TI 2019-029/IV
Tinbergen Institute Discussion Paper

Constructing and using double-adjusted alphas to analyze mutual fund performance

Erik Kole[1]
Reza Brink[2]

[1] Erasmus University Rotterdam, Tinbergen Institute
[2] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam

Constructing and using double-adjusted alphas to analyze mutual fund performance

Reza Brink Erik Kole*

Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam

April 16, 2019

Abstract

We propose a new approach for estimating mutual fund performance that simultaneously controls for both factor exposure and firm characteristics. This double-adjusted alpha is motivated by the recent findings that traditional Fama-French style factor models do not fully adjust returns for the anomalies related to the factors. We formulate a hierarchical Bayesian model which separates the part of the traditional alpha that can be related to firm and asset characteristics from the true alpha. Our Bayesian approach is straightforward, has theoretical advantages over the traditional two-pass estimation and leads to higher precision. Our double-adjusted alphas produce a different ranking of mutual funds than the traditional alphas. We show that as a consequence, the double-adjusted alphas lead to stronger evidence of persistence of mutual fund performance. On the other hand, we find that the link between selectivity and alpha is driven by the effect of characteristics. Finally, we show that fund flows are mostly driven by the true skill part of the return and hardly by the effect of characteristics. We conclude that good measurement of the true outperformance of mutual funds is crucial for understanding skill.

Keywords: Mutual fund performance, Double-adjusted performance, Firm characteristics, Hierarchical Bayes

JEL codes: G11, G23, C11

*Corresponding author. Address: Burg. Oudlaan 50, Room ET-31, P.O. Box 1738, 3000DR Rotterdam, The Netherlands, Tel. +31 10 408 12 58. E-mail addresses kole@ese.eur.nl.

1 Introduction

Evalutions of the performance of mutual funds commonly resort to factor models such as the Capital Asset Pricing Model (CAPM), the three-factor model of Fama & French (1993) or the model of Carhart (1997). Recently, the beta-pricing relation underlying these models has been put under scrutiny, as the empirical asset pricing literature has looked beyond a unilateral explanation of cross-sectional variation in equity returns by factor exposures. Instead, considerable evidence indicates that firm and asset characteristics such as market capitalization, book-to-market ratios, and past returns substantially help explaining the cross section of equity returns, see for example Avramov & Chordia (2006); Brennan et al. (1998); Chordia et al. (2015).

The conclusion that both factors and characteristics should be used to explain equity returns implies that both should also be used in the evaluation of the performance of mutual funds and the determination of their alpha. Chen et al. (2018) show that using either factor models or characteristics to adjust performance leads to materially different values for alpha, and a different alpha-based ranking of mutual funds. Busse et al. (2017) propose a double-adjusted alpha as a performance measure. In the first step, returns of a mutual fund are adjusted for factor exposures, and in the second for characteristics of the fund's holdings. They show by means of simulation that their double-adjusted alpha is less prone to errors in performance evaluation than the traditional single-adjusted alpha approach that only adjusts returns for factor exposures.

The improvement in performance evaluation offered by using both factors and characteristics comes at the cost of a more complex analysis. Typically a two-pass procedure is employed, where factor exposures are estimated in the first pass by time-series regressions, and the rewards to factor exposures and characteristics in the second pass by a cross-sectional regression (cf. Busse et al., 2017; Chordia et al., 2015). The pricing errors or alphas, which are taken as measure for skill by Busse et al. (2017) also follow from this second pass. Because the factor exposures are estimated, the second pass suffers from an errors-in-variable bias. Chordia et al. (2015) show how to correct for this bias, whereas Jegadeesh et al. (2016) use an instrumental variable approach to solve it. Using the intercepts of the first pass as dependent variables in a regression on characteristics as in Busse et al. (2017) does not suffer from an errors-in-variables bias, but induces heteroskedasticity and should address that these intercepts are not observed.

In this paper, we show how to circumvent the complexities arising from the two-pass estimation by a hierarchical Bayesian approach to estimating a double-adjusted alpha. The model that we propose for asset returns consists of three layers. In the first layer, we relate daily asset returns to risk factors by beta-coefficients and an intercept that all vary on a monthly basis. This intercept corresponds with the traditional alpha. In the next layer, we relate this monthly intercept to asset characteristics, a time-random effect and an asset-specific component. The sum of the two latter terms gives the double-adjusted alpha. In the third layer, we specify that the time-random effect and the coefficients for the asset characteristics for each month are drawn from a given distribution with constant mean and variance. The asset-specific component is drawn from a distribution with a zero mean and a variance that varies per month. Our model is similar to Cederburg & ODoherty (2015), though we allow for more risk factors in the first layer. Moreover, they combine monthly returns with yearly varying coefficients. Cosemans et al. (2015) use a similar setup to model time-variation in betas related to firm characteristics.

As in Busse et al. (2017), our double-adjusted alpha is constructed by adjusting for exposure to risk factors and the effect of characteristics, but our approach comes with several advantages. First, we can estimate it in a single pass, which leads to efficiency gains and inferences for the double adjusted alphas that account for the estimation uncertainty from both the time series and the cross-sectional domain. Second, when estimating the parameters for mutual fund returns, the time-random effect in the double-adjusted alpha measures the out- or underperformance of the average mutual fund in a particular month. The mean of the distribution from which the time-random effects are drawn measures the extent to which the average mutual fund can deliver outperformance. We call this mean the aggregate alpha. Third, because we use a Bayesian approach, we can specify a prior distribution for it. For example, if we believe that the model uses the correct risk factors and characteristics, and that the average mutual funds does not possess skill, the prior distribution of the aggregate alpha should be centered on zero. The posterior distribution then shows the support for this hypothesis in the data.

We use the hierarchical Bayes approach to construct double-adjusted alphas for actively managed equity mutual funds in the U.S. over the period 2001–2016. We take the risk factors from the five-factor model of Fama & French (2015), augmented with a momentum factor as in Carhart (1997). Jordan & Riley (2016) provide evidence favoring the addition of the profitability and investment factors next to the standard market, size, value and momentum factors. We also include the corresponding characteristics: marktet capitalization, book-to-market ratio, momentum, operating profitability and asset growth. We aggregate the firm characteristics of a fund's holdings to compute aggregate monthly fund-level characteristics. Similar to the evidence on stock returns, we find that mutual fund factor betas and holdings characteristics are correlated, but the correlations are modest in magnitude (e.g., the average of the absolute value is roughly 0.4), implying that factor betas and characteristics do not convey identical information. When we conduct monthly cross-sectional regressions of mutual fund returns on factor betas and characteristics as in Avramov & Chordia (2006), we find that both account for about half of the model explained variation. This highlights the importance of including both and constructing double-adjusted alphas.

Our results confirm that characteristics should be taken into account when evaluating mutual fund performance. The posterior distribution of the coefficient for each characteristic when included in isolation strongly indicates values different from zero, except for profitability. Including all characteristics simultaneously yields similar estimates to the univariate results. This result means that funds can exhibit higher relative performance based on standard factor model alpha by passively loading on characteristics, even when the factor model explicitly adjust returns for those characteristics. When we use the frequentist two-pass procedure as in Busse et al. (2017), we do not find an effect for the size, momentum and investments characteristics. In our Bayesian approach, the posterior mean of the aggregate alpha is negative, and the 95% posterior density does not include zero. The two-pass procedure yields estimates that are close to this posterior mean, but it's not significant, indicating that our Bayesian approach leads to higher precision.

Next, we analyze the impact of our double-adjusted performance measure on relative mutual fund performance. We find that the median change in percentile performance ranking is roughly 9%. That is, a fund ranked in the median percentile according to the traditional six-factor alpha would be ranked in the 41th or 59th percentile based on our double-adjusted measure. Furthermore, we find that a large number of funds exhibit dramatic changes in percentile ranks, with 10 (5) percent of funds exhibiting a mean change in percentile ranking greater than 22.99% (27.24%).

We continue our analysis by showing that replacing the traditional alpha by our more refined double-adjusted alpha has consequences for our understanding of the causes and consequences of skill in the mutual fund industry. First, we show that the changes in relative performance alter inferences on the persistence of mutual fund performance as documented in Bollen & Busse (2004); Carhart (1997). We create decile portfolios of mutual funds ranked on their traditional alpha and their double-adjusted alpha based on the past 24 months. In both cases we observe that the top decile beats the bottom decile in terms of alpha over periods of one quarter, one year and three years. However, when we rank on double-adjusted alpha, significance levels are higher with t-statistics of the difference well over 5, the difference is also present in returns, and persistence lasts up to six years after portfolio formation. These results confirm the findings of Busse et al. (2017), though we examine a shorter period. The part of the traditional alpha that is related to characteristics does not predict differences in returns or alpha, showing that that part of the traditional alpha should be excluded from a measure of skill.

Second, we show that skill can no longer be linked to the selectivity of a fund. Amihud & Goyenko (2013) document that the R^2 of the first-pass regression of fund returns on the returns to factor portfolios is negatively related to alpha. They argue that low R^2 values show that fund managers deviate from factor portfolios, and interpret the relation with alpha as a sign of skill. We find that low R^2 values are related to the part of the traditional alpha that is related to characteristics, but not to the double-adjusted alpha that is a better measure for skill. It means that as commonly found in finance literature low values of R^2 indicate that factor models do not explain returns well, but they cannot be directly interpreted as a measure for selectivity related to skill. Our results are a bit stronger than those reported by Busse et al. (2017) over a longer time sample period.

Third, we investigate the relation between fund flows, factor exposures and alpha as in Barber et al. (2016). They argue that truly sophisticated investors should adjust the returns of mutual funds for all factor exposures and effects of characteristics, and should select funds only based on their true alpha. In their analysis they only take factor exposures into account, but do not investigate the effect of characteristics. We complement their analysis by showing that fund flows are most responsive to the double-adjusted alpha. The effect of the double-adjusted alpha on fund flows is ten times larger than the part of alpha that is related to characteristics. The effect of the characteristic part is mainly driven by size and momentum. We interpret this result as evidence that at least part of the investment community can really distinguish skilled fund managers.

Our findings contribute to the literature in two ways. First, we show how to efficiently estimate double-adjusted alphas based on a hierarchical Bayesian model that combines both time series and cross-sectional information, but without the need for a two-pass procedure. We theoretically argue and empirically show that our approach leads to more precise estimates of the double-adjusted alpha and the effect of characteristics than the two-pass approach that is used by Busse et al. (2017); Chordia et al. (2015).

Second, we confirm and extend the conclusion of Busse et al. (2017) that the effect of characteristics present in traditional alphas contaminates further analyses related to mutual fund performance. Conform their results, rankings of mutual funds change substantially, leading to stronger evidence of persistence in mutual fund performance. We also find that the relation between time series R^2

and skill disappears. As an extension, we show that the relation between alpha and fund flows survives when we use double-adjusted alphas. Taken together, we conclude that the method to measure alpha is crucial in the analysis of skill in mutual funds.

This paper proceeds as follows. Section 2 describes the data set, the construction of the fund-level characteristics, and the cross-sectional regressions of mutual fund returns on factor betas and characteristics. Section 3 describes our hierarchical Bayes approach to obtain our double-adjusted performance measure. Section 4 examines the implications of this new performance measure on previous findings in the mutual fund literature. Section 5 contains robustness checks. Finally, Section 6 concludes.

2 Data

This section describes our data set and sample selection criteria, followed by summary statistics for the mutual fund sample including the firm characteristics obtained from the holdings of mutual funds. As a preliminary indication of the importance of characteristics, we run cross-sectional regressions of individual mutual fund returns on their factor betas and characteristics using the traditional two-pass procedure of Fama & MacBeth (1973) with rolling estimation of factor betas. To address the inherent errors-in-variables (EIV) bias, we use the EIV-corrected estimator of Chordia et al. (2015).

2.A Data Selection

The mutual fund database is constructed by combining the Center for Research in Security Prices (CRSP) Survivor-Bias-Free U.S. Mutual Fund database with the Thomson Reuters Mutual Fund Holdings S12 database, formerly known as CDA/Spectrum. To combine these databases, we rely on the MFLINKS database provided by Russ Wermers on Wharton Research Data Services (WRDS). The main focus is on U.S. equity actively managed mutual funds, for which the data on holdings is the most complete and reliable; we eliminate balanced, bond, money market, international, index, and sector funds, as well as funds not invested primarily in common stocks (for details on our selection, see Appendix A).

We obtain both monthly and daily mutual fund returns from the CRSP mutual fund database. Additional stock-level information on the fund holdings are retrieved from the CRSP monthly stock file and the Compustat database. Consistent with previous literature, this study only considers common stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange and NASDAQ; we excluding real estate trusts, foreign companies, closed-end funds and primes (we only retain shares codes 10 or 11). The final database for the empirical analysis consists of monthly panel data on mutual fund holdings at the intersection of CRSP, Compustat and Thomson Reuters, spanning the period from January 2001 to December 2016.

