Heuristic Fund Allocation Decisions*

Indraneel Chakraborty   Alok Kumar
Tobias Mühlhofer   Ravi Sastry

February 12, 2020

Abstract

We demonstrate that fund investors employ a heuristic benchmark model to estimate alphas and allocate capital. This can result in observational equivalence to CAPM driven investment decisions. The benchmark estimator trades off bias against precision, accommodating finite sample constraints. The estimator does not require knowledge of the (unknown) true model. Under a diffuse prior on alpha, the benchmark estimator dominates factor model estimates by the MSE criterion in 88% of our sample. We derive a sufficient condition under which the benchmark estimator dominates, even when the true pricing model is known. Fund investors employ a sophisticated approach to capital allocation.

JEL Classification: G11

Keywords: heuristics, mutual fund flows, sector funds, sophistication, model uncertainty.

*We thank Sandro Andrade, Kerry Back, Brad Barber, Bruce Carlin, James Choi, Alan Crane, Stefanos Delikouras, Daniel Dorn, David Feldman (discussant), Chotibhak Jotikasthira (discussant), Matti Keloharju, Shimon Kogan, Veronika Pool, Clemens Sialm, Matt Spiegel, Sheridan Titman, Jules van Binsbergen, Scott Yonker (discussant); seminar participants at the University of Melbourne, University of Miami, University of Texas at Dallas, Stockholm Business School, Rice University; and conference participants at the Miami Behavioral Finance Conference and 2019 FIRN Annual Conference for helpful comments and suggestions. William Bazley and Sarah Khalaf provided excellent research assistance. Indraneel Chakraborty: University of Miami, Email: i.chakraborty@miami.edu. Alok Kumar: University of Miami, Email: akumar@miami.edu. Tobias Mühlhofer: University of Miami, Email: txm467@miami.edu. Ravi Sastry: University of Melbourne, Australia, Email: ravi.sastry@unimelb.edu.au.

Recent work has shown that investors are not utilizing multifactor alphas to evaluate mutual fund managers and allocate capital.[1] We argue that this behavior should be interpreted as a sophisticated model-free benchmark-driven heuristic rather than a naive decision rule or a revealed preference for any particular pricing model. We posit a *benchmark model* that substitutes the return of a single passive benchmark for one or more factors from the true pricing model. We demonstrate, empirically and analytically, that fund alphas estimated with respect to a benchmark model are often preferable to estimates generated by the true model.[2] Although such a procedure can increase bias in estimated alphas, it can dramatically improve precision as well — by an order of magnitude, on average, in our sample. For investors, the net effect is more accurate estimates of alpha and, consequently, better allocation decisions.

It is natural for investors to acknowledge and make some trade-off between bias and precision when estimating alphas, rather than focusing exclusively on minimizing bias.[3] Such a balance is consistent with Merton (1969), which shows that the sufficient statistic for portfolio formation is an asset's information ratio, i.e. the ratio of its alpha to its idiosyncratic volatility. Intuitively, adjusting for idiosyncratic volatility—as Merton (1969) prescribes in a setting where return dynamics are known—is similar to adjusting for estimation error, which is what we advocate. In either case, the benefit of a positive alpha is attenuated in the presence of noise. Investors' motives are thus different from those of researchers, who seek to understand the underlying risk factors and associated betas driving risk premia. For investors, who are not concerned with betas, the benchmark model sidesteps the factor zoo while *improving* estimates of alpha.

The benchmark model is especially relevant since the academic literature itself has not reached consensus regarding the true asset pricing model.[4] Given such model uncertainty, Pástor and Stambaugh (2002) show that investors can utilize passive assets (including benchmarks) in lieu of

[1] Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) show that, conditional on CAPM alphas, multifactor alphas do not have additional explanatory power with respect to aggregate fund flows.

[2] We will formally adopt a mean squared error (MSE) criterion in our analysis.

[3] If zero bias is truly the objective, then only the true model will suffice, and investors must contend with the many competing approximations (e.g. CAPM, Fama-French-Carhart, etc.) to the unknown true model. Once precision is allowed to influence decisions, we show that the true model often turns out to be irrelevant to their decision-making.

[4] See Barillas and Shanken (2018) for a recent study demonstrating that the posterior probabilities associated with various models are never close to one. If a single model is to be selected, it cannot be on the basis of empirical fit alone.

the unknown risk factors:

> Loosely speaking, if you believe that *some* pricing model holds exactly and want a fund's alpha with respect to it, you need not identify the model. The appropriate estimate of alpha is then simply the estimated intercept in a regression of the fund's return on all of the passive assets (p. 319).

Since it is not feasible to run a regression on "all of the passive assets," as suggested by this thought experiment, sophisticated investors may employ a second-best heuristic that reflects its spirit. A sensible approach that narrows the set of passive assets is to focus on salient benchmarks. Specifically, in the case of sector funds, Pástor and Stambaugh (2002) recommend the use of sector benchmarks:

> It is likely that future research could refine the selection of non-benchmark assets and further increase the precision of estimated alphas. A different set of non-benchmark assets could be specified for each fund, so that the assets have a high correlation with the specific fund at hand. With sector funds, for example, a passive index for the same sector could be included (p. 326).

The benchmark model that we propose, consisting of the market factor as well as the passive benchmark return associated with a given mutual fund, follows these insights to their logical conclusion.

While Pástor and Stambaugh (2002) can be understood as normative—it describes an optimal Bayesian procedure that explicitly incorporates model uncertainty but that is infeasible for most investors—our analysis is descriptive. We show that the simple two-"factor" benchmark alphas correspond to observed investor behavior, providing higher explanatory power for fund flows than CAPM alphas.[5] Moreoever, among sector funds, we show that these benchmark alphas usually have lower mean squared error than more conventional Fama-French-Carhart alphas.[6]

Our focus on the space of sector funds allows us to distinguish between the use of risk factors and passive benchmarks as explanatory variables. Like general equity mutual funds, sector funds

[5] We also show that the benchmark alphas have significantly higher explanatory power than benchmark-adjusted returns; investors are not simply return-chasing.

[6] Section 3.2 shows that under a diffuse prior on alpha, the benchmark alpha dominates the FFC alpha for 88% of the funds in our sample, according to the MSE criterion.

self-designate a passive benchmark.[7,8] But, crucially, these sector benchmarks have sufficiently low correlation with the Fama-French-Carhart factors that it is possible in realistic sample sizes to discriminate between investor reliance upon the benchmark and investor reliance upon the factors. In contrast, general equity funds designate benchmarks corresponding to their styles (e.g. small growth) and, as these are long-only portfolios diversified across industries, they tend to have very high correlation with the market itself.[9] As a result, statistical power in the space of general equity funds for the tests that we propose is quite low. Among sector funds, however, there is sufficient power to show that investors are responding to the two-"factor" benchmark alphas rather than CAPM alphas.

To establish the viability of the benchmark model, we conduct three related analyses. Our first analysis consists of empirical tests that show that aggregate sector fund flows are consistent with the utilization of the benchmark model. The intuition behind our tests is that if investors evaluate funds according to their CAPM alphas, then the sector benchmark should not have incremental explanatory power for fund flows. We find that the opposite holds in the data, across a range of specifications.[10] Adding the sector benchmark as a "factor" dramatically improves the correspondence between estimated alphas and fund flows. We can therefore reject the null hypothesis that investors assess mutual funds according to the univariate regression implied by the CAPM. Moreover, we can reject the hypothesis that investors are simply return-chasing, i.e. allocating capital to funds with the highest performance relative to their benchmark.[11]

In our model-free bootstrap analysis, we determine the empirical distributions of the benchmark alpha and the Fama-French-Carhart alpha for each of the funds in our sample of $N = 759$ actively

[7]Sensoy (2009) shows that investors in general equity mutual funds are influenced by the manager-selected benchmark. To the extent that there is substantial within-sector variation of factor loadings across competing sector benchmarks, there is scope for sector fund managers to play a similar game. However, our empirical results with respect to exogenously constructed sector benchmarks suggest that mis-matched benchmarks are not of first-order concern, as aggregate flows are robustly predicted by outperformance with respect to these exogenous benchmarks.

[8]A good benchmark will have factor loadings that are similar to the fund in question, even as these loadings vary over time. Thus, a single benchmark can act as an observable dynamic projection of a fund's systematic exposures to an arbitrary number of (potentially unidentified) risk factors. As the *benchmark's* idiosyncratic volatility approaches zero, the (conditional) relevance of the risk factors approaches zero as well.

[9]See Table I in Cremers, Petajisto, and Zitzewitz (2012) for a comprehensive summary of mutual fund benchmarks.

[10]Our principal specification follows the semi-parametric methodology of Berk and van Binsbergen (2016).

[11]In the benchmark model we posit, both β_{market} and β_{sector} are free parameters. The CAPM imposes the restriction that $\beta_{\text{sector}} = 0$, while the return-chasing model effectively imposes $\beta_{\text{market}} = 0$ and $\beta_{\text{sector}} = 1$.

managed sector funds. We find that for 94% of the funds, the variance of $\widehat{\alpha}_{\text{benchmark}}$ is smaller than the variance of $\widehat{\alpha}_{\text{factor}}$. Variance is reduced, on average, by an order of magnitude.[12] As a consequence of this dramatic increase in precision, the benchmark alpha will often dominate the FFC alpha by the MSE criterion, even though the former may have a larger bias. Under a diffuse prior on the true α, the realized empirical distributions imply that $\widehat{\alpha}_{\text{benchmark}}$ dominates $\widehat{\alpha}_{\text{factor}}$ for 88% of the funds in our sample. This fraction remains above 50% as long as the prior on alpha has a standard deviation of at least 1.7% per year.[13] Only under a dogmatic prior, as the standard deviation goes to zero, does this fraction fall below 50%. Even then, the benchmark estimator dominates for at least 46% of the funds in the sample.

Finally, we provide a theoretical analysis that demonstrates that the benefits of the benchmark model revealed by the bootstrap are widely applicable beyond our particular sample. We derive a sufficient condition under which the MSE of the alpha estimated with respect to the true (possibly unknown) pricing model is larger than the MSE of the alpha estimated with respect to the benchmark model. We find that the benchmark alpha is more likely to dominate the multifactor alpha when (a) the fund's loadings on non-market factors are closer to zero, (b) the R^2 of the fund with respect to the multifactor model is lower, and (c) the R^2 of the benchmark with respect to the multifactor model is higher. We calibrate the sufficient condition according to the realized moments of the Fama-French-Carhart factors, and determine the boundary between the regions of four-factor dominance and benchmark dominance.

In summary, the theoretical analysis provides econometric foundations for the use of the benchmark estimator of alpha instead of the usual estimator of alpha that includes all factors as independent regressors. When the sufficient condition is satisfied, the benchmark estimator will dominate the usual estimator by the MSE criterion, even when the true pricing model is known and all factors are observable. The calibration exercise shows that this condition is likely to be satisfied for many funds. Moreover, the bootstrap analysis shows that this condition is satisfied for nearly all (88%) of the sector funds in our sample. And our empirical flow regressions show that investors

[12]This is the average over all funds, including the 6% for which variance increases. A histogram of this highly skewed distribution of variance ratios is shown in Panel B of Figure 1. The median variance ratio is 3.0.

[13]The prior mean is assumed to be zero.

are responding accordingly.

Our focus on sector funds has several motivations. As noted previously, our empirical tests will have much higher power among sector funds, given the covariance structure of common factor returns and general equity benchmarks. Additionally, even though sector funds are a large segment of the mutual fund industry with 10–15% of the total assets invested in active mutual funds, previous work has not focused on this part of the fund market. From the perspectives of Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) who focus entirely on general equity funds, our analysis is an out-of-sample test.

A concern with our empirical strategy is that sector fund investors and managers might be different from their general equity counterparts.[14] For example, sector fund investors may be attempting to bet on momentum (Moskowitz and Grinblatt, 1999) or sector fund managers may be better informed about their sectors (Kacperczyk, Sialm, and Zheng, 2005). To address this concern regarding external validity, we conduct an experiment and test our hypothesis out-of-sample using micro-data. An advantage of the experimental setting is that we can conduct within-respondent comparisons. The experimental evidence also suggests that when benchmark information is provided, respondents utilize it to make capital allocation decisions. While these results support our hypothesis, we recognize that experimental subjects are not representative of mutual fund investors.

