Department of Finance
Faculty of Business and Economics

Working Paper Series

Investing in Mutual Funds: Exploiting the Cross-sectional Predictability in Fund Performance

Xin Gao and Federico Nardari

Working Paper No. 17/20
July 2020

Investing in Mutual Funds:

Exploiting the Cross-sectional Predictability in Fund Performance*

Xin Gao† Federico Nardari‡

July 21, 2020

*We thank Sean Anthonisz, Hank Bessembinder, Tom George, Kris Jacobs, Spencer Martin, Rauli Susmel, Takeshi Yamada, Stijn van Nieuwerburgh, Anna von Reibnitz, Russ Wermers and seminar participants at Australian National University, the University of Houston, Lehigh University, Massey University, UMass Boston, University of Sydney and participants at the 2016 FMA meeting (Atlanta, GA). All remaining errors are the authors' responsibility.

†Assistant Professor of Finance, Jack Welch College of Business, Sacred Heart University; 3135 Easton Turnpike, Fairfield, CT; gaox3@sacredheart.edu; (203) 371-7959.

‡Corresponding author; Professor of Finance, Faculty of Business and Economics, The University of Melbourne; 198 Berkeley St, Level 11, Victoria 3010, Australia; federico.nardari@unimelb.edu.au; +61 (0)3 9035 4133.

Investing in Mutual Funds:
Exploiting the Cross-sectional Predictability in Fund Performance

Abstract

We investigate the economic value generated by U.S. active equity mutual funds from the perspective of a utility-maximizing investor. The proposed optimal portfolio strategy, which *jointly* exploits the information conveyed by fund characteristics and by macroeconomic indicators about future fund performance, outperforms commonly used passive benchmarks as well as active strategies that exploit fund characteristics but without relying on the principles of portfolio theory. Our findings indicate that investing in active mutual funds can be value adding if investors: a) account for the predictability in performance originating from fund-level information; b) adopt an optimal portfolio approach, as opposed to simpler strategies based on sorting or equal-weighting.

JEL classification: G11; G12

Keywords: Mutual funds; Portfolio allocation; Fund characteristics; Time-varying predictability; Investor Utility.

1 Introduction

With $18 trillion in assets under management, mutual funds serve as a significant component of U.S. investors' holdings. Among all types of funds, about 3,200 actively managed domestic equity funds with over $5 trillion in total net assets constitute the largest portion of the average U.S. investor's investment portfolio at year-end of 2017.[1] Given such a large cross-section of funds, how should investors effectively invest in the universe of active equity mutual funds, or should they at all?

Despite numerous efforts, the literature has not been able to provide a definitive answer to this question. Early studies, e.g. Jensen (1968), Malkiel (1995), Gruber (1996) and Carhart (1997), find that active equity mutual funds, on average, persistently underperform the aggregate stock market and other passive benchmarks after fees and expenses, suggesting that these funds possess little skill and, hence, that investors would be better off allocating to passively managed vehicles. More recent work, such as Fama and French (2010) and Barras, Scaillet, and Wermers (2010), comes to similar conclusions. On the other hand, some studies report that a small proportion of fund managers does exhibit superior skill in. e.g., stock picking, market timing, volatility timing or earning forecasting.[2] This evidence suggests that investors may try to invest with those managers who were identified as skilled *ex-ante*. Nevertheless, Berk and Van Binsbergen (2015) argue that because of competition in capital markets, managerial skill does not translate into net gains to investors. Similarly, Pástor, Stambaugh, and Taylor (2015) show that, although the active management industry has become more skilled over time, investors do not benefit from such improvement due to increasing competition as the industry size grows. Finally, severe issues plague the reliance on commonly used risk-adjusted measures (alphas) for measuring investors gains. [3]

[1] See Investment Company Fact Book 2018 at `http://www.icifactbook.org`.

[2] See, among others, Grinblatt and Titman (1989); Busse (1999); Bollen and Busse (2005); Kosowski, Timmermann, Wermers, and White (2006); Jiang, Yao, and Yu (2007); Kacperczyk, Van Nieuwerburgh, and Veldkamp (2013); Jiang, Verbeek, and Wang (2014).

[3] One set of issues with alpha pertains to the choice of the appropriate passive benchmark: see, e.g., Cremers, Fulkerson, and Riley (2019) and the comprehensive literature review therein, Berk and Van Binsbergen (2015), Pástor et al. (2015) and Cremers, Petajisto, and Zitzewitz (2012). For a discussion of additional in-

The above overview of the literature suggests that: a) neither buying the average mutual fund nor investing with *ex-post* skilled fund managers would be an effective solution for investors who seek to select a portfolio of mutual funds; b) alpha, whether from a factor model, a passive market index or a self-declared benchmark, is a rather problematic measure of investors benefits. In this paper, we reexamine the choice between active and passive funds but take the point of view of an investor who aims to maximize her utility within a portfolio choice framework. Novel to the literature, we construct optimal portfolio strategies that exploit the cross-sectional predictability in fund performance induced by fund-level characteristics and by macroeconomic information, and evaluate their performance out-of-sample and net of fees and expenses using a large cross-section of active US equity funds over the period 1996:01–2015:12. [4]

A number of papers find that fund-level information, such as past performance, fund size, recent cash inflows and the degree of activeness in management, can help predict the future relative performance of funds.[5] These studies further show that the predictive power of each individual characteristic could be exploited using a sorting rule to build a long-short portfolio that produces a positive benchmark-adjusted return, suggesting that such fund-specific information is economically valuable. However, the profitability of the hypothetical long-short portfolio may not truly reflect the economic gains that the characteristics could generate for an investor, because short-selling a mutual fund is very hard, if not impossible, in practice. One could, then, eliminate short positions and follow a long-only strategy that recursively selects a small group of top-ranked funds based on a particular characteristic.

ferential problems see, among many others, Grønborg, Lunde, Timmermann, and Wermers (2020). Another source of concern comes from the work of Back, Crane, and Crotty (2018) as they point out that active funds, while seeking higher alpha, are prone to generating negative (co)skewness, which may be undesirable for investors in terms of utility.

[4] As discussed below, only a handful of papers cast the choice among active mutual funds within an optimal portfolio problem. But none of these studies relies on fund-level information to characterize predictability in performance and, hence, to assess the impact of such predictability on utility-maximizing portfolios of funds.

[5] Related studies include Hendricks, Patel, and Zeckhauser (1993), Goetzmann and Ibbotson (1994), Chen, Hong, Huang, and Kubik (2004), Gruber (1996), Zheng (1999), Cremers and Petajisto (2009), Kacperczyk, Sialm, and Zheng (2005), Kacperczyk, Sialm, and Zheng (2008), Amihud and Goyenko (2013) and Jordan and Riley (2015).

Albeit simple and intuitive, such an approach may not be optimal for, at least, four reasons: (i) it does not take full advantage of conditioning information available to investors, as it focuses on a single characteristic; (ii) it does not account for the documented time variation in the degree of cross-sectional predictability in fund returns; (iii) it requires a subjective choice about the number of groups funds are sorted into; (iv) as the strategy typically equally weights the selected funds, it does not align with the principles of portfolio theory and, more generally, with utility maximization. We detail these issues in what follows.

The advantages of integrating information from multiple cross-sectional predictors of asset returns are recently investigated outside the mutual fund literature.[6] The added value stems from the diversification across multiple return predictors whose predictive ability can vary over time. We contend that similar advantages may be present when exploiting the cross-sectional performance predictability of mutual funds. Funds with distinct characteristics may differ in their trading strategies and therefore in their performance. For instance, previous studies find that hot-hand and high alpha funds are likely to be momentum followers in the equity market.[7] In addition, highly active funds tend to exhibit strong stock selection skill and perform relative well during recessions and periods of poor business conditions.[8] Such distinct trading behavior and abilities could lead to wide spreads in the cross-section of fund performance during different periods of time. Recent studies provide empirical evidence that relative fund performance does change over time and appears to be dependent on the macroeconomic environment and on market conditions.[9] From the mutual fund in-

[6]For example, Asness, Moskowitz, and Pedersen (2013) average momentum and value portfolios within and across multiple asset classes. Brandt, Santa-Clara, and Valkanov (2009) optimally combine size, value and momentum in an equity-only portfolio. Barroso and Santa-Clara (2015) bundle the carry trade with other characteristics, such as momentum, reversal, real exchange rate and current account to build optimal currency portfolios.

[7]See, e.g., Grinblatt, Titman, and Wermers (1995); Carhart (1997).

[8]See, among others, Moskowitz (2000); Kosowski (2011); Glode (2011). Sun, Wang, and Zheng (2009) offer two plausible economic rationales for this performance counter-cyclicality: (i) the information opaqueness provides a better profit opportunity for informed managers in the down market; (ii) as noise traders withdraw from the market, professional money managers are more likely to succeed by trading on signals about the fundamentals of firms.

[9]E.g. Glode, Hollifield, Kacperczyk, and Kogan (2012), Kacperczyk et al. (2013) and Banegas, Gillen, Timmermann, and Wermers (2013)

vestor's perspective, a single characteristic portfolio approach that maintains a constant tilt toward funds with a particular characteristic could result in factor or sector biases, and, consequently, inferior performance. For instance, persistently chasing hot-hand and/or high alpha funds would have loaded up on the technology sector prior to the Dot-com collapse in 2000. In a similar vein, systematically buying highly active funds, which typically pick value stocks, might miss profit opportunities offered by momentum stocks, particularly in bull markets. On the contrary, a composite portfolio strategy with exposures to multiple fund characteristics is poised to take advantage of diversification opportunities. However, using multiple predictors and estimating optimal portfolio weights increases the estimation error faced by the investor/econometrician, making unclear whether from an empirical standpoint the optimized composite strategy compares favorably with simpler alternatives.

To summarize, the above discussion suggests that, in order to effectively capitalize on the cross-sectional predictability of mutual funds returns, one needs to identify what fund information to use as performance predictors, as well as how to optimally weigh predictive signals over time from a utility-maximizing perspective. As such, the investor/econometrician also needs to deal with issues such as restrictions on short-sales and leverage. To the best of our knowledge, the literature has been silent on these issues.

Empirically, this paper brings these issues to an out-of-sample optimal portfolio choice framework and provides new evidence on the economic value generated by investing in active equity mutual funds. Our main findings can be summarized as follows.

First, we find that the proposed composite strategy significantly outperforms passive investments based on a set of low-cost index funds that maintain constant exposures to different segments of the market, after fees and expenses and without relying on short-sales or leverage.[10] Quantitatively, a moderately risk-averse investor characterized by a relative risk aversion coefficient of 5 (RRA=5) is willing to pay up to 5.11% per annum in certainty

[10] As this paper is primarily interested in net investment outcomes experienced by investors in the funds, all performance measures are computed based on funds' net returns, i.e. after fees and expenses. As in, among others, Avramov and Wermers (2006) and Banegas et al. (2013), brokerage commissions and taxes are ignored in the results.

equivalent return (CEQ) in order to switch from a broad equity index fund to the optimal composite strategy that exploits the predictive power of multiple fund characteristics. The switching could earn the investor an increase in the Sharpe ratio from 0.51 to 0.82. Second, the strategies based on the maximization of expected utility generate economic gains compared to their counterparts that rely on a sorting rule, suggesting that the sorting-based portfolio approach understates the value added by exploiting the predictive power of fund characteristics. Third, the optimal composite strategy consistently delivers superior performance in comparison with its peers that are formed using the same optimized portfolio approach but only rely on a single characteristic. Fourth, motivated by recent asset allocation literature on the superiority of naïve diversification over more complex optimized approaches [See, e.g. DeMiguel, Garlappi, and Uppal (2009)], we construct and evaluate a naïve composite strategy that simply takes equal-weighted long positions in the top single characteristic sorted portfolios. Despite its simplicity and parsimony the simple averaging rule fails to fully capitalize on the predictability out of sample, as the naïve composite strategy significantly underperforms its optimized counterparts. Fifth, we show that, consistent with theoretical developments in Kacperczyk, van Nieuwerburgh, and Veldkamp (2016), the superior performance of the optimal composite strategy materializes almost exclusively during periods of relatively high market volatility. Finally, utilizing fund stock holdings, we show that the funds included in the optimal composite strategy differ from those selected by alternative strategies especially in terms of their stock picking ability. On the other hand, the exposures of the optimal strategy to the size, value and momentum styles seem to contribute only marginally to its superior performance.

This paper makes several contributions to the literature. First, this study is related to the large body of research that compares the value of active versus passive investments for investors. We provide novel out-of-sample evidence that investing in actively managed equity mutual funds is value adding relative to a set of index-based passive investments, if the cross-sectional predictability in fund performance originating from fund-level attributes

is properly exploited.

Second, our work relates to studies that look at mutual funds performance from the point of view of a utility-maximizing investor. These papers, though, either do not allow for performance predictability of any sort (see Baks, Metrick, and Wachter (2001) and Pástor and Stambaugh (2002)), or focus on the predictability of fund alphas (and betas) by macroeconomic variables, e.g. Avramov and Wermers (2006), and, hence, do not consider fund-specific information as part of the investor's conditioning set. We complement this area of research by building optimal portfolio strategies that exploit the cross-sectional predictability in fund performance induced by fund specific information *and* by macroeconomic variables. We show that integrating the predictive information in multiple fund characteristics into a composite investment strategy leads to considerably larger economic gains than relying separately on each characteristic. Although some literature discounts the role of those fund-level attributes by pointing out that they are subsumed by fund exposures to assumed benchmarks [11], our analysis indicates that the attributes do contain value-adding information from a utility perspective. We also find that accounting for macroeconomic information provides only a rather marginal improvement over and above the gains induced by fund-specific information.