2.B Mutual Fund Holdings Characteristics

For each stock in a fund's portfolio, we obtain firm characteristics, including market capitalization, book-to-market ratio, momentum return, operating profitability and asset growth. Market capitalization (Mcap) is defined as the product between the previous month-end stock price and the previous month-end total shares outstanding. Book-to-Market (B/M) is the ratio between the most recently available book value of equity and the previous month-end market capitalization.[1] Momentum (Mom12) is the past twelve-month cumulative return over the period from month $t-12$ to $t-2$, where the most recent month is excluded to avoid short-term reversal effects. Operating profitability (Profit) is the current revenues minus costs of goods sold, interest expense, selling, general, and administrative expenses, divided by book equity for the last fiscal year $t-1$. Asset growth (Invest) is the percentage change in total assets from fiscal year $t-2$ to fiscal year $t-1$. We assume that all the accounting variables, e.g., book value of equity, operating profitability and asset growth, are publicly available six months after the fiscal year-end.

For each fund in our sample, we use individual stock holdings to compute the monthly fund-level market capitalization, book-to-market ratio, momentum return, operating profitability, and asset growth. We weight each firm characteristic according to its current portfolio weight and calculate a fund's portfolio-weighted average characteristic. For each characteristic, values greater than the 0.995 percentile or less than the 0.005 percentile are set equal to the 0.995 and the 0.005 percentiles each month.

Table 1 reports statistics on the characteristics and the factor betas. Each month, we calculate the cross-sectional mean, standard deviation and percentiles for each characteristic and factor beta. We report the time series averages of the monthly cross-sectional distribution for characteristics in Panel C and for factor betas in Panel D. The factor betas are estimated from a six-factor model which augments the Carhart (1997) four-factor model with the profitability and investment factors

[1] We supplement the book values from Compustat with hand-collected data provided by Moody's. It includes the data used in Davis et al. (2000) and contains data ranging from 1926 to 2001. The data is available on http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

from Fama & French (2015).[2] The factor betas are derived from rolling-window regressions using daily returns of the past two years.

Panel E of Table 1 reports the cross-sectional correlations between the factor betas and the characteristics averaged across all sample months. Conform to expectations, the betas for SMB and CMA are negatively correlated with their underlying characteristics, while the betas for HML, WML, and RMW are positively correlated with their underlying characteristics. The correlations vary in magnitude, with the highest correlations for the size and value factors (-0.82 and 0.75), the momentum and investment factors yield slightly lower correlations (0.57 and -0.47), and a modest correlation of 0.14 for the profitability factor. These correlations indicate that the factor betas and the characteristics do not convey identical information on the expected return, especially for the profitability factor. As a consequence, regressing fund returns on the risk factors may not fully adjust performance for the main anomalies.

2.C The Importance of Characteristics

We examine the role of factor betas and firm characteristics in explaining the returns of mutual funds. For this purpose we run monthly cross-sectional regressions of individual mutual fund returns on their factor betas and characteristics following the traditional two-pass procedure of Fama & MacBeth (1973) with rolling estimation of factor betas. The monthly regressions are of the form:

$$R_t = \gamma_{0t} + \gamma_{1t}\hat{B}_{t-1} + \gamma_{2t}Z_{t-1} + \xi_t, \ \ t = 1, ..., T, \tag{1}$$

where R_t is a $N_t \times 1$ vector of excess fund returns, $\hat{B}_{t-1}$ is a $N_t \times K$ matrix of estimated factor betas from a K-factor asset pricing model using the past two years of daily ending with month $t-1$, and N_t is the number of funds in month t. The $N_t \times L$ matrix Z contains lagged characteristics including the logarithm of market capitalization (Mcap), the logarithm of book-to-market ratio (B/M), the logarithm of one plus the past twelve-month cumulative return (Mom12), operating profitability (Profit), and asset growth (Invest). We employ the errors-in-variables (EIV)-corrected estimator of Chordia et al. (2015) to obtain coefficient estimates of $\hat{\Gamma}_t = (\hat{\gamma}_{0t}, \hat{\gamma}_{1t}, \hat{\gamma}_{2t})$ in Eq.(1). We describe the EIV-corrected estimator in Appendix B.

We present the results for the CAPM, the Fama & French (1993) three-factor model, the Carhart (1997) four-factor model, the Fama & French (2015) five-factor model and a six-factor model with

[2] The risk factor returns for all factor models are provided on Ken French's Website.

all factors. Table 2 reports the estimated coefficients on both factor betas and characteristics with the EIV bias correction and without. We find that all factor premia except for the value premium are insignificant when their underlying characteristics are added to the model. Momentum return (Mom12) is significant in all factor models which exclude the momentum factor with t-statistics just above 2. Lou (2012) and Vayanos & Woolley (2013) argue that winning funds, by scaling up their existing holdings that are concentrated in past winning stocks, drive up the returns of past winning stocks and thereby enhance the subsequent return of past winning stocks. This self-enhancing mechanism may lead to higher momentum returns for funds.

We use the measure of Lindeman & Lindeman (1980)[3] to evaluate the relative contributions of factor betas and characteristics to the explanatory power of the joint model. We find that in each factor model, almost 50% of the model R^2 is attributed to the characteristics. The relative contributions of the size and value betas decrease by roughly 50% when Mcap and B/M are added to the joint model. The RMW and CMA betas contribute roughly twice as much as their respective characteristics.

To sum up, the factor betas and characteristics both account for roughly half of the model explained variation in fund returns. These findings are consistent with those of Chordia et al. (2015), which find that in a joint model, firm characteristics explain a majority of cross-sectional variation in stock returns. In the remainder of this paper, we will explore the implications of the equal explanatory power of factor betas and characteristics for the performance evaluation of mutual funds.

3 Double-adjusted Mutual Fund Performance Measure

There are many alternatives to evaluate the performance of mutual funds. A common performance measure is a fund's alpha, the return adjusted for exposures to risk factors that drives most of the fund's return. The intercept (alpha) is then interpreted as the abnormal return reflecting the ability of the fund manager. Alternatively, funds are evaluated relative to the characteristic-based benchmark approach of Daniel, Grinblatt, Titman, and Wermers (DGTW; 1997). The DGTW measure controls for the main anomalies through characteristic-sorted portfolio returns. We will examine this measure in Section 5.B.

[3] The measure of Lindeman & Lindeman (1980) uses sequential sums of squares from the linear model. In particular, the relative contribution of regressor j is measured by averaging over all possible permutations of p regressors the increase of R^2 when regressor j is added to the model based on the other regressors entered before j in the model.

We develop a performance measure which adjusts the alphas from a linear factor model for fund exposures to characteristics. We adopt a system of equations in which we simultaneously model conditional factor model alphas and analyze the cross-sectional relation between fund alphas and characteristics. The model takes on the following structure:

$$R_{i\tau,t} = \alpha_{it} + \beta'_{it} F_{\tau,t} + \epsilon_{i\tau,t}, \;\; \epsilon_{i\tau,t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_{it}}), \; \tau = t - \mathcal{T} + 1, ..., t, \; i = 1, ..., N_t \tag{2}$$

$$\alpha_{it} = \delta_{0t} + \delta'_{1t} Z_{it-1} + \eta_{it}, \;\; \eta_{it} \sim \mathcal{N}(0, \sigma^2_{\eta t}), \; t = 1, ..., T, \tag{3}$$

where $R_{i\tau,t}$ and $F_{\tau,t}$ are daily excess returns of fund i and daily factor returns, respectively. The superscript τ is used to index the daily returns in the rolling window ending in month t and $\mathcal{T}$ is the length of the rolling window, that is, two years ($\mathcal{T} \approx 500$ trading days).

The vector $\delta_t = [\delta_{0t} \; \delta_{1t}]$' contains the coefficients that measures for a given month t the average abnormal returns of mutual funds and the relation between alphas (α_{it}) and characteristics (Z_{it-1}). We follow Cederburg & ODoherty (2015) in adding a third layer to the model hierarchy:[4]

$$\delta_t = \bar{\delta} + v_t, \;\; v_t \sim \mathcal{N}(0, \Sigma_\delta), \; t = 1, ..., T, \tag{4}$$

where the monthly coefficients $\{\delta_t\}_{t=1}^T$ are drawn from a multivariate normal distribution centered at $\bar{\delta}$ with covariance matrix Σ_δ. This layer of the model means that the coefficients δ_t can be interpreted as time-random effects in a panel-data model.

We treat $\bar{\delta}$ as an unknown parameter and aggregate the evidence about δ_t across each month in our sample to estimate $\bar{\delta}$. If the factor model prices the mutual funds, we should not find a significant (systematic) relation between alphas and characteristics, such that the elements of $\bar{\delta}$ should be centered around zero. If an element of $\bar{\delta}$ deviates from zero, there is evidence that the factor model inadequately adjusts for exposure to a certain characteristic, such that fund managers can generate alpha by passively loading on this characteristic. In the remainder of this paper, we analyze $\bar{\delta}$ when assessing the importance of adjusting alpha for passive exposures to characteristics.

Eq.(2) represents the factor model for each fund i based on daily fund returns from the rolling window. In our empirical analysis we use a six-factor model which augments the Carhart (1997) four-factor model with the profitability and investment factors from Fama & French (2015). In Eq.(3),

[4] Cederburg & ODoherty (2015) propose a hierarchical Bayes approach to model CAPM alphas conditioned on a set of firm characteristics. Our model extends their work by including daily returns and by evaluating a larger set of risk factors. Another example of hierarchical Bayes is presented in Cosemans et al. (2015), which specify a hierarchical prior on the parameters in their conditional factor beta model.

we relate the cross-section of alphas to the characteristics from the previous month. In Z we include the logarithm of market capitalization (Mcap), the logarithm of book-to-market ratio (B/M), the logarithm of one plus the past twelve-month cumulative return (Mom12), operating profitability (Profit), and asset growth (Invest). Before including them in the regressions, each characteristic is standardized by subtracting the cross-sectional mean in each month. In each rolling window we obtain estimates of α_{it} and δ_t and we roll the window a month at a time.

Using the definitions in Busse et al. (2017), double-adjusted performance measure of fund i in month t is defined as

$$\alpha^*_{it} = \alpha_{it} - \delta'_{1t} Z_{it-1} = \delta_{0t} + \eta_{it}. \tag{5}$$

In this way we control for both the exposures to risk factors through α_{it} and the effects of firm characteristics by subtracting the component of alpha attributable to passive exposures to firm characteristics. Characteristic-driven performance is defined as

$$\alpha^{char}_{it} = \alpha_{it} - \alpha^*_{it} = \delta'_{1t} Z_{it-1}. \tag{6}$$

The errors in Eqs.(2) and (3) are assumed to be independent and normally distributed, such that excess returns are conditionally independent across funds and within rolling windows. Similarly, alphas are conditionally independent across funds and the monthly relations δ_t are independent conditional on $\bar{\delta}$. The main advantage of the specification of independent errors in returns and alphas is that we avoid the estimation of large variance-covariance matrices to capture autocorrelations and cross-sectional correlations.

3.A Model Estimation

We adopt a hierarchical Bayes approach to estimate Eqs.(2), (3) and (4) simultaneously.[5] This approach yields several advantages over the alternatives, as the Bayesian approach is computationally more attractive than maximum likelihood estimation, and yields better finite sample properties than General Method of Moments (GMM).[6] Another main advantage of our model is that we estimate all model parameters simultaneously. Conversely, a majority of previous studies employ a two-pass

[5]The first examples of Bayesian inference in asset pricing models are presented in McCulloch & Rossi (1990) and Harvey & Zhou (1990). More recently, Cremers (2006) propose Bayesian tests for the mean-variance efficiency of a given portfolio. Avramov & Chao (2006) tests the international Capital Asset Pricing Model (ICAPM) with time-varying risk premia using posterior probabilities.

[6]Ferson & Foerster (1995) demonstrate that GMM has poor finite sample properties in the context of latent variable asset pricing models.

procedure to estimate the relation between alphas and characteristics (e.g., Brennan et al. (1998), Avramov & Chordia (2006) and Busse et al. (2017)). Brennan et al. (1998) find that characteristics (e.g., size, book-to-market, momentum) remain significantly related to expected returns even after the risk-adjustment by a factor model with risk factors based on those same characteristics. The two-step procedure includes the estimation of a factor model before regressing the resulting estimates of alpha on the firm characteristics. In the first step alphas are estimated with error such that the variance of these estimates equal the variance of the true alphas plus a measurement error term. Consequently, the standard errors of the estimated coefficients in the second-step regression are overstated. This measurement error may lead to insignificant coefficients even if the data conveys a significant relation between alphas and characteristics. A simultaneous estimation of all model parameters mitigates this measurement error problem.

3.A.1 Prior Distributions

To conduct Bayesian estimation techniques, we need to specify prior distributions for the model parameters. A prior distribution incorporates a researcher's belief about the parameter of interest, often founded on empirical research or economic theory. We specify proper, but relatively uninformative, priors for all parameters.[7]

Regarding the loadings on characteristics, the expression in Eq.(4) implies the following hierarchical prior on δ_t

$$\delta_t|\bar{\delta}, \Sigma_\delta \sim \mathcal{N}(\bar{\delta}, \Sigma_\delta), \tag{7}$$

which is a normal distribution with location parameter $\bar{\delta}$ and covariance matrix Σ_δ. We treat $\bar{\delta}$ and Σ_δ as unknown parameters and assume the following uninformative priors

$$\bar{\delta} \sim \mathcal{N}\left(d, \Sigma_d\right), \tag{8}$$

$$\Sigma_\delta \sim IW\left([\psi_\delta S_\delta], \psi_\delta\right), \tag{9}$$

where IW denotes the inverted Wishart distribution with degrees of freedom ψ_δ and scale matrix $[\psi_\delta S_\delta]$. We specify an uninformative prior for $\bar{\delta}$ by setting d equal to the zero vector and Σ_d to $100I$, where I is the identity matrix. Setting d equal to the zero vector implies that alphas are not associated with characteristics, which contradicts previous empirical findings. However, by setting

[7] Hobert & Casella (1996) advocate the use of proper priors in estimating hierarchical models. They find that using improper priors may lead to improper posterior distributions which is difficult to detect from posterior draws.

the diagonal elements of Σ_d to large values we minimize the influence of the prior on the posterior distribution of $\bar{\delta}$. ψ_β is set to the dimension of Σ_δ, since this value gives the lowest weight to the prior (see, Gelman et al. (2014)). The matrix S_δ is the prior mean of Σ_δ and is set to I.