Two recent papers argue that it is Morningstar ratings and not CAPM or multifactor alphas that drive fund flows (Evans and Sun, 2018; Ben-David, Li, Rossi, and Song, 2019). These findings, in the space of general equity funds, are not inconsistent with ours. In particular, Evans and Sun (2018) includes a passive benchmark as an explanatory variable, in addition to the fund's Morningstar rating. Our argument relies upon the ability of a benchmark to capture risk exposures that are unspanned by the market factor. If a fund's market beta is sufficiently highly correlated with its Morningstar rating, employing the latter is a distinct but reasonable heuristic.

Although we demonstrate that investors' decisions are consistent with the two-factor benchmark alpha, our purpose in labeling this a heuristic is to emphasize that mutual fund investors are not necessarily performing multivariate regression analysis. In the well-known example from Friedman

[14] Although, in a fully rational world where some factor model (e.g. the CAPM) holds, investors in any fund should base their allocations on funds' alphas with respect to this model.

and Savage (1948), an expert billiard player behaves "as if" he knew the physics required to predict the consequences of his shots, but is actually making decisions via some other heuristic process. So too, in our setting, we draw no conclusions about investors' actual decision-making process, which is unobservable in any event. A disaggregated sample would surely reveal a multitude of decision rules to be in use, including CAPM and return-chasing. But, given the limitations of our aggregated data and the corresponding representative agent analysis, the benchmark alpha is the best explanatory model.

Our paper does not address the question of which risk-adjustment procedure investors *should* employ. But, in a setting with adequate statistical power (sector funds), we show that investors' aggregate fund flows correspond to the two-"factor" benchmark alpha. This heuristic is consistent with belief in some (set of) multifactor model(s). As the mapping from behavior to beliefs is not one-to-one, investors' true beliefs are unknowable. But in practice, we show that they utilize benchmark assets to estimate managerial skill — sidestepping the need to identify the true pricing model and the associated factor loadings — and that this is a sophisticated response to a challenging problem.

1 Data

This section provides details regarding the construction of our data.

1.1 Sample selection

We use a monthly panel of sector fund returns as our main dataset with a sample period of Jan 1999–Dec 2016. Observations from Dec 1998 are included to calculate lagged measures. To identify our fund universe, we utilize the Center for Research in Security Prices (CRSP) Survivorship-Bias Free Mutual Fund Database. Within that sample of mutual funds, our analysis focuses on sector funds.

In the case of general equity funds, passive performance benchmarks are highly correlated with the overall market return. This makes distinguishing between a CAPM alpha and a benchmark-adjusted alpha an empirical challenge. In contrast, sector funds provide a valuable empirical setting. In addition to a broad market benchmark, sector funds provide information regarding sector-specific benchmarks which are less correlated with general equity benchmarks, e.g. the market, size, and value factors. (See Table B.1 in the Appendix). Hence, if investors utilize benchmark information to evaluate managers, sector funds provide a useful laboratory. Compared to previous research that focuses on general equity mutual funds, sector funds are out of sample.

To identify sector funds, we utilize Lipper Objective Codes. In our sample, the sectors with significant fund presence are *Health and Biotechnology*, *Natural Resources*, *Real Estate*, *Science and Technology*, *Telecommunications*, and *Utilities*.

1.2 Variable construction

The benchmarks in our analysis are constructed by forming a value-weighted portfolio of all sector funds within a given sector. This approach is similar in spirit to the benchmark selection procedure in Berk and van Binsbergen (2015), as it corresponds to investors' opportunity set. This also avoids concerns regarding strategic benchmark choice, as described by Sensoy (2009) in the context of general equity funds.[15]

Fund flows are calculated following Berk and van Binsbergen (2016). Over a horizon of length T, this is:

$$F_{p,t} = q_{p,t} - q_{p,t-T}\left(1 + R^V_{p,t}\right),$$

where $R^V_{p,t}$ is defined to be the return on the relevant sector benchmark.[16] We define the signal of managerial skill on fund p with respect to pricing model j to be:

$$\alpha^j_{p,t+1} = R^e_{p,t+1} - R^j_{p,t+1},$$

i.e., the difference between the fund's excess return and its risk-adjusted return. If model j is a

[15]In unreported results, we create alternative benchmarks following Hartzell, Mühlhofer, and Titman (2010) and find similar results in all empirical tests.

[16]In Berk and van Binsbergen (2016), $R^V_{p,t}$ is the projection of fund p's returns on the space of available Vanguard index funds, i.e. the passive alternative investment opportunity.

linear factor (beta) model, then

$$R^j_{p,t+1} = F^j_{t+1}\,\widehat{\beta}^j_p \tag{1}$$

and

$$\alpha^j_{p,t} = \prod_{s=t-T+1}^{t} \left(1 + R^e_{p,s} - F^j_s \widehat{\beta}^j_p\right) - 1. \tag{2}$$

We compute this signal of skill conditional on two simple models: a single factor model (CAPM) and a two-factor model that includes both the market return and the sector's benchmark return.

While one could estimate (1) once for each fund, over the entire sample, we employ rolling regressions to estimate the factor loadings in (1) and, in turn, the managerial skill in (2). This approach corresponds to the information set actually available to investors in real time and avoids any look-ahead bias.

1.3 Summary statistics

Panel A of Table 1 reports summary statistics for our data. We begin by showing the number of active share classes that we identify in each sector. This number ranges from 63 for Telecommunications to 473 for Real Estate. The table also reports distributional statistics for the number of share classes that exist in a given year.

Our empirical analysis in Section 2 consists of a set of flow-performance regressions. Although share classes within a given fund may have identical performance, their investor bases and capital flows are necessarily distinct. Thus, the share class is the appropriate level of analysis for flow regressions. Our bootstrap analysis in Section 3 employs only returns data, and is therefore conducted at the fund/portfolio level.

Panel B of Table 1 provides descriptive statistics regarding fund and benchmark performance. The benchmark portfolio for each sector is the value-weighted portfolio of all stocks within the sector. Data are at a monthly frequency. The average sector fund has a positive mean raw return of 0.68%, but close to zero alpha with respect to the Fama-French-Carhart four factor model. Figure B.1 reports the four-factor risk-adjusted three-year rolling alphas of sector funds in the sample. The unit of observation is fund-month. Table B.1 reports correlations between monthly value weighted sector returns in our sample with monthly returns of the market index.

2 Empirical analysis of sector fund flows

This section presents our first empirical result: investors utilize benchmark returns to estimate alpha — rather than explicitly adopting the CAPM or any other pricing model — when making capital allocation decisions. Section 2.1 compares alternative specifications from the literature to provide justification for our empirical tests. Section 2.2 demonstrates, using the framework from Berk and van Binsbergen (2016), that the benchmark model has greater explanatory power for aggregate fund flows than the CAPM. Section 2.3 utilizes a hybrid approach that draws from the strengths of the approaches of Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) to again establish that in our dataset, investors utilize benchmark returns along with market returns to allocate capital. Section 2.4 shows that the benchmark model also has greater explanatory power for aggregate fund flows than a return-chasing model in which investors' allocation decisions are based on the simple difference between a fund's realized return and the return on the associated passive benchmark.

2.1 Discussion of alternative specifications

Barber, Huang, and Odean (2016) and Berk and van Binsbergen (2016) both establish that fund flows are consistent with investors utilizing the CAPM. We consider both specifications, which are largely consistent but with some important differences. The primary specification for aggregate fund flows from Barber, Huang, and Odean (2016) is:

$$F_{p,t} = a + \sum_i \sum_j b_{i,j} D_{i,j,p,t} + cX_{p,t} + \mu_t + \epsilon_{p,t}, \tag{3}$$

where flows for fund p and time t are projected onto a set of control variables, X, and a collection of indicators, D, corresponding to the fund's alpha decile rankings according to two competing models, i and j. For example, if the estimated alpha according to model i is in decile 4 while the estimated alpha according to model j is in decile 7, then $D_{4,7,p,t} = 1$ and $D_{i \neq 4, j \neq 7,p,t} = 0$.

This specification abstracts away from the numerical estimates of alpha and considers only relative rankings. It is therefore robust to non-linearities in the fund flow-performance relationship.

To assess the relative predictive powers of models of i and j, the authors compare $\sum_{i>j} b_{i,j}$ and $\sum_{i<j} b_{i,j}$. If the difference is statistically significant, then it can be reasonably concluded that one model is more consistent with investors' decisions than the other.[17] Note, however, that the $b_{i,j}$'s are assumed not to depend on the fund; the exact (non-linear) flow-performance relationship is common to all funds.

The methodology proposed by Berk and van Binsbergen (2016) is robust to this assumption as well, allowing heterogeneous (nonlinear) flow-performance relationships. They estimate:

$$\operatorname{sign}\left(F_{p,t}\right) = \gamma_0 + \gamma_1 \left(\frac{\operatorname{sign}\left(\alpha^i_{p,t}\right)}{\operatorname{Var}\left(\operatorname{sign}\left(\alpha^j_{p,t}\right)\right)} - \frac{\operatorname{sign}\left(\alpha^j_{p,t}\right)}{\operatorname{Var}\left(\operatorname{sign}\left(\alpha^j_{p,t}\right)\right)} \right) + \epsilon_{p,t}, \tag{4}$$

where the notation has been slightly modified to allow straightforward comparison with (3). $\gamma_1 > 0$ if and only if model i is a better pricing model than model j with respect to explaining fund flows. By considering only the signs of both flows and alphas, (4) makes no assumptions about the functional form of the flow-performance relationship, either within or across funds. Positive alphas should be associated with positive flows, and vice-versa.

Overall, the two approaches are largely consistent. We utilize two specifications below. The first approach is that of Berk and van Binsbergen (2016), and a second hybrid approach that builds on the strengths of Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016).

2.2 Berk and van Binsbergen (2016) specification

To compare two models, i and j, (4) regresses the sign of capital flows on the difference in the standardized signs of fund outperformance with respect to models i and j. Essentially, the semi-parametric approach of Berk and van Binsbergen (2016) identifies the better model by focusing on the cases when the two models disagree regarding the direction of fund flows. For example, in

[17] An alternative test could impose a monotonicity constraint, such that the b coefficients are weakly monotonic in both i and j. This would still allow for a nonlinear flow-performance relationship, but it would disallow some locally suboptimal behaviors (for example, negative marginal flows for improving rank) that would otherwise be admissible.

certain time periods, model i might suggest that the manager has a positive alpha and therefore investors should invest more with her. In contrast, if for the same time period, model j suggests that the manager has a negative alpha, we would expect investors to withdraw (some) of their capital from the fund.

The fund flow data allow us to observe which model better captures investor decisions in these cases when the models disagree. The main variable of interest is the second term in (4), i.e. "Diff of Signs," which is the difference in signs of fund outperformance measured by the two-factor sector-plus-market model and the one-factor market model. A statistically and economically significant positive coefficient suggests that the two-factor benchmark model is better than the CAPM at predicting investors' allocations.

Table 2 reports the results. Column (1) provides estimates for a pooled OLS specification. Column (2) adds fund size to address search cost differences (Sirri and Tufano, 1998) and differences in returns to scale (Berk and Green, 2004; Stambaugh, 2014; Pástor, Stambaugh, and Taylor, 2015). Column (3) includes sector-fixed effects to address concerns regarding time-invariant differences in sectors. Column (4) includes fund-fixed effects that address cross-sectional variation in time-independent differences in funds.

Column (1) reports that the coefficient on "Diff of Signs" is positive and significant. This result suggests that investors use the sector benchmark along with market returns to determine whether the manager outperformed in a specific period, and that this model explains fund flow signs better than the CAPM. Column (2) also reports similar results. The point estimates in columns (3) and (4) remain similar in magnitude (and significance) to the previous columns, suggesting that time-invariant sector-specific or fund-specific differences are not driving the results. Such differences can include variation in how funds are marketed (Del Guercio and Reuter, 2014).

2.3 Hybrid specification

As an alternative test, we build on the strengths of both Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) and develop a hybrid specification. We compute decile ranks for funds with respect to both the CAPM and the benchmark model, as in Barber, Huang, and Odean

(2016). As the cells have non-linear flow effects, Barber, Huang, and Odean (2016) compare the differences in a pairwise manner and report the sum of coefficient differences.

We sidestep the issue of non-linear flow effects by utilizing the sign of fund flows, as in Berk and van Binsbergen (2016). This approach allows for robust aggregation across funds. As we can aggregate across funds, we create an indicator variable $D_{i,j,p,t}$ that equals one when the decile rank according to model i is higher than the decile rank according to model j for fund p in time t. This approach effectively collapses the upper off-diagonal 45 cells of the 10×10 matrix in Barber, Huang, and Odean (2016) into a single statistic.