Third, this paper adds to the literature on the efficacy (or, lack thereof) of optimized versus simpler portfolio formation approaches. The sorting-based rule, for instance, has been extensively used to form portfolios capturing the cross-sectional predictability in asset returns. We show that, based on the same conditioning information, the portfolio approach based on the maximization of expected utility is superior to the sorting method in terms of generating economic value for investors. Moreover, recent studies, e.g. DeMiguel et al. (2009), advocate that naïve diversification and, more specifically, the 1/N rule tend to outperform more complex optimized methods out-of-sample. In contrast, we show that, at least within the universe of US active equity mutual funds, optimization-based portfolio strategies

[11] For instance, Carhart (1997) shows that the Hot Hand strategy fails to deliver abnormal returns once momentum is controlled for. Subsequent literature contends that the smart money effect also becomes insignificant in the presence of momentum

outperform out-of-sample those constructed using the 1/N rule.

From an industry perspective our analysis speaks especially to large institutional investors, such as funds of mutual funds (FoMFs) and to investment consultants advising a Defined Contribution (DC) plan .[12] Fixed weighting schemes across mutual fund styles (e.g. aggressive, growth and income or large-, mid- and small-cap) are widely used as the basis for making portfolio allocation decisions by FoMFs.[Brands and Gallagher (2005); Elton, Gruber, and Blake (2006).] However, style-based approaches have been questioned by academic researchers who argue that mutual fund style classifications do a poor job on forecasting differences in future performance and consequently such misclassification has an adverse effect on investors' ability to build diversified portfolios.[Brown and Goetzmann (1997); diBartolomeo and Witkowski (1997); Chan, Chen, and Lakonishok (2002).] The characteristic-based portfolio approach proposed in this paper offers an alternative to aid FoMFs managers in making mutual fund investment decisions. Similar considerations apply to the fiduciary of a DC plan seeking to select from the rather large menu of active funds.

The rest of this paper proceeds as follows: Section 2 introduces mutual fund characteristics and their measures. Section 3 discusses methodologies for constructing optimal composite strategies. Section 4 introduces the data. Section 5 reports all empirical results. Section 6 concludes.

2 Fund characteristics and measures

In this section, we introduce the set of mutual fund characteristics that have been extensively studied in the literature and their measures. We focus on six fund characteristics that have been shown to be associated with future fund performance: the degree of activeness in management, historical alpha, past returns (hot-hand), net money inflows, fund size and

[12] FoMFs have emerged as a popular investment vehicle in recent times. According to the 2014 Investment Company Fact Book, the FoMFs industry grew from as little as $1.4 billion in early 1990s to nearly $1.6 trillion in AUM at the year-end 2013.

past return volatility.[13]

1. Activeness. To measure the degree of activeness in management, Amihud and Goyenko (2013) propose an intuitive and easily calculable measure, which is defined as $1 - R^2$, where R^2 is estimated by regressing the fund's returns on benchmark returns (e.g. the Fama-French-Carhart 4-Factors). Thus, lower R^2 indicates greater degree of activeness, which in turn predicts better performance. In this paper, we use the $1 - R^2$ computed from Eq. (1) as the active characteristic.

2. Alpha. We follow Pástor et al. (2015) and estimate alpha by regressing a fund's net return on the returns of the fund's benchmark index designated by Morningstar.[14]

$$r_{i,t} = \alpha_i + \beta_i \cdot r_{\mathrm{BMK},t} + e_{i,t} \tag{1}$$

where $r_{i,t}$ denote fund i's net return during period t; $r_{\mathrm{BMK},t}$ is the return of Morningstar designated benchmark index; α_i is the benchmark alpha. The estimation is performed recursively with a 12-month rolling sample at the end of each month t.

3. Net money inflows. Gruber (1996) and Zheng (1999) document that funds that

[13]We acknowledge that other characteristics that are related to mutual funds' cross-sectional differences in performance have been proposed in the literature, e.g. the Return Gap of Kacperczyk et al. (2008), Active Shares of Cremers and Petajisto (2009) and industry- or sector-concentration of Kacperczyk et al. (2005). These findings are certainly of interest from the performance evaluation point of view. However, constructing these predictors entails creating portfolios from funds' quarterly stock holdings, which may be difficult for many investors to obtain in a timely manner. Also, the quarterly frequency of observation would make for a relatively small sample, possibly weakening the effectiveness of the empirical tests. Furthermore, relying on a mutual fund' self-reported quarterly positions may distort our view of the fund's actual performance due to potential reporting biases, such as missing cash and bond holdings in the database [Wermers (2000)] and/or window dressing and tax-motivated trading [Moskowitz (2000)]. Therefore, when constructing portfolios, we focus only on the six characteristics discussed earlier, as they can be relatively accurately measured using the information available to investors at the monthly frequency. Nevertheless, it is important to note that the analytical framework provided in this paper can be easily extended to include additional cross-sectional predictors of mutual fund performance.

[14]Since the work of Carhart (1997) a common choice in the literature is using Fama-French-Carhart 4 factors for the risk adjustment. On the other hand, issues arising in the choice of the benchmark have been raised numerous times. For instance, Cremers, Petajisto, Zitzewitz et al. (2013) argue that the Fama-French factor model produces biased assessments of fund performance and recommend using index-based benchmarks, as such benchmarks better explain the cross-section of mutual fund returns. Similar and additional concerns are put forward by Pástor et al. (2015) and by Berk and Van Binsbergen (2015). Further, Pástor et al. (2015) note that Morningstar chooses benchmarks based on funds' holdings rather than their reported objective, so the Morningstar benchmark does not suffer from the cherry-picking bias.

experience recent positive net money inflows tend to outperform their less popular peers subsequently. The authors term this phenomenon as “smart money" effect. Following the literature, we calculate the Flow characteristic as the normalized new money net flow, which is calculated as the quarterly net cash flow divided by the TNA at the end of the previous quarter:

$$\text{Flow}_{i,t} = \frac{\text{netflow}_{i,t}}{\text{TNA}_{i,t-1}} \tag{2}$$

where $\text{netflow}_{i,t}$ is fund i's net cash inflows during time t; $\text{TNA}_{i,t-1}$ is the fund's total net assets at the end of time $t-1$.

4. Hot-hand. Hendricks et al. (1993) and Goetzmann and Ibbotson (1994) first find that recently top-performing funds tend to continue to be good performers in the near term and a portfolio strategy purchasing the good performing funds and selling the underperforming ones could earn positive risk-adjusted returns. Following Hendricks et al. (1993), we compute 12-month compounded prior net returns as the hot-hand characteristic.

5. Fund size. Chen et al. (2004), Yan (2008) and McLemore (2016) provide empirical evidence that fund size is inversely related to performance. Following the literature, we compute the size characteristic as:

$$\text{Size}_{i,t} = 1 - \frac{\text{TNA}_{i,t}}{\text{IndSize}_t} \tag{3}$$

where IndSize_t denotes the industry size, computed as the sum of TNAs across all funds at time t; $\text{TNA}_{i,t}$ is fund i's TNA at time t. To be consistent with other characteristics, for which better performing funds have higher measures, we use one minus the fund-to-industry size ratio.

6. Volatility. Jordan and Riley (2015) show that a mutual fund's past return volatility is a strong predictor for future fund performance. We calculate the standard deviation of a fund's monthly returns during the past 12-month window as the volatility characteristic.

3 Optimal portfolio strategies and performance measures

To optimally integrate the information content from individual fund characteristics in a portfolio strategy, we employ the parametric portfolio policy proposed by Brandt et al. (2009) (hereafter, BSV). The BSV approach is especially suited here for several reasons. First, it directly models portfolio weights as a linear function of fund characteristics. In such a way, it effectively avoids the extremely sensitive estimation of the moments of the future distribution of fund returns, which often involves substantial sampling errors and leads to extreme portfolio weights.[15] Second, it mitigates the dimensionality problem: the number of parameters in the BSV approach depends only on the number of characteristics rather than the number of assets as in traditional portfolio methods.[16] Third, it implicitly captures the impacts of fund characteristics on expected returns, variances, covariances and even higher-order moments of the return distribution.[17] This is nontrivial in the portfolio choice problem. For example, a specific characteristic might be found to be positively associated with funds' expected returns. Naturally, portfolio methods focusing on mean returns, such as the sorting-based rule, would favor the funds possessing this characteristic. On the contrary, those funds may not look attractive to the BSV approach, if the characteristic is also associated to undesirable moments of the resulting portfolio's returns, such as volatility and negative skewness. In this paper, we provide the first application of the BSV parametric portfolio policy in the context of optimal mutual fund portfolio selection.

[15] Brandt (2009) provides a comprehensive discussion on the issue of estimation errors in asset allocation context.

[16] For example, in a classic MV portfolio optimization with N assets, it requires modeling N first and $(N^2+N)/2$ second moments of returns.

[17] Brandt et al. (2009): "To better understand this point, we can approximate the expected utility of the investor with a Taylor series expansion around the portfolio's expected return $E[r_{p,t+1}] \approxeq u(E[r_{p,t+1}]) + \frac{1}{2}u''(E[r_{p,t+1}])E[(r_{p,t+1} - E[r_{p,t+1}])^2] + \frac{1}{6}u'''(E[r_{p,t+1}])E[(r_{p,t+1} - E[r_{p,t+1}])^3] + \cdots$. This expansion shows that, in general, the investor cares about all the moments of the distribution of the portfolio return. Since the portfolio return is given by $r_{p,t+1} = \sum_{i=1}^{N_t} f(x_{i,t};\theta) r_{i,t+1}$, the moments of its distribution depend implicitly on the joint distribution of the returns and characteristics of all assets. The coefficients θ affect the distribution of the portfolio return by changing the weights given to the returns of the individual assets in the overall portfolio."

3.1 Estimating optimal portfolio weights: unconditional case

Following BSV, we optimize mutual fund portfolios from the perspective of a risk-averse investor characterized by a CRRA utility function:

$$U(r_p) = \frac{(1+r_p)^{1-\gamma}}{1-\gamma} \tag{4}$$

The investor's problem is to recursively choose the portfolio weights $\omega_{i,t}$ at time t to maximize her conditional expected utility of the portfolio return $r_{p,t+1}$ at time $t+1$:

$$\max_{\{\omega_{i,t}\}_{i=1}^{N_t}} \quad E_t[U(r_{p,t+1})] = E_t\left[U\left(\sum_{i=1}^{N_t} \omega_{i,t} r_{i,t+1}\right)\right] \tag{5}$$

where $r_{p,t+1}$ and $r_{i,t+1}$ are the portfolio's and fund i's returns at time $t+1$, respectively; N_t is the number of mutual funds available to trade at time t. Fund i's weight, $\omega_{i,t}$, is parametrized as a linear function of the fund's standardized characteristics:

$$\omega_{i,t} = f(z_{i,t};\theta_t) = \frac{1}{N_t}\theta_t^\top z_{i,t} \tag{6}$$

where θ is a $K \times 1$ vector of parameters to be estimated and K is the number of fund characteristics; $z_{i,t}$ is a $K \times 1$ vector of fund i's standardized characteristic measures with zero mean and unit standard deviation across all funds at time t.

Rewriting the investor's problem by plugging Eq. (6) and Eq. (4) into Eq. (5) and replacing the expectation with its sample analog one gets:

$$\max_{\theta_t} \quad \frac{1}{\tau(1-\gamma)} \sum_{t'=t-\tau}^{t-1} \left[1 + \sum_{i=1}^{N_{t'}} \left(\frac{1}{N_{t'}}\theta_t^\top z_{i,t'} r_{i,t'+1}\right)\right]^{1-\gamma} \tag{7}$$

The optimal parameter vector θ_t^* is obtained by solving Eq. (7) numerically over the rolling sample period $[t-\tau, t-1]$. With the estimated parameter vector, the desired portfolio weight on fund i for time $t+1$ is calculated as $\omega_{i,t}^* = \frac{1}{N_t}\theta_t^{*\top} z_{i,t}$. Note that we only use the

information available up to time t to derive the optimal weights for time $t+1$. By repeating this procedure at each time $t \in [T_0, T-1]$, where T is the last sample observation, we can obtain the out-of-sample unconditional optimal portfolio policy $\{\omega_t^*\}_{t=T_0,\cdots,T-1}$. [18]

3.2 Estimating optimal portfolio weights: conditional case

While able to incorporate the predictive information in fund characteristics, the unconditional BSV approach assumes implicitly constant impact of fund characteristics on the optimal portfolio policy. To further incorporate the dependence of fund performance predictability on macroeconomic and market conditions that has been documented in the literature, we follow Bredendiek, Ottonello, and Valkanov (2016) to estimate optimal portfolio weights conditioning on a state variable. More specifically, let d_t be a dummy variable that equals one if the economy is in the high uncertainty state at time t and zero, otherwise. Then, we include two interaction terms $\omega_{i,t} \times d_t$ and $\omega_{i,t} \times (1-d_t)$ in the objective function in Eq. (7), where $\omega_{i,t}$ is the weight function defined in Eq. (6), to obtain two sets of θ estimates. With the two estimated parameters θ_t^{1*} and θ_t^{0*}, the desired portfolio weight on fund i for time $t+1$ is calculated as $\omega_{i,t}^* = \frac{1}{N_t}\theta_t^{1*\top} z_{i,t}$ if the economic state is high ($d_t = 1$) and as $\omega_{i,t}^* = \frac{1}{N_t}\theta_t^{0*\top} z_{i,t}$, otherwise. Note that we only use the information available up to time t to derive the optimal weights for time $t+1$. By repeating this procedure at each time $t \in [T_0, T-1]$, we can obtain the out-of-sample conditional optimal portfolio policy $\{\omega_t^*\}_{t=T_0,\cdots,T-1}$. [19]

[18] Since we are interested in long-only investment strategies for mutual funds. we follow BSV's recommendation to truncate negative weights of the unconstrained optimal portfolio at zero and then re-scale the truncated weights so that they sum up to one. Plazzi, Torous, and Valkanov (2010) and Barroso and Santa-Clara (2015) also follow this approach in order to satisfy the constraints.