Regarding the fund-specific parameters, the hierarchical structure of our model implies the following prior on α_{it}

$$\alpha_{it}|\delta_t, \sigma^2_{\eta t} \sim \mathcal{N}(\delta_{0t} + \delta'_{1t} Z_{it-1}, \sigma^2_{\eta t}), \tag{10}$$

in which the prior mean is dependent on characteristics. The posterior of α_{it} combines information from both the fund-specific time series of returns and from the cross-section of alphas. The relative importance of both components depends on the rolling window size $\mathcal{T}$ and the cross-sectional residual variance $\sigma^2_{\eta t}$. We specify a conjugate prior for β_{it} as

$$\beta_{it} \sim \mathcal{N}(\bar{\beta}, \Sigma_\beta), \tag{11}$$

where we set $\bar{\beta}$ equal to the sample mean of rolling window estimates of fund betas (see Panel D of Table 1). We define Σ_β equal to $10I$ to let the data speak for itself.

We complete the prior specification of our model by specifying conjugate priors for the idiosyncratic variances $\sigma^2_{\epsilon_{it}}$ and $\sigma^2_{\eta t}$ as

$$\sigma^2_{\epsilon_{it}} \sim IG2(v_\epsilon, s_\epsilon) \tag{12}$$

$$\sigma^2_{\eta t} \sim IG2(v_\eta, s_\eta), \tag{13}$$

where $IG2$ denotes the inverted Gamma-2 distribution. We specify relative uninformative priors by setting the degrees of freedom parameters v_ϵ and v_η equal to 3 and the scale parameters s_ϵ and s_η equal to 10^{-3}.

3.A.2 Markov Chain Monte Carlo Gibbs Sampler

We employ the Markov Chain Monte Carlo (MCMC) Gibbs sampler[8] to estimate the model. An attractive feature of MCMC techniques is that samples of random drawings of the model parameters θ can be generated from the joint posterior indirectly, without the need to specify the exact form of this joint distribution directly. The Gibbs sampler uses an iterative procedure to create Markov chains by simulating from full conditional posteriors instead which are typically much easier to

[8] For an extensive reading on the application of the Gibbs sampler in several econometric models, we refer to De Pooter et al. (2006).

derive. In particular, the parameter vector θ is partitioned into B blocks $(\theta^{(1)},...,\theta^{(B)})$. At each iteration of the Gibbs sampler each block is sampled from its posterior distribution conditioned on the other blocks and the data. After the Markov chains have converged, the sets of draws that are obtained from the conditional posteriors can be effectively considered as samples from the joint posterior. We provide the steps in the Gibbs sampler and the derivation of the full conditional posterior distributions in Appendix D.

To check the convergence of the Gibbs sampler, we turn to standard diagnostics tests, including the partial means test of Geweke (1992) and the Gelman-Rubin statistic described in Gelman et al. (2014). Both tests indicate that the Markov Chains have converged after 2500 iterations. Therefore, in our empirical analysis we run 5000 iterations of the Gibbs sampler and discard the first 2500 iterations as burn-in period. The remaining draws are used to derive posterior results. The m^{th} draw of double-adjusted alpha and characteristic-driven performance are given by

$$\alpha_{it}^{*\,(m)} = \alpha_{it}^{(m)} - \delta_{1t}'^{\,(m)} Z_{it-1} \quad \text{and} \quad \alpha_{it}^{char\,(m)} = \delta_{1t}'^{\,(m)} Z_{it-1} \tag{14}$$

The posterior mean of draws $\{\alpha_{it}^{*\,(m)}\}_{m=2501}^{5000}$ and $\{\alpha_{it}^{char\,(m)}\}_{m=2501}^{5000}$ constitute the final estimates of double-adjusted alpha and characteristic-driven performance, respectively.

3.B Estimation Results

Table 3 summarizes the relation between six-factor alphas and fund-level characteristics. In Panel A, we report the posterior estimates of the distribution of $\bar{\delta}$, which measures the systematic relation between alphas and characteristics over the entire sample period. When we estimate the model for each characteristic in isolation, we find that market capitalization (Mcap) and book-to-market (B/M) are negatively associated with alpha, while momentum return (Mom12) and asset growth (Invest) are positively associated with alpha. Alphas are unrelated to the profitability characteristic. If the characteristics are correlated with each other and offer little unique information about alphas then studying each characteristic in isolation will overstate the failings of the factor model. If we specify alpha as a linear combination of all characteristics, Mcap, B/M, Mom12 and Invest remain significantly associated with alphas. The constant term in Table 3 measures the outperformance of the average mutual fund relative to the Fama-French benchmarks, which we refer to as the aggregate alpha. Over our sample period we find that the average mutual fund earns negative risk-adjusted returns.

Recall that the six-factor model adjusts returns for exposures to factors based on Mcap, B/M, Mom6, Profit, and Invest. The results in Panel A suggest that the six-factor model under-adjusts for exposures to the size and momentum factors, while the model appears to over-adjust for influences related to the value and investment factors. In other words, funds holding small (high-momentum) stocks show higher abnormal return despite the risk-adjustment to the SMB (WML) factor. Conversely, funds holding positions in high book-to-market (conservative investment) stocks under perform, contradictory to previous findings on value and investment effects. Similar conclusions are drawn by Huij & Verbeek (2009), which find that fund managers following a value-orientated strategy earn a substantially lower premium than those projected by the hypothetical hedge portfolio HML, while the momentum premium earned by funds is larger than that projected by the WML factor.

In Panel B we report the estimates from a standard OLS two-step procedure. In each month t, we estimate the six-factor model for each fund i. We then regress cross-sectionally the estimated alphas on a constant and lagged characteristics. In the style of Fama & MacBeth (1973) we report the time series averages of the cross-sectional coefficients. The estimation results using OLS lead to several differences in comparison to the Bayesian estimation. While the inferences regarding B/M, Mom12 and Invest are similar between the two estimation methods, alpha becomes unrelated to Mcap using the two-step OLS approach. Moreover, if we consider all characteristics simultaneously we find that all characteristics but B/M become unrelated with alphas using the two-step OLS approach.

Inferences for the hierarchical Bayes and two-step OLS approach may differ due to way information from both the cross-section and time series is combined. In the hierarchical Bayes approach, the estimated factor model alphas are a weighted average of fund-specific time series information and cross-sectional information, with the precision (the inverse of the variance) as weights. Conversely, the two-step approach directly uses parameter estimates into the second-stage regression, thereby not exploiting information from the cross-section of funds, which makes the two-step approach more prone to the influence of outliers. As these outliers are marked by considerable uncertainty, the hierarchical Bayes approach is more suitable to combine multiple sources of information with varying degrees of precision.

Figure 1 graphs the time series plots of the posterior mean and 95% credible interval of $\{\delta_t\}_{t=1}^{T}$, obtained from estimating the model using all characteristics. As expected, we find the relation between six-factor alpha and a given characteristic to be varying across our sample. For instance, we find that δ_t for the momentum characteristic is negative in times when momentum stocks perform poorly (e.g., the internet bubble (post 2001) and the financial crisis of 2007-2008). Moreover, we find that the monthly posterior distributions for the size, value and momentum characteristics exhibit higher precision in comparison to the posteriors of the other two characteristics.

Finally, we examine whether our results are robust to the risk factors included in Eq.(2). In Table 4 we add the posterior results from the model using alphas from the CAPM, the Fama-French three-factor model, the Carhart (1997) four-factor model and the Fama & French (2015) five-factor model. The systematic relation between alphas and characteristics are similar across factor models. The posterior means of B/M, Mom12 and Invest vary the most across the factor models, indicating that cross-correlations are the largest between the value, momentum and investment effects, both through the risk factors and the characteristics. Interestingly, we find that the posterior mean of B/M changes sign when the quality factors are considered. The value factor in the three- and four-factor models under-adjusts for the value factor, as alphas are positively related to funds holding value stocks. Conversely, alphas from the five- and six-factor models are positively related to funds holding growth stocks. This contradiction is possibly caused by the interaction between the value factor and the quality factors, which is a negative correlation between the two. All in all, we conclude that the traditional Fama-French risk factors (jointly) insufficiently adjust returns for the main anomalies.

4 Impact on Relative Mutual Fund Performance

In Section 2.C, we show the importance of incorporating characteristics in explaining the returns of mutual funds. The results in the previous section demonstrate an important flaw of the traditional factor model alphas, insofar as they attribute skill to passive exposures to characteristics. Therefore, we propose a mutual fund performance measure which adjusts performance for exposures to both factor betas and characteristics.

In this section, we examine the extent to which our new double-adjusted alpha affects inference relating to relative mutual fund performance. We replicate the work of several mutual fund performance studies using six-factor alpha as our baseline measure of performance, computed using the traditional time series regression approach.[9] Because we use traditional alpha as our baseline performance measure, we can easily extend previous analysis with the two components of alpha, our double-adjusted measure and the portion of alpha associated with characteristics (see Eqs.(5) and (6)). The decomposition of alpha will indicate the extent to which both components contribute to the main findings of previous studies.[10]

We begin by calculating the difference in fund percentile performance rankings between using double-adjusted alpha in comparison to those based on alpha. Next, central to the mutual fund performance literature are studies that analyze persistence in fund performance, e.g., Carhart (1997). Other studies examine the relation between fund performance and certain fund features, such as a fund's factor model R-squared (Amihud & Goyenko, 2013) and a fund's investors cash flows (Barber et al., 2016). We plug in our estimates of the posterior mean of both components of alpha in the analysis of these papers. We use the estimation methods used in these papers to accommodate a fair comparison between inferences based on traditional alpha and those based on double-adjusted alpha.

4.A Relative Mutual Fund Performance Rankings

We examine the degree to which our new performance measure alters the performance rankings of funds. In each month in our sample, we sort funds into percentiles based on six-factor alpha (α)

[9]Even though the hierarchical Bayes model also provides estimates of fund alphas, we choose the time series regression approach as this is standard in previous literature. Moreover, the Bayesian estimate of standard alpha makes use of cross-sectional information. To highlight the difference made by adjusting performance for characteristics, we compare our double-adjusted alpha to a performance measure which is not influenced by these characteristics.

[10]We follow Busse et al. (2017) in reassessing previous findings in the mutual fund literature using the double-adjusted performance measure. In addition to the analyses in this paper, they also reevaluated the relation between fund performance and the industry concentration (Kacperczyk et al., 2005), the return gap (Kacperczyk et al., 2008), and active share (Cremers & Petajisto, 2009).

and on double-adjusted six-factor alpha (α^*). Given a number of N_t funds in month t, we obtain pairs of percentile ranks for each fund $\{(P_{1t}^{\alpha}, P_{1t}^{\alpha^*}), ..., (P_{N_t t}^{\alpha}, P_{N_t t}^{\alpha^*})\}$. Using the percentile rank pairs we compute Kendall's tau coefficient, which is a correlation coefficient between two rankings. We find a time series average of Kendall's tau equal to 0.70, which indicates differences between the two rankings.

To examine the impact of double-adjusted alpha on performance rankings in more detail, we compute the difference between the performance ranks P_t^{α} and $P_t^{\alpha^*}$ for each fund in each month t. Table 5 presents the time series average distribution (across all months in our sample) of the difference in percentile performance rankings. We find that the median change in percentile rankings is roughly 9%, i.e., a fund which is ranked in the median percentile according to alpha is ranked in either the 41th or 59th percentile based on double-adjusted alpha. Moreover, many funds exhibit dramatic changes in percentile ranking, with 10 (5) percent of funds exhibiting a mean change in percentile ranking of at least 22.99% (27.24%). Thus, we find that adjusting traditional alpha for characteristics materially impacts the performance ranking of funds.

4.B Mutual Fund Return Persistence

Mutual fund persistence is well documented in the finance literature. Early works include Grinblatt & Titman (1992), Brown & Goetzmann (1995) and Wermers (1997), which document significant persistence in mutual fund rankings based on returns after adjustment for risk. This persistence lasts over horizons of one to three years, and they attribute the persistence to "hot hands" or common investment strategies. Carhart (1997) add stock momentum as an additional risk factor and find a significant difference in abnormal returns of more than 4% on an annual basis. Drawing on Carhart's findings, Bollen & Busse (2004) find an average abnormal return of 39 basis points for the top quintile in the post-ranking quarter. This post-ranking abnormal return disappears when evaluated over longer periods.

