consistency* funds still earn statistically zero abnormal returns net of fees captured by the expense ratio. This does not mean they lack skill, however. Berk and Green (2004) argue that investors will invest in the funds of skilled managers until their fees exactly offset their before-costs alpha. Indeed, high-*consistency* funds earn significantly positive abnormal returns before expense ratios (though after trading costs), evidence of skill among these funds. In fact, these are the only funds in the entire Table to accomplish this feat. Panel B shows that using the same metrics, highly active managers underperform their low-activeness counterparts by about 2% per year. Panel C confirms the null relationship over this time between conviction and performance.

Following Cremers and Petajisto (2009) and Amihud and Goyenko (2013), Panel A of Table 9 examines the performance, measured by CPZ alphas, of portfolios formed by double sorting funds into five groups based on *Consistency*, and then, within each consistency group, into five subgroups based on lagged alpha. This double sorting accounts for the fact that fund performance exhibits significant persistence. If some managers have skill to beat their benchmarks, we expect persistence in their performance. Based on the reasoning that motivates *Consistency* as a performance predictor, we would also expect high-*Consistency* funds to perform especially well following strong performance based on the following intuition. High *Consistency* indicates that a given fund manager is coming to the same conclusions as other managers following a similar strategy, presumably validating these conclusions ex ante. Observing high *Consistency* and strong past performance improves on this predictive content by evincing that a manager is not only coming to the same conclusions as the other same-strategy managers, but also that these managers were successful in the past.

Panel A shows that relative performance persistence is significant in all *Consistency* groups, with the high-prior-month-alpha funds outperforming low-alpha funds by 4.6% per year in both high- and low-*Consistency* groups. However, only the high-*Consistency* funds exhibit persistence in superior performance, with high-*Consistency* funds with high prior-month alphas earning significant positive post-ranking alphas of 4.2% per year. Said differently, in our sample period, (high) *Consistency* helps investors identify the "needle in the haystack", that is, the small minority of mutual funds that generate positive abnormal returns.

Panels B and C repeat the same exercise as Panel A, but using *Activeness* and *Conviction*, respectively, in lieu of *Consistency*. Like Amihud and Goyenko (2013), Panel B shows persistence in relative performance increasing with *Activeness*. Funds with high prior-month alphas have 1.8% per year higher post-ranking alphas than funds with low alphas conditional on having low activeness (high R^2); and this spread increases to 7.0% for funds with high-activeness. However, in our sample, we do not find a significantly positive alpha on funds with high activeness and high prior-month alpha. Panel C shows qualitatively similar results for conviction as Panel B shows for activeness.

Overall, the evidence in Table 9 shows that, unlike *Activeness* and *Conviction*, *Consistency* does not significantly affect persistence in relative performance. However, high-Consistency funds perform better on average than low-*Consistency* funds so that high-*Consistency* funds with high past alphas continue to earn significantly positive alphas for investors, even net of costs.

Table 9: Four-factor alphas of double-sorted fund portfolios based on alpha and *Consistency*, *Activeness*, or *Conviction*

Each month, we sort mutual funds into five equal-weighted portfolios, denoted 1 (Low) to 5 (High), based on *Consistency* (Panel A), *Activeness* (Panel B), or *Conviction* (Panel C). Then, within each portfolio, we further sort funds into five equal-weighted sub-portfolios based on prior-month alpha with respect to the four-factor model of Cremers, Petajisto, and Zitzewitz (2013). This table presents four-factor alphas of each of the resulting twenty-five portfolios. The sample is January 2007 through December 2019.

Panel A: Portfolios formed on *Consistency* and alpha							
		Ranking on Alpha					
		1 (Low)	2	3	4	5 (High)	5-1
Ranking on *Consistency*	1 (Low)	-5.86	-3.14	-2.63	-0.01	-1.24	4.62
		(-2.85)	(-2.55)	(-2.52)	(-0.01)	(-0.91)	(2.29)
	2	-3.96	-1.75	-0.85	-0.81	-0.41	3.56
		(-3.42)	(-2.42)	(-1.66)	(-1.68)	(-0.48)	(2.86)
	3	-2.96	-1.11	-0.30	-0.23	0.23	3.19
		(-3.49)	(-2.07)	(-0.67)	(-0.50)	(0.35)	(3.26)
	4	-2.54	-1.31	0.10	0.87	-0.16	2.38
		(-2.93)	(-2.15)	(0.18)	(1.55)	(-0.21)	(1.98)
	5 (High)	-0.44	-1.30	-0.56	1.37	4.16	4.60
		(-0.28)	(-1.18)	(-0.57)	(1.29)	(2.47)	(2.10)
	5-1	5.42	1.84	2.07	1.39	5.39	-0.02
		(2.68)	(1.32)	(1.76)	(1.17)	(2.65)	(-0.01)