[19] To estimate the BSV coefficients in Eq. (7), we use a rolling sample window of 24 months, as it is commonly done for evaluating mutual fund performance in the literature. See, for example, Jiang et al. (2007), Amihud and Goyenko (2013), Del Guercio and Reuter (2014), Pástor et al. (2015). For the sake of robustness, we also entertain other window lengths, including 12, 36, 48, 60 months. we find that all the baseline results are qualitatively unchanged and that the performance of optimal composite strategies tend to decrease with the window length, consistent with the well documented short-lived superior performance of mutual funds in the literature.

3.3 Performance measures and significance tests

We employ three metrics to measure the out-of-sample performance of investment strategies: the certainty equivalent excess return (*CER*) based on the CRRA utility function, the Sharpe ratio (*Sharpe*) and the 4-Factor alpha (*Alpha*),

Given our focus on the investor's utility in measuring economic gains, CER seems to be a natural choice.[20] For a given strategy, CER is defined as the riskless return an investor is willing to accept instead of facing the uncertain return generated by such strategy.[21]

Given its popularity in performance evaluation of both delegated portfolios and asset allocation strategies, we also consider the Sharpe Ratio. Moskowitz (2000), among many others, advocates looking at the second moment of fund returns in mutual fund performance evaluation, as active managers could add value by reducing the volatility of their managed portfolios, meanwhile delivering similar mean returns to the market.

Following several hints from the extant literature, we see CER as more meaningful than Sharpe ratio within the context of our analysis. This is for two main reasons. First, several studies point out that skewness and kurtosis in returns matter to risk averse investors and that portfolio strategies accounting for the dynamics of such moments generate rather different allocations than those induced by, say, mean-variance preferences and lead to economically significant gains.[22] Second, while well suited within a world of I.I.D. Gaussian returns, Sharpe ratio may generate misleading conclusions when applied to predictability-based strategies that exploit time variation in investment opportunities [See Farinelli and Tibiletti (2008)].[23]

[20] In the asset allocation literature, CER has been commonly employed as a performance measure. See, among others, Brandt et al. (2009); DeMiguel et al. (2009); Cenesizoglu and Timmermann (2012); Barroso and Santa-Clara (2015).

[21] Under the CRRA utility, CER is computed as: $\text{CER}_{\tau+1:T} = [(1-\gamma)\bar{U}_{\tau+1:T}(W_t)]^{\frac{1}{1-\gamma}} - 1$, where $\bar{U}_{\tau+1:T}(W_t)$ is the average of realized CRRA utility values.

[22] On the theoretical side see e.g. Arditti (1967), Kraus and Litzenberger (1976) and Scott and Horvath (1980). On the empirical side see Bekaert, Erb, Harvey, and Viskanta (1998), Harvey and Siddique (1999), Ang and Bekaert (2002), Patton (2004), Guidolin and Timmermann (2008), Harvey, Liechty, Liechty, and Müller (2010) and Timmermann (2006), among others.

[23] As noted by Bianchi and Guidolin (2014), with predictable returns and CRRA utility, one cannot generally approximate preferences as mean-variance objectives: hence, a higher Sharpe ratio yielded by any

Despite its drawbacks *Alpha* is still the most widely used performance measure in the mutual fund literature to date; it measures a portfolio's mean returns adjusted for a chosen risk benchmark. We, hence, report Carhart Alpha as a measure of strategy performance but acknowledge its potentially limited informative content.

To establish the statistical significance level of the performance measures, we employ a bootstrap procedure, following Bredendiek et al. (2016). We generate 1,000 sample paths from the original fund data set. Specifically, to form a sample path, we randomly draw (with replacement) the same number of observations on fund returns and characteristics as the number of funds in the original sample for a given month. Then, we implement all portfolio strategies and compute their out-of-sample performance measures along each sample path. Based on the empirical distribution of performance measures, we make statistical inferences . For example, to assess the statistical significance of the outperformance of the composite optimized strategy over an alternative strategy, we first compute the t-statistic of the outperformance as $(\overline{\text{CER}}_{\text{com}} - \overline{\text{CER}}_{\text{bmk}})/\sigma_{\text{bmk}}$, where $\overline{\text{CER}}_{\text{com}}$ ($\overline{\text{CER}}_{\text{bmk}}$) is the average CER of the composite optimized strategy (the alternative strategy) across 1,000 sample paths and $\sigma(\text{CER}_{\text{bmk}})$ is the bootstrap standard error of CER for the alternative strategy. Then, we calculate the two-tailed p-value from the t-statistic under the null hypothesis that CER_{com} is equal to CER_{bmk}.

4 Data

The mutual fund data used in this paper come from three main sources: Center for Research in Security Prices (CRSP) Survivorship Bias-free U.S. Mutual Fund Database, Morningstar Direct and Thomson Reuters Mutual Fund and Investment Company Common Stock Holdings Database. In particular, we use CRSP and Morningstar to collect the data on fund returns, share-level total net assets (TNA), fees and expense ratios, investment objectives

given strategy does not necessarily imply an increase in the investor's welfare, especially if this is achieved by altering the higher-order moments of portfolio returns away from what they should be under multivariate Gaussian returns.

and other variables. In addition, we collect the benchmark return data for each fund from Morningstar.[24] Portfolio holdings data are obtained from Thomson Reuters.

Next, we follow the procedures specified in the data appendix of Pástor et al. (2015) to merge the fund data from CRSP and Morningstar.[25] Briefly, we first use funds' tickers, CUSIPs and names to match funds that are available in both datasets. We then double check the matching accuracy by comparing the matched funds' TNAs and returns across the two databases. To minimize the impacts caused by data errors, we follow Berk and Van Binsbergen (2015) to correct the discrepancies in the data. If some observations are not fixable, we set them to missing. For instance, we set fund returns to missing if they still differ across the two datasets by more than 10 bps after the correcting process.

To obtain a clean set of active U.S. equity mutual funds, we screen funds based on investment objective and policy codes and fund names. As a result, we eliminate bond funds, money market funds, international funds, index funds, funds of funds, industry funds, real estate funds, target retirement funds, and funds that are not invested primarily in U.S. equities. To eliminate the return bias of small-sized funds, we also exclude fund-month observations with lagged TNA below $10 million from the sample.[26] After obtaining the clean sample of active U.S. equity mutual funds, we follow the methodology of Wermers (2000) to use the MFLINKS table from WRDS to merge the fund data and Thomson Reuters holdings data. The resulting sample set contains a total number of 3,255 unique active U.S. equity mutual funds during the period January 1980 – December 2015. We refer to this data set as the large sample.

Following Pástor and Stambaugh (2002), Avramov and Wermers (2006) and Geczy, Stambaugh, and Levin (2005), we further restrict the sample to no-load funds, as their returns are

[24] Morningstar assigns each fund into a category based on the fund's holdings and designates a benchmark portfolio to each fund category.

[25] The data appendix can be downloaded at `http://faculty.chicagobooth.edu/lubos.pastor/research/Data_Appendix_Aug_2013_V3.pdf`

[26] We considered other variations that have been used in the literature, such as $5 million and $15 million. Our empirical results are not sensitive to the changes of the threshold value.

more representative of what investors in the fund actually receive. [27] Finally, for a fund to be included in our sample we require that it has non-missing values for all six characteristic measures specified in Section 2. After this screening procedure, the final sample consists of 944 funds. We label the final data set as the small sample. [28]

Figure 1 depicts the number of funds in the two sample sets over time.

[Figure 1]

The dashed (blue) line represents the large sample that includes all active U.S. equity mutual funds. The number of funds increases from 170 in January 1980 to 1,364 in December 2015 and reaches the maximum of 2,086 in September 2008. The solid (red) line represents the small sample that excludes funds that charge load fees or those with missing one of the six characteristic measures. The number of funds starts with 1 in January 1980, then peaks at 817 in June 2008 and finally ends with 647 in December 2015. Comparing the two samples, we see that there is a substantial loss of funds during the first 10 years. The loss is due to two reasons: (i) a large number of mutual funds report TNA only quarterly or yearly before January 1991, so the characteristic measures, such as Size and Flow, could not be constructed. (ii) The recorded expense ratios across the two databases often disagree and exhibit large jumps during this period. Due to the limited number of funds in the early years, we only use the data in the small sample from January 1991 to December 2015 as our final sample for empirical analysis.

Following Bredendiek et al. (2016) we use three indices as proxies for macroeconomic and market conditions: 1. the credit spread index (GZ Spread) of Gilchrist and Zakrajšek (2012)[29]; 2. the macroeconomic uncertainty index (MU) of Jurado, Ludvigson, and Ng

[27] Avramov and Wermers (2006) describe in more detail the advantages of considering no-load funds from an investor's portfolio perspective.

[28] It is worth noting that we repeat our analysis including load funds, and find that our baseline results hold. We acknowledge, however, that these results do not account for sales charges.

[29] The data is at `http://people.bu.edu/sgilchri/Data/data.htm`

(2015)[30]; 3. the volatility index (VIX) from the Chicago Board Options Exchange. Generally, the three variables have been shown to have conceptual and empirical ability to capture and, to some extent, anticipate fluctuations in the state of the economy. For each index, we further create a dummy variable, which equals one if the index level is higher than one standard deviation above its historical mean and zero otherwise.

5 Empirical results

We, first, verify that the fund characteristics illustrated in section 2 do have predictive power for subsequent fund performance in our sample. Next, in section 5.2 we compare the out-of-sample performance of the optimal portfolio strategies introduced in section 3.1 with the performance of various active and passive alternatives. Last, in section 5.3 we conduct several robustness checks.

5.1 Cross-sectional predictability

We first reexamine the predictive ability of each individual fund characteristic for funds relative performance. At each month-end we sort funds into ten decile portfolios according to the rankings of the individual fund characteristics and, then, report their performance over the following month. Following, e.g., Pástor and Stambaugh (2002) and Avramov and Wermers (2006), each decile portfolio is constructed as an equally weighted combination of the funds in that decile. Six groups of single characteristic decile portfolios are, thus, formed; we label them as the Alpha, Hot, Flow, Size, Active and Vol. decile portfolios.

Table 1 reports the out-of-sample performance of the top, bottom and long-short decile portfolios for each fund characteristic. Decile 10 (1) represents the portfolio of funds with the highest (lowest) values of the given characteristic while 10-1 denotes the portfolio that longs Decile 10 and shorts Decile 1. Performance is measured out-of-sample for the investment

[30]The data is at `http://sydneyludvigson.com/data-and-appendixes`

period 1996:01–2015:12.

[Table 1]

The main observation from Table 1, is that all three performance measures show large dispersion and declining trends from the top to bottom decile.[31] Further, we find that the long-short portfolios produce positive and significant performance measures in almost all cases.[32] In particular, the Hot Hand and Alpha portfolios yield the highest CERs at around 2.5% per year. The also generate the highest Sharpe ratios at around 0.5 and Carhart-alphas above 3.5% These results lend support to prior literature that relates fund characteristics to cross-sectional variations in performance and suggests the possibility of abnormal returns from strategies that exploit such predictability. As discussed earlier, though, mutual fund investors face short-selling restrictions. Consequently, the profitability of the long-short portfolios in Table 1 may not truly reflect the economic gains that could be earned by a real-world investor. A more meaningful assessment for the investor should, therefore, be based on the performance of long-only strategies, i.e. the top-decile portfolios. By restricting our attention to strategies with long-only positions (see first row of Table 1 we still find evidence of significantly positive CERs for all portfolios, the highest being the Hot hand and the Volatility portfolios at about 2.5% annually. The Sharpe Ratios are comparable and in some cases higher than for the long-short portfolios, ranging between 0.52 and 0.63. Noticeably, the alphas generated from long-only portfolios are significantly lower than those from the long-short strategies and only three out of six are statistically significant at conventional levels.

[31]For brevity we report only the top and bottom deciles: we do find, in accordance to most literature, that performance declines monotonically across deciles from 10 to 1.The complete results are available from the authors.

[32]Statistical significance for the performance measures is assessed using the bootstrap procedure outlined in section 3.3.

From now on, we will refer to the long-only top-decile portfolios as *single characteristic sorted strategies*.

As funds in these portfolios are ranked according to their characteristics measures, the correlations in the rankings can offer a first indication of the potential diversification benefits (or, lack thereof) for portfolios that combine top-decile funds across multiple characteristics. At the basic level, rank correlations illustrate whether there are significant overlaps in the composition of the six top deciles of funds. Table 2 reports the rank correlation matrix of the six fund characteristics. The correlations are computed using the decile ranks for each month and then averaged across the sample period, 1995-2015. The general observation is that the characteristic measures are not strongly related to each other, as evidenced by an average absolute value across all pairs of 0.11 with a standard deviation of 0.14. The absolute values of 12 out of the 15 correlation coefficients are lower than 0.13 and statistically insignificant. The other 3 pairs, Flow/Alpha, Hot-hand/Alpha and Hot-hand/Flow, are significantly positively correlated, although the correlation never exceeds 0.45.

The above characterizations are interesting from a diversification perspective. In what follows we examine whether the potential diversification benefits translate into investment results.

5.2 Comparing strategies

In this subsection, we assess the out-of-sample performances of the composite optimized strategies outlined in Section 3.1 and 3.2, and we compare them with the results from various active and passive alternatives: these represent the baseline results of our empirical analyses. In an attempt to determine the sources of the differences in performance we, then, further characterize the comparison by: a) assessing how the alternative strategies perform across different market and economy-wide conditions; b) assessing the attributes (such as average fund size, expenses and net flows) and investment styles of the portfolios generated by alternative strategies.