Thus, we conduct a joint test of the superiority of the benchmark model over the CAPM. The specification is as follows:

$$\text{sign}\left(F_{p,t}\right) = a + b_{i,j} D_{i,j,p,t} + cX_{p,t} + \eta_p + \epsilon_{p,t},$$

Table 3 reports the results. Columns (1) and (2) report the estimates where we divide outperformance into deciles as in Barber, Huang, and Odean (2016). We also conduct robustness tests in columns (3) and (4) where we divide the funds into terciles. This is a more stringent test because the disagreement has to be stronger: top tercile by benchmark model rank and bottom tercile by the CAPM. Columns (1)–(4) confirm that when the two models disagree with respect to the quantile rank of alpha, investors allocate capital in the direction of the benchmark model's alpha.

2.4 Are investors simply chasing returns?

So far, we have presented evidence that investors' allocation decisions are consistent with the estimated alphas from the two-factor "benchmark model." Some readers may have in mind the alternative explanation that investors are simply chasing funds that are producing higher non-risk-adjusted returns than the benchmark. In that scenario, investors are not doing anything sophisticated, but simply using sector fund returns net of sector benchmark returns to make investment decisions. Note that such a prior is at odds with the findings of Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016), as both show that the CAPM dominates return-chasing as an explanation for aggregate fund flow (among general equity funds).

Nevertheless, to address this alternative, we need to show that fund flows can be explained better by the (risk-adjusted) alphas estimated from a two-factor benchmark model than by raw outperformance implied by the return-chasing model. Table 4 reports the results of such a test. The positive coefficients in the specifications for the difference of signs variable suggest that risk-adjusted alpha from the two factor model is better able to explain fund flows than fund return net of the sector benchmark return. Thus, investors are not simply chasing fund returns. In addition, the results also provide supportive evidence for Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016), as they show that investors are adjusting for risk and not using raw returns.

3 Model-free analysis

The empirical analysis in Section 2 establishes that investors evaluate funds with respect to both the market return and the sector benchmark, not just the market return as suggested by a belief in the CAPM. This section shows that this behavior — using the benchmark alpha instead of a factor model alpha — minimizes MSE, *regardless of the true pricing model*, when investors' priors on alpha are not excessively dogmatic.[18]

3.1 Methodology

We start by calculating fund-level MSE as a function of the true (unknown) alpha. To do so, we obtain model-free estimates of bias and variance conditional on the benchmark model and the four-factor model.

Let R_j denote the $T_j \times 1$ vector of excess returns associated with each active sector fund, $j \in \{1, \ldots, 759\}$, in our sample.[19] Let B_j denote the $T_j \times 1$ vector of excess returns associated with the passive sector benchmark, and let F_j denote the contemporaneous $T_j \times 4$ matrix of Fama-

[18] By dogmatic, we mean a prior belief that alpha is very close to zero. Under such a prior, it can be the case that the factor model alpha obtains lower MSE, but it is more important to observe that in such cases, one may as well assert that $\widehat{\alpha} = 0$ and dispense with the regression entirely.

[19] While the analysis in Section 2 is conducted at the share class level, as investor bases and fund flow responses are not common across share classes of a single fund, the bootstrap analysis in this Section is conducted at the fund level, since only returns data are used. The results are not sensitive to this choice.

French-Carhart factor returns.

The bootstrapped benchmark estimator of alpha is computed by running the following regression,

$$R_j^{(n)} = \begin{pmatrix} \mathbb{1} & MktRf^{(n)} & B_j^{(n)} \end{pmatrix} \begin{pmatrix} \widehat{\alpha}_{\text{benchmark},j}^{(n)} \\ \hdashline \widehat{\beta}_{\text{benchmark},j}^{(n)} \\ \hdashline \widehat{\delta}_{\text{benchmark},j}^{(n)} \end{pmatrix} + \nu_j^{(n)},$$

where the (n) superscript on an observed variable indicates the n^{th} bootstrapped realization of the underlying variable. Analogously, the bootstrapped factor model estimator of alpha is computed by running the following regression,

$$R_j^{(n)} = \begin{pmatrix} \mathbb{1} & F^{(n)} \end{pmatrix} \begin{pmatrix} \widehat{\alpha}_{\text{factor},j}^{(n)} \\ \hdashline \widehat{\beta}_{\text{factor},j}^{(n)} \\ \widehat{\gamma}_{\text{factor},j}^{(n)} \end{pmatrix} + u_j^{(n)}.$$

In all cases, samples are drawn from the full matrix, $\begin{pmatrix} R_j & B_j & F_j \end{pmatrix}$, so all contemporaneous relationships in the data are preserved. Regardless of T_j, all bootstrapped samples have a length of 36 (months) to facilitate straightforward comparisons across funds.

For each fund, this bootstrap procedure is repeated 10,000 times to provide model-free estimates of $\mathbb{E}\left[\widehat{\alpha}_{\text{benchmark},j}\right]$, $\text{Var}\left(\widehat{\alpha}_{\text{benchmark},j}\right)$, $\mathbb{E}\left[\widehat{\alpha}_{\text{factor},j}\right]$, and $\text{Var}\left(\widehat{\alpha}_{\text{factor},j}\right)$. Conditional on the true (unknown) alpha, the MSE for a given estimator (i.e., "benchmark" or "factor") is

$$\text{MSE}_j\left(\alpha_j\right) = \left(\alpha_j - \mathbb{E}\left[\widehat{\alpha}_j\right]\right)^2 + \text{Var}\left(\widehat{\alpha}_j\right). \tag{5}$$

3.2 Results

Before examining MSEs, we first consider the variances of the benchmark and factor model estimators. Figure 1 shows $\text{Var}\left(\widehat{\alpha}_{\text{factor}}\right)$ and $\text{Var}\left(\widehat{\alpha}_{\text{benchmark}}\right)$ for each active sector fund in the sample. In Panel A, each point in the scatterplot represents one fund. 94% of the funds are above the dashed 45° line, indicating that $\text{Var}\left(\widehat{\alpha}_{\text{factor}}\right) > \text{Var}\left(\widehat{\alpha}_{\text{benchmark}}\right)$ in nearly all cases. This asymme-

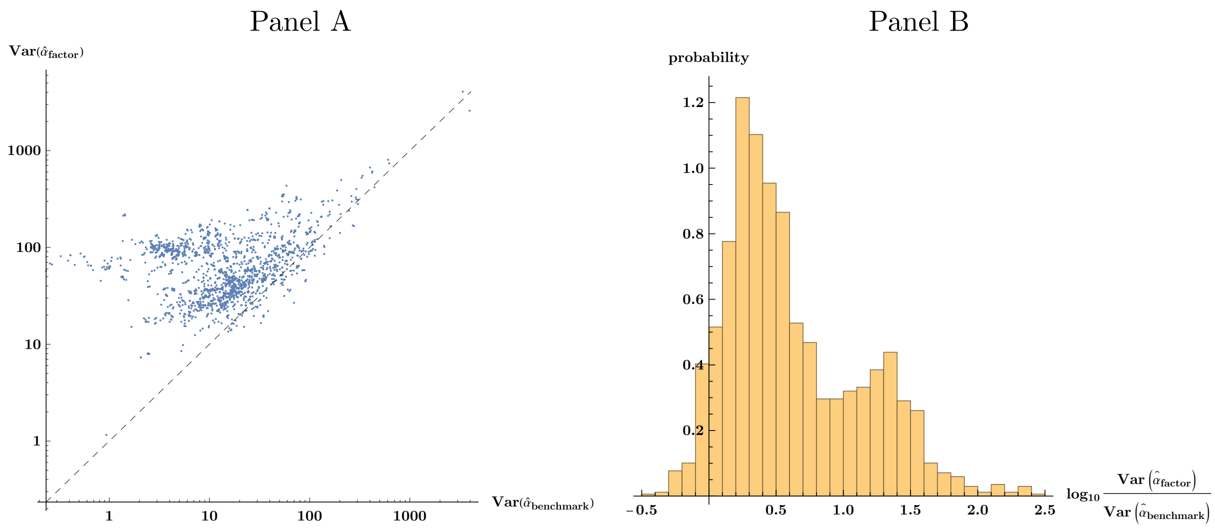

Figure 1: Bootstrapped variances of estimated alphas under the factor model and the benchmark model. Panel A shows a scatterplot of $\text{Var}(\widehat{\alpha}_{\text{factor}})$ versus $\text{Var}(\widehat{\alpha}_{\text{benchmark}})$ for each of $N = 1,687$ active sector funds. 94% of the funds are above the dashed 45° line, indicating that $\text{Var}(\widehat{\alpha}_{\text{factor}}) > \text{Var}(\widehat{\alpha}_{\text{benchmark}})$. Panel B shows a histogram of the common logarithm of the variance ratio for each fund. On average, $\text{Var}(\widehat{\alpha}_{\text{factor}})$ is 9.7 times larger than $\text{Var}(\widehat{\alpha}_{\text{benchmark}})$. The median variance ratio is 3.0. Estimated alphas are measured in units of percent per year, i.e. a value of 5 indicates an annualized alpha of 5%.

try is expected, as the benchmark estimator ideally reduces the dimensionality of the regression without a proportional reduction in R^2. Conversely, the higher-dimensional factor model would be expected to result in lower bias than the benchmark estimator; it is this trade-off between bias and variance that results in the benchmark estimator's MSE dominance in most cases.[20] Panel B of Figure 1 shows a histogram of the common log of the ratio of the two variances. Positive values correspond to funds where $\text{Var}(\widehat{\alpha}_{\text{factor}}) > \text{Var}(\widehat{\alpha}_{\text{benchmark}})$. The distribution has both a positive mean and substantial positive skewness. On average, $\text{Var}(\widehat{\alpha}_{\text{factor}})$ is almost 10 times larger than $\text{Var}(\widehat{\alpha}_{\text{benchmark}})$.

Figure 2 shows the bootstrapped distributions of $\widehat{\alpha}_{\text{benchmark}}$ and $\widehat{\alpha}_{\text{factor}}$ for one particular fund. In this case, $|\mathbb{E}[\widehat{\alpha}_{\text{benchmark}}]| > |\mathbb{E}[\widehat{\alpha}_{\text{factor}}]|$, which is to be expected if the higher dimensional model reduces bias and if the true alpha is close to zero. Also as expected, we see that $\text{Var}(\widehat{\alpha}_{\text{benchmark}}) < \text{Var}(\widehat{\alpha}_{\text{factor}})$. The implications for MSE naturally depend on the true value of α. Using (5), we compute the MSE for each estimator as a function of α. The green line in Figure 2 shows the

[20] If the Fama-French-Carhart model were the true model, then the bias of $\widehat{\alpha}_{\text{factor}}$ would be zero by definition. But we make no assumption about the true pricing model, one way or the other.

common log of the ratio of the two MSEs. Positive values imply that the benchmark estimator is preferred according to the MSE criterion. In this case, we find that the critical value of α is -0.3% per year. Below this value, $\widehat{\alpha}_{\text{benchmark}}$ dominates. Above this value, $\widehat{\alpha}_{\text{factor}}$ dominates.

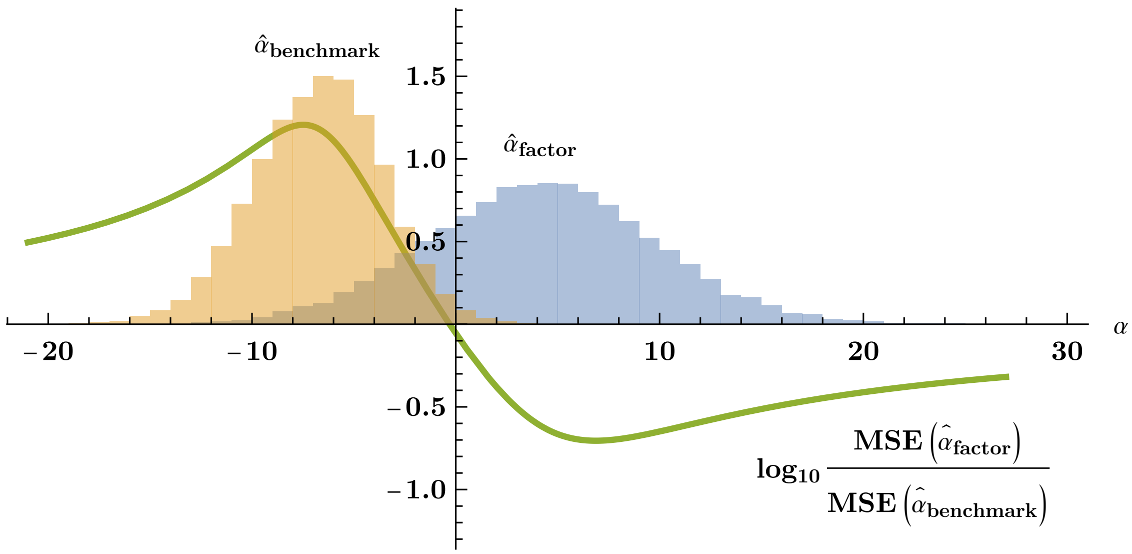

Figure 2: Bootstrapped alpha distributions and implied MSE ratio under the factor model and the benchmark model for one fund in our sample. The horizontal axis represents the true (unknown) alpha. Alpha is measured in units of percent per year, i.e. a value of 5 indicates an annualized alpha of 5%. Each histogram depicts the results of 10,000 replications. The distribution of $\widehat{\alpha}_{\text{factor}}$ has a mean of 4.2% and a standard error of 5.6%. The distribution of $\widehat{\alpha}_{\text{benchmark}}$ has a mean of -6.8% and a standard error of 3.1%. Consistent with Figure 1, $\text{Var}(\widehat{\alpha}_{\text{factor}})$ is 3.2 times larger than $\text{Var}(\widehat{\alpha}_{\text{benchmark}})$ for this fund. In expectation, $\widehat{\alpha}_{\text{factor}}$ is (1.6 times) closer to zero than $\widehat{\alpha}_{\text{benchmark}}$. The solid green line shows the common logarithm of the ratio of MSEs based on the bootstrapped alpha distributions. If the true alpha is less than -0.3% annually, then $\text{MSE}(\widehat{\alpha}_{\text{factor}}) > \text{MSE}(\widehat{\alpha}_{\text{benchmark}})$ and the benchmark estimator is preferred.