Panel B: Portfolios formed on *Activeness* and alpha							
		Ranking on Alpha					
		1 (Low)	2	3	4	5 (High)	5-1
Ranking on *Activeness*	1 (Low)	-1.96	-0.58	-0.36	0.09	-0.16	1.80
		(-2.95)	(-1.23)	(-0.91)	(0.19)	(-0.28)	(2.19)
	2	-2.29	-1.15	-0.74	-0.23	0.27	2.56
		(-2.93)	(-2.13)	(-1.60)	(-0.53)	(0.45)	(3.23)
	3	-2.72	-1.04	-0.50	-0.36	-0.20	2.52
		(-3.00)	(-1.75)	(-1.03)	(-0.79)	(-0.31)	(2.63)
	4	-2.67	-1.38	-1.13	-0.26	-0.06	2.61
		(-2.76)	(-1.97)	(-1.67)	(-0.44)	(-0.08)	(2.37)
	5 (High)	-5.69	-3.68	-1.19	-0.77	1.28	6.97
		(-3.47)	(-2.98)	(-1.19)	(-0.77)	(0.82)	(2.97)
	5-1	-3.73	-3.10	-0.82	-0.86	1.44	5.17
		(-2.77)	(-2.86)	(-0.97)	(-0.92)	(0.98)	(2.36)

Table 9: (Continued)

Panel C: Portfolios formed on *Conviction* and alpha							
		Ranking on Alpha					
		1 (Low)	2	3	4	5 (High)	5-1
Ranking on *Conviction*	1 (Low)	-2.52	-0.81	-0.71	-0.04	-0.21	2.31
		(-2.72)	(-1.38)	(-1.31)	(-0.09)	(-0.33)	(2.28)
	2	-2.27	-0.85	-0.20	-0.30	-0.26	2.01
		(-2.63)	(-1.38)	(-0.45)	(-0.66)	(-0.39)	(2.11)
	3	-3.41	-1.60	-0.67	-0.44	0.47	3.89
		(-3.15)	(-2.81)	(-1.29)	(-0.87)	(0.64)	(3.31)
	4	-3.44	-1.56	-0.44	-0.27	-0.11	3.33
		(-3.34)	(-2.28)	(-0.94)	(-0.60)	(-0.12)	(2.68)
	5 (High)	-4.11	-2.73	-0.75	-0.59	1.33	5.44
		(-3.45)	(-3.32)	(-1.34)	(-1.15)	(1.30)	(3.57)
	5-1	-1.59	-1.92	-0.05	-0.55	1.54	3.13
		(-1.82)	(-3.40)	(-0.11)	(-1.27)	(1.82)	(2.53)

4. Conclusion

In this paper, we propose a novel predictor of mutual fund returns called strategy consistency, the degree to which fund managers invest in similar stocks as the group of all managers following the same self-declared strategy. We find that, from 2007 to 2019, a period that was especially harsh for active managers, high-consistency funds significantly outperform low-consistency funds, and this performance is not subsumed by measures of mutual fund activeness and conviction. Overall, our results show that investors can use consistency to help identify the small minority of funds expected to deliver superior future performance.

About the authors

C. Thomas Howard is the co-founder, chief investment officer, and Director of Research at AthenaInvest. Building upon the Nobel Prize winning research of Daniel Kahneman, Howard is a pioneer in the application of behavioral finance for investment management. He is a professor emeritus at the Reiman School of Finance, Daniels College of Business, University of Denver, where he taught courses and published articles in the areas of investment management and international finance. He is the author of Behavioral Portfolio Management. Howard holds a BS in mechanical engineering from the University of Idaho, an MS in management science from Oregon State University, and a PhD in finance from the University of Washington.

Andrew L. Detzel an Associate Professor in the Department of Finance, Insurance and Real Estate in the Hankamer School of Business at Baylor University. He holds a PhD in Finance from University of Washington Seattle, a MS in Mathematics from the University of Oregon, and a BS in Mathematics from California State University. Detzel is a member of the American Finance Association and the Macro Finance Society. His research covers several areas in Asset Pricing, including macro-finance, return forecasting, capital market frictions, and volatility. Detzel's papers can be found in leading finance and business journals such as Journal of Finance, Journal of Financial Economics, Review of Finance, and Management Science.

References

Abis, S. and Anton Lines, 2022, Do mutual funds keep their promises? Columbia Working Paper.