5.2.1 Out-of-sample performance

We compare the performance of the proposed composite optimized strategies with that of four sets of competing strategies. The first set includes four low-cost Vanguard index funds.[33] The idea here is to use actually traded investment vehicles, rather than non-traded indexes, to represent the passive investment opportunities available to real-world investors. The second set consists of the six single characteristic sorted strategies that we introduced and analyzed in Section 5.1. The third set contains six single characteristic optimized strategies, each of which exploits the predictive power of only one fund characteristic and is optimized using the BSV portfolio approach introduced in Section 3.1. Although both aim to capitalize on the cross-sectional performance predictability based on a particular fund characteristic, the optimization-based approach is different from the sorting-based method as the former takes into account an investor's preferences and risk appetite (and, more generally, portfolio theory), whereas the latter is predicated on simplicity as it equal weights the funds ranked in decile 10. The fourth alternative is a naïve composite strategy that allocates equally across the six single characteristic sorted portfolios in the second set.[34] Like the optimization-based strategies, this strategy makes use of the predictive information in all the fund characteristic signals; on the other hand it does not involve portfolio weight estimation, as there is evidence, e.g. DeMiguel et al. (2009), showing that simple portfolio rules can outperform more complex optimized peers out-of-sample due to estimation error.

For all the investment strategies except the passive ones, portfolio weights are computed recursively at the end of each month starting from December 1995 and through November 2015. Hence, the first allocation is made at the end of December 1995 using historical information known at that point in time, the second allocation at the end of January 1995 including the additional information that becomes available over that month, and so on,

[33] We choose the Vanguard index family because of its well-known cost efficiency and market-leading role in the index fund industry.

[34] In the equity market, Asness et al. (2013) construct an equal-weighted 50/50 momentum and value combination strategy and find that the resulting portfolio outperforms each individual one.

..., with the last portfolio formed at the end on November 2015. The month-t realized net excess return on each portfolio is calculated by multiplying the desired weights estimated in month $t-1$ by the month-t realized net excess returns (i.e. after fees and expenses, and above the risk-free rate) of funds included in the portfolio. This recursive scheme generates 252 monthly excess returns for each strategy.

The out-of-sample investment results for all the strategies are presented in Table 3. On the columns, we report the first four moments (Mean, Std, Skew, Kurt) of the portfolio net excess return distributions and the three risk-adjusted performance measures: 4-Factor alphas (Alpha), Sharpe ratios (Sharpe) and Certainty Equivalent Returns (CERs) for a moderately risk averse investor (RRA=5) as a base case. In robustness checks (see below in section 5.3.3) we experiment with a wide range of relative risk aversion coefficients. All metrics are reported in annualized terms. Panel A, B, C and D of the table provide the results for the passive, single characteristic sorted, single characteristic optimized and composite strategies, respectively. In addition to the estimated return moments and performance measures, the table reports bootstrapped p-values (see section 3.3) for the null hypothesis that the estimate under a given strategy equals the corresponding estimate under the composite optimized strategy.

[Table 3]

First, we compare the composite optimized strategy with the passive vehicles (Panel D vs. Panel A in Table 3). Panel A shows that index funds with focus on different segments of the market exhibit tangible heterogeneity in their return distributions and, especially, across the CERs and 4-factor alpha measures. While the S&P500 and the Value Index generated an average annual excess return around 7.5% with a 15% volatility, the Mid-Cap and Small-Cap

funds returned about 9% with a volatility around 19.5%. Although the Sharpe Ratios of all passive vehicles were comparable, the CER for the former funds were higher than for the latter ones by 2.5-3% per year. Comparing the composite optimized strategy with the passive benchmarks we see that the former consistently outperforms the latter across all performance measures. For instance, a moderately risk-averse investor (RRA=5) who holds the S&P500 would be willing to pay up to 5.11% (6.75 - 1.64) per year in CER to switch to the composite optimized strategy. By doing so, she would also increases her Sharpe ratio from 0.51 to 0.82 and earn an extra annual 4-Factor alpha of 4.87%. Similar or larger differences arise when considering the other three passive strategies. In terms of returns, the gains for the composite strategy appears to first come from the ability to generate higher average return (around 13% per year vs. 8-9% for the index funds) while keeping a comparable portfolio volatility (around 15% per year). In addition, the composite optimal portfolio returns display more favorable skewness and kurtosis from the point of view of the power utility investor. They have essentially zero skewness (estimated at 0.03), whereas the returns of all passive portfolios are negatively skewed (between -0.51 and -0.75). And their kurtosis is higher.

The above results represent strong evidence that, when one relies on fund specific information to predict cross-sectional differences in fund performance, investing in active equity mutual funds can be value adding relative to passively holding index funds, even when short-selling and leveraging are precluded.

Next, we compare the composite optimized strategy with the single sorted strategies (Panel D vs. Panel B in Table 3) analyzed above in section 5.1. These are the strategies that, according to previous literature, can generate abnormal risk-adjusted returns. Yet, the composite optimized strategy we propose outperforms all of them along the three performance metrics, generating an increase in CERs ranging between 418 basis points per year (vs. the Hot- Hand strategy) and 515 basis points per year (vs. the Alpha strategy), all statistically and economically significant. A similar picture emerges from comparing the

Sharpe Ratios: no single sorted strategy generates a Sharpe Ratio above 0.65 while the composite optimized strategy tops them all at 0.82. Finally, the 4-factor alpha of the composite optimized strategy, 5.19% per year, is also much larger than those produced by the single sorted approaches, none of which exceeds 1.52%.

We, next, compare the composite optimized strategy with the single characteristic optimized strategies. (see Panel D vs. Panel C). The strategies in this comparison rely on the portfolio approach specified in Section 3.1, but they differ in the number of fund characteristics utilized in deriving optimal portfolio weights. Therefore, the differences in their investment outcomes should speak more directly to the additional profitability (or, lack thereof) induced by exploiting multiple rather than single fund-specific signals. One initial observation from the sample moments is that the composite strategy is able to produce an average return comparable to that of the best single characteristic optimized strategy (i.e. Hot) but with lower volatility. It is also interesting that most single characteristic strategies, except Hot, have negative skewness, while the returns of the optimal composite portfolio are slightly positively skewed. These results suggest that combining fund characteristics in a single investment strategy could achieve more desirable risk-return profiles for risk-averse investors through risk reduction without sacrificing high mean returns. Moving to the portfolio performance measures, the results appear to be clearly outlined: the composite strategy consistently delivers superior performance over all single characteristic peers regardless of metrics or risk aversion levels. Quantitatively, the composite strategy, benchmarked on the best-performing single characteristic portfolio — Hot, still outperforms by 1.79% per year in CER (RRA=5) or 90 bps in 4-Factor alpha, and 11% in Sharpe ratio. These results highlight the advantages of diversifying across fund characteristics and, more importantly, that such qualities can translate into significant economic gains for investors when exploited in investment strategies.

As we consider two methods for forming portfolios, the sorting rule and the utility maximization rule, we, next, compare the two approaches controlling for the fund specific informa-

tion they use (see Panel B vs. Panel C). In general, the optimization-based strategies appear to outperform their sorting-based counterparts in all cases. Although the difference in performance is rather small in terms of Sharpe Ratio, the CERs reveal very sizable gains for the optimized strategies, with differences ranging between 82 basis points (Activeness strategy) and 176 basis points per year (Size strategy) For instance, a moderately risk-averse investor who adopts the Alpha sorted strategy would be willing to pay up to 1.2% per annum in CER to switch to the Alpha optimized strategy. These results show that, given the same *ex-ante* fund specific information, the optimization-based method tends to be more effective than the sorting rule from the investor's perspective. Despite not being the main focus of this paper, it is still interesting to point out that all single characteristic strategies (both sorted and optimized), except the one based on size sorting, outperform the top-performing index fund in terms of both Sharpe ratio and CER (Panel B and C vs. Panle A). The results indicate that even though inferior to the optimal composite strategy, portfolio strategies that utilize the predictions from single fund characteristics can perform better than passive investment strategies using index funds, highlighting the promise of predictability-based active investment strategies in the U.S. equity mutual fund space.

Finally, we compare the results within the composite strategies category (see Panel D). In this comparison we are motivated by recent findings in the asset allocation literature indicating the superiority of simple portfolio formation approaches, such as the 1/N rule of DeMiguel et al. (2009), over utility-based optimization methods. We, hence, compare the out-of-sample performance of a strategy that equally combines six single characteristic sorted portfolios with that of the composite optimized strategy. The results are clear-cut: the naïve composite portfolio produces a much lower mean return and similar volatility compared with the composite optimized portfolio. In addition, the naïve strategy has a negative and lower skewness of -0.53, which mirrors the findings by Brown, Hwang, and In (2013) that naïve diversification relative to optimal diversification comes with increases in tail risk and reduced upside potential associated with the concave payoff. Moving to the performance measures,

the naïve strategy substantially underperforms its optimized counterpart across all metrics, suggesting that when applied within the space of equity mutual funds, the BSV approach based on expected utility maximization is preferable to the 1/N rule advocated in other asset universes, such as individual US equities.

To summarize, the above analyses demonstrate that simultaneously incorporating the information conveyed by multiple fund characteristics into a composite portfolio strategy and relying on portfolio theory within a utility-maximization framework generates significant economic value, as the resulting portfolio outperforms: a) several passive benchmarks; b) single characteristic sorted and optimized strategies that rely on the predictions from only one fund characteristic and c) the naïve composite portfolio that averages all single characteristic sorted portfolios without involving estimation of portfolio weights. Noticeably, the superior performance is realized out of sample, after subtracting funds' fees and expenses as well as precluding short-sales and leverage.

5.2.2 Macro vs. Fund-specific information

We want to investigate whether different macroeconomic conditions affect the optimal allocation of mutual funds and, in particular, whether the economy-wide information leads to utility gains when added to the fund specific information we analyzed in the previous sub-sections.

As illustrated in the Data section, we interact each fund-specific characteristics with a dummy variable that equals one in economic downturns and zero otherwise. We use three alternatives for defining changes in economic regimes: the credit spread index (GZ) of Gilchrist and Zakrajšek (2012), the macroeconomic uncertainty index (MU) of Jurado et al. (2015) and the VIX index of the Chicago Board Option Exchange (CBOE).

We focus on the differences in performance between the optimized composite strategy that only relies on fund-specific characteristics, which we labelled as unconditional, and the optimized composite strategy that also takes advantage about economy-wide information,

which we labeled as conditional.

[Table 4]

The results are reported in Table 4. Regardless of the proxy used for measuring economic conditions, there is essentially no improvement in the performance of the optimized portfolios with respect to the unconditional composite strategy. Whether one looks at Sharpe ratio, CER or 4-factor alpha the conditional strategy is performing worse or insignificantly better as indicated by the high bootstrapped p-values.

We conclude that the fund-specific information appears to be sufficient in capturing potential time variations in the cross-sectional predictability of fund performance and, hence, in determining the funds to be included in the optimal portfolio depending on economy and market wide conditions.

Given the above results and conclusions, for the remainder of the analysis we will continue to focus on the composite optimized strategy that relies only on fund specific information.

5.2.3 Portfolio weights

The superior performance of optimal composite strategies might be the results of unreasonably large bets on very few funds that perform well in short periods. To address this concern, we show the number of funds included in the optimal composite portfolio in Panel A of Figure 2 (blue solid line) and the time-series evolution of the maximum weight allocated to a given fund in Panel B of Figure 2 (red solid line).

[Figure 2]

Panel A shows that the number of funds included in the optimal portfolio rises steadily from about 50 at the start of the out-of-sample period (1995) to about 150 in the second half of 2002. In the second half of the sample the optimal number of funds remain fairly stable within the 150 to 200 range. There are two noticeable exceptions: the drop to zero during the stock market downturn of 2002-2003 and the longer lasting drop during the Great Financial Crisis of 2008-2009. The plot in Panel B shows that the highest allocation to a single funds by the optimal unconditional strategy very rarely exceeds 5%, especially in the more recent portion of the sample.

We also compute summary statistics for the cross-sectional *average* and *maximum* weights. For the cross-sectional average weights, i.e. Avg$\{w_{i,t}\}$, the strategy produces a time-series average, minimum and maximum of 0.29% , 0.17% and 0.63%, respectively. For the cross-sectional maximum weights, i.e. Max$\{w_{i,t}\}$, the time-series mean of the maximum weights is low (3.05%) and that the all-time maximum weight on a single fund is only 12.41%.

Overall, the above evidence shows that the optimal portfolios do not involve extremely large weights on individual funds. Therefore, the outperformance of optimal composite strategies is unlikely to be driven by taking extreme bets on very few funds at certain times.

5.2.4 Characterization of alternative strategies: evidence from volatility states

We follow Grønborg et al. (2020) to define the aggregate risk in month t as $|\beta\sigma_{m,t}|$, where β is the average market beta across all funds and $\sigma_{m,t}$ is the realized volatility of the stock market based on daily returns. The risk measure is estimated monthly based on a 12-month rolling window. We sort all months over the entire sample period into terciles based on

aggregate risk levels. A given month is defined as high, medium or low risk state if it is in the top, middle or bottom tercile, respectively. The results are reported in Table 5.

[Table 5]

First, all strategies, active or passive, record a much better performance during periods of low volatility. With rare exceptions, performance deteriorates monotonically from low to medium to high volatility states. For the passive strategies this result is not surprising, as most periods of elevated volatility are associated with market downturns.

More relevant for our analysis, though, is that all active strategies perform very closely to the passive benchmarks during low and medium volatility times but significantly better during stretches of high volatility. The single sorted and single optimized strategies actually do not even keep on par with the broad market indices during high-vol periods, as it can be seen by comparing CERs and Sharpe ratios. No active strategy comes close to the market (i.e., S&P500) Sharpe Ratio and Certainty Equivalent Return. And none of them beats the S&P 500 by meaningful margin in medium volatility times either. It is, instead, during turbulent periods that the active strategies seem to find their edge over the passive ones, the only exception being the volatility strategy. All other single characteristic strategies (sorted or optimized) return somewhere between 100 and 650 basis points per year in certainty equivalent terms above the passive indices when volatility is high.