Since the true value of α is unknown, we must evaluate MSEs in the context of some prior belief regarding α. For simplicity, assume that the prior on α is normal with zero mean and standard deviation σ,

$$\alpha \sim \mathcal{N}\left(0, \sigma^2\right). \tag{6}$$

Then the *expected* MSE conditional on σ can be computed by integrating (5) against the normal density implied by (6). For the fund in Figure 2, this results in the expected MSE ratio shown in Figure 3. Again, positive values imply that $\widehat{\alpha}_{\text{benchmark}}$ dominates according to the MSE criterion *in expectation*. For this fund, if the prior standard deviation of α is greater than 1.1% per year,

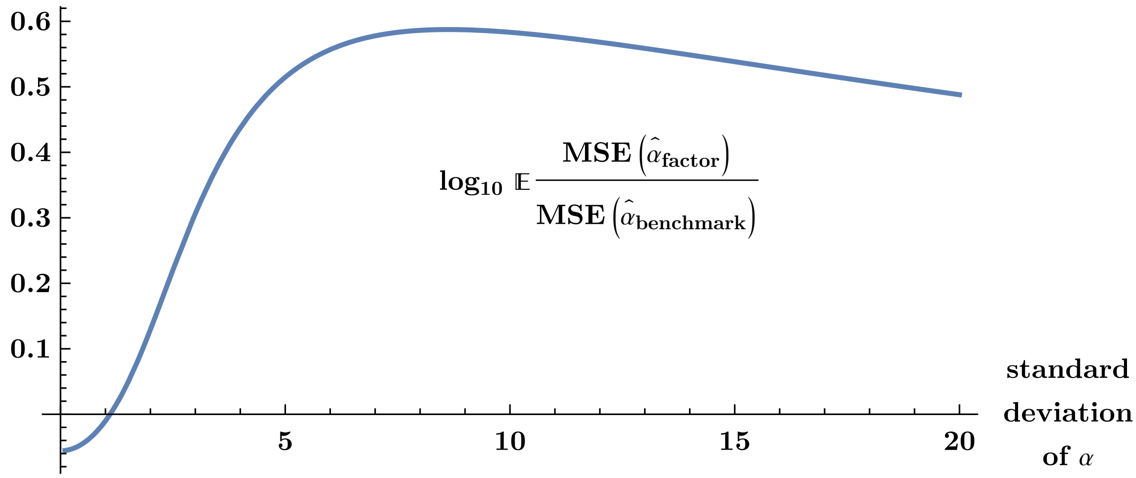

Figure 3: Expected MSE ratio as a function of the prior standard deviation of alpha for one fund. The horizontal axis represents the prior standard deviation of the true (unknown) alpha. Alpha is measured in units of percent per year, i.e. a value of 5 indicates an annualized alpha of 5%. Under a dogmatic prior that $\alpha = 0$, implying that the prior standard deviation of alpha is zero, $\text{MSE}(\widehat{\alpha}_{\text{factor}})$ is *less* than $\text{MSE}(\widehat{\alpha}_{\text{benchmark}})$ by 12%, corresponding to a log-ratio of -0.06. (The intercept in this Figure is equivalent to the intercept in Figure 2.) For prior standard deviations greater than 1.1%, $\text{MSE}(\widehat{\alpha}_{\text{factor}})$ is greater than $\text{MSE}(\widehat{\alpha}_{\text{benchmark}})$ for this fund. The prior mean on alpha is fixed at zero.

the benchmark estimator is preferred.

This same analysis is repeated for all funds, generating a critical σ for each fund in the sample. As long as $\text{Var}(\widehat{\alpha}_{\text{factor}}) > \text{Var}(\widehat{\alpha}_{\text{benchmark}})$, there will always be some value of σ above which the benchmark estimator will dominate. (As shown in Figure 1, this condition holds in 94% of the funds in the sample.) Furthermore, if $|\mathbb{E}[\widehat{\alpha}_{\text{benchmark}}]|$ happens to also be lower than $|\mathbb{E}[\widehat{\alpha}_{\text{factor}}]|$, the benchmark estimator will dominate for *all* values of σ.

The aggregate results for the critical σ are shown in Figure 4. The vertical axis shows the fraction of funds for which the benchmark estimator dominates with respect to the MSE criterion. The horizontal axis shows σ, the prior standard deviation of α. Under a default uninformative prior, as $\sigma \to \infty$, the benchmark estimator dominates for 88% of the funds in the sample. This fraction remains above 50% as long as $\sigma > 1.7\%$ per year. Only under a dogmatic prior, as $\sigma \to 0$, does this fraction fall below 50%. But even then, the benchmark estimator dominates for at least 46% of the funds in the sample.

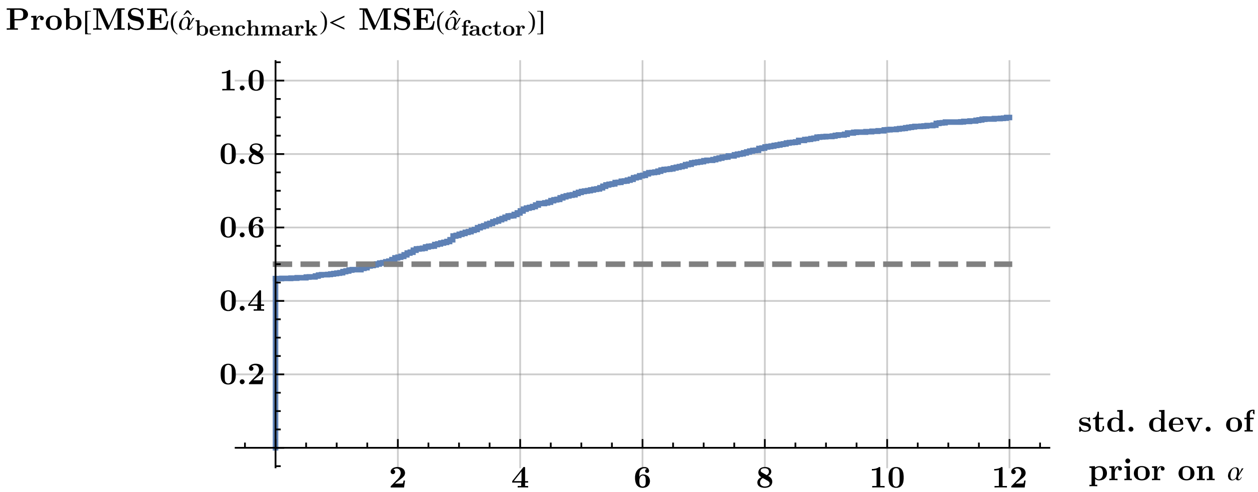

Figure 4: MSE dominance as a function of the prior standard deviation of alpha. The vertical axis shows the fraction of all active sector funds for which the benchmark estimator dominates. The horizontal axis shows the prior standard deviation of α. Alpha is measured in units of percent per year, i.e. a value of 5 indicates an annualized alpha of 5%. For approximately 46% of the funds, the benchmark estimator dominates for any $\sigma \geq 0$. As σ increases, the fraction of funds increases as well. This fraction does not converge to 1 as σ grows. Thus, there is an implied asymptote at 88%. That is, for 12% of the funds in the sample, the factor model estimator dominates for any $\sigma \geq 0$. If the prior on α is sufficiently diffuse — greater than 1.7% per year in this sample of sector funds — the benchmark estimator will be preferred overall to the factor model estimator.

4 Theory of the benchmark estimator

In Section 3, we demonstrated the superiority of the benchmark estimator with respect to the Fama-French-Carhart alpha without making any assumptions about the true pricing model. In this section, we consider the hypothetical case in which the true pricing model is known. We derive a sufficient condition for the dominance of the benchmark estimator with respect to the alpha estimated under the true pricing model. We show that when the R^2 of the passive benchmark with respect to the true pricing model is high—which it is likely to be—the benchmark estimator is likely to dominate.

The heuristic approach adopted by investors and documented in Section 2 is not a sign of low sophistication. Rather, it is a well-founded and pragmatic resolution of a complex challenge—the estimation of alpha—to which academics have failed to provide a consensus recommendation. While the analysis of Section 3 shows the value of the benchmark heuristic in the sample of sector funds, this section establishes the dominance of the benchmark estimator (by the MSE criterion) for any asset, even when the true pricing model is known and all factors are observable.

4.1 Assumptions

We begin by making some assumptions regarding fund and benchmark returns:

Assumption 1 *Fund returns follow a factor structure.*

Fund returns, $R \in \mathbb{R}^{T\times 1}$, are generated according to

$$R_t = \alpha + \left(\begin{array}{c:c} F_t & Z_t \end{array} \right) \left(\begin{array}{c} \beta \\ \hdashline \gamma \end{array} \right) + u_t; \quad \mathbb{E}\left[uu^{'}\right] = \sigma_u^2 I, \tag{7}$$

where the factors have been partitioned into two sets, $F \in \mathbb{R}^{T\times p}$ and $Z \in \mathbb{R}^{T\times q}$, to facilitate our analysis. ■

F represents the "consensus" factors that appear in both the true pricing model and the benchmark model. In our empirical work, we take F to be the market factor ($p = 1$), but the theoretical analysis here is general. Z represents the (possibly unknown) factors that appear in the true pricing model but that will be replaced by the benchmark return in the benchmark model.

Assumption 2 *Existence of a passive benchmark.*

There is a passive benchmark asset with observable returns $B \in \mathbb{R}^{T\times 1}$ generated according to

$$B_t = \left(\begin{array}{c:c} F_t & Z_t \end{array} \right) \left(\begin{array}{c} \beta_B \\ \hdashline \gamma_B \end{array} \right) + e_t; \quad \mathbb{E}\left[ee^{'}\right] = \sigma_e^2 I. \tag{8}$$

Note that, unlike the fund return in (7), the benchmark return has zero alpha by definition. ■

Assumption 3 *Joint normality.*

All random variables in (7) and (8) are jointly normal,

$$\begin{pmatrix} F_t^{'} \\ Z_t^{'} \\ u_t \\ e_t \end{pmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(\begin{pmatrix} \lambda_F \\ \lambda_Z \\ 0 \\ 0 \end{pmatrix}, \left(\begin{array}{c:c:c:c} \Sigma_{FF} & & & \\ \hdashline \Sigma_{ZF} & \Sigma_{ZZ} & & \\ \hdashline 0 & 0 & \sigma_u^2 & \\ \hdashline 0 & 0 & 0 & \sigma_e^2 \end{array}\right)\right).$$

■

While Assumptions 1 and 2 are standard, Assumption 3 is required in order to explicitly consider the case of stochastic regressors (factor returns). As investors cannot specify factor realizations in advance, there is little value in considering the alternative, and much simpler, case of fixed regressors.

4.2 Analysis

We now consider the benchmark specification for the estimation of alpha and establish the conditions under which this deviation from the factor specification—i.e., the specification that corresponds to the true factor model—improves alpha estimates by the MSE criterion.