We examine return persistence in several performance measures: six-factor alpha (α), double-adjusted six-factor alpha (α^*) and characteristic-driven performance (α^{char}). If our new performance measure, which accounts for characteristics, is a better indicator of true skill, we expect double-adjusted alpha to account for most of the persistence in alpha, such that double-adjusted alpha persists for a longer period. In this case, there should be no distinct return pattern using characteristic-driven performance. That is, if true skill goes beyond the premiums which are passively associated with characteristics.

We replicate the studies on return persistence.[11] Each quarter-end, we sort funds into deciles using the aforementioned performance measures; decile 10 contains the best performing funds and decile 1 contains the worst performing funds. We hold the sorted portfolios for up to six years and compute the equal-weighted return in each decile in the post-ranking period. To deal with overlapping decile portfolios formed in different quarters, we compute the equal-weighted return across overlapping periods. We examine the post-ranking performance by concatenating the returns of all post-ranking periods for each decile and estimate six-factor alpha using the resulting time series of post-ranking (monthly) returns for each decile.

Table 6 presents the persistence results. Panel A shows strong evidence of persistence in standard six-factor alpha. In the short term, the average difference in alpha between the top and bottom deciles is 0.28 (t = 3.62) in the following quarter and 0.27 (t = 3.85) in the year thereafter, both of which significant from a statistical and an economic perspective. This abnormal return difference remains statistically significant in the second and third year after formation, followed by a decrease in the top-minus-bottom post ranking abnormal return from the fourth year onwards. Consistent with the findings of Carhart (1997), we find that the majority of the return spread is driven by past losing funds in the bottom deciles. In untabulated results, we find significant negative alphas in the bottom decile with t-statistics below -3.

The results in Panel B show strong evidence that the double-adjusted six-factor alpha predicts future fund performance. We find a distinct return pattern across deciles sorted on our double-adjusted performance measure, as post-ranking alpha increases nearly monotonically across deciles and indicates a sizable return spread of 0.19 (t = 5.31) in the post-ranking quarter. The top-minus-bottom portfolios yield statistical significant alphas between 0.05 and 0.17, which persist up to six years. In Figure 2, we analyze return persistence on the long term. We graph the cumulative six-factor alpha of the top-minus-bottom portfolios, where the three plots lines represent alternative sorts on six-factor alpha, double-adjusted six-factor alpha, and characteristic-driven performance. The plots indicate an upward trend in the cumulative performance of both alpha and double-adjusted alpha, which lasts for the entire holding period of ten years. The results in Panel B of Table 6 combined with the plots from Figure 2 suggest that double-adjusted alpha reflects true skill, which pays off in the long run such that persistence in this component of alpha causes most of the

[11]This paper adopts a similar approach to the one used in Carhart (1997). The only difference is the frequency of rebalancing, which is set to a quarterly basis in this work. Busse et al. (2017) analyzed persistence in their double-adjusted performance measure. They adopt an annual rebalancing of portfolios and found new evidence of persistence in mutual fund skill based on their double-adjusted performance measure.

persistence in alpha. Our results are in line with the persistence analysis in Busse et al. (2017). Using their double-ajdusted alpha they find evidence of persistence in fund four-factor alphas up to nine years after the initial ranking. We show that persistence is also present when a shorter sample period is used.

In Panel C of Table 6 we turn to the component of alpha driven by passively loading on characteristics. The average return difference in alpha in the first year after formation is 0.16 with insignificant t-statistics, followed by a slight reversal from the second year onwards. The graph in Figure 2 shows that sorts on characteristic-driven performance does not exhibit a distinct return pattern after formation in the long term. These results indicate that returns associated with characteristics do not convey information about future fund performance. This supports the claim that evidence of return predictability in alpha is mostly accounted for by the component adjusted for passive loadings on characteristics. By removing this noisy component, which does not detect skill, double-adjusted alpha is a more precise measurement of skill, leading to stronger evidence of persistence in mutual fund performance.

4.C Mutual Fund's R-squared

Recent studies show that fund performance is positively affected by fund selectivity or active fund management, which is measured by the deviation of fund holdings from a diversified benchmark portfolio. This selectivity measure requires data on the holdings of funds and knowledge on the benchmark indices of funds, which are difficult to obtain. In addition, the benchmark portfolio is not always accurately defined. Amihud & Goyenko (2013) propose a simple and intuitive measure of mutual fund selectivity:[12] the fund's R-squared (R^2), the proportion of the return variance that is explained by the benchmark portfolios, estimated from a multi-factor model. Amihud & Goyenko (2013) hypothesize that a low R^2 corresponds to high stock selectivity, as a low R^2 indicates that mutual returns are not well explained by the Fama-French risk factors. In this respect, they find that R^2 has a negative and significant predictive effect on fund performance. We replicate the work of Amihud & Goyenko (2013) by examining the relation between a fund's factor model R^2 and fund performance. Each month, we calculate a fund's R^2 from a six-factor model over a 24-month estimation period using daily returns. Following Amihud & Goyenko (2013), since the distribution

[12] Amihud & Goyenko (2013) define fund selectivity as 1 - $R^2 = \frac{\text{RMSE}^2}{\text{TotalVariance}} = \frac{\text{RMSE}^2}{\text{SystemeticRisk+RMSE}^2}$. RMSE is the idiosyncratic volatility, which is the volatility of the factor model residuals. SystemeticRisk is the return variance that is due to the risk factors. Fund selectivity is greater if the fund's idiosyncratic volatility is higher relative to its total variance, meaning that the fund's volatility is less driven by factor-based (systematic) volatility.

of R^2 is negatively skewed, we apply the log transformation of R^2: $\tilde{R}^2 = \log\left[\frac{\sqrt{R^2}}{(1-\sqrt{R^2})}\right]$. We test whether fund selectivity explains fund performance by the following panel regression

$$\begin{aligned}\text{Performance}_{it} = c_{0t} + c_{1t}\tilde{R}^2_{it} + c_{2t}\text{ExpRatio}_{it-1} + c_{3t}\text{Turnover}_{it-1} + c_{4t}\log(\text{TNA})_{it-1} \\ + c_{5t}\log(\text{FundAge})_{it-1} + v_{it},\end{aligned} \tag{15}$$

where Performance_{it} refers to fund i's six-factor alpha (α), double-adjusted six-factor alpha (α^*), or characteristic-driven performance (α^{char}), all estimated over a rolling window of 24 months ending in month t. We relate fund performance to the contemporaneous log-transformed R-squared, $\tilde{R}^2$, which is measured over the same rolling window period. Lagged control variables are included in the regression comprising ExpRatio, the total expenses divided by a fund's Total Net Assets (TNA); Turnover, defined as the minimum of aggregated sales or aggregated purchases divided by TNA; TNA in logarithm; and FundAge in logarithm, computed as the difference in years between the current date and the date the fund was first offered. The model residuals are given by v_{it}. We estimate this regression in each month in our sample.

Table 7 reports the averaged cross-sectional regression coefficients along with Fama-MacBeth t-statistics with the Newey & West (1986) correction for time series correlation with 12 lags. Firstly, similar to the findings of Amihud & Goyenko (2013), we find that alpha is higher for funds with lower R-squared, that is, funds with high stock selectivity. The control variables yield no significant estimates. When we consider double-adjusted alpha we find no significant relation with $\tilde{R}^2$. However, the results indicate a significant (inverse) relation between the characteristic-driven component of alpha and R-squared, with t-statistics well above two. A potential reason might be that characteristics can help explain fund returns in cases where the Fama-French factors fail to do so, such that we find a strong inverse relation between characteristic-driven performance and R-squared. Thus, we find that inverse relation between alpha and R-squared does not indicate a positive relation between fund performance and fund selectivity, but rather indicates the poor explanatory power of the Fama-French style factor models. This calls into question the use of R-squared as a measure of fund selectivity.

4.D Mutual Fund Flows

There is no shortage of literature on mutual fund flows. The first stream of literature (e.g., Warther (1995), Edelen & Warner (2001), Brown et al. (2003)) have focused on aggregate fund flows to the equity market, documenting a positive correlation with contemporaneous stock returns. Early works of Ippolito (1992), Gruber (1996) and Sirri & Tufano (1998) find that capital flows to and from funds are strongly related to past fund performance. The general consensus is that the flow-performance relation is positive, asymmetric and convex; the inflows generated by positive returns is of greater magnitude than the outflows due to negative returns. The establishment of a convex flow-performance relation is generally robust to different performance measures varying from raw returns to multi-factor model alphas.

More recently, a second stream of literature goes beyond the simple flow-performance relations.[13] A recent paper of Barber et al. (2016) investigate which risk factors are used by investors to adjust raw returns when evaluating fund performance. They run linear regressions of fund flows on fund alphas obtained from several Fama-French factor models and conclude that CAPM alphas are the best predictor of flows among all factor model alphas. In additional analysis, they decompose fund returns into alpha and returns resulting from factor tilts. They conclude that alpha generates the largest flow response, closely followed by a fund's momentum related return, while flows are least sensitive to fund returns traced to market beta.

We aim to replicate the work of Barber et al. (2016) using our decomposition of traditional alpha. Using fund flows, we investigate whether investors tend toward the double-adjusted component α^* or toward the characteristic-driven component α^{char} when assessing fund managers. For this purpose, we follow the majority of the prior literature on fund flows and calculate net capital flow[14] to fund i during month t as

$$\text{flow}_{it} = \frac{\text{TNA}_{it} - \text{TNA}_{it-1}(1 + R_{it})}{\text{TNA}_{it-1}}, \tag{16}$$

where TNA_{it} is fund i's total net assets at the end of month t and R_{it} is the monthly net return of fund i in month t.[15] The variable flow reflects the percentage growth of a fund that is due to new investment (under the assumption of dividends being reinvested in the fund).

Similar to Barber et al. (2016), we further decompose characteristic-driven performance into

[13] A mutual fund's investment style is an important source of information to investors. Cooper et al. (2005) document an increase in fund flows to funds adopting the current hot investment style in their names. They find that these inflows are similar across funds who alter their positions matching their new name and those who do not, suggesting that investors are irrationally influenced by cosmetic effects. Guo (2016) use a fund's reported holdings to determine a fund's investment style and find that funds adopting a more volatile implementation of style strategy garner higher inflows.

[14] Frazzini & Lamont (2008) and Lou (2012) correct for fund mergers when calculating fund flows. As fund mergers are quite rare, this study ignores them.

[15] Fund flows are dropped if the percentage difference in TNA in between two months is greater than 200% or less than -50%. These extreme flows are rare and are typically related to structural changes within funds, e.g., mergers.

alpha related to each individual characteristic. Recall that characteristic-driven performance (see Eq.(6)) is defined as

$$\begin{aligned}\alpha_{it}^{char} &= \delta_{1t}' Z_{it-1} \\ &= \delta_{1t}^{\text{Mcap}} \text{Mcap}_{it-1} + \delta_{1t}^{\text{B/M}} \text{B/M}_{it-1} + \delta_{1t}^{\text{Mom12}} \text{Mom12}_{it-1} + \delta_{1t}^{\text{Profit}} \text{Profit}_{it-1} + \delta_{1t}^{\text{Invest}} \text{Invest}_{it-1}\end{aligned} \tag{17}$$

which is the sum of lagged (standardized) characteristics multiplied by the corresponding cross-sectional premia. With this decomposition, we gauge whether fund flows are distributed differently across the characteristic components of alpha by estimating the following panel regression

$$\begin{aligned}\text{flow}_{i,t+1:t+12} = c_{0t} &+ c_{1t}\alpha_{it}^* + c_{2t}[\delta_{1t}^{\text{Mcap}} \text{Mcap}_{it-1}] + c_{3t}[\delta_{1t}^{\text{B/M}} \text{B/M}_{it}] + c_{4t}[\delta_{1t}^{\text{Mom12}} \text{Mom12}_{it-1}] \\ &+ c_{5t}[\delta_{1t}^{\text{Profit}} \text{Profit}_{it-1}] + c_{6t}[\delta_{1t}^{\text{Invest}} \text{Invest}_{it-1}] \\ &+ c_{7t}\text{ExpRatio}_{it} + c_{8t}\text{Turnover}_{it} + c_{9t}\log(\text{TNA})_{it} + c_{10t}\log(\text{FundAge})_{it} + v_{it},\end{aligned} \tag{18}$$

where the dependent variable is the average monthly fund flow of fund i in the following year. The independent variables include double-adjusted alpha α_{it}^* and alpha related to a fund's size, value, momentum, profitability and investment characteristics (see Eq.(17)). We include the same control variables as in Eq.(15). The model residuals are given by v_{it}. We estimate this regression at each year-end. We expect that managerial skill attracts fund, such that there is no capital directed towards abnormal return related to known characteristics. That is, we expect that the parameters related to the characteristic components to be indistinguishable from zero. Contrary evidence implies that either investors are not considering known factors when assessing fund performance, or alpha provides an incomplete risk-adjustment for these characteristics.

Table 8 reports the cross-sectional regression coefficients averaged across time along with Fama-MacBeth t-statistics with the Newey & West (1986) correction for time series correlation with 3 lags. In the first two columns, we specify future fund flows as a linear combination of double-adjusted alpha and characteristic-driven performance, with and without the set of control variables. We find strong positive relations between both performance measures and fund flows, where the double-adjusted component garners the most inflows. Among the set of control variables, we find positive but statistically insignificant coefficients for ExpRatio and log(FundAge). Of interest are the estimated sensitivities of flows to the characteristic components of alpha. Generally, fund returns related to characteristics do not garner the same magnitude of flows as double-adjusted alpha does. The coefficients on the value and profitability characteristics are both statistically and economically insignificant. However, we do find evidence of investors tending to size and momentum characteristics, which yield significant parameter estimates of about half of that of double-adjusted alpha (in absolute value). Thus, in aggregate, we find that the flow-performance relation is mostly driven by double-adjusted alpha and that investors allocate a portion of their investment toward funds with favourable characteristics in the size and momentum dimensions.