In comparing single sorted and single optimized strategies, we do not observe any clear pattern. Whether one looks at Sharpe ratios or at CERs, there performance is rather similar state by state, with the possible exception of the Size strategy (about 120 basis per year points better for the single optimized than the single sorted during high-vol months).

Finally, and most importantly, consider the optimized composite strategy. Although it

is still the case that its performance deteriorates, and by a large margin, moving from low to medium volatility periods, the strategy performs extremely well when volatility is high. Its Sharpe Ratio of 0.71 is significantly higher than for every other strategy and more than four times the market SR. Even more impressive is the CER during high vol times: about 10% per year and, hence, between 3.5 and 9% higher than for any active strategy that relies on a single characteristic and almost 11% higher than for all passive benchmarks. It is also noticeable that the annualized performance of the composite optimized is very comparable between high and low vol periods.

To summarize, the superior performance of the active strategies is entirely concentrated during periods of high market or economic volatility while very little difference exists between active and passive approaches during more normal times. Furthermore, it is during those turbulent periods that the optimized composite strategy is far superior to all other active strategies.

Overall, the findings about the state-dependent nature of the investment performances are reminiscent of those reported by Kosowski (2011), Glode (2011) and Grønborg et al. (2020). Our findings are also consistent with the theoretical and empirical analysis in Kacperczyk et al. (2016). In their equilibrium setup skilled fund managers have the incentive to allocate more attention to their portfolio decisions when uncertainty about asset payoffs is higher; i.e., information is most valuable about the most uncertain outcomes and when the price of risk is elevated. Skilled (or, informationally advantaged) funds take, therefore, larger positions in risky assets because they are less uncertain than unskilled investors about their payoffs: these larger positions yield, on average, higher returns when the price of risk is higher. It follows that the outperformance of skilled over unskilled investors rises in periods of higher aggregate uncertainty.

It appears that exploiting the cross-sectional predictability in performance induced by multiple fund-specific signals allows one to identify the more skilled funds that reap the benefits of their superior skill when the rewards are higher.

5.2.5 Characterization of alternative strategies: evidence from fund holdings

Having assessed large differences in performance across strategies and across economic states, we now aim to gain additional insights on the funds that each strategy selects. By using fund-level portfolio holdings and returns we can gauge along what dimensions the portfolios generated by alternative strategies differ in an economically important manner.

We, first, consider the following portfolio (i.e., strategy) level attributes: total net assets (TNA, in millions of dollars), annual expense ratio (ExpR), annual turnover ratio (Turnover) and net monthly cash flow (NetFlow, as percentage of lagged TNA). The first four columns of Table 6 reports the attributes for all the active strategies and for the average fund. For each strategy the reported figures are calculated by taking cross-sectional averages of the fund-level attributes at a given point in time, then averaging these cross-sectional means across time. For the average fund and for the single sorted strategies the cross-sectional means are calculated by equally-weighting the attributes for either all funds available at a given point in time (Panel A) or for all funds in the top decile for a given characteristic (Panel B). For the optimized strategies (Panels C and D) the individual fund attributes are given the weights generated by the utility maximization procedure. Starting with TNA, the best performing strategy (i.e., the composite optimized approach) selects funds that are typically $25 - 30\%$ smaller than the average fund as well as of the fund typically selected by the Alpha, Hot Hand and Volatility strategies, whether they rely on the sorting rule or on utility maximization. The average fund size selected by the composite optimized strategy is, on the other hand, comparable to what selected by the Activeness strategy and much larger than what generated by the Size strategy. In terms of expense ratio, the optimized composite strategy does not appear to select substantially different funds from most of the other strategies. Most expense ratios range between 1.05 (Alpha optimized strategy) and 1.30, with the composite optimized strategy (1.24) and the average fund (1.19) somewhere in the middle of the range. At the extremes stand the Volatility optimized strategy (0.96) and the sorted size strategy (1.47). The annualized turnover figures show that the composite

optimized approach selects funds with lower turnover (0.73 on average) relative to those chosen by the single sorted strategies (turnover between 0.82 and 0.89), and slightly higher turnover than funds in the single optimized strategies (between 0.66 and 0.73). The exception is the Volatility strategy, which chooses funds with much lower turnover (0.41). Funds in the optimized composite strategy receive much higher average lagged inflows (6.96% of TNA per year) than the average equity fund in the sample (0.96% per year) but comparable to those received by the funds in the Alpha and Hot Hand strategies. Still, the funds in the optimized composite portfolio do not possess high similarities with those identified by the Flow (or, smart-money) strategy, which display a net flow of about 18% per year.

[Table 6]

Overall, the funds selected by the composite strategy have smaller size, lower turnover and higher inflows than the average active fund. On one hand, the evidence that these attributes have a relevant impact on investors' utility is consistent with extant literature showing outperformance (typically measured in terms of alpha) by those types of funds.[35] On the other hand, using multiple fund-specific signals generates higher utility by not excessively concentrating the portfolio of funds along one specific dimension (e.g., high net flows or low turnover). Interestingly, fund fees, as measured by the expense ratio, do not appear to be a separating factor for a utility maximizing investor who relies on a diversified set of fund-specific information.

Next, we look at the non-parametric style exposures and characteristic selectivity measure first introduced by Daniel, Grinblatt, Titman and Wermers (DGTW 1997). We consider

[35] Among several others, see Carhart (1997), Chen et al. (2004).

exposures to size (Size), book-to-market (BTM) and momentum (MOM).[36] We calculate the characteristic selectivity measure over 3-month and 12-month periods and report its annualized value in percentage (CS_{1m} and CS_{12m}).[37] The individual fund CS measures and style exposures are averaged across funds as illustrated above for the fund attributes. We report the results in the five right-most columns of Table 6.

Starting with the CS measures, the strategy that relies on multiple fund-specific information and on utility maximization exhibits much higher levels than all other strategies. For instance, while the average fund displays an annualized CS measure of about 1% at the monthly horizon, the optimized composite strategy produces a measure of 4.22%, which is statistically and economically much larger. Although some of the active single sorted and single optimized strategies (namely, Activeness, Alpha, Hot Hand and Flow) improve upon the performance of the average fund, none of these strategies' CS measure exceeds 2.85%, leaving them substantially below the performance recorded by the optimized composite strategy. When only one source of fund-specific information is utilized, the reliance on utility maximization rather than on the sorting rule does not lead to economically appreciable improvements in the CS measures for any of the strategies, with the possible exception of the Volatility strategy at the 1-month frequency. Once more, it appears that it is the combination of multiple signals *and* a portfolio optimization approach that originate the important performance gains.

Moving to the style exposures (columns SIZE, BTM and MOM in Table 6), we can get a sense for the type of stocks that funds selected by different strategies invest in. Funds in the best performing (i.e, composite optimized) strategy hold, on average, smaller cap stocks

[36] We follow DGTW in creating 125 portfolios on June 30, after sorting all CRSP stocks, conditionally, into quintiles based on their size, book-to-market, and momentum characteristics. For a given fund exposure to, e.g., size is the quintile size portfolio number in which each stock is sorted into for a given year, weighted across all stocks held by the fund each month. The BTM and MOM style exposures of each fund are computed analogously.

[37] The characteristic adjusted return for each stock, from July 1 to June 30 of the following year, is the return on the stock minus the return on the value-weighted portfolio to which that stock belongs. The CS measure for a given fund is computed as the portfolio-weighted characteristic-adjusted return during each month of that funds existence.

than those in the single sorted strategies although the differences appear to be not very large, with the possible exception of the funds in the Volatility strategy. Funds selected by the single optimized strategies invest in stocks slightly smaller than those held by the funds in the composite strategy but, again, not importantly so. Even smaller differences across strategies emerge from book-to-market styles. Finally, the funds in the composite strategy seem to allocate less heavily to momentum stocks than funds in the single sorted portfolios and slightly more than funds in the single optimized portfolios. Overall, the unconditional (i.e., full sample) style analysis seems to point to lower exposures to size, book-to-market and momentum for the optimized strategies than for the single sorted strategies, although the differences between the composite optimize strategy and the others do no appear to be economically large

Next, we look at the evolution over time of the style exposures and plot them in Figure 3. First, it is apparent that all active strategies induce much higher changes in the exposures to SIZE, BTM and MOM than simply holding the average active fund (orange lines in the Figure). Second, the optimized strategies (blue lines for single optimized, red lines fro composite optimized) have, generally, lower and more volatile exposures than the single sorted strategies (green lines). But the differences between optimized and sorted strategies appear to be almost entirely due to two well-identifiable periods: the bear market between late 2000 and early 2003 and the Great Financial Crisis (GFC) of 2008-2009. A further look at the optimal portfolio weights in Figure 2 reveals the cause of the marked differences: during the big equity market downturns the optimized strategies reallocate to the riskless asset, leaving the investors unexposed to the market or to the style factors.[38] For most of the months outside of those downturns the differences is style exposures across strategies are relatively minor. We also compute rank correlations between the returns of each strategy and the returns of SIZE, BTM and MOM. In unreported results we find that the differences

[38] Although there are several months in which the optimized portfolios are entirely out of equities, their style exposures do not necessarily reduce to zero as, in any given month, they are computed as rolling 12-month averages of the end-of-month exposures.

across strategies are not apparent, especially when excluding the above mentioned short periods.

Given that the optimized strategies have access to the risk-free asset while the sorted ones (and, of course, the passive ones) do not and given that the optimal portfolios are characterized by periods with zero weights on risky assets, we recalculate the performance measures for all strategies excluding those months where the optimal portfolio consist of a 100% investment in the risk-free security. We find that all the main results remain essentially unchanged.[39]

We conclude that the differences in performance and, hence, in investors' utility across strategies are not meaningfully determined by style exposures and/or by the possibility of allocating (portions of) the portfolio to a safe asset. Rather, it appears that the reliance on multiple fund specific signals and on utility maximization allows the investor to choose funds which are better able to outperform their characteristic-based benchmarks, as illustrated by the results for the CS measures.

5.3 Robustness

In this section, we undertake additional checks to see how sensitive the baseline results are to changes in several choices made throughout the empirical analysis presented above.

5.3.1 Number of funds

In the baseline results, a single characteristic sorted strategy allows the investor to recursively take equal-weighted long positions in the top 10% funds, ranked by one of their characteristics. By doing so each portfolio in the sorting-based strategies contains about 160 funds on average over the entire out-of-sample period. The fund universe for the single-characteristic optimized strategies is, then, the same top 10%, while the composite optimized strategy is based on the union of the six top 10% sets. For robustness, we now change this assump-

[39]These results are reported in an online Appendix and are available from the authors upon request.

tion and let the investor pick the top 5% and 20% funds instead. As a result, each sorted portfolio contains about 80 and, respectively, 300 funds on average. Table 7 provides the out-of-sample investment results under those two different choices of the sorting quantiles.

[Table 7]

The figures from performance measures show that reducing the initial investable universe boosts investment outcomes, as evidenced by higher and more significant CERs as well as higher Sharpe ratio and 4-Factor alpha. Interestingly, we can also observe that the performance increases associated with reduced-size portfolios seem to be more pronounced for the optimal composite strategy.

Similar considerations arise from the results based on the top 20% of funds although the CER spreads in favor of the optimal composite strategy are not as large.

Overall, the reported values show that the superior performance of the optimal composite strategy still remains economically large relative to the alternatives, confirming that the baseline results are robust to different initial choices about the numbers of funds available for forming the investor's portfolios.

5.3.2 Sub-period Analysis

The out-of-sample investment period in the previous analysis (1996-2015) covers very different market conditions: the bull market of 1990s, the Asian financial crisis and Dot-Com bust in the earlier years; the Great Financial Crisis (GFC) crisis and European debt crises in the later period. In addition, the universe of investable active equity funds has undergone a dramatic growth from the early to the later portion of the sample, as shown in Figure

1. It is important to assess whether our main results hold across such different time periods. To do so, we split the entire out-of-sample period into two spans, 1996:01–2004:12 and 2005:01–2015:12, and report the results on Table 8

[Table 8]

Looking at the first investment sub-period (left-hand side of the table), we find that the active strategies still outperform the passive ones by a wide margin. Compared to the full sample results, the comparison between single sorted and singe optimized strategies is more mixed: it is no longer the case that those based on portfolio optimization always produce the better performances. Indeed, only the Hot Hand and, to a lesser extent the Size and Alpha strategies seem to benefit from using the utility-based approach. More importantly, though, it is still the case that our main object of interest, the optimal composite strategy, produces remarkably larger economic gains and substantially outperforms all alternative strategies in both Sharpe ratios and CERs.

Turning to the second sub-period (right-hand side of the table), we find that all mutual fund strategies deliver significantly lower performance than during the first sub-period, possibly reflecting the large market-wide losses during the GFC. The optimized strategies appear to significantly outperform the passive ones as well as the sorted ones. For instance, the optimal composite strategy yields a CER that is about 230 basis points per year higher than those of both the broad market index and the best performing single sorted strategy (which, for this second sub-period turned out to be the Vol. strategy). So, the composite strategy dominates the passive indices, all the sorting based strategies and the naive composite approach, although it does not significantly beat all the single optimized approaches.

Overall, the sub-period analysis confirms the main message from the full period: namely,

relying on multiple fund-specific signals *and* on an optimal portfolio approach consistently outperforms passive strategies as well as active strategies based on a sorting or a 1/N-type rule.