The usual estimator for α is the OLS estimate of (7), which requires that all factors are known and observable,

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \\ \widehat{\gamma} \end{pmatrix} = \left(\begin{pmatrix} \mathbb{1} & F & Z \end{pmatrix}^{'} \begin{pmatrix} \mathbb{1} & F & Z \end{pmatrix}\right)^{-1} \begin{pmatrix} \mathbb{1} & F & Z \end{pmatrix}^{'} R. \tag{9}$$

The benchmark estimator replaces the factor returns matrix, Z, with the vector of benchmark returns, B,

$$\begin{pmatrix} \widehat{\alpha}_{\text{benchmark}} \\ \widehat{\beta}_{\text{benchmark}} \\ \widehat{\delta}_{\text{benchmark}} \end{pmatrix} = \left(\begin{pmatrix} \mathbb{1} & F & B \end{pmatrix}^{'} \begin{pmatrix} \mathbb{1} & F & B \end{pmatrix}\right)^{-1} \begin{pmatrix} \mathbb{1} & F & B \end{pmatrix}^{'} R. \tag{10}$$

Use of (10) instead of (9) can be motivated in several ways. If the true pricing model is unknown, then (9) is not feasible to begin with. F can then be taken to represent the *observable* risk factors that are, by consensus, assumed to be included in the true pricing model.[21] Z then represents the factors that are unknown, unobservable, or otherwise subject to disagreement. Substituting the observable vector B for the matrix Z (of potentially unknown dimension) leads to the feasible estimator in (10), whose econometric value will be established in Proposition 1, below.[22]

Even when the true pricing model is known, and Z is observable, it may be advantageous to estimate (10) rather than (9). In finite samples of the size commonly used in performance evaluation—e.g. 36 or 60 months—(9) is the best linear unbiased estimator (BLUE) of alpha, but *not* the estimator that minimizes mean squared error (MSE). An investor assessing funds will rationally want estimates of alpha that are as "close" as possible to the true values; bias per se is not the concern.

More technically, a rational investor will optimize her estimate of alpha with respect to a loss function, which in turn is determined by her preferences. Any continuous utility function implies a loss function that trades off bias against variance. In particular, quadratic utility—as often assumed in portfolio optimization problems—implies quadratic loss and minimizing MSE is the unique utility maximizing response. To rationalize adoption of the BLUE, we would need lexicographic loss with respect to bias and variance, and this implies discontinuous utility.

Thus, regardless of whether the true pricing model is known, there are strong theoretical arguments in favor of minimizing MSE rather than bias, and there are realistic conditions under which (10) dominates (9) by this measure. The following proposition provides a sufficient condition for $\widehat{\alpha}_{\text{benchmark}}$ to dominate $\widehat{\alpha}$ with respect to the MSE criterion.

Proposition 1 *MSE dominance of the benchmark estimator.*

A sufficient condition for the MSE dominance of the benchmark estimator in (10) with respect to

[21] In our setting, we take F to be the excess return on the market. But the analysis is general, and F can be of arbitrary dimension.

[22] Note that, unlike (9), (10) cannot be used to generate estimates of γ. If the true objects of interest are the factor loadings, then some implementation of (9) that allows for model uncertainty will be required. In the context of fund evaluation, however, factor loadings are nuisance parameters, and alpha is the object of interest.

the BLUE estimator in (9) is given by

$$\left(\frac{T-p-3}{\sigma_u^2}\right)\left(\gamma'\Sigma_{ZZ\cdot F}\gamma-\frac{\left(\gamma'\Sigma_{ZZ\cdot F}\gamma_B\right)^2}{\gamma_B'\Sigma_{ZZ\cdot F}\gamma_B+\sigma_e^2}\right)<1, \tag{11}$$

where $\Sigma_{ZZ\cdot F}$ denotes the conditional variance, $\text{Var}\left(Z_t \middle| F_t\right)$,

$$\Sigma_{ZZ\cdot F}=\text{Var}\left(Z_t \middle| F_t\right)=\Sigma_{ZZ}-\Sigma_{ZF}\Sigma_{FF}^{-1}\Sigma_{FZ}.$$

Recall from Assumption 1 that p is the number of columns in F.

Proof. See Appendix A. ■

When (11) is satisfied, then investors should rationally evaluate funds and allocate capital according to $\widehat{\alpha}_{\text{benchmark}}$. Since the condition depends on unobservable parameters, it cannot be independently verified and in practice must be computed with respect to investors' beliefs.[23] Nevertheless, employing the factor model estimator from (9) as a default is an implicit assertion that the sufficient condition *does not* hold. The relevance of investors' beliefs is not easily circumvented.

Note, however, that (11) is written in terms of quadratic forms in the unobserved factor loadings. Each of these forms is a scalar that corresponds to a variance. To evaluate (11), it is not required for investors to form beliefs over each factor loading and each term of the covariance matrix of the factors. It is sufficient to form beliefs over the systematic variance that can be explained by the excluded factors, Z, conditional on F, as well as the idiosyncratic variances with respect to the true pricing model.[24]

In the following corollaries, we provide sufficient conditions for special cases in which (1) the fund return is regressed only on the benchmark return, (2) the fund and the benchmark have identical factor loadings, and (3) the benchmark is omitted entirely.

[23] As we will show in Section 4.3, the sufficient condition often obtains using realistic parameter values calibrated to the Fama-French-Carhart 4-factor model.

[24] See Teräsvirta (1987) and Thursby (1989) for further discussion of the formation of beliefs over the latent parameters in a more general framework.

Corollary 1 *MSE dominance of the "pure" benchmark estimator.*

Suppose that F is omitted from the estimation, such that (10) becomes a univariate regression of the fund return on the benchmark return. Then the sufficient condition for the MSE dominance of the benchmark estimator is

$$\left(\frac{T-3}{\sigma_u^2}\right)\left(\Theta'\Sigma\,\Theta - \frac{\left(\Theta'\Sigma\,\Theta_B\right)^2}{\Theta_B'\Sigma\,\Theta_B + \sigma_e^2}\right) < 1,$$

where $\Theta = \left(\begin{array}{c;{2pt/2pt}c} \beta' & \gamma' \end{array}\right)'$, $\Theta_B = \left(\begin{array}{c;{2pt/2pt}c} \beta_B' & \gamma_B' \end{array}\right)'$, and $\Sigma = \left(\begin{array}{c;{2pt/2pt}c} \Sigma_{FF} & \Sigma_{FZ} \\ \hdashline \Sigma_{ZF} & \Sigma_{ZZ} \end{array}\right)$. ■

Corollary 2 *MSE dominance of the "ideal" benchmark estimator.*

Suppose that the passive benchmark is ideal, in the sense that $\gamma_B = \gamma$. Then the sufficient condition for the MSE dominance of the benchmark estimator is

$$(T-p-3)\cdot\frac{\sigma_e^2}{\sigma_u^2}\cdot\frac{\gamma'\Sigma_{ZZ\cdot F}\gamma}{\gamma'\Sigma_{ZZ\cdot F}\gamma + \sigma_e^2} < 1.$$

Note that $\frac{\gamma'\Sigma_{ZZ\cdot F}\gamma}{\gamma'\Sigma_{ZZ\cdot F}\gamma+\sigma_e^2}$ can be interpreted as the R^2 arising from a regression of the component of B that is orthogonal to F, $B_{\perp F}$, on the component of Z that is orthogonal to F, $Z_{\perp F}$. ■

Corollary 3 *MSE dominance of the omitted variables estimator.*

Suppose that the passive benchmark is omitted entirely, and the fund returns are regressed only on the observed factors, F. Then the sufficient condition for the MSE dominance of the omitted variables estimator is

$$\left(\frac{T-p-q-2}{\sigma_u^2}\right)\gamma'\Sigma_{ZZ\cdot F}\gamma < 1.$$

Proof. This is a direct application of the results of Kinal and Lahiri (1983), but can also be understood as the limit of the Proposition 1 result as $\gamma_B \to 0$. ■

We note that Corollary 3 encapsulates one of the key results of Jegadeesh and Mangipudi (2019): omitting a true factor can improve the estimation of alpha in finite samples.

4.3 Calibration

To illustrate the applicability of the benchmark estimator more concretely, we now present the implications of Proposition 1 under the assumptions that the Fama-French-Carhart model holds, and that factor returns are serially uncorrelated and jointly normal, with parameters equal to the realized sample moments in Table 5.

We consider the usual estimator of alpha, computed using the full regression of fund returns on all 4 factors, as well as the benchmark estimator of alpha, in which fund returns are regressed on the market return as well as the benchmark return. Proposition 1 depends on the variance of Z (the size, value, and momentum factors) conditional on F (the excess market return). Based on the values in Table 5, this conditional variance is given by

$$\Sigma_{FF \cdot Z} = \begin{pmatrix} 10.2 & -2.5 & 3.5 \\ -2.5 & 9.8 & -3.2 \\ 3.5 & -3.2 & 25.0 \end{pmatrix}.$$

For simplicity of exposition and graphical interpretation, we make the following assumptions:

1. The conditions of Corollary 2 hold: $\gamma_B = \gamma$. This is not as strong as it seems, since the relevance of a passive benchmark implies that $\gamma_B \approx \gamma$. That is, we expect differences in returns between an active fund and its benchmark to be largely attributable to managerial choices other than systematic deviations in risk exposures.

2. The sample size is $T = 36$ months.

3. The loadings on the market factor are $\beta = \beta_B = 1$.

4. The non-market loadings are $\gamma = \gamma_B = \begin{pmatrix} \kappa & \kappa & \kappa \end{pmatrix}'$. That is, the loadings on all of the omitted factors are equal to a single scalar, κ. Although somewhat restrictive, this allows a cleaner examination of the value of the heuristic benchmark estimator for funds with "high"

versus "low" non-market factor exposures.[25]

Finally, it is more intuitive to think in terms of R^2 rather than idiosyncratic variance directly, so we can express the sufficient condition from Corollary 2 in terms of R^2_{fund} and $R^2_{\text{benchmark}}$ rather than σ_u^2 and σ_e^2. Note that these R^2s are the values with respect to the true pricing model, in this case assumed to be the 4-factor Fama-French-Carhart model.[26]

Calibrating according to the sample moments and making the assumptions described above, we ultimately obtain

$$\frac{843\,\kappa^2\,R^2_{\text{fund}}(1-R^2_{\text{benchmark}})}{(27.3\,\kappa^2-7.1\kappa+12.4)\left(1-R^2_{\text{fund}}\right)\left(1-R^2_{\text{benchmark}}\right)+26.3\,\kappa^2 R^2_{\text{benchmark}}\left(1+R^2_{\text{fund}}\right)}<1. \tag{12}$$

Although a bit unwieldy, this condition depends on only three parameters. When it is satisfied, the benchmark estimator of alpha will dominate the full model estimator by the MSE criterion. When the expression on the left-hand-side is exactly equal to one, we obtain the boundary between the regions of dominance. The results for three representative values of $R^2_{\text{benchmark}} \in \{0.90, 0.95, 0.99\}$ are shown in Figure 5. These values of $R^2_{\text{benchmark}}$ are high, but this is what we would expect for a well-diversified benchmark with respect to the true asset pricing model.[27]

Figure 5 shows the regions of the parameter space, $\left(R^2_{\text{benchmark}}, R^2_{\text{fund}}, \kappa\right)$, for which the benchmark estimator in (10) dominates the "full" estimator in (9), after calibrating to the observed sample moments in Table 5 and making the simplifying assumptions described above. All panels show the R^2 of an active fund with respect to the 4-factor model on the vertical axis and the common (non-market) factor loading κ on the horizontal axis. The dashed lines show the values

[25] We already know that when non-market loadings are zero, the benchmark specification will dominate, as it includes only 1 irrelevant regressor while the full specification includes 3. And at the other extreme, sufficiently high factor exposures will make the full specification dominate. We are interested in exploring the boundary between the two regimes.

[26] Furthermore, these are the true R^2s implied by the underlying factor structure of returns, not the estimated R^2s arising from any particular regression or set of regressions. In sample sizes relevant to our setting, estimated R^2s are not only noisy but biased. Cramer (1987) provides a detailed analysis along with the following recommendation: "R^2 should not be quoted for samples of less than fifty observations."

[27] Intuitive arguments justifying the APT often invoke the existence of well-diversified portfolios with zero idiosyncratic variance. We don't consider this boundary case explicitly, but, under the conditions of Corollary 2, $R^2_{\text{benchmark}} = 1 \implies \text{MSE}\left(\widehat{\alpha}_{\text{benchmark}}\right) < \text{MSE}\left(\widehat{\alpha}\right)$, regardless of R^2_{fund} or the true factor loadings. This result obtains because, in this limiting case, the benchmark estimator is unbiased and has strictly lower variance than the full 4-factor estimator.