Alexander, G. J., G. Cici, and S. Gibson, 2007, Does motivation matter when assessing trade performance? An analysis of mutual funds, *Review of Financial Studies* 20: 125–50.

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

Antón, M., Cohen, R., and C. Polk, 2020, Best Ideas, HBS Working Paper.

Armstrong, W. J., Genc, E., and M. Verbeek, 2019, Going for Gold: An Analysis of Morningstar Analyst Ratings, *Management Science* 65: 2310–2327.

Barras, L., Scaillet, O. and Wermers, R. (2010), False discoveries in mutual fund performance: measuring luck in estimated alphas, *Journal of Finance* 65: 179–216.

Ben-David, I., Li, J., Rossi, A., and Song, Y., 2020, Non-Fundamental Demand and Style Returns, Fisher College of Business Working Paper.

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

Berk, J. B. and J. H. van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118: 1–20.

Berk, J. B. and J. H. van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119: 1–23.

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

Brown, Stephen J. and W. N. Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics* 43: 373-399.

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

Chan, L. K. C., H. Chen, and J. Lakonishok, 2002, On mutual fund investment styles, *Review of Financial Studies* 15: 1407–1437.

Chan, L. K. C., S. G. Dimmock, and J. Lakonishok, 2009, Benchmarking money manager performance: issues and evidence, *Review of Financial Studies* 22: 4553–4599.

Chen, H-L., N. Jegadeesh, and R. 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: 343–68.

Cremers, M., 2017, Active Share and the Three Pillars of Active Management: Skill, Conviction, and Opportunity, *Financial Analysts Journal*, 73: 61–79.

Cremers, Martijn, and Ankur Pareek, 2016. Patient capital outperformance: The investment skill of high active share managers who trade infrequently, *Journal of Financial Economics* 122: 288–306.

Cremers, 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, M., A. Petajisto, and E. Zitzewitz, 2013, should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review* 2: 1–48.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52: 1035–58.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53: 1839–1885.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 2001, Overconfidence, arbitrage, and equilibrium asset pricing, Journal of Finance 56: 921–965

Del Guercio, D, and P.A. Tkac, 2008, Star power: The effect of Morningstar ratings on mutual fund flow, *Journal of Financial and Quantitative Analysis* 43: 907–936.

De Miguel, V., Gil-Bazo, J., Nogales, F., and A.A.P. Santos, 2021, Can machine learning help to select portfolios of mutual funds?, London Business School Working Paper.

Doshi, E., R. Elkamhi, and M. Simutin, 2015, Managerial Activeness and Mutual Fund Performance, *Review of Asset Pricing Studies* 5: 156–184.

Evans, R. B., 2010, Mutual fund incubation, *Journal of Finance* 65: 1581–1611.

Fama, E. F., and K. French, 1993, Common Risk Factors in the Returns on Bonds and Stocks, *Journal of Financial Economics* 33: 3–56.

Fama, E.F., K. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51: 55–84.

Fama, E. F., and K. French, 2010, Luck versus Skill in Mutual Fund Returns, *Journal of Finance* 61: 2551–2595.

Fama, E. F., and J. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81: 607–636.

Ferson, W. E., and R. W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51: 425–461.

Grinblatt, M., and M. Keloharju, 2001, Sensation seeking, overconfidence, and trading activity, *Journal of Finance* 64: 549–578.

Grinblatt, M., and M. Keloharju, 2009, What makes investors trade?, *Journal of Finance* 56: 589–616.

Grinblatt, M., and S. Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business* 62: 393–416.

Grinblatt, M., and S. Titman, 1993, Performance measurement without benchmarks: An examination of mutual fund returns, *Journal of Business* 66: 47–68.

Howard, C., 2010, The Importance of Investment Strategy: Working Paper

Huij, J. and Verbeek, M., 2009, On the use of multifactor models to evaluate mutual fund performance, *Financial Management* 38: 75–102.

Jones, C., and H. Mo, 2021, Out-of-Sample performance of mutual fund predictors, *Review of Financial Studies* 34: 149–193.

Kacperczyk, M., C. Sialm, and L. Zheng. 2005. On the industry concentration of actively managed equity mutual funds. *Journal of Finance* 60: 1983–2011.

Malkiel, B. G., 1995, Returns from Investing in Equity Mutual Funds, 1971–1991, *Journal of Finance* 50: 549–572.

McLean, R. D., and Pontiff, J. (2016), Does academic research destroy stock return predictability?, *Journal of Finance*, 71: 5-32.

Pástor, L. and R. F. Stambaugh 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111: 642–85.

Sirri, E.R., and P. Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53:1589–1622