5.3.3 Risk Aversion

We reassess the relative performances of all strategies for the full sample period by varying the assumed degree of investors' risk aversion. Specifically, we rerun the analysis with coefficient of relative risk aversion $= 2, 3, 4, 6, 7$ and 10. We do not find any appreciable differences in our main findings and, hence, takeaways.[40]

6 Conclusion

We revisit the long standing issue of whether investing in the universe of actively managed mutual funds may be value enhancing for an investor. Relying on a large sample of U.S. domestic no-load equity funds, we employ a utility-based metric to construct optimal portfolios that exploit the cross-sectional predictability in performance coming from multiple fund characteristics as well as from economy-wide information. We find that the resulting optimal composite strategy significantly outperforms *out-of-sample* and after fees and expenses (i) a set of passive investments represented by low-cost index funds, (ii) active strategies that rely on sorting-based trading rules to utilize the predictive power of a single characteristic, (iii) active strategies that rely on an optimal portfolio approach but exploit only the predictive power of a single characteristic, (iv) a naïve active strategy that equally weighs single characteristic sorted portfolios. We also find that, once fund-specific information is accounted for, economy-wide and market-wide information have a rather limited impact on investors' utility. We further show that the optimal composite strategy generates most of, if not all, its utility gains during periods of high market or economic volatility, especially when compared to the passive alternatives. Finally, an analysis based on fund-level portfolio

[40] These results are also reported in an online Appendix and are available from the authors upon request.

holdings demonstrates that the outperformance of the optimal composite strategy can be largely attributed to its ability to identify funds with superior selectivity skills with respect to their (characteristic-based) benchmarks.

The findings shed, first, new light on the value of investing in active equity funds relatively to passive vehicles. We show that investors may derive substantial utility gains, of the order of $4-5\%$ on an annual basis for a moderate degree of risk aversion, from allocations to active funds that exploit predictability in performance implied by fund-level characteristics. Second, the results demonstrate the economic advantages of exploiting the joint predictive power of multiple fund attributes relatively to trading rules based on a single characteristic. Although some literature discounts the role of those fund-level attributes by pointing out that they are subsumed by fund exposures to assumed benchmarks, we find that the attributes do contain value-adding information from a utility perspective. Third, the significantly improved performance of the optimized strategies relatively to simpler approaches, such as the sorting-based rule or the 1/N policy, offers support to portfolio theory, whose relevance has been questioned from an applied perspective. Fourth, the results lend further empirical verification to the model of mutual funds' attention allocation recently proposed by Kacperczyk et al. (2016), suggesting not only that more skilled funds seem to allocate extra attention to their investment decisions during relatively high volatile periods, but also that there is a potential for investors to reap some benefits from such higher attention and skill.

This paper suggests several avenues for future research. First, the framework proposed in the present paper could be applied to other mutual fund classes (e.g., fixed income, non-U.S. equities, commodities) and other active investment vehicles (e.g., hedge funds). In particular, an analysis of portfolio strategies involving mutual funds across asset classes would fill a significant void and could offer further insights into the value of active management. In addition, this study could also be extended to include a multi-period investment objective, within which the time-varying nature of fund performance predictability might play an even

larger role. Such extensions could be of interest to, among others, target date funds.

References

Amihud, Y., and R. Goyenko. 2013. Mutual fund's R2 as predictor of performance. *Review of Financial Studies* p. hhs182.

Ang, A., and G. Bekaert. 2002. International asset allocation with regime shifts. *Review of Financial studies* 15:1137–1187.

Arditti, F. D. 1967. Risk and the required return on equity. *Journal of Finance* pp. 19–36.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen. 2013. Value and momentum everywhere. *The Journal of Finance* 68:929–985.

Avramov, D., and R. Wermers. 2006. Investing in mutual funds when returns are predictable. *Journal of Financial Economics* 81:339–377.

Back, K., A. Crane, and K. Crotty. 2018. Skewness Consequences of Seeking Alpha. *Review of Financial Studies* 32:4720–4761.

Baks, K. P., A. Metrick, and J. Wachter. 2001. Should investors avoid all actively managed mutual funds? A study in Bayesian performance evaluation. *The Journal of Finance* 56:45–85.

Banegas, A., B. Gillen, A. Timmermann, and R. Wermers. 2013. The cross section of conditional mutual fund performance in European stock markets. *Journal of Financial Economics* 108:699–726.

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

Barroso, P., and P. Santa-Clara. 2015. Beyond the carry trade: Optimal currency portfolios. *Journal of Financial and Quantitative Analysis* 50:1037–1056.

Bekaert, G., C. Erb, C. Harvey, and T. Viskanta. 1998. Distributional characteristics of emerging market returns and asset allocation. *The Journal of Portfolio Management* 24:102–116.

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

Bianchi, D., and M. Guidolin. 2014. Can long-run dynamic optimal strategies outperform fixed-mix portfolios? Evidence from multiple data sets. *European Journal of Operational Research* 236:160–176.

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

Brands, S., and D. R. Gallagher. 2005. Portfolio selection, diversification and fund-of-funds: a note. *Accounting & Finance* 45:185–197.

Brandt, M. 2009. Portfolio choice problems. *Handbook of financial econometrics* 1:269–336.

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies* p. hhp003.

Bredendiek, M., G. Ottonello, and R. I. Valkanov. 2016. Corporate Bond Portfolios and Macroeoconomic Conditions. *Available at SSRN: http://dx.doi.org/10.2139/ssrn.2886825* .

Brown, S. J., and W. N. Goetzmann. 1997. Mutual fund styles. *Journal of financial Economics* 43:373–399.

Brown, S. J., I. Hwang, and F. In. 2013. Why optimal diversification cannot outperform naive diversification: Evidence from tail risk exposure. *Available at SSRN* 2242694.

Busse, J. A. 1999. Volatility timing in mutual funds: Evidence from daily returns. *Review of Financial Studies* 12:1009–1041.

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

Cenesizoglu, T., and A. Timmermann. 2012. Do return prediction models add economic value? *Journal of Banking & Finance* .

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

Chen, J., H. Hong, M. Huang, and J. D. Kubik. 2004. Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review* pp. 1276–1302.

Cremers, K. M., J. A. Fulkerson, and T. B. Riley. 2019. Challenging the Conventional Wisdom on Active Management: A Review of the Past 20 Years of Academic Literature on Actively Managed Mutual Funds. *Financial Analysts Journal* 75:8–35.

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

Cremers, K. M., A. Petajisto, and E. Zitzewitz. 2012. Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation. *Critical Finance Review* 2:1–48.

Cremers, M., A. Petajisto, E. Zitzewitz, et al. 2013. Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation. *Critical Finance Review* 2:001–048.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance* 52:1035–1058.

Del Guercio, D., and J. Reuter. 2014. Mutual fund performance and the incentive to generate alpha. *The Journal of Finance* 69:1673–1704.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22:1915–1953.

diBartolomeo, D., and E. Witkowski. 1997. Mutual fund misclassification: Evidence based on style analysis. *Financial Analysts Journal* pp. 32–43.

Elton, E. J., M. J. Gruber, and C. R. Blake. 2006. The adequacy of investment choices offered by 401 (k) plans. *Journal of Public Economics* 90:1299–1314.

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

Farinelli, S., and L. Tibiletti. 2008. Sharpe thinking in asset ranking with one-sided measures. *European Journal of Operational Research* 185:1542–1547.

Geczy, C., R. F. Stambaugh, and D. Levin. 2005. Investing in socially responsible mutual funds. *Available at SSRN 416380* .

Gilchrist, S., and E. Zakrajšek. 2012. Credit spreads and business cycle fluctuations. *American Economic Review* 102:1692–1720.

Glode, V. 2011. Why mutual funds underperform. *Journal of Financial Economics* 99:546–559.

Glode, V., B. Hollifield, M. T. Kacperczyk, and S. Kogan. 2012. Time-Varying Predictability in Mutual Fund Returns. *Available at SSRN 1364351* .

Goetzmann, W. N., and R. G. Ibbotson. 1994. Do winners repeat? *The Journal of Portfolio Management* 20:9–18.

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

Grinblatt, M., S. Titman, and R. Wermers. 1995. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American economic review* pp. 1088–1105.

Grønborg, N. S., A. Lunde, A. Timmermann, and R. Wermers. 2020. Picking funds with confidence. *Journal of Financial Economics* .

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

Guidolin, M., and A. Timmermann. 2008. International asset allocation under regime switching, skew, and kurtosis preferences. *Review of Financial Studies* 21:889–935.

Harvey, C. R., J. C. Liechty, M. W. Liechty, and P. Müller. 2010. Portfolio selection with higher moments. *Quantitative Finance* 10:469–485.

Harvey, C. R., and A. Siddique. 1999. Autoregressive conditional skewness. *Journal of financial and quantitative analysis* 34.

Hendricks, D., J. Patel, and R. Zeckhauser. 1993. Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988. *The Journal of Finance* 48:93–130.

Jensen, M. C. 1968. The performance of mutual funds in the period 1945–1964. *The Journal of finance* 23:389–416.

Jiang, G. J., T. Yao, and T. Yu. 2007. Do mutual funds time the market? Evidence from portfolio holdings. *Journal of Financial Economics* 86:724–758.

Jiang, H., M. Verbeek, and Y. Wang. 2014. Information content when mutual funds deviate from benchmarks. *Management Science* 60:2038–2053.

Jordan, B. D., and T. B. Riley. 2015. Volatility and mutual fund manager skill. *Journal of Financial Economics* 118:289–298.

Jurado, K., S. C. Ludvigson, and S. Ng. 2015. Measuring uncertainty. *American Economic Review* 105:1177–1216.

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

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

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp. 2013. Time-Varying Fund Manager Skill. *The Journal of Finance* .

Kacperczyk, M., S. van Nieuwerburgh, and L. Veldkamp. 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84:571–626.

Kosowski, R. 2011. Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions. *The Quarterly Journal of Finance* 1:607–664.

Kosowski, R., A. Timmermann, R. Wermers, and H. White. 2006. Can mutual fund stars really pick stocks? New evidence from a bootstrap analysis. *The Journal of finance* 61:2551–2595.

Kraus, A., and R. H. Litzenberger. 1976. Skewness preference and the valuation of risk assets*. *The Journal of Finance* 31:1085–1100.

Malkiel, B. G. 1995. Returns from investing in equity mutual funds 1971 to 1991. *The Journal of Finance* 50:549–572.

McLemore, P. 2016. Do mutual funds have decreasing returns to scale? evidence from fund mergers. *Journal of Financial and Quantitative Analysis* pp. 1–51.

Moskowitz, T. J. 2000. Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses: Discussion. *The Journal of Finance* 55:1695–1703.

Pástor, L., and R. F. Stambaugh. 2002. Investing in equity mutual funds. *Journal of Financial Economics* 63:351–380.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2015. Scale and skill in active management. *Journal of Financial Economics* 116:23–45.

Patton, A. 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics* 2:130–168.

Plazzi, A., W. Torous, and R. Valkanov. 2010. Expected returns and expected growth in rents of commercial real estate. *Review of Financial Studies* p. hhq069.

Scott, R. C., and P. A. Horvath. 1980. On the direction of preference for moments of higher order than the variance. *Journal of Finance* pp. 915–919.

Sun, Z., A. Wang, and L. Zheng. 2009. Do active funds perform better in down markets? New evidence from cross-sectional study. *New Evidence from Cross-Sectional Study (September 1, 2009)* .

Timmermann, A. 2006. Forecast combinations. *Handbook of economic forecasting* 1:135–196.

Wermers, R. 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *The Journal of Finance* 55:1655–1703.

Yan, X. S. 2008. Liquidity, investment style, and the relation between fund size and fund performance. *Journal of Financial and Quantitative Analysis* 43:741–767.

Zheng, L. 1999. Is money smart? A study of mutual fund investors' fund selection ability. *The Journal of Finance* 54:901–933.

Table 1. Fund performance predictability by fund characteristics

This table presents the performance of decile mutual fund portfolios that are formed based on rankings of each individual fund characteristic over the period of January 1995–December 2015. Decile 10 represents the "best-performing" portfolio implied by the corresponding fund characteristic, and Decile 1 is the "worst-performing" portfolio. 10-1 shows the differences in performance between the best-performing and worst-performing portfolios. We report three performance metrics: the four-factor alpha (Alpha), Sharpe ratio (Sharpe) and the certainty equivalent excess return (CER) based on the CRRA utility function with a relative risk aversion coefficient of five (RRA=5). All metrics are reported in annualized terms. Figures in parentheses are bootstrapped p-values.

	Activeness			Alpha			Flow		
Decile	Alpha(%)	Sharpe	CER(%)	Alpha(%)	Sharpe	CER(%)	Alpha(%)	Sharpe	CER(%)
10	0.54	0.57	2.42	1.49	0.56	1.30	0.79	0.58	2.24
	(0.142)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.025)	(0.000)	(0.000)
1	-1.29	0.41	-0.61	-2.08	0.39	-1.41	-1.08	0.42	-0.61
	(0.000)	(0.000)	(0.016)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.066)
10-1	1.84	0.33	1.14	3.57	0.55	2.60	1.87	0.44	1.84
	(0.000)	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

	Hot-hand			Size			Volatility		
Decile	Alpha(%)	Sharpe	CER(%)	Alpha(%)	Sharpe	CER(%)	Alpha(%)	Sharpe	CER(%)
10	1.52	0.63	2.58	0.17	0.52	1.67	-0.10	0.53	2.49
	(0.000)	(0.000)	(0.000)	(0.701)	(0.000)	(0.000)	(0.546)	(0.000)	(0.000)
1	-2.44	0.22	-5.55	-0.83	0.45	0.08	-2.35	0.33	-7.55
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.894)	(0.000)	(0.000)	(0.000)
10-1	3.96	0.52	2.49	1.00	0.23	0.48	2.26	-0.08	-10.44
	(0.000)	(0.000)	(0.000)	(0.012)	(0.058)	(0.297)	(0.000)	(0.000)	(0.000)

Table 2. Rank Correlations of Fund Characteristics

This table reports percentile rank correlation matrix of the six fund characteristics. The sample period is from January 1992December 2015. At the end of each month, we independently compute each fund's percentile ranks based on the six characteristics. The correlation matrix is computed using the six percentile ranks for each month and then averaged across the sample period. Figures in parentheses are p-values.