Panel A: $R^2_{\text{benchmark}} = 0.90$ Panel B: $R^2_{\text{benchmark}} = 0.95$ Panel C: $R^2_{\text{benchmark}} = 0.99$

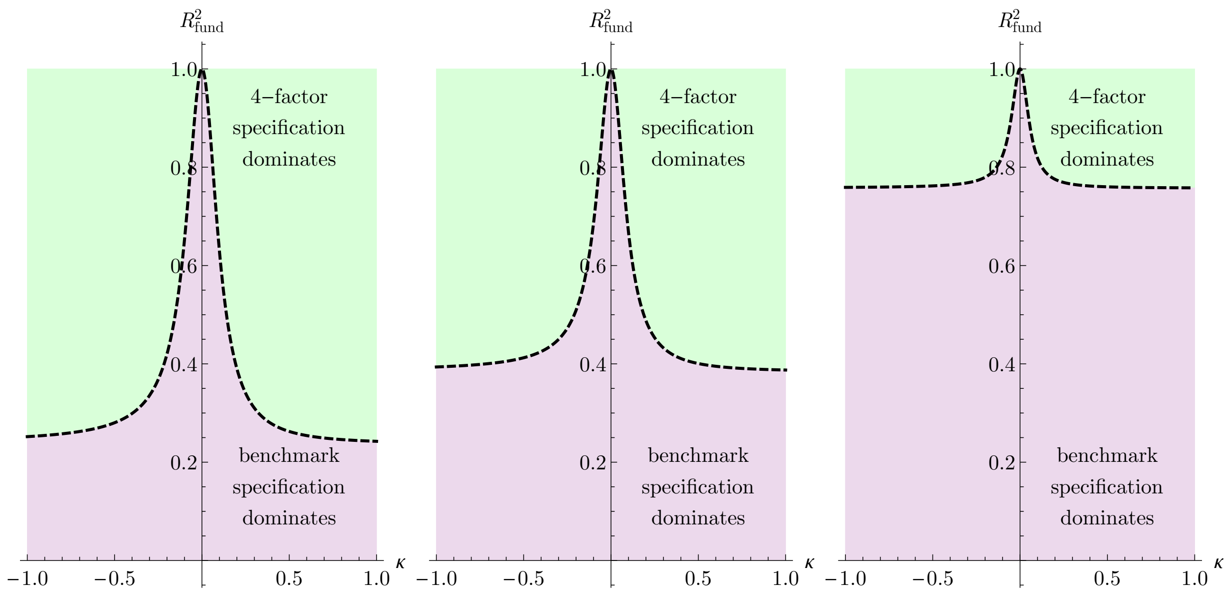

Figure 5: This figure shows the regions of the parameter space for which the benchmark estimator in (10) dominates the factor estimator in (9), after calibrating to observed sample moments and making the simplifying assumptions described in Section 4.3. All panels show the R^2 of an active fund with respect to the 4-factor model on the vertical axis and the common (non-market) factor loading κ on the horizontal axis. The dashed lines show the values for which the sufficient condition in (12) is exactly one. Below the lines, the benchmark estimator produces alpha estimates with lower MSE than the full model estimator. Panels A, B, and C correspond to the cases when the *benchmark*'s R^2 with respect to the 4-factor model is 90%, 95%, and 99%, respectively. When κ is close to zero, the benchmark estimator invariably dominates, since all non-market regressors are superfluous. When the loading on these factors is large, the 4-factor specification will eventually dominate. Note, however, that when $R^2_{\text{benchmark}}$ is sufficiently high (Panel C), dominance of the factor estimator requires a reasonably high value of R^2_{fund} as well, regardless of κ.

for which the sufficient condition in (12) is exactly one. Below the lines, the benchmark estimator produces alpha estimates with lower MSE than the full model estimator. Panels A , B, and C correspond to the cases when the *benchmark*'s R^2 with respect to the 4-factor model is 90%, 95%, and 99%, respectively. When κ is close to zero, the benchmark estimator invariably dominates, since all non-market regressors are superfluous. When the loading on these factors is large, the 4-factor specification—which holds by assumption in this example—will eventually dominate. Note, however, that when $R^2_{\text{benchmark}}$ is sufficiently high (Panel C), dominance of the factor estimator requires a reasonably high value of R^2_{fund} as well, regardless of κ.

In this section, we have provided econometric foundations for the use of the benchmark estimator of alpha instead of the usual estimator of alpha that includes all factors as independent regressors. Under the conditions established in Proposition 1, the benchmark estimator will dominate the factor model estimator, even when the true pricing model is known and all factors are observable. The calibration exercise shows that this condition is likely to be satisfied for many funds. Moreover, the bootstrap analysis in Section 3 shows that this condition *is* satisfied for nearly all of the sector funds in our sample. And the flow regressions of Section 2 show that investors are responding accordingly.

5 Experimental evidence

Our empirical results are based on fund level data. This section presents an out-of-sample micro-level test of our argument that benchmark performance affects investment choices. A benefit is that we are able to control for respondent characteristics and conduct within respondent comparisons. An important caveat, which is generally applicable, is that experimental subjects are not representative of mutual fund investors.

We conduct an online survey on the Amazon Mechanical Turk (mTurk) platform.[28] The survey recruited a sample of 149 individuals living in the U.S.[29] We collect demographic characteristics that include age, gender, race, and education. In addition, we collect financial information such as income, employment status, industry of employment, investing experience, and risk tolerance. Summary statistics of the complete set of respondent characteristics are provided in Panel A of Table B.2.

The randomized control trial assigns the participants the task of allocating \$100 into three funds: a fund investing in government securities, a general equity mutual fund, and a sector fund. Each respondent is provided five allocation scenarios where the performance numbers are randomly drawn from uniform distributions. In each scenario respondents are given performance figures

[28] Recent studies that use the same platform include Horton, Rand, and Zeckhauser (2011) and Kuziemko, Norton, Saez, and Stantcheva (2015), among others. Casler, Bickel, and Hackett (2013) compared the online responses of participants recruited via Amazon's Mechanical Turk (MTurk), social media, and face-to-face behavioral testing, and found that mTurk respondents are more diverse, yet their behavioral test results are indistinguishable from the other groups of participants.

[29] The survey on average lasted approximately 5 minutes 18 seconds. Participants received 75 cents, which is competitive in the mTurk marketplace.

for the government fund, the general equity mutual fund, the market index, and the sector fund. The returns are drawn from independent uniform distributions. Market and sector returns are drawn from an interval of $[0\%, 20\%]$, while the return on government securities ranges from 0% to 4%. Further, only the treatment group is provided performance figures for the sector index. Figure B.2 reports the screens visible to respondents before they allocate capital. Given their respective performance figures, respondents are then asked to allocate their capital. The summary statistics of allocation decisions are also provided in Panel A of Table B.2.

Using the allocation decisions of survey respondents, we conduct two simple tests. First, we test a relevance condition: whether the presence of information regarding benchmarks affects allocation decisions at all. Panel B of Table B.2 shows that survey respondents allocate more to sector funds in the presence of benchmark information.[30] Next, we test *how* the presence of benchmark information affects investment decisions. Panel C of Table B.2 reports the results for the subset of survey respondents who received benchmark performance data. Column (1) reports that survey takers respond in a statistically and economically significant manner to the hypothetical fund outperformance with respect to the hypothetical sector benchmark.

The results obtain after controlling for variation in demographic characteristics in column (2) and financial characteristics in column (3) as well. Column (4) includes individual investor fixed effects. The results suggest that benchmark information is utilized in allocation decision by respondents.

The results suggest that inherent differences between sector fund investors and other investors are not driving the allocation results. In sum, the experiment provides evidence that investors utilize benchmark information in addition to market returns.

[30]Column (1) of Panel B in Table B.2 reports that in presence of the treatment of benchmark information, survey investors allocate on average 6.3% of their portfolio to sector funds. This result remains similar after including fund and market performance metrics in column (2), demographic controls (age, income, education, marital status, gender, race and state of residence) in column (3), and financial controls (home owner status, employment status, industry of employment, investing experience, investing experience in mutual funds and risk aversion) in column (4). Thus, the presence of information regarding benchmark returns positively affects allocation to sector funds on average. In column (4), our most exhaustive specification, investors allocate approximately 5.1% of their portfolio to sector funds in the presence of benchmark information.

6 Conclusion

With approximately 8.5 trillion dollars invested in U.S equity funds in 2019, capital allocation decisions across these funds are an important issue. A concern among academics and policymakers is that U.S. households, which own a large portion of these assets, are making sub-optimal allocation decisions. Researchers before us have identified several cases of such sub-optimal behavior (see Hirshleifer, 2001; Barber and Odean, 2008; Choi, Laibson, and Madrian, 2010; Cohen and Lou, 2012, among others).

Furthermore, recent work has found that investor are not using prominent multi-factor models to estimate fund alpha when making allocation decisions. Our paper shows that this is not a sign of unsophisticated behavior. We document that investors use a benchmark model for capital allocation. This approach is appropriate given the uncertainty regarding the true asset pricing model. We further show that, in our sample, the benchmark model improves the precision of alpha estimates by an order of magnitude over multi-factor estimates. Under a diffuse prior on alpha, the benchmark model generates superior alpha estimates in 88% of the funds. Finally, we provide conditions under which the dominance of the benchmark alpha will obtain, even if the true pricing model were known and all factors were observable.

Our results demonstrate that retail funds investors do not make a naive decisions when they employ benchmarks. These investors apply a thoughtful heuristic to address a challenging problem.

References

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

Barber, Brad M., and Terrance Odean, 2008, All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *Review of Financial Studies* 21, 785–818.

Barillas, Francisco, and Jay Shanken, 2018, Comparing Asset Pricing Models, *Journal of Finance* 73, 715–754.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2019, What Do Mutual Fund Investors Really Care About?, available at ssrn.com/abstract=3292317.

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

Berk, Jonathan B., and Jules H. van Binsbergen, 2015, Measuring Skill in the Mutual Fund Industry, *Journal of Financial Economics* 118, 1–20.

Berk, Jonathan B., and Jules H. van Binsbergen, 2016, Assessing Asset Pricing Models using Revealed Preference, *Journal of Financial Economics* 119, 1 – 23.

Casler, Krista, Lydia Bickel, and Elizabeth Hackett, 2013, Separate but Equal? A Comparison of Participants and Data Gathered via Amazon's MTurk, Social Media, and Face-to-face Behavioral Testing, *Comput. Hum. Behav.* 29, 2156–2160.

Choi, James J., David Laibson, and Brigitte C. Madrian, 2010, Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds, *The Review of Financial Studies* 23, 1405–1432.

Cohen, Lauren, and Dong Lou, 2012, Complicated firms, *Journal of Financial Economics* 104, 383 – 400.

Cramer, J. S., 1987, Mean and Variance of R^2 in Small and Moderate Samples, *Journal of Econometrics* 35, 253–266.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2012, Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation, *Critical Finance Review* 2, 1–48.

Del Guercio, Diane, and Jonathan Reuter, 2014, Mutual Fund Performance and the Incentive to Generate Alpha, *Journal of Finance* 69, 1673–1704.

Evans, Richard B., and Yang Sun, 2018, Models or Stars: The Role of Asset Pricing Models and Heuristics in Investor Risk Adjustment, available at ssrn.com/abstract=3287997.

Friedman, Milton, and L. J. Savage, 1948, The Utility Analysis of Choices Involving Risk, *Journal of Political Economy* 56, 279–304.

Hartzell, Jay, Tobias Mühlhofer, and Sheridan Titman, 2010, Alternative Benchmarks for Evaluating Mutual Fund Performance, *Real Estate Economics* 38, 121–154.

Hirshleifer, David, 2001, Investor Psychology and Asset Pricing, *Journal of Finance* 56, 1533–1597.

Horton, John J., David G. Rand, and Richard J. Zeckhauser, 2011, The Online Laboratory: Conducting Experiments in a Real Labor Market, *Experimental Economics* 14, 399–425.

Jegadeesh, Narasimhan, and Chandra Sekhar Mangipudi, 2019, What Do Fund Flows Reveal About Asset Pricing Models and Investor Sophistication?, available at ssrn.com/abstract=3090541.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the Industry Concentration of Actively Managed Equity Mutual Funds, *Journal of Finance* 60, 1983–2011.

Kinal, Terrence, and Kajal Lahiri, 1983, Specification Error Analysis with Stochastic Regressors, *Econometrica* 51, 1209–1219.