In sum, we have replicated parts of the analysis from several papers on mutual fund performance. First, we find stronger evidence of return persistence when using double-adjusted alpha. Second, we find that the inverse relation between alpha and R-squared is mostly driven by the characteristic-component of alpha. Third, fund flows are more responsive to double-adjusted alpha rather than characteristic-driven performance. These findings show that adjusting alphas for characteristics materially impacts previous findings on mutual fund performance. As our results together with those of Busse et al. (2017) indicate that double-adjusted alpha better reflects true skill, we believe that this performance measure should be used to reassess previous studies and in future studies on mutual fund performance.

5 Robustness Checks

5.A Overlapping Rolling Windows

In a recent paper, Britten-Jones et al. (2011) examine the impact of overlapping dependent variables on the inference from standard Fama-MacBeth cross-sectional regressions. The overlap induces an autocorrelation pattern in the standard errors of the cross-sectional regressions and commonly used methods (e.g., White or common Newey-West standard errors) to deal with this autocorrelation are inadequate and can lead to misleading estimates of the confidence intervals associated with coefficient estimates obtained from finite samples.[16] As the rolling windows overlap in our model in Eqs.(2), (3) and (4), we check the robustness of our results in Section 3.B. Specifically, we re-estimate our model using different frequencies of the cross-sectional regressions in Eq.(3).

Table 9 reports the results. In Panel A, we estimate the cross-sectional regressions at each quarter-end rolling the window three months at a time. The results are very similar compared

[16] Britten-Jones et al. (2011) transform original regressions into an equivalent representation in which the dependent variables are non-overlapping to remove the autocorrelation. Their method is easily applicable within standard frequentist analyses and they show that conventional inference procedures (OLS-, White-, Newey-West- standard errors) are asymptotically valid when applied to the transformed regression.

to the results using monthly cross-sectional regressions (see Table 3).[17] All characteristics except operating profitability remain significantly associated with six-factor alphas. Moreover, we find an increase in the posterior standard deviations for the momentum, profitability and investment characteristics. In Panel B, we estimate the cross-sectional regressions at a semi-annual frequency. We use less periods to estimate the systematic relation between alphas and characteristics, which leads to a further increase in posterior standard deviations. Consequently, we find more tenuous relations between six-factor alphas and the momentum and investment characteristics.

Another factor to consider is the frequency of mutual fund portfolio disclosure. The Thomson Reuters database provides quarterly snapshots of fund portfolios, which we keep constant between quarters to create a monthly time series of fund holdings (see Appendix A). Consequently, the variation in fund-level characteristics within quarters is only caused by the monthly variation in firm characteristics of each fund position. The staleness in our holdings data might compound the autocorrelation in the standard errors described above. While our analysis would benefit from portfolio disclosure at a higher frequency, the potential effects of frequent mutual fund portfolio disclosure remains the focus of a longstanding debate among practitioners, regulators, and academics.

5.B Characteristic-based Benchmark Approach

Daniel et al. (1997) develop a new measure of mutual fund performance which use benchmarks based on the firm characteristics of mutual fund holdings. The benchmarks are constructed from the returns of passive portfolios sorted on characteristics that are matched with the characteristics of the evaluated fund's holdings. Specifically, the DGTW characteristic selectivity (CS) measure is calculated as

$$CS_{it} = \sum_{j=1}^{N_{it}} w_{jt-1}(R_{jt} - R_t^{b_{jt}}), \tag{19}$$

where w_{jt-1} is the portfolio weight on stock j at the end of month $t-1$, R_{jt} is the month t return of stock j, $R_t^{b_{jt}}$ is the month t return of the characteristic-based benchmark portfolio that is matched to stock j, and N_{it} is the number of holdings of fund i in month t. We use the DGTW CS measure to adjust fund returns for size, value, momentum, profitability and investment effects. In Appendix E, we describe the construction of the characteristic-based benchmark portfolios in greater detail.

To test whether this characteristic-based performance measure fully adjusts fund returns for the

[17] We repeated the analysis of Section 4 using double-adjusted alpha obtained from quarterly cross-sectional regressions. In untabulated results, we find that inference is qualitatively similar to that based on double-adjusted alpha from monthly cross-sectional regressions.

main anomalies, we regress cross-sectionally the DGTW CS measure on the factor betas from the Fama-French six-factor model. Particularly, we conduct the following panel regression

$$CS_{i,t-23:t} = c_{0t} + \sum_{k=1}^{6} c_{kt}\beta_{ikt} + v_{it}, \tag{20}$$

where the dependent variable is the average DGTW CS measure of fund i across a 24-month period ending in month t. The independent variables include the six-factor betas of fund i estimated over the same 24-month period using daily fund returns. The model residuals are given by v_{it}. Beginning with the 24$^{\text{th}}$ month of our sample, we estimate the monthly cross-sectional regressions over the period December 2002 to December 2016, consisting of 169 months.

Table 10 reports the results. In Panel A, we only adjust returns for size, value and momentum effects. When considered in isolation, we find significant relations between the DGTW CS measure and the market, momentum and profitability betas. That is, the characteristic-based measure under-adjusts for exposures to the momentum and profitability factors. When we consider all factor betas together, the profitability beta is no longer significantly related with characteristic-adjusted fund performance. In Panel B, we adjust returns for all characteristics and the results are qualitatively similar.

In sum, we conclude that the characteristic-based measure of Daniel et al. (1997) does not fully adjust fund performance for the main anomalies. Controlling only for factor betas, as in Fama-French factor models, or only for characteristics, as in DGTW, may overlook the other effect, and in so doing materially impact estimates of fund manager skill.

6 Conclusion

We propose a new approach to evaluate the performance of mutual funds. It combines traditional factor models as in Carhart (1997); Fama & French (1993) with the effects of firm and asset characteristics, conform the recent insight that these information sources complement each other in explaining asset returns (Avramov & Chordia, 2006; Brennan et al., 1998; Chordia et al., 2015). Busse et al. (2017) shows that the firm and asset characteristics resulting from mutual fund holdings should also be included in analyses of mutual fund performance.

Our approach consists of a hierarchical Bayesian model with three layers in the style of Cederburg & ODoherty (2015). The first layer specifies a factor model for daily returns, whereas the second splits the intercept in the effect of firm and asset characteristics, and the remainder which we call the double-adjusted alpha. The third layer adds monthly time variation to it. The double-adjusted alpha for each asset and each point in time can be split in an aggregate time-random effect and a time- and asset-specific component. Though our model combines the same information as Busse et al. (2017) to construct the double-adjusted alpha, the Bayesian approach offers a couple of advantages. Whereas Busse et al. (2017) uses a two-pass estimation, our Bayesian approach estimates the time series and cross-sectional parts of the model in one pass which leads to efficiency gains and automatically accounts for the estimation uncertainty that otherwise should be included in the second part. More generally, the posterior mean of the time-random component of alpha shows to which extent the aggregate mutual fund can deliver outperformance, while the prior can be used to specify the belief in outperformance.

We apply our approach to analyse the performance of U.S. mutual funds over the period 2001–2016, with the risk factors from Fama & French (2015) and a momentum factor, as well as the corresponding characteristics. We show that both factor exposures and the effect of characteristics are needed to explain mutual fund performance, and that our Bayesian approach leads to more precise estimates than the frequentist approach. Moreover, we show that the ranking of mutual funds based on the posterior mean of the double-adjusted alpha differs substantially from the ranking that results from the traditional alphas based on factor models.

We then show how our new inferences on mutual fund performance impact research into the skill of mutual funds. Replacing the traditional alpha by our double-adjusted one leads to stronger evidence for persistence in the performance of mutual funds. However, when we use the double-adjusted alpha, we do not find a relation with the time-series R^2 of the factor regression anymore as in Amihud & Goyenko (2013), indicating that the R^2 cannot be interpreted as a measure of selectivity related to skill. We extend Barber et al. (2016) by showing that fund flows are most responsive to the double-adjusted alpha, and not so much to the part that is related to characteristics. We interpret this as evidence that at least part of the investment community also adjusts returns for the part that can be explained by well-known return drivers.

Our findings contribute to the literature in two ways. First, we show how to efficiently estimate a double-adjusted alpha in a single-pass using Bayesian techniques. We argue theoretically and show empirically that our approach improves upon the more standard two-pass estimation methodology by offering better and more precise estimates. Second, we show that these improved inferences are relevant for studies related to the performance and skill of mutual funds. Some of the results from earlier studies turn out to be stronger, while other results vanish. Taken together, it means that we have to be careful in the investigation of mutual fund performance, and should make sure that we include the necessary information in the right way. Our paper shows how a straightforward Bayesian model can accomplish that.

References

Amihud, Y., & Goyenko, R. (2013). Mutual fund's r 2 as predictor of performance. *The Review of Financial Studies*, *26*(3), 667–694.

Avramov, D., & Chao, J. C. (2006). An exact bayes test of asset pricing models with application to international markets. *The Journal of Business*, *79*(1), 293–324.

Avramov, D., & Chordia, T. (2006). Asset pricing models and financial market anomalies. *The Review of Financial Studies*, *19*(3), 1001–1040.

Barber, B. M., Huang, X., & Odean, T. (2016). Which factors matter to investors? evidence from mutual fund flows. *The Review of Financial Studies*, *29*(10), 2600–2642.

Bollen, N. P., & Busse, J. A. (2004). Short-term persistence in mutual fund performance. *The Review of Financial Studies*, *18*(2), 569–597.

Brennan, M. J., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, *49*(3), 345–373.

Britten-Jones, M., Neuberger, A., & Nolte, I. (2011). Improved inference in regression with overlapping observations. *Journal of Business Finance & Accounting*, *38*(5-6), 657–683.

Brown, S. J., & Goetzmann, W. N. (1995). Performance persistence. *The Journal of finance*, *50*(2), 679–698.

Brown, S. J., Goetzmann, W. N., Hiraki, T., Shirishi, N., & Watanabe, M. (2003). *Investor sentiment in japanese and us daily mutual fund flows* (Tech. Rep.). National Bureau of Economic Research.

Busse, J., Jiang, L., & Tang, Y. (2017). Double-adjusted mutual fund performance.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, *52*(1), 57–82.

Cederburg, S., & ODoherty, M. S. (2015). Asset-pricing anomalies at the firm level. *Journal of Econometrics*, *186*(1), 113–128.

Chen, Z., Liu, B., Wang, H., Wang, Z., & Yu, J. (2018). Characteristics-based factors.

Chordia, T., Goyal, A., & Shanken, J. A. (2015). Cross-sectional asset pricing with individual stocks: betas versus characteristics.

Cooper, M. J., Gulen, H., & Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. *The Journal of Finance*, *60*(6), 2825–2858.

Cosemans, M., Frehen, R., Schotman, P. C., & Bauer, R. (2015). Estimating security betas using prior information based on firm fundamentals. *The Review of Financial Studies*, *29*(4), 1072–1112.

Cremers, K. M. (2006). Multifactor efficiency and bayesian inference. *The Journal of Business*, *79*(6), 2951–2998.

Cremers, K. M., & Petajisto, A. (2009). How active is your fund manager? a new measure that predicts performance. *The Review of Financial Studies*, *22*(9), 3329–3365.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance*, *52*(3), 1035–1058.

Davis, J. L., Fama, E. F., & French, K. R. (2000). Characteristics, covariances, and average returns: 1929 to 1997. *The Journal of Finance*, *55*(1), 389–406.

De Pooter, M., Segers, R., & van Dijk, H. (2006). *Gibbs sampling in econometric practice* (Tech. Rep.).

Edelen, R. M., & Warner, J. B. (2001). Aggregate price effects of institutional trading: a study of mutual fund flow and market returns. *Journal of Financial Economics*, *59*(2), 195–220.

Evans, R. B. (2004). Does* alpha really matter? evidence from mutual fund incubation, termination and manager change.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3–56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1 - 22. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304405X14002323 doi: https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, *81*(3), 607–636.

Ferson, W. E., & Foerster, S. R. (1995). Method of moments: Tests of latent variable models.

Frazzini, A., & Lamont, O. A. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of financial economics*, *88*(2), 299–322.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (Vol. 2). CRC press Boca Raton, FL.

Grinblatt, M., & Titman, S. (1992). The persistence of mutual fund performance. *The Journal of Finance*, *47*(5), 1977–1984.

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

Guo, X. (2016). Mutual fund style volatility and fund flows.

Harvey, C. R., & Zhou, G. (1990). Bayesian inference in asset pricing tests. *Journal of Financial Economics*, *26*(2), 221–254.

Hobert, J. P., & Casella, G. (1996). The effect of improper priors on gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, *91*(436), 1461–1473.

Huij, J., & Verbeek, M. (2009). On the use of multifactor models to evaluate mutual fund performance. *Financial Management*, *38*(1), 75–102.

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

Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., & Wang, J. L. (2016). Empirical tests of asset pricing models with individual assets: resolving the errors-in-variables bias in risk premium estimation.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, *48*(1), 65–91.