	Activeness	Alpha	Flow	Hot-hand	Size
Alpha	0.04				
	(0.689)				
Flow	0.02	0.34			
	(0.881)	(0.000)			
Hot-hand	0.02	0.45	0.33		
	(0.927)	(0.020)	(0.000)		
Size	0.13	-0.06	0.02	-0.05	
	(0.139)	(0.425)	(0.697)	(0.620)	
Volatility	-0.04	0.04	0.05	-0.02	0.02
	(0.799)	(0.829)	(0.712)	(0.958)	(0.814)

Table 3. Out-of-sample performance of mutual fund portfolio strategies

This table reports the out-of-sample performance of mutual fund portfolio strategies over the period of January 1995–December 2015. For each strategy, we report the rst four sample moments of realized excess returns (Mean, StdDev, Skew and Kurt), and three performance metrics: the 4-Factor alpha (Alpha), Sharpe ratio (Sharpe) and certainty equivalent excess returns (CER) based on the CRRA utility function with a relative risk aversion coefficient of five (RRA=5). All metrics are reported in annualized terms. Specifically, Panel A shows the results for four passive funds covering different segments of the equity market, Panel B shows the results for six single characteristic sorted strategies, each of which takes equal-weighted long positions in the top 10% funds sorted on a specic lagged characteristic measure; Panel C displays the results for six single characteristic optimized strategies formed based on a single characteristic predictive signal and using the methodology specied in Section 3.1; Panel D reports the results for composite strategies which combine predictive signals from six fund characteristics. Figures in parentheses are the bootstrapped p-values under the null hypothesis that the performance of the competing strategy is equal to that of the composite optimized strategy.

Portfolio Strategy	Mean(%)	Std(%)	Skew	Kurt	Alpha(%)	Sharpe	CER(%)
Panel A: Passive							
S&P 500 Index	7.67	15.05	-0.67	4.10	0.32	0.51	1.64
Mid-Cap Index	8.97	19.38	-0.63	4.48	-0.56	0.46	-1.31
Small-Cap Index	8.96	19.41	-0.51	4.37	-0.74	0.46	-1.20
Value Index	7.51	15.27	-0.75	4.58	0.15	0.49	1.23
Panel B: Single Sorted							
Activeness	8.76	15.27	-0.81	5.11	0.54	0.57	2.42
	(0.000)	(0.000)	(0.057)	(0.000)	(0.000)	(0.000)	(0.000)
Alpha	10.29	18.43	-0.40	4.56	1.49	0.56	1.30
	(0.000)	(0.000)	(0.798)	(0.004)	(0.000)	(0.000)	(0.000)
Flow	9.86	17.00	-0.42	4.70	0.79	0.58	2.24
	(0.000)	(0.000)	(0.218)	(0.000)	(0.000)	(0.000)	(0.000)
Hot-hand	12.34	19.61	0.11	6.06	1.52	0.63	2.57
	(0.000)	(0.000)	(0.600)	(0.000)	(0.000)	(0.000)	(0.000)
Size	8.07	15.41	-0.75	4.47	0.17	0.52	1.67
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Volatility	6.65	12.50	-0.81	4.82	-0.10	0.53	2.49
	(0.000)	(0.000)	(0.000)	(0.381)	(0.000)	(0.000)	(0.000)
Panel C: Single Optimized							
Activeness	7.22	12.23	-0.72	5.39	1.70	0.59	3.26
	(0.000)	(0.000)	(0.411)	(0.124)	(0.004)	(0.007)	(0.010)
Alpha	9.58	16.53	-0.19	5.61	2.96	0.58	2.50
	(0.031)	(0.000)	(0.008)	(0.001)	(0.029)	(0.000)	(0.000)
Flow	9.17	14.78	-0.12	5.76	2.36	0.62	3.58
	(0.028)	(0.222)	(0.006)	(0.002)	(0.028)	(0.018)	(0.041)
Hot-hand	12.13	18.56	0.47	7.39	3.50	0.65	3.66
	(0.285)	(0.000)	(0.000)	(0.000)	(0.063)	(0.006)	(0.014)
Size	7.83	12.32	-0.61	5.36	2.76	0.64	3.83
	(0.000)	(0.000)	(0.481)	(0.034)	(0.016)	(0.023)	(0.026)
Volatility	5.83	9.98	-0.63	5.30	1.33	0.58	3.25
	(0.000)	(0.000)	(0.711)	(0.050)	(0.008)	(0.008)	(0.001)
Panel D: Composite							
Naïve	8.60	15.69	-0.71	4.43	0.15	0.55	1.98
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimized	13.03	15.81	0.03	5.04	5.19	0.82	6.75
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Table 4. Out-of-sample performance of mutual fund portfolio strategies: incorporating macroeconomic information

This table reports the out-of-sample performance of mutual fund portfolio strategies conditioning on two macroeconomic states. The macroeconomic states are defined based on three variables: 1. the credit spread index (GZ Spread) of Gilchrist and Zakrajšek (2012); 2. the macroeconomic uncertainty index (MU) of Jurado et al. (2015); 3. the CBOE volatility index (VIX). The evaluation period is from January 1995 to December 2015. For each strategy, we report three performance metrics: the 4-Factor alpha (Alpha), Sharpe ratio (Sharpe) and certainty equivalent excess returns (CER) based on the CRRA utility function with a relative risk aversion coefficient of five (RRA=5). All metrics are reported in annualized terms. Specifically, Panel A shows the performance of the unconditional composite optimized strategy which combine predictive signals from six fund characteristics based on optimization. Panel B presents the performance of three conditional composite optimized strategies based on the three macroeconomic indicators. Figures in parentheses are the bootstrapped p-values under the null hypothesis that the performance of the competing strategy is equal to that of the composite optimized strategy.

Portfolio Strategy	Mean(%)	Std(%)	Skew	Kurt	Alpha(%)	Sharpe	CER(%)
Panel A: Unconditional							
Composite optimized	13.03	15.81	0.03	5.04	5.19	0.82	6.75
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Panel B: Conditional							
GZ Spread	13.32	15.64	0.11	5.24	5.58	0.85	7.22
	(0.865)	(0.879)	(0.742)	(0.778)	(0.796)	(0.898)	(0.919)
MU	12.99	15.74	0.04	5.09	5.18	0.83	6.78
	(0.768)	(0.395)	(0.732)	(0.670)	(0.510)	(0.481)	(0.518)
VIX	12.83	15.90	-0.01	4.60	5.34	0.81	6.47
	(0.955)	(0.618)	(0.610)	(0.737)	(0.828)	(0.773)	(0.838)

Table 5. Out-of-sample performance of mutual fund portfolio strategies by risk states

This table reports the out-of-sample performance of mutual fund investment strategies in high, medium and low aggregate market risk states over the period of January1995–December 2015. We follow Grønborg et al. (2020) to define the aggregate risk in month t as $|\beta\sigma_{m,t}|$, where β is the average market beta across all funds and $\sigma_{m,t}$ is the realized volatility of the stock market based on daily returns. The risk measure is estimated monthly based on a 12-month rolling window. We sort all months over the entire sample period into terciles based on aggregate risk levels. A given month is defined as high, medium or low risk state if it is in the top, middle or bottom tercile, respectively. We report three performance metrics: the 4-Factor alpha (Alpha), Sharpe ratio (Sharpe) and certainty equivalent excess returns (CER) based on the CRRA utility function with a relative risk aversion coefficient of five (RRA=5). All metrics are annualized, and Alpha and CER are expressed as percentages. Figures in parentheses are the bootstrapped p-values under the null hypothesis that the performance of the competing strategy is equal to that of the composite optimized strategy.

	High Risk			Medium Risk			Low Risk		
Portfolio Strategy	Alpha(%)	Sharpe	CER(%)	Alpha(%)	Sharpe	CER(%)	Alpha(%)	Sharpe	CER(%)
Panel A: Passive									
S&P 500 Index	0.54	0.16	-0.60	-0.40	0.80	6.24	-0.47	1.62	11.94
Mid-Cap Index	-0.98	0.22	-0.70	0.79	0.77	5.98	0.61	1.19	10.41
Small-Cap Index	-1.48	0.26	0.21	0.40	0.70	5.04	0.05	1.04	8.97
Value Index	-0.20	0.13	-1.18	-0.98	0.60	3.82	-0.05	1.79	13.13
Panel B: Single Sorted									
Activeness	1.99	0.38	3.41	-2.17	0.57	3.41	-0.14	1.22	9.57
	(0.000)	(0.000)	(0.000)	(0.067)	(0.000)	(0.000)	(0.002)	(0.003)	(0.000)
Alpha	2.87	0.37	3.09	0.43	0.84	6.90	1.24	1.12	9.49
	(0.000)	(0.000)	(0.000)	(0.284)	(0.000)	(0.000)	(0.019)	(0.000)	(0.000)
Flow	1.58	0.36	2.97	-0.51	0.80	6.11	0.74	1.31	10.74
	(0.000)	(0.000)	(0.000)	(0.184)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hot-hand	4.16	0.49	6.18	-3.93	0.74	5.50	0.74	1.11	9.81
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Size	0.58	0.25	1.10	-0.99	0.68	4.76	0.10	1.33	10.50
	(0.000)	(0.000)	(0.000)	(0.603)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Volatility	-0.31	0.19	0.53	0.29	0.83	6.05	-0.75	1.53	9.95
	(0.000)	(0.000)	(0.000)	(0.605)	(0.000)	(0.000)	(0.000)	(0.095)	(0.000)
Panel C: Single Optimized									
Activeness	1.87	0.31	2.37	-1.34	0.55	3.17	-0.14	1.23	9.44
	(0.017)	(0.021)	(0.009)	(0.809)	(0.000)	(0.000)	(0.017)	(0.251)	(0.024)
Alpha	3.63	0.35	3.04	0.32	0.84	6.96	1.44	1.10	9.40
	(0.037)	(0.003)	(0.008)	(0.098)	(0.987)	(0.721)	(0.987)	(0.000)	(0.019)
Flow	2.58	0.36	3.23	-1.05	0.76	5.72	0.57	1.29	10.61
	(0.039)	(0.018)	(0.020)	(0.981)	(0.135)	(0.145)	(0.049)	(0.888)	(0.764)
Hot-hand	5.93	0.51	6.68	-4.51	0.71	5.15	1.21	1.08	9.73
	(0.041)	(0.019)	(0.040)	(0.000)	(0.008)	(0.020)	(0.571)	(0.000)	(0.037)
Size	2.61	0.31	2.38	-0.99	0.68	4.76	0.10	1.33	10.50
	(0.024)	(0.009)	(0.003)	(0.875)	(0.001)	(0.002)	(0.003)	(0.426)	(0.489)
Volatility	-0.17	0.13	0.22	0.35	0.84	6.02	-0.76	1.53	9.83
	(0.000)	(0.000)	(0.000)	(0.009)	(0.702)	(0.078)	(0.000)	(0.000)	(0.000)
Panel D: Composite									
Naïve	0.84	0.30	1.99	-0.96	0.75	5.55	-0.10	1.28	10.07
	(0.000)	(0.000)	(0.000)	(0.060)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimized	8.75	0.71	10.06	-2.00	0.76	5.81	1.23	1.24	10.44
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Table 6. Attributes and Styles of mutual fund portfolio strategies

This table reports two sets of portfolio attributes of mutual fund strategies. The first set includes fund size (TNA) measured in million dollars, expense ratio (ExpR, annual), turnover ratio (Turnover, annual) and net cash inflow (Netflow, as percentage of lagged TNA). For each strategy the reported figures are calculated by taking cross-sectional averages of the fund-level attributes at a given point in time, then averaging these cross-sectional means across time. The second set includes the holdings based style exposures to size, value and momentum (SIZE, BTM and MOM) and Characteristic Selectivity (CS) measures at 3- and 12-month horizons. The latter set follows Daniel, Grinblatt, Titman, and Wermers (1997).

Portfolio Strategy	TNA($M)	ExpR(%)	Turnover	Netflow(%)	SIZE	BTM	MOM	CS_{3m}(%)	CS_{12m}(%)
Panel A: Passive									
The average fund	1284.83	1.19	0.81	0.96	3.91	2.96	3.18	0.98	1.31
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: Single Sorted									
Activeness	970.30	1.35	0.89	1.74	3.68	3.04	3.13	1.93	2.09
	(0.695)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Alpha	1332.84	1.29	0.84	4.85	3.44	2.90	3.31	2.17	1.77
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Flow	570.94	1.25	0.82	12.35	3.75	2.97	3.24	2.04	1.75
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hot-hand	1311.53	1.25	0.84	5.17	3.52	2.94	3.39	2.85	1.94
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.060)	(0.000)
Size	30.85	1.47	0.87	2.75	3.93	2.98	3.13	1.05	1.63
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Volatility	1261.39	1.17	0.60	1.24	4.31	3.06	2.95	0.52	0.85
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Single Optimized									
Activeness	938.93	1.22	0.71	1.78	2.98	2.50	2.54	2.01	2.09
	(0.547)	(0.493)	(0.701)	(0.000)	(0.000)	(0.014)	(0.000)	(0.000)	(0.000)
Alpha	1305.52	1.10	0.71	5.60	2.58	2.22	2.60	2.21	1.54
	(0.000)	(0.000)	(0.196)	(0.005)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Flow	531.90	1.05	0.63	18.13	2.74	2.19	2.41	1.80	1.53
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hot-hand	1268.90	1.12	0.73	6.33	2.83	2.39	2.80	2.82	1.82
	(0.000)	(0.000)	(0.006)	(0.592)	(0.000)	(0.000)	(0.710)	(0.080)	(0.000)
Size	33.17	1.24	0.66	3.07	2.92	2.23	2.35	1.09	1.43
	(0.000)	(0.774)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Volatility	1254.17	0.96	0.41	1.30	3.24	2.35	2.19	0.80	0.80
	(0.000)	(0.000)	(0.000)	(0.000)	(0.509)	(0.000)	(0.000)	(0.000)	(0.000)
Panel D: Composite									
Naïve	925.85	1.27	0.80	3.21	3.82	2.98	3.18	1.42	1.54
	(0.493)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimized	914.07	1.24	0.73	6.96	3.07	2.55	2.84	4.22	3.11
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Table 7. Out-of-sample performance of mutual fund portfolio strategies: alternative investment universes

This table reports the out-of-sample performance of mutual fund investment strategies based on alternative investment universes. To form the investment universe, we sort all funds into 5 (20) portfolios based on each of the six fund characteristics and then select all funds that are in at least one of the six top groups. We report three performance metrics: the 4-Factor alpha (Alpha), Sharpe ratio (Sharpe) and certainty equivalent excess returns (CER) based on the CRRA utility function with a relative risk aversion coefficient of five (RRA=5). All metrics are annualized, and Alpha and CER are expressed as percentages. Figures in parentheses are the bootstrapped p-values under the null hypothesis that the performance of the competing strategy is equal to that of the composite optimized strategy.