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva, 2015, How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments, *American Economic Review* 105, 1478–1508.

Merton, Robert C., 1969, Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case, *Review of Economics and Statistics* 51, 247–257.

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do Industries Explain Momentum?, *Journal of Finance* 54, 1249–1290.

Pástor, L'uboš, and Robert F. Stambaugh, 2002, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 63, 315–349.

Pástor, L'uboš, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and Skill in Active Management, *Journal of Financial Economics* 116, 23 – 45.

Sensoy, Berk A., 2009, Performance Evaluation and Self-Designated Benchmark Indexes in the Mutual Fund Industry, *Journal of Financial Economics* 92, 25 – 39.

Sirri, Erik R., and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *Journal of Finance* 53, 1589–1622.

Stambaugh, Robert F., 2014, Presidential Address: Investment Noise and Trends, *Journal of Finance* 69, 1415–1453.

Teräsvirta, Timo, 1987, Usefulness of Proxy Variables in Linear Models with Stochastic Regressors, *Journal of Econometrics* 36, 377–382.

Thursby, Jerry G., 1989, Expected MSE of Errors-in-Variables and Omitted Variables Models, *Economics Letters* 30, 211–216.

Weber, Elke U., Ann-Rene Blais, and Nancy E. Betz, 2002, A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors, *Journal of Behavioral Decision Making* 15, 263–290.

Table 1: Summary Statistics

Panel A: Summary Statistics for Funds by Sector at the Share Class Level

The table reports summary statistics for all funds in our sample at the share class level by sector. For each sector, we provide distributional statistics for the number of funds active each year, the total net assets in millions of USD, and the sector level total net assets. Numbers are rounded to the nearest integer. The data sample period is 1999–2016.

	Mean	**Stdev**	**1st Quartile**	**Median**	**3rd Quartile**
Health and Biotechnology, $N = 274$					
Number of Funds Active in a Given Year	140	40	110	143	180
Total Net Assets (Millions of $)	386	1,487	7	35	170
Sector TNA (Millions of $)	51,632	27,224	31,915	46,543	57,267
Natural Resources, $N = 222$					
Number of Share Classes Active in a Given Year	107	29	87	112	127
Total Net Assets (Millions of $)	310	959	8	45	205
Sector TNA (Millions of $)	31,287	20,167	7,613	35,390	44,621
Real Estate, $N = 473$					
Number of Share Classes Active in a Given Year	250	71	202	278	303
Total Net Assets (Millions of $)	291	1,223	8	40	172
Sector TNA (Millions of $)	69,132	48,351	24,059	63,373	106,259
Science and Technology, $N = 471$					
Number of Share Classes Active in a Given Year	244	64	195	220	303
Total Net Assets (Millions of $)	239	830	4	24	125
Sector TNA (Millions of $)	55,426	30,733	35,864	43,986	70,372
Telecommunications, $N = 63$					
Number of Share Classes Active in a Given Year	37	8	35	37	43
Total Net Assets (Millions of $)	149	422	1	7	99
Sector TNA (Millions of $)	5,095	2,178	3,362	4,880	6,520
Utilities, $N = 184$					
Number of Share Classes Active in a Given Year	99	14	92	97	105
Total Net Assets (Millions of $)	264	659	8	41	194
Sector TNA (Millions of $)	25,134	8,381	18,227	24,333	31,227
Combined TNA across all Sectors (Millions of $)	237,706	101,609	157,173	224,775	282,297

Panel B: Descriptive Statistics of Mutual Fund Performance Data

This table reports summary statistics of mutual fund performance data for the sample period (years 1999 to 2016). Data are at monthly frequency. α is computed as the difference between realized excess returns in month t minus realized excess returns to a set of benchmarks the same month, each multiplied by its respective Beta. For $fund - sec$ the subject returns are the returns to a fund, and benchmarks are the fund's sector benchmark and the S&P 500. For $sec - S\&P$ the subject returns are the return to the fund's sector benchmark, and the benchmark is the S&P 500. Fund size represents the size of the fund in logarithm of Total Net Assets.

	mean	sd	min	max	N
Fund Flow (Next Month)	-0.0007	0.3094	-9.9958	9.8027	119,304
S&P$_t$	0.0059	0.0410	-0.1674	0.1096	121,631
Sector Benchmark Return	0.0071	0.0547	-0.2943	0.2675	121,631
Fund Return	0.0068	0.0612	-0.7502	2.9521	121,489
α(fund - sec)$_t$	-0.0002	0.0209	-0.6243	2.9129	120,707
α(sec - S&P)$_t$	0.0011	0.0318	-0.1339	0.1612	121,631
4 factor α_t	0.0002	0.0341	-0.4128	2.8031	120,707
(S&P - Risk-free)$_t$	0.0047	0.0411	-0.1679	0.1096	121,631
Sign of Index-Adjusted Fund Flow	-0.1461	0.9892	-1.0000	1.0000	121,107
Difference of Signs	-0.0474	1.2151	-3.1500	3.1500	120,568
Fund Size (log TNA)	3.7517	2.2835	-2.3026	10.5039	119,738

Table 2: Model Comparison: Berk and van Binsbergen (2016) Approach

Dependent variable is the sign of Index-Adjusted Fund Flow. This table shows results for a fund-level panel regression of the sign of the index-adjusted flow of funds to a mutual fund, on the difference in signs of fund outperformance measured by the two-factor sector-plus-market model and the one-factor market (CAPM) model. Data is at fund-month level. Standard errors are clustered at the fund level.

	(1)	(2)	(3)	(4)
	Sign of Index-Adjusted Fund Flow			
Difference of Signs	0.0659***	0.0669***	0.0624***	0.0521***
(Two factor − CAPM)	(16.57)	(16.61)	(16.21)	(15.51)
Fund Size		0.0183***	0.0157***	0.0285***
		(4.43)	(3.81)	(3.29)
Constant	-0.236***	-0.309***	-0.301***	-0.357***
	(-24.74)	(-17.62)	(-16.82)	(-10.47)
Sector FE	No	No	Yes	No
Fund FE	No	No	No	Yes
Observations	119989	118686	118686	118686
Adjusted R^2	0.008	0.009	0.017	0.005

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 3: Model Comparison: Hybrid Specification

Dependent variable is the sign of index-adjusted fund flow. This table shows results for a fund-level panel regression of the sign of the index-adjusted flow of funds to a mutual fund, on the indicator variable that captures fund outperformance measured by the two-factor sector-plus-market model compared to the one-factor market (CAPM) model. The first two columns have fund performance ranked into deciles as in Barber, Huang, and Odean (2016) and the next two have fund performance ranked into terciles. Data is at fund-month level. Standard errors are clustered at the fund level.

	(1)	(2)	(3)	(4)
	Sign of Index-Adjusted Fund Flow			
Rank(α_m)$\leq$ Rank(α_s)	0.170***	0.175***	0.178***	0.183***
	(13.96)	(14.22)	(13.62)	(13.91)
Fund Size		0.0315***		0.0216**
		(3.42)		(2.22)
Constant	-0.351***	-0.479***	-0.353***	-0.442***
	(-55.70)	(-12.96)	(-52.90)	(-11.40)
Fund FE	Yes	Yes	Yes	Yes
Observations	97403	96462	59809	59312
Adjusted R^2	0.008	0.009	0.009	0.010

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 4: Are Investors Chasing Sector Returns?

Dependent variable is the sign of fund flow. This table shows results for a fund-level panel regression of the sign of the flow of funds to a mutual fund, on the difference in signs of fund outperformance measured by the two-factor sector-plus-market model and fund outperformance measured as the raw return net of sector benchmark (return chasing against sector benchmark approach). Data is at fund-month level. Standard errors are clustered at the fund level.

	(1)	(2)	(3)	(4)
	Sign of Fund Flow			
Difference of Signs	0.0187***	0.0190***	0.0124***	0.0128***
(Two factor − Raw sector)	(4.34)	(4.37)	(2.91)	(3.64)
Fund Size		0.00651	0.00277	-0.0465***
		(1.28)	(0.55)	(-5.30)
Constant	-0.218***	-0.242***	-0.227***	-0.0331
	(-19.09)	(-11.81)	(-10.92)	(-0.96)
Sector FE	No	No	Yes	No
Fund FE	No	No	No	Yes
Observations	118977	117732	117732	117732
Adjusted R^2	0.000	0.001	0.011	0.002

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 5: Data for Calibration

Monthly means, volatilities, and correlations for the Fama-French-Carhart factors estimated from data for the period January 2000–January 2019. Data are obtained from WRDS.

Factor	Mean [%]	Volatility [%]	Correlation Matrix			
Mkt $-$ R_f	0.42	4.37	1.00			
SMB	0.27	3.31	0.27	1.00		
HML	0.26	3.13	-0.06	-0.26	1.00	
UMD	0.17	5.35	-0.36	0.11	-0.17	1.00

Appendix

A Proof of Proposition 1

The benchmark estimator in (10) can be interpreted as a restricted estimation of the full model,

$$R_t = \alpha_{\text{benchmark}} + \left(\begin{array}{c:c:c} F_t & B_t & \psi_t \end{array} \right) \begin{pmatrix} \beta_{\text{benchmark}} \\ \delta_{\text{benchmark}} \\ \eta_{\text{benchmark}} \end{pmatrix} + u_t, \tag{13}$$

where the scalar $\psi_t = F_t (\beta - \beta_B) + Z_t (\gamma - \gamma_B) - e_t$ and $\eta_{\text{benchmark}} = 0$. Observe that setting $\beta_{\text{benchmark}} = 0$ and $\delta_{\text{benchmark}} = \eta_{\text{benchmark}} = 1$ in (13) recovers the true data generating process in (7).

Now that the benchmark estimator has been recast as an omitted variables specification, and given Assumption 3, the sufficient condition from Kinal and Lahiri (1983, p. 1215) implies that

$$\left(\frac{T-p-3}{\sigma_u^2} \right) \Xi < 1 \implies \text{MSE}\left(\widehat{\alpha}_{\text{ov}}\right) < \text{MSE}\left(\widehat{\alpha}\right), \tag{14}$$

where

$$\Xi = \text{Var}\left[\psi_t\right] - \text{Cov}\left[\psi_t, \begin{pmatrix} F_t' \\ B_t \end{pmatrix}\right] \left(\text{Var} \begin{pmatrix} F_t' \\ B_t \end{pmatrix} \right)^{-1} \text{Cov}\left[\psi_t, \begin{pmatrix} F_t' \\ B_t \end{pmatrix}\right]'. \tag{15}$$

Note that, compared to the expression in Kinal and Lahiri (1983), the degrees of freedom in (14) have been adjusted to reflect the inclusion of a constant in the regression.

From Assumption 3, we have

$$\begin{aligned} &\text{Var}\left[\psi_t\right] = \\ &\qquad (\beta - \beta_B)' \Sigma_{FF} (\beta - \beta_B) + (\gamma - \gamma_B)' \Sigma_{ZZ} (\gamma - \gamma_B) + 2 (\beta - \beta_B)' \Sigma_{FZ} (\gamma - \gamma_B) + \sigma_e^2, \end{aligned} \tag{16}$$

$$\text{Var} \begin{pmatrix} F_t' \\ B_t \end{pmatrix} = \left(\begin{array}{c:c} \Sigma_{FF} & \\ \hdashline \beta_B' \Sigma_{FF} + \gamma_B' \Sigma_{ZF} & \beta_B' \Sigma_{FF} \beta_B + \gamma_B' \Sigma_{ZZ} \gamma_B + 2 \beta_B' \Sigma_{FZ} \gamma_B + \sigma_e^2 \end{array} \right), \tag{17}$$

and

$$\begin{aligned} &\text{Cov}\left[\psi_t, \begin{pmatrix} F_t' \\ B_t \end{pmatrix}\right]' = \\ &\left(\begin{array}{c} \Sigma_{FF} (\beta - \beta_B) + \Sigma_{FZ} (\gamma - \gamma_B) \\ \hdashline \beta_B' \Sigma_{FF} (\beta - \beta_B) + \gamma_B' \Sigma_{ZZ} (\gamma - \gamma_B) + \gamma_B' \Sigma_{ZF} (\beta - \beta_B) + \beta_B' \Sigma_{FZ} (\gamma - \gamma_B) - \sigma_e^2 \end{array} \right). \end{aligned} \tag{18}$$

To compute the inverse of the partitioned variance matrix in (17), we make use of the following formula.