Jordan, B., & Riley, T. (2016). Skill and persistence in mutual fund returns: Evidence from a six-factor model.

Kacperczyk, M., Sialm, C., & Zheng, L. (2005). On the industry concentration of actively managed equity mutual funds. *The Journal of Finance*, *60*(4), 1983–2011.

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

Lindeman, R. H., & Lindeman, R. H. (1980). *Introduction to bivariate and multivariate analysis* (Tech. Rep.).

Litzenberger, R. H., & Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of financial economics*, *7*(2), 163–195.

Lou, D. (2012). A flow-based explanation for return predictability. *Review of financial studies*, *25*(12), 3457–3489.

McCulloch, R., & Rossi, P. E. (1990). Posterior, predictive, and utility-based approaches to testing the arbitrage pricing theory. *Journal of Financial Economics*, *28*(1-2), 7–38.

Newey, W. K., & West, K. D. (1986). *A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.* National Bureau of Economic Research Cambridge, Mass., USA.

Nydick, S. W. (2012). The wishart and inverse wishart distributions. *Electron. J. Statist.*, *6*, 1–19.

Raponi, V., Robotti, C., & Zaffaroni, P. (2017). Testing beta-pricing models using large cross-sections.

Shanken, J. (1992). On the estimation of beta-pricing models. *The review of financial studies*, *5*(1), 1–33.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *The journal of finance*, *53*(5), 1589–1622.

Theil, H., & Theil, H. (1971). *Principles of econometrics* (Tech. Rep.).

Vayanos, D., & Woolley, P. (2013). An institutional theory of momentum and reversal. *Review of Financial Studies*, *26*(5), 1087–1145.

Warther, V. A. (1995). Aggregate mutual fund flows and security returns. *Journal of financial economics*, *39*(2-3), 209–235.

Wermers, R. (1997). Momentum investment strategies of mutual funds, performance persistence, and survivorship bias. *Unpublished Working Paper, University of Colorado,[downloaded from http://bus. colorado. edu/faculty/wermers/.]*.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817–838.

Figure 1: Time series plots of the posterior means of the time-varying relation between six-factor alphas and characteristics δ_t

The figure presents estimation results of the model in Eqs.(2), (3) and (4) examining the cross-sectional relation between six-factor alphas and all characteristics simultaneously. In each plot we report posterior results for a given characteristic across our sample period February 2001 to December 2016. The solid line is the posterior mean of δ_t and the dashed lines represent the 95% credible interval.

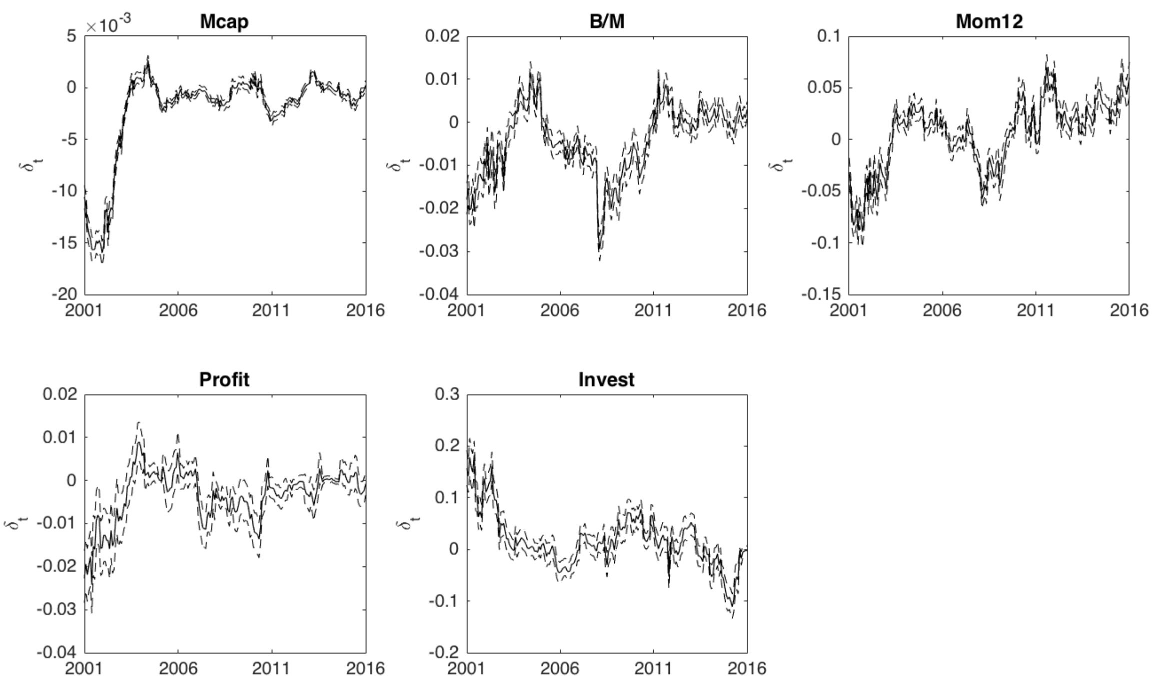

Figure 2: Cumulative alphas for top-minus-bottom portfolios

This figure presents the cumulative post-ranking Fama-French six-factor alpha for top-minus-bottom portfolios sorted on one of the following performance measures: six-factor alpha (α), double-adjusted six-factor alpha (α^*), or characteristic-driven performance (α^{char}). These performance measures are calculated using rolling windows with a window size of 24 months. Portfolios are rebalanced every quarter-end and held up to 40 quarters. The horizontal axis shows the post-ranking holding period in quarters.

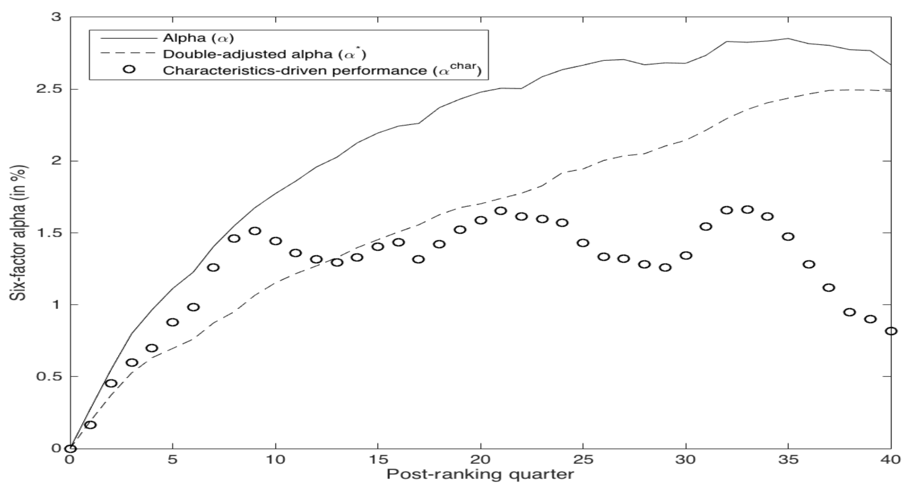

Table 1: Summary statistics of equity mutual fund sample

This table presents the summary statistics for the equity mutual funds sample over the period January 2001 to December 2016. Panel A reports statistics on the sample size. Panel B reports additional information on the funds. Panel C reports the cross-sectional distribution of characteristics averaged across all sample months. We weight each firm characteristic according to its current portfolio weight and calculate a fund's portfolio-weighted average characteristic. Market capitalization (Mcap) is the product between the previous month-end stock price and the previous month-end total shares outstanding. Book-to-Market (B/M) is the ratio between the most recently available book value of equity and the previous month-end market capitalization. Momentum (Mom12) is the past twelve-month cumulative return excluding the most recent month. Operating profitability (Profit) is the current revenues minus costs of goods sold, interest expense, selling, general, and administrative expenses, divided by book equity for the last fiscal year $t-1$. Asset growth (Invest) is the percentage change in total assets from fiscal year $t-2$ to fiscal year $t-1$. Panel D reports the time series averages of the cross-sectional distributions of six-factor betas, which are estimated from rolling time series regressions using the past two years of daily fund returns. Panel E reports the time series averages of the cross-sectional correlations between factor betas and characteristics.

	Mean	25% percentile	Median	75% percentile	Standard deviation
Panel A: Observations					
Number of distinct funds	2,871				
Number of fund-report dates	92,903				
Number of fund-month dates	314,362				
Number of distinct stocks	7,952				
Panel B: Fund characteristics					
Fund age	22.90	15.76	19.96	25.38	12.33
Fund monthly net return (in %)	0.51	-2.08	0.99	3.60	5.12
TNA (total net assets) (in millions)	1,341.43	56.00	220.00	863.10	5,335.74
Expense ratio (in %)	1.24	0.97	1.19	1.45	0.63
Turnover ratio (in %)	83.19	34.00	62.00	105.00	101.68
Panel C: Fund holdings characteristics					
Mcap (in millions)	44,475.74	3,792.98	43,016.63	78,966.25	39,731.13
B/M	0.46	0.32	0.44	0.56	0.16
Mom12 (in %)	17.12	9.15	14.89	22.86	12.13
Profit	0.46	0.35	0.42	0.50	0.24
Invest	0.09	0.07	0.09	0.12	0.04
Panel D: Rolling window fund six-factor betas					
β_{MKT}	0.98	0.94	0.99	1.03	0.10
β_{SMB}	0.20	-0.08	0.07	0.46	0.34
β_{HML}	0.00	-0.16	0.01	0.16	0.21
β_{WML}	0.02	-0.05	0.01	0.09	0.12
β_{RMW}	-0.06	-0.17	-0.02	0.09	0.21
β_{CMA}	-0.04	-0.15	-0.02	0.10	0.21

Table 1 (Continued)

Panel E: Cross-correlations between six-factor betas and characteristics

	β_{SMB}	β_{HML}	β_{WML}	β_{RMW}	β_{CMA}	Mcap	B/M	Mom12	Profit	Invest
β_{MKT}	0.096	-0.093	0.113	-0.204	-0.173	-0.023	-0.114	0.170	-0.004	0.077
β_{SMB}		0.095	0.043	-0.142	-0.075	-0.816	0.110	0.283	-0.255	-0.064
β_{HML}			-0.315	0.610	0.328	-0.042	0.746	-0.236	-0.079	-0.452
β_{WML}				-0.231	-0.270	-0.054	-0.469	0.568	0.103	0.191
β_{RMW}					0.361	0.246	0.367	-0.283	0.136	-0.274
β_{CMA}						0.049	0.354	-0.215	-0.004	-0.470
Mcap							0.049	0.354	-0.215	-0.004
B/M								-0.363	-0.189	-0.486
Mom12									-0.003	0.054
Profit										0.049

Table 2: Cross-sectional regressions of mutual fund returns

This table presents the time series averages of risk premia (γ) estimated using the cross-section of mutual fund (monthly) returns following the Fama & MacBeth (1973) procedure. The monthly regressions are of the form:

$$R_t = \gamma_{0t} + \gamma_{1t}\hat{B}_{t-1} + \gamma_{2t}Z_{t-1} + \xi_t, \ \ t = 1, ..., T.$$

We employ the CAPM, the Fama & French (1993) three-factor model, the Carhart (1997) four-factor model, the Fama & French (2015) five-factor model and a six-factor model combining all factors. Factor betas ($\hat{B}$) are estimated from rolling time series regressions using daily returns from the past two years. The characteristics (Z) are the logarithm of market capitalization (Mcap), the logarithm of book-to-market ratio (B/M), the logarithm of one plus the cumulative past twelve-month cumulative return (Mom12), operating profitability (Profit) and asset growth (Invest). Each characteristic is winsorized at the 0.5% and the 99.5% levels. To address the EIV bias, we employ the EIV-corrected estimator of Chordia et al. (2015). Risk premia (in percent per month) are fitted using OLS, both with EIV-correction and without. Fama & MacBeth (1973) t-statistics are reported in parenthesis. Estimates significant at the 5% are in bold font. Risk premia are estimated over the period February 2001 until December 2016.