	Top 20%			Top 5%		
Portfolio Strategy	Alpha	Sharpe	CER	Alpha	Sharpe	CER
Panel A: Passive						
S&P 500 Index	0.32	0.51	1.64	0.32	0.51	1.64
Mid-Cap Index	-0.56	0.46	-1.31	-0.56	0.46	-1.31
Small-Cap Index	-0.74	0.46	-1.20	-0.74	0.46	-1.20
Value Index	0.15	0.49	1.23	0.15	0.49	1.23
Panel B: Single Sorted						
Activeness	0.64	0.57	2.22	-0.38	0.51	1.50
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Alpha	0.68	0.54	1.22	1.87	0.56	0.97
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Flow	0.32	0.56	1.86	0.97	0.59	2.40
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hot-hand	0.89	0.61	2.50	1.61	0.62	2.07
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Size	0.21	0.53	1.64	0.11	0.52	1.69
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Volatility	-0.03	0.54	2.46	0.19	0.55	2.87
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Single Optimized						
Activeness	0.17	0.48	1.51	1.91	0.58	3.09
	(0.047)	(0.005)	(0.006)	(0.016)	(0.002)	(0.002)
Alpha	2.18	0.54	2.05	2.39	0.59	2.92
	(0.156)	(0.000)	(0.000)	(0.009)	(0.004)	(0.007)
Flow	1.12	0.55	2.22	-0.07	0.47	1.05
	(0.185)	(0.011)	(0.026)	(0.016)	(0.010)	(0.020)
Hot-hand	2.47	0.61	3.11	3.76	0.66	3.53
	(0.230)	(0.001)	(0.003)	(0.081)	(0.018)	(0.016)
Size	3.14	0.66	4.16	3.25	0.69	4.54
	(0.132)	(0.014)	(0.017)	(0.012)	(0.030)	(0.026)
Volatility	2.03	0.63	3.72	1.54	0.60	3.32
	(0.006)	(0.003)	(0.000)	(0.001)	(0.004)	(0.000)
Panel D: Composite						
Naïve	-0.14	0.53	1.57	0.28	0.56	2.22
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimized	4.91	0.81	6.54	7.39	0.89	8.30
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Table 8. Out-of-sample performance of mutual fund portfolio strategies: subperiods

This table reports the out-of-sample performance of mutual fund investment strategies in subperiods. We consider two sample periods: January 1995 : December 2004 and January 2005 : December 2015. For each subperiod, we evaluate the performance of all portfolio strategies and report three performance metrics: 4-Factor alpha (Alpha), Sharpe ratio (Sharpe) and certainty equivalent excess returns (CER) based on the CRRA utility function with a relative risk aversion coefficient of five (RRA=5). All metrics are annualized, and Alpha and CER are expressed as percentages. Figures in parentheses are the bootstrapped p-values under the null hypothesis that the performance of the competing strategy is equal to that of the composite optimized strategy.

	Jan. 1995–Dec. 2004			Jan. 2005–Dec. 2015		
Portfolio Strategy	Alpha	Sharpe	CER	Alpha	Sharpe	CER
Panel A: Passive						
S&P 500 Index	0.99	0.56	2.38	-0.16	0.46	0.98
Mid-Cap Index	-1.88	0.48	-1.79	0.40	0.45	-0.86
Small-Cap Index	-3.37	0.50	-0.74	0.44	0.43	-1.61
Value Index	-0.47	0.57	2.42	-0.07	0.41	0.16
Panel B: Single Sorted						
Activeness	1.50	0.88	7.10	-2.06	0.31	-1.76
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Alpha	3.20	0.68	2.79	-0.35	0.43	-0.04
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Flow	1.86	0.75	4.98	-0.97	0.40	-0.24
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hot-hand	3.76	0.82	6.00	-1.23	0.41	-0.50
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Size	0.93	0.70	4.41	-1.44	0.36	-0.78
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Volatility	-1.29	0.66	4.16	-0.68	0.42	0.97
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel C: Single Optimized						
Activeness	-1.38	0.71	4.61	0.97	0.46	2.04
	(0.001)	(0.006)	(0.003)	(0.547)	(0.144)	(0.266)
Alpha	5.36	0.70	3.55	0.08	0.44	1.55
	(0.356)	(0.001)	(0.001)	(0.248)	(0.049)	(0.110)
Flow	3.16	0.75	5.32	0.80	0.48	2.01
	(0.098)	(0.004)	(0.015)	(0.599)	(0.280)	(0.380)
Hot-hand	7.50	0.89	7.77	-1.48	0.35	-0.02
	(0.024)	(0.053)	(0.039)	(0.095)	(0.023)	(0.041)
Size	2.42	0.76	5.33	1.40	0.50	2.49
	(0.498)	(0.023)	(0.012)	(0.696)	(0.281)	(0.427)
Volatility	-1.77	0.58	3.21	1.83	0.59	3.28
	(0.000)	(0.000)	(0.000)	(0.372)	(0.627)	(0.510)
Panel D: Composite						
Naïve	0.51	0.72	4.69	-1.17	0.39	-0.45
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Optimized	5.96	1.01	10.20	1.69	0.59	3.32
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

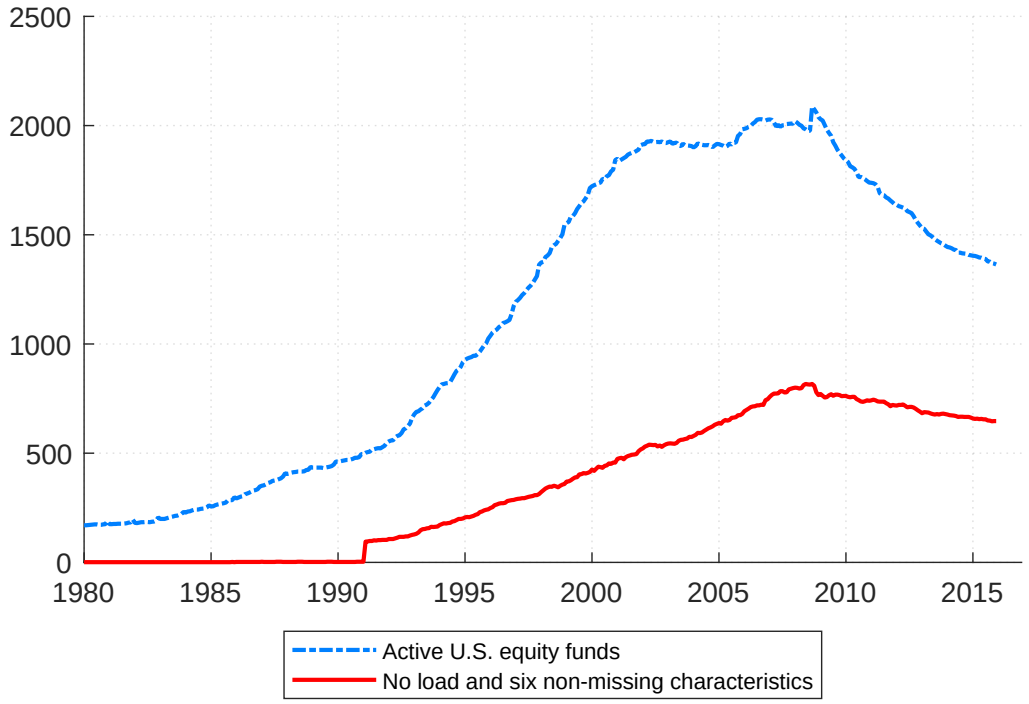

Figure 1. Sample property: the number of funds

This figure shows how the number of funds included in the two sample sets changes over time. The dashed (blue) line represents the number of active U.S. equity mutual funds. The solid (red) line shows the number of the funds charging no-load fees and with all six non-missing characteristic measures.

Figure 2. Properties of composite optimized strategy

This figure displays portfolio properties of the composite optimized strategy. Panel A plots the number of holdings in the portfolio over time and the average number of holdings. Panel B plots the time series of the maximum portfolio weights on an individual fund and its average.

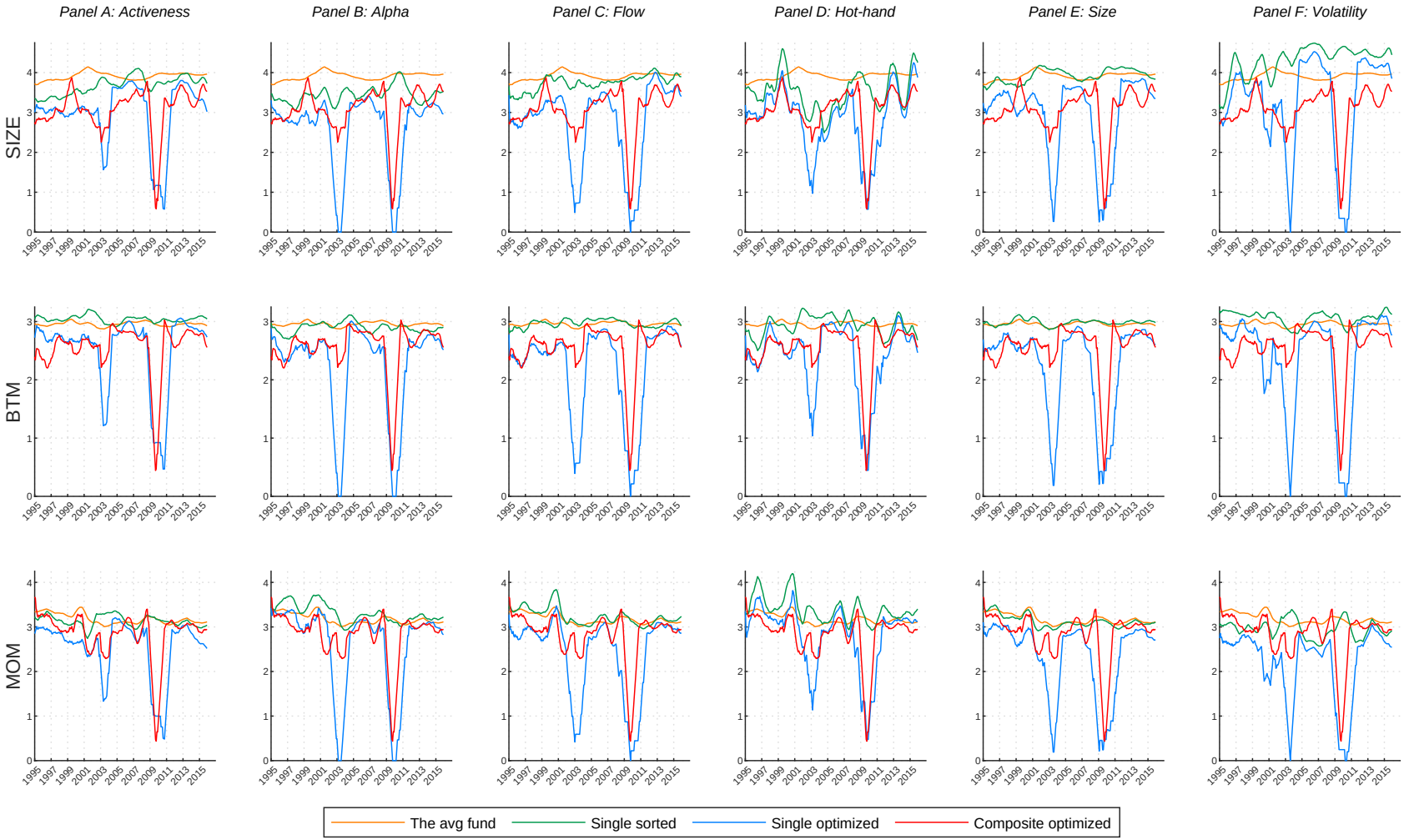

Figure 3. DGTW style exposures

This figure displays DGTW style exposures of alternative mutual fund portfolio strategies. Every year, all CRSP stocks are sorted, conditionally, into quintiles based on their size, book-to-market, and momentum characteristics. For a given fund, exposure to, e.g., SIZE is the quintile size portfolio number in which each stock is sorted into for a given year, weighted across all stocks held by the fund each month. The BTM and MOM fund-level style exposures of each fund are computed analogously. The strategy-level exposure is computed by averaging the fund-level exposures with the weights dictated by each strategy.

THE UNIVERSITY OF MELBOURNE

fbe.unimelb.edu.au/finance