Lemma 1 *Banachiewicz inversion formula***.** Let M denote an invertible partitioned matrix,

$$M = \left(\begin{array}{c|c} P & Q \\ \hline R & S \end{array}\right).$$

If P and $M/P \equiv S - RP^{-1}Q$, are invertible, then

$$M^{-1} = \left(\begin{array}{c|c} P^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) + \left(\begin{array}{c} -P^{-1}Q \\ \hline I \end{array}\right) (M/P)^{-1} \left(\begin{array}{c|c} -RP^{-1} & I \end{array}\right).$$

■

Applying this to (17) gives

$$\mathrm{Var}\begin{pmatrix} F_t' \\ B_t \end{pmatrix}^{-1} = \left(\begin{array}{c|c} \Sigma_{FF}^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) + \frac{1}{\gamma_B'\Sigma_{ZZ\cdot F}\gamma_B + \sigma_e^2} \left(\begin{array}{c} \beta_B + \Sigma_{FF}^{-1}\Sigma_{FZ}\gamma_B \\ \hline -1 \end{array}\right) \left(\begin{array}{c} \beta_B + \Sigma_{FF}^{-1}\Sigma_{FZ}\gamma_B \\ \hline -1 \end{array}\right)'. \quad (19)$$

where $\Sigma_{ZZ\cdot F} \equiv \Sigma_{ZZ} - \Sigma_{ZF}\Sigma_{FF}^{-1}\Sigma_{FZ}$, the conditional variance of Z. Substituting the results from (16), (19), and (18) into (15) gives the full expression for Ξ,

$$\begin{aligned}\Xi = {} & (\beta - \beta_B)' \Sigma_{FF} (\beta - \beta_B) + (\gamma - \gamma_B)' \Sigma_{ZZ} (\gamma - \gamma_B) + 2 (\beta - \beta_B)' \Sigma_{FZ} (\gamma - \gamma_B) + \sigma_e^2 - \\ & \left(\begin{array}{c} \Sigma_{FF} (\beta - \beta_B) + \Sigma_{FZ} (\gamma - \gamma_B) \\ \hline \beta_B'\Sigma_{FF} (\beta - \beta_B) + \gamma_B'\Sigma_{ZZ} (\gamma - \gamma_B) + \gamma_B'\Sigma_{ZF} (\beta - \beta_B) + \beta_B'\Sigma_{FZ} (\gamma - \gamma_B) - \sigma_e^2 \end{array}\right)' \times \\ & \left[\left(\begin{array}{c|c} \Sigma_{FF}^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) + \frac{1}{\gamma_B'\Sigma_{ZZ\cdot F}\gamma_B + \sigma_e^2} \left(\begin{array}{c} \beta_B + \Sigma_{FF}^{-1}\Sigma_{FZ}\gamma_B \\ \hline -1 \end{array}\right) \left(\begin{array}{c} \beta_B + \Sigma_{FF}^{-1}\Sigma_{FZ}\gamma_B \\ \hline -1 \end{array}\right)'\right] \times \\ & \left(\begin{array}{c} \Sigma_{FF} (\beta - \beta_B) + \Sigma_{FZ} (\gamma - \gamma_B) \\ \hline \beta_B'\Sigma_{FF} (\beta - \beta_B) + \gamma_B'\Sigma_{ZZ} (\gamma - \gamma_B) + \gamma_B'\Sigma_{ZF} (\beta - \beta_B) + \beta_B'\Sigma_{FZ} (\gamma - \gamma_B) - \sigma_e^2 \end{array}\right)\end{aligned}$$

After some simplification, we obtain

$$\Xi = \gamma^{'}\Sigma_{ZZ\cdot F}\gamma - \frac{\left(\gamma^{'}\Sigma_{ZZ\cdot F}\gamma_B\right)^2}{\gamma_B^{'}\Sigma_{ZZ\cdot F}\gamma_B + \sigma_e^2}.$$

This and (14) gives the result.

B Additional figures and tables

Table B.1: Cross-correlation table

The table reports correlations between monthly value weighted sector returns in our sample with monthly returns of the market index. Sample period is from 1999 to 2016.

Variables	S&P 500 Index	Health and Biotech	Natural Resources	Real Estate	Science and Tech.	Telecom	Utilities
S&P 500 Index	1.000						
Health and Biotech	0.753	1.000					
Natural Resources	0.675	0.477	1.000				
Real Estate	0.605	0.522	0.426	1.000			
Science and Tech.	0.848	0.613	0.571	0.441	1.000		
Telecom	0.885	0.633	0.589	0.502	0.908	1.000	
Utilities	0.823	0.659	0.728	0.545	0.623	0.736	1.000

Table B.2: Experimental Evidence

Panel A: Descriptive Statistics of Allocation Experiment

The table reports summary statistics of fund and index performance data for the experiment. The table also reports summary statistics of demographic and financial experience information collected during the survey. Risk Aversion is measured using the approach of Weber, Blais, and Betz (2002), who provide the survey questions in their Appendix C, Questions I and G. The Domain-Specific Risk-Taking (DOSPERT) scale is a psychometric scale that assesses risk taking in five content domains, one of which is financial decisions.

	mean	sd	min	max	N
Sector Fund Allocation	0.35	0.26	0.00	1.00	745
Market Fund Allocation	0.45	0.28	0.00	1.00	745
Treatment	0.41	0.49	0.00	1.00	745
Sector Fund Return	0.10	0.06	0.00	0.20	745
Market Fund Return	0.10	0.06	0.00	0.20	745
Government Fund Return	0.02	0.014	0.00	0.04	745
Market Index Return	0.10	0.06	0.00	0.20	745
Sector Index Return	0.10	0.06	0.00	0.20	305
Government Index Return	0.02	0.014	0.00	0.04	745
Age	31–35	N/A	18–20	61–65	745
Income ('000s)	43.83	31.21	0.00	190.00	745
Gender	0.50	0.50	0.00	1.00	745
Investor?	0.60	0.49	0.00	1.00	730
Investor in Mutual Fund?	0.47	0.50	0.00	1.00	690
Risk Aversion Measure	3.84	0.52	2.13	4.75	745

Panel B: Sector Fund Allocation with or without Benchmark Information

Dependent variable: Sector Fund Allocation, next period. This table shows regressions of sector allocation decision, on an intercept, returns to the market fund, sector fund, market index, and the random treatment which is information regarding the benchmark index performance. The table utilizes allocation data from a panel of survey respondents.

	Sector Fund Allocation			
	(1)	(2)	(3)	(4)
Treatment	0.0630***	0.0629***	0.0669***	0.0511*
	(3.33)	(3.59)	(3.80)	(2.37)
Sector Fund		1.442***	1.426***	1.422***
		(9.81)	(9.54)	(7.85)
Market Fund		-0.972***	-0.957***	-0.816***
		(-6.54)	(-6.45)	(-4.55)
Market Index		-0.230	-0.176	-0.211
		(-1.56)	(-1.19)	(-1.12)
Constant	0.328***	0.307***	0.285***	0.281*
	(26.10)	(10.39)	(5.74)	(2.14)
Demographic Controls	No	No	Yes	Yes
Financial Controls	No	No	No	Yes
Observations	745	745	735	545
Adjusted R^2	0.013	0.175	0.183	0.152

t statistics in parentheses, standard errors are robust to heteroscedasticity.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel C: Experimental Evidence regarding Information Availability

Dependent variable: Sector Fund Allocation, next period. This table shows regressions of sector allocation decision, on an intercept, returns to the market index, the difference between returns to the sector benchmark and returns to the market index (Sec − Mkt), and the difference between returns to the fund and returns to the sector benchmark ($Fund - Sec$).

	Sector Fund Allocation			
	(1)	(2)	(3)	(4)
Mkt	1.652***	1.602***	1.732***	1.674***
	(4.64)	(4.49)	(3.78)	(3.86)
Sec − Mkt	1.719***	1.691***	1.903***	1.816***
	(5.97)	(6.03)	(5.49)	(5.41)
Fund − Sec	1.144***	1.134***	1.281***	1.201***
	(5.10)	(5.23)	(5.17)	(4.51)
Constant	0.227***	0.197**	0.396	0.215***
	(6.42)	(2.88)	(1.77)	(4.81)
Demographic Controls	No	Yes	Yes	No
Financial Controls	No	No	Yes	No
Individual FE	No	No	No	Yes
Observations	305	305	240	240
Adjusted R^2	0.092	0.125	0.160	0.311

t statistics in parentheses, standard errors are robust to heteroscedasticity.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

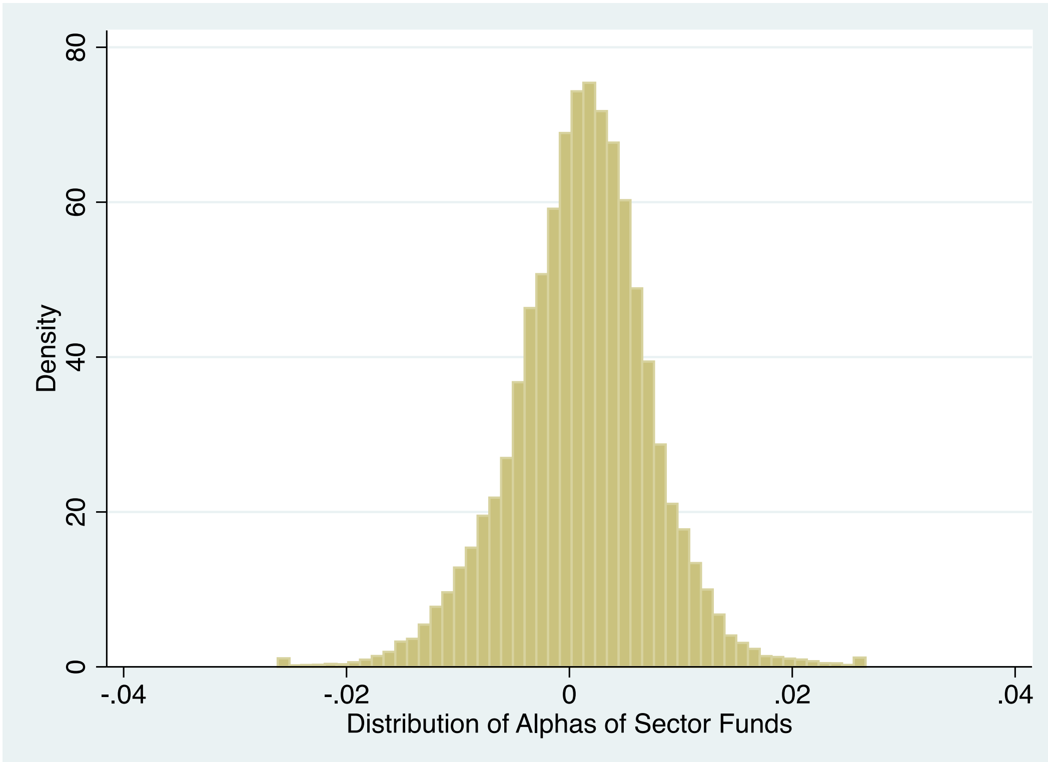

Figure B.1: Risk-Adjusted Out-performance of Sector Funds

This figure reports the four factor risk-adjusted three year rolling alpha of sector funds in the sample. Unit of observation is fund-month and the sample period is from Jan 1999–Dec 2016.

Vignette

You have saved $100 and have decided to invest the money in order to fund your lifestyle in the future.

The following screens will provide information about various investment options which are available to you. After reviewing the information, you will allocate your money into the investments as you prefer.

>>

(a) A brief description viewed by respondents before the task.

The below information reports the returns of the funds and benchmark indices from the last period. Please allocate your $100.

Fund	Return (%)
Stock Market Fund	4
Sector Fund	6
Risk-free Government Securities Fund	4

Benchmark Index	Return (%)
Stock Market Index	1
Sector Index	0
Risk-free Government Securities Index	3

Please allocate your funds into the investments as you desire.

Stock Market Fund	0
Sector Stock Fund	0
Risk-free Government Securities Fund	0
Total	0

(b) An allocation example for the treatment group.

The below information reports the returns of the funds and benchmark indices from the last period. Please allocate your $100.

Fund	Return (%)
Stock Market Fund	19
Sector Fund	18
Risk-free Government Securities Fund	3

Benchmark Index	Return (%)
Stock Market Index	7
Risk-free Government Securities Index	2

Please allocate your funds into the investments as you desire.

Stock Market Fund	0
Sector Stock Fund	0
Risk-free Government Securities Fund	0
Total	0

(c) An allocation example for the control group.

Figure B.2: Experiment Screens

The figure reports the description respondents view before the task, and allocation examples for the treatment and control group. The control group does not receive a benchmark index return.