	CAPM		FF 3FM		Carhart 4FM		FF 5FM		FF 6FM	
	OLS	EIV	OLS	EIV	OLS	EIV	OLS	EIV	OLS	EIV
Cnst	**1.013**	**1.020**	0.685	0.628	**0.887**	**0.857**	**0.651**	0.638	**0.769**	**0.759**
	(2.46)	(2.49)	(1.88)	(1.67)	(2.61)	(2.38)	(2.00)	(1.93)	(2.47)	(2.37)
β_{MKT}	-0.071	-0.080	-0.011	-0.011	-0.029	-0.033	0.130	0.177	0.100	0.133
	(-0.24)	(-0.26)	(-0.04)	(-0.03)	(-0.09)	(-0.10)	(0.46)	(0.61)	(0.35)	(0.46)
β_{SMB}			0.098	0.120	0.049	0.061	0.099	0.099	0.074	0.073
			(0.62)	(0.66)	(0.30)	(0.33)	(0.66)	(0.59)	(0.48)	(0.42)
β_{HML}			**0.347**	**0.373**	**0.372**	**0.412**	0.220	0.190	0.266	0.252
			(2.12)	(2.09)	(2.04)	(1.99)	(1.45)	(1.07)	(1.61)	(1.30)
β_{WML}					0.176	0.227			0.114	0.195
					(1.14)	(1.19)			(1.01)	(1.13)
β_{RMW}							0.190	0.224	0.225	0.271
							(1.17)	(1.16)	(1.29)	(1.32)
β_{CMA}							0.152	0.141	0.173	0.164
							(1.13)	(0.90)	(1.33)	(1.09)
Mcap	-0.065	-0.065	-0.047	-0.043	-0.055	-0.052	-0.050	-0.051	**-0.054**	-0.054
	(-1.97)	(-1.97)	(-1.46)	(-1.24)	(-1.79)	(-1.56)	(-1.82)	(-1.77)	(-1.96)	(-1.89)
B/M	0.004	0.003	-0.074	-0.082	0.072	-0.085	-0.025	-0.010	-0.031	-0.017
	(0.04)	(0.03)	(-0.98)	(-1.08)	(-1.08)	(-1.26)	(-0.36)	(-0.07)	(-0.50)	(-0.27)
Mom12	**0.783**	**0.787**	**0.863**	**0.859**	0.478	0.446	**0.794**	**0.774**	0.513	0.524
	(2.11)	(2.12)	(2.34)	(2.36)	(1.78)	(1.67)	(2.17)	(2.13)	(1.85)	(1.81)
Profit	0.066	0.065	0.047	0.046	0.019	0.016	0.030	0.025	0.003	0.000
	(1.18)	(1.17)	(0.94)	(0.94)	(0.40)	(0.34)	(0.68)	(0.59)	(0.07)	(0.02)
Invest	-0.573	-0.560	-0.389	-0.371	-0.391	-0.355	-0.279	0.257	-0.230	-0.222
	(-1.69)	(-1.67)	(-1.43)	(-1.38)	(-1.53)	(-1.40)	(-1.14)	(-0.97)	(-0.98)	(-0.87)

Table 3: Six-factor alpha vs. characteristics

This table presents the results of the estimation of the model in Eqs.(2), (3) and (4). We estimate this model in each month during the period February 2001 to December 2016, using an estimation period of two years to estimate the six-factor model in Eq.(2), rolling the window a month at a time. The characteristics (Z) are the logarithm of market capitalization (Mcap), the logarithm of book-to-market ratio (B/M), the logarithm of one plus the past twelve-month cumulative return (Mom12), operating profitability (Profit) and asset growth (Invest). Each characteristic is standardized by subtracting the cross-sectional mean each month. We estimate the model for each characteristic in isolation and for all characteristics in a joint model. Panel A presents the posterior mean and standard deviation for the aggregate-level parameters in $\bar{\delta}$, based on the posterior distribution of the parameters constructed from 5000 iterations of the Gibbs sampler with the first 2500 iterations discarded as a burn-in period. Panel B presents Fama & MacBeth (1973) estimates and Fama & MacBeth (1973) t-statistics with the Newey & West (1986) correction of 12 lags. In Panel A, estimates in bold font indicate that the 95% credible interval of the posterior distribution does not include zero. In Panel B, estimates in bold font indicate significance at the 5% level.

Panel A: Simultaneous Bayesian estimation						
Cnst	**-0.003**	**-0.003**	**-0.003**	**-0.003**	**-0.003**	**-0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mcap	**-0.002**					**-0.002**
	(0.000)					(0.000)
B/M		**-0.003**				**-0.004**
		(0.000)				(0.000)
Mom12			**0.019**			**0.005**
			(0.002)			(0.003)
Profit				-0.011		-0.003
				(0.002)		(0.002)
Invest					**0.024**	**0.017**
					(0.004)	(0.004)
Panel B: Two-step OLS approach						
Cnst	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(-1.40)	(-1.40)	(-1.40)	(-1.40)	(-1.40)	(-1.40)
Mcap	-0.001					-0.002
	(1.81)					(-1.60)
B/M		**-0.005**				**-0.006**
		(-2.45)				(-2.29)
Mom12			**0.016**			0.002
			(4.66)			(0.18)
Profit				-0.011		**-0.005**
				(-1.57)		(-2.27)
Invest					**0.040**	0.023
					(2.27)	(1.26)

Table 4: Multi-factor model alphas vs. characteristics

This table presents the results of the estimation of the model in Eqs.(2), (3) and (4) based on several Fama-French models. We employ the CAPM, the Fama & French (1993) three-factor model, the Carhart (1997) four-factor model, the Fama & French (2015) five-factor model and a six-factor model combining all factors. We estimate the model in each month during the period February 2001 to December 2016, using an estimation period of two years to estimate the factor model in Eq.(2), rolling the window a month at a time. The characteristics (Z) are the logarithm of market capitalization (Mcap), the logarithm of book-to-market ratio (B/M), the logarithm of one plus the past twelve-month cumulative return (Mom12), operating profitability (Profit) and asset growth (Invest). Each characteristic is standardized by subtracting the cross-sectional mean each month. We report the posterior mean and standard deviation for the aggregate-level parameters in $\bar{\delta}$, based on the posterior distribution of the parameters constructed from 5000 iterations of the Gibbs sampler with the first 2500 iterations discarded as a burn-in period. Estimates in bold font indicates that the 95% credible interval of the posterior distribution does not include zero.

	CAPM	FF 3FM	Carhart 4FM	FF 5FM	FF 6FM
Cnst	0.000	**-0.004**	**-0.003**	**-0.003**	**-0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mcap	**-0.003**	**-0.001**	**-0.001**	**-0.002**	**-0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
B/M	**0.006**	**0.002**	**0.001**	**-0.003**	**-0.003**
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Mom12	**0.015**	**0.008**	**0.004**	**0.007**	**0.005**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Profit	0.003	0.001	0.001	-0.004	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Invest	**-0.046**	**-0.023**	**0.020**	**0.014**	**0.017**
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)

Table 5: Change in performance percentile rankings

This table presents the differences between the performance percentile ranks based on double-adjusted six-factor alpha (α^*) relative to six-factor alpha (α). These performance measures are calculated using rolling windows with a window size of 24 months (see section **??**). Each month in our sample period we compute the difference between the percentile ranking based on alpha (P^{α}) and the percentile ranking based on double-adjusted alpha (P^{α^*}). We report the time series average distribution of the difference between performance rankings. The sample period is February 2001 to December 2016.

Percentile	5	10	25	50	75	90	95
Rank (%)	-22.93	-17.89	-8.99	-0.02	9.05	17.84	23.07
Abs. Rank (%)	0.99	1.49	4.08	9.03	15.98	22.99	27.24

Table 6: Mutual fund performance persistence

This table presents the returns of decile porfolios sorted by six-factor alpha (Panel A), double-adjusted six-factor alpha (Panel B) and characteristic-driven performance (Panel C). These performance measures are calculated using rolling windows with a window size of 24 months. Portfolios are rebalanced every quarter-end and are held for up to six years. To deal with overlapping portfolios, we follow Jegadeesh & Titman (1993) to take the equal-weighted return across overlapping portfolios formed in different quarters. Two different returns are reported: the excess (monthly) return over the risk-free rate and the Fama-French six-factor alpha. T-statistics, shown in parenthesis, are computed using White's standard errors. Estimated significant at the 5% level are in bold font. The sample period is February 2001 to December 2016.

	Qtr 1		Qtr 1-4		Qtr 5-12		Qtr 13-24	
Decile	Excess	6F alpha	Excess	6F alpha	Excess	6F alpha	Excess	6F alpha
Panel A: Six-factor alpha (α)								
1	0.43	-0.27	0.44	-0.26	0.49	-0.18	0.57	-0.12
2	0.46	-0.20	0.45	-0.21	0.49	-0.15	0.56	-0.10
3	0.44	-0.19	0.45	-0.19	0.49	-0.15	0.56	-0.10
4	0.47	-0.16	0.47	-0.16	0.49	-0.13	0.56	-0.10
5	0.49	-0.13	0.49	-0.13	0.50	-0.12	0.56	-0.10
6	0.53	-0.09	0.52	-0.11	0.54	-0.10	0.57	-0.09
7	0.54	-0.08	0.54	-0.09	0.55	-0.08	0.57	-0.09
8	0.57	-0.06	0.56	-0.07	0.56	-0.08	0.59	-0.08
9	0.57	-0.05	0.57	-0.04	0.59	-0.06	0.59	-0.08
10	0.58	0.01	0.58	0.01	0.57	-0.06	0.58	-0.08
10-1	0.15	**0.28**	0.14	**0.27**	0.08	**0.12**	0.01	0.05
	(1.63)	(3.62)	(1.61)	(3.85)	(1.14)	(2.17)	0.16	(1.01)
Panel B: Double-adjusted six-factor alpha (α^*)								
1	0.43	-0.21	0.45	-0.20	0.51	-0.14	0.57	-0.12
2	0.48	-0.18	0.49	-0.17	0.52	-0.13	0.57	-0.10
3	0.47	-0.18	0.50	-0.17	0.52	-0.13	0.58	-0.09
4	0.48	-0.17	0.49	-0.15	0.52	-0.13	0.56	-0.10
5	0.51	-0.14	0.50	-0.15	0.51	-0.13	0.58	-0.09
6	0.53	-0.11	0.52	-0.12	0.54	-0.11	0.58	-0.09
7	0.55	-0.07	0.52	-0.11	0.53	-0.11	0.57	-0.09
8	0.51	-0.11	0.53	-0.09	0.53	-0.09	0.56	-0.10
9	0.54	-0.04	0.53	-0.06	0.54	-0.07	0.57	-0.08
10	0.57	-0.02	0.56	-0.03	0.55	-0.05	0.59	-0.07
10-1	**0.14**	**0.19**	**0.11**	**0.17**	0.04	**0.08**	0.02	**0.05**
	(3.81)	(5.31)	(3.17)	(5.37)	(1.34)	(2.94)	(0.69)	(2.05)

Table 6 (Continued)

	Qtr 1		Qtr 1-4		Qtr 5-12		Qtr 13-24	
Decile	Excess	6F alpha	Excess	6F alpha	Excess	6F alpha	Excess	6F alpha
Panel C: Characteric-driven performance (α^{char})								
1	0.48	-0.18	0.44	-0.22	0.47	-0.15	0.58	-0.09
2	0.47	-0.16	0.46	-0.17	0.47	-0.14	0.56	-0.09
3	0.46	-0.16	0.45	-0.16	0.48	-0.13	0.55	-0.10
4	0.44	-0.16	0.45	-0.15	0.48	-0.13	0.55	-0.10
5	0.42	-0.18	0.44	-0.15	0.50	-0.12	0.55	-0.11
6	0.45	-0.14	0.46	-0.13	0.52	-0.11	0.55	-0.11
7	0.52	-0.10	0.52	-0.10	0.55	-0.09	0.58	-0.10
8	0.59	-0.06	0.57	-0.08	0.59	-0.06	0.61	-0.08
9	0.62	-0.05	0.62	-0.05	0.60	-0.08	0.61	-0.07
10	0.63	-0.02	0.65	-0.02	0.60	-0.08	0.60	-0.07
10-1	0.15	0.16	0.20	0.20	0.13	0.07	0.02	0.02
	(0.97)	(1.28)	(1.44)	(1.91)	(1.11)	(0.84)	(0.20)	(0.32)

Table 7: Relation between mutual fund performance and R^2

This table presents the time series averages of monthly cross-sectional regressions of mutual fund performance measures on fund selectivity, measured by the contemporaneous log-transformed R-squared ($\tilde{R}^2$). As (annualized) performance measures we employ six-factor alpha (α), double-adjusted six-factor alpha (α^*) and characteristic-driven performance (α^{char}). These performance measures are calculated using rolling windows with a window size of 24 months. We estimate the regressions with and without a set of control variables. Fama & MacBeth (1973) t-statistics with the Newey & West (1986) correction of 12 lags are reported in parenthesis. Estimates significant at the 5% are in bold font. The monthly regressions cover the period February 2001 until December 2016.

	α		α^*		α^{char}	
Cnst	**4.770**	**4.778**	**-0.518**	**-0.540**	**1.073**	**1.110**
	(2.01)	(1.98)	(-2.12)	(-2.20)	(2.76)	(2.61)
$\tilde{R}2$	**-1.148**	**-1.269**	-0.018	-0.021	**-0.314**	**-0.306**
	(-2.12)	(-2.36)	(-1.02)	(-1.18)	(-2.67)	(-2.55)
ExpRatio		8.112		0.848		-0.548
		(1.30)		(1.62)		(-0.37)
Turnover		-0.081		0.001		-0.002
		(-1.78)		(0.32)		(-0.29)
Log(TNA)		-0.014		-0.001		-0.002
		(-1.07)		(-0.23)		(-0.45)
Log(FundAge)		0.015		0.006		-0.010
		(0.48)		(1.45)		(-0.60)

Table 8: Response of fund flows to components of alpha

This table presents the time series averages of annual cross-sectional regressions of the one-year ahead average of monthly fund flows on double-adjusted six-factor alpha (α^*) and characteristic-driven performance (α^{char}). These performance measures are calculated using rolling windows with a window size of 24 months. We also estimate the regressions in which we further decompose α^{char} into alpha related to each individual characteristic. We estimate the regressions with and without a set of control variables. Fama & MacBeth (1973) t-statistics with the Newey & West (1986) correction of 3 lags are reported in parenthesis. Estimates significant at the 5% are in bold font. We estimate these annual regressions at the end of each year, covering 16 years from 2001 to 2016.

Cnst	**0.014**	0.004	**0.014**	0.005
	(6.69)	(0.58)	(6.48)	(0.69)
α^*	**4.061**	**4.044**	**4.065**	**4.047**
	(9.75)	(9.80)	(9.75)	(9.81)
α^{char}	**0.396**	**0.366**		
	(2.41)	(2.16)		
$\delta^{\text{Mcap}} \cdot$ Mcap			**-1.476**	**-1.455**
			(-2.19)	(-2.11)
$\delta^{\text{B/M}} \cdot$ B/M			0.275	0.215
			(0.31)	(0.24)
$\delta^{\text{Mom12}} \cdot$ Mom12			**1.778**	**1.720**
			(3.06)	(2.91)
$\delta^{\text{Profit}} \cdot$ Profit			0.338	0.221
			(0.37)	(0.22)
$\delta^{\text{Invest}} \cdot$ Invest			1.776	1.694
			(0.85)	(0.87)
ExpRatio		0.481		0.468
		(1.42)		(1.40)
Turnover		0.000		0.000
		(-0.96)		(-1.07)
Log(TNA)		0.000		0.000
		(-0.82)		(-0.82)
Log(FundAge)		0.002		0.002
		(1.75)		(1.78)

Table 9: Rolling window frequency robustness

This table presents the results of the estimation of the model in Eqs.(2), (3) and (4) using different frequencies of the cross-sectional regressions in Eq.(3). We estimate this model during the period February 2001 to December 2016, using an estimation period of two years to estimate the six-factor model in Eq.(2), rolling the window three months at a time between quarter-ends in Panel A and six months at a time in Panel B. The characteristics (Z) are the logarithm of market capitalization (Mcap), the logarithm of book-to-market ratio (B/M), the logarithm of one plus the past twelve-month cumulative return (Mom12), operating profitability (Profit) and asset growth (Invest). Each characteristic is standardized by subtracting the cross-sectional mean each month. We estimate the model for each characteristic in isolation and for all characteristics in a joint model. We presents the posterior mean and standard deviation for the aggregate-level parameters in $\bar{\delta}$, based on the posterior distribution of the parameters constructed from 5000 iterations of the Gibbs sampler with the first 2500 iterations discarded as a burn-in period. Estimates in bold font indicate that the 95% credible interval of the posterior distribution does not include zero.

Panel A: Quarterly cross-sectional regressions						
Cnst	**-0.003**	**-0.003**	**-0.003**	**-0.003**	**-0.003**	**-0.002**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mcap	**-0.002**					**-0.002**
	(0.000)					(0.000)
B/M		**-0.003**				**-0.004**
		(0.000)				(0.001)
Mom12			**0.019**			**0.005**
			(0.003)			(0.005)
Profit				-0.011		-0.003
				(0.003)		(0.002)
Invest					**0.022**	**0.016**
					(0.007)	(0.007)
Panel B: Semi-annual cross-sectional regressions						
Cnst	**-0.003**	**-0.003**	**-0.003**	**-0.003**	**-0.003**	**-0.003**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mcap	**-0.002**					**-0.002**
	(0.000)					(0.000)
B/M		**-0.003**				**-0.003**
		(0.001)				(0.001)
Mom12			**0.017**			0.006
			(0.004)			(0.007)
Profit				-0.011		-0.003
				(0.004)		(0.002)
Invest					**0.017**	0.012
					(0.009)	(0.009)

Table 10: DGTW CS measure vs. six-factor betas

This table presents the time series averages of monthly cross-sectional regressions of the DGTW characteristic selectivity (CS) measure on the Fama-French six-factor betas. We calculate the average DGTW CS measure across the past 24 months. We estimate the six-factor model over the same 24-month period using daily fund returns. In Panel A, the DGTW CS measure controls for the size, value and momentum characteristics. In Panel B, the DGTW CS measure also controls for the profitability and investment characteristics. Fama & MacBeth (1973) t-statistics with the Newey & West (1986) correction of 12 lags are reported in parenthesis. Estimates significant at the 5% are in bold font. Starting with the 24th month in our sample, the monthly regressions cover the period December 2002 to December 2016.

Panel A: Size, value and momentum characteristics							
Cnst	**0.225**	0.003	0.001	-0.004	0.013	0.003	**0.207**
	(3.30)	(0.31)	(0.06)	(-0.40)	(1.81)	(0.32)	(2.89)
RMRF	**-0.227**						**-0.211**
	(-3.09)						(-2.71)
SMB		-0.008					-0.008
		(-0.35)					(-0.31)
HML			0.090				0.071
			(1.67)				(1.65)
WML				**0.207**			**0.323**
				(2.31)			(3.89)
RMW					**0.116**		0.063
					(2.03)		(1.18)
CMA						0.010	0.001
						(0.11)	(0.02)
Panel B: All characteristics							
Cnst	**0.250**	0.006	0.007	0.000	0.016	0.011	**0.218**
	(3.54)	(0.59)	(0.68)	(0.01)	(1.79)	(1.20)	(3.31)
RMRF	**-0.246**						**-0.223**
	(-3.10)						(-3.08)
SMB		0.014					0.013
		(0.76)					(0.60)
HML			0.046				0.024
			(1.06)				(0.48)
WML				**0.237**			**0.351**
				(2.59)			(4.11)
RMW					**0.092**		0.067
					(2.23)		(1.27)
CMA						0.019	0.041
						(0.29)	(0.91)

Appendices

Appendix A: Mutual Fund Selection

Our sample contains U.S. equity actively managed funds at the intersection of the CRSP Survivor-Bias-Free U.S. Mutual Fund database with the Thomson Reuters Mutual Fund Holdings S12 database. We use the MFLINKS database available from Wharton Research Data Services (WRDS) to combine both databases. Our final database contains mutual fund holdings spanning the period from January 2001 to December 2016.

A.1 CRSP Mutual Fund Database

Since we wish to capture active mutual funds that invest primarily in U.S. equities, we follow Kacperczyk et al. (2008) and Lou (2012), by eliminating balanced, bond, money market, international, index, and sector funds, as well as funds that do not primarily invest in U.S. common equity. We base our mutual fund selection on identifiers provided in the CRSP mutual fund database. Specifically, we select funds with the following Lipper Objective codes: EI, EIEI, EMN, FLX, G, GI, I, LCCE, LCGE, LCVE, LSE, MC, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, SCVE, SESE, SG. If a fund's Lipper Objective code is missing, we pick funds with the following Strategic Insight Objective codes: SCG, GRO, AGG, ING, GRI, GMC. If both codes are missing for a fund, we select funds with the following Wiesenberger Fund Type codes: SCG, AGG, G, G-S, S-G, GRO, LTG, I, I-S, IEQ, ING, GCI, G-I, G-I-S, G-S-I, I-G, I-G-S, I-S-G, S-G-I, S-I-G, GRI, MCG.

Since some funds misreport their objective code, we require funds to hold at least 80% and at most 105% in common stocks, on average. Index funds are eliminated based on the CRSP index fund flags (provided since 2003) and by screening fund names. In particular, funds are dropped if the fund name contains the following strings: INDEX, IND, INDX, IDX, IDX, MKT, MARKET, SP, SP, MSCI, NYSE, RUSSELL, NASDAQ, ISHARES, DOWJONES, SPDR, ETF, 100, 400, 500, 600, 1000, 1500, 2000, 3000, 5000. Following Kacperczyk et al. (2008), to address the incubation bias[18] we delete observations for which the date of the observation is prior to the reported fund start-date and we delete observations with missing fund names. Other data from the CRSP mutual fund database include the total net assets (TNA), net returns (both monthly and daily), fees, and

[18] Evans (2004) and Kacperczyk et al. (2008) detect a form of survival bias in the CRSP mutual fund database, which stems from fund families sugarcoating their past performance. Fund families incubate private funds and only report the returns of the surviving incubated funds and do not disclose the past performance of terminated funds.

other qualitative fund data. We aggregate all different share classes belonging to a single fund at each point in time into one observation. Regarding the quantitative attributes of funds, we sum the TNA and we take a weighted average of the fund returns, expense ratio, turnover ratio and fees, using the lagged TNAs of each individual share class as weights. Regarding the qualitative attributes of funds (e.g., fund name, CRSP objective code, year of origin), the data of the oldest fund is retained.

A.2 Thomson Reuters Database

Mutual fund holdings are provided by Thomson Reuters and are compiled from mandatory SEC fillings[19] and voluntary disclosures. From this database we exclude funds with the following objective codes: International, Municipal Bonds, Bond & Preferred, Balanced and Metals. Every fund files the SEC form at the end of a quarter (the file date), which is often in the same quarter as the report date; the date for which the holdings are actually held (adjusted for stock splits[20]). To create a monthly time series of fund holdings, we keep reported holdings constant between report dates (e.g., holdings reported at the end of September are valid in October, November and December). A majority of funds report holdings on a quarterly basis, while a small number of funds have gaps between report dates of more than two quarters. To fill these gaps (of no more than two quarters), we impute holdings of missing quarters using the most recently available report date, assuming that these funds adopt a buy-and-hold strategy. In the final database about 65% of the funds disclose their holdings quarterly, 34% semi-annually, and 1% on a less frequent basis.

A.3 MFLINKS

To combine the CRSP mutual fund database with the Thomson Reuters database, we use the MFLINKS provided by Russ Wermers on Wharton Research Data Services (WRDS). MFLINKS maps CRSP fund identifiers to Thomson Reuters fund identifiers, covering approximately 98% of the domestic equity mutual funds. We manage to link about 92% of the target universe in the CRSP mutual fund database to holdings data from Thomson Reuters. To ensure a reliable linkage between the two databases, we require that the TNAs reported by both databases do not differ by more than a factor of two. Finally, funds with less than 10 identified stock positions and less than

[19]Investment companies, which include mutual funds, insurance companies, banks, pension funds, and numerous other institutions are often called 13f institutions. These institutions are required to fill in a form with the Securities and Exchange Commission (SEC) on a semiannual basis.

[20]Adjustments are made using the cumulative adjustment factor for shares in the CRSP monthly file.

$5 million assets under management are excluded.

The final mutual fund database contains 2,871 distinct mutual funds including 92,903 fund-report dates and 314,362 fund-month observations. Table A1 presents the number of funds at the end of each year along with the TNA and number of holdings reported by Thomson Reuters. There is a rising trend in both the number of funds, the average fund size, and the market share held.

Table A1: End of year summary statistics of the equity mutual fund sample

This table presents summary statistics for the mutual fund database as of December each year. *#Holdings* is the number of reported holdings in Thomson Reuters. *#Distinct stocks* contains the number of unique stocks held by the funds in the sample and the aggregated percentage market share held. TNA is the total net assets under management reported by CRSP, expressed in millions USD.

Year	#Funds	#Holdings		#Distinct stocks		TNA($M)	
		Mean	Median	Mean	%Market	Mean	Median
2001	1,126	110	73	4,802	7.83	730.69	165.20
2002	1,236	120	75	4,301	9.57	759.20	141.00
2003	1,221	115	78	4,383	10.10	903.92	164.85
2004	1,137	113	75	4,392	12.57	1,192.07	192.30
2005	1,061	124	77	4,056	11.71	1,420.41	238.80
2006	967	122	76	4,021	12.14	1,700.04	298.60
2007	1,071	113	75	4,211	11.96	1,649.38	276.30
2008	1,192	120	75	4,012	11.98	972.76	165.20
2009	1,114	116	80	3,793	12.82	1,387.80	233.45
2010	1,294	122	78	3,691	12.91	1,520.00	299.10
2011	1,226	122	78	3,470	12.17	1,440.66	290.80
2012	1,335	123	77	3,398	12.79	1,579.65	361.80
2013	1,176	114	77	3,365	12.44	2,153.20	519.85
2014	1,170	119	78	3,424	11.58	2,383.97	540.90
2015	797	111	70	3,320	9.56	2,584.96	584.60
2016	697	113	75	3,180	8.45	2,798.21	648.10

Appendix B: Errors-in-variables (EIV) - corrected estimator

The cross-sectional regression in Eq.(1) is inherently subject to the EIV bias, since the explanatory variables are estimations resulting from the first-pass time series regressions. Since $\hat{B}_{t-1}$ is estimated with error, the OLS-estimator of Γ will be biased downwards.

Chordia et al. (2015) propose the following bias-corrected estimator of Γ_t

$$\hat{\Gamma}_t^{EIV} = \left[\hat{X}'_{t-1}\hat{X}'_{t-1} - \sum_{i=1}^{N_t} M'\hat{\Sigma}_{\beta_{it-1}} M \right]^{-1} \hat{X}'_{t-1} R_t, \tag{A.1}$$

where $\hat{X}_{t-1} = [1_{N_t}\ \hat{B}_{t-1}\ Z_{t-1}]$ contains the lagged regressors, 1_{N_t} is a $N_t \times 1$ vector of ones, M is a $K \times (1{+}K{+}L)$ matrix defined as $M = [0_{K\text{x}1}\ 1_{K\text{x}K}\ 0_{K\text{x}L}]$, and $\hat{\Sigma}_{\beta_{it-1}}$ is the heteroskedasticity-consistent $K \times K$ covariance matrix estimator of White (1980) for the estimation of the factor model parameters β_{it-1}. The matrix M ensures that the bias-correction only affects the $K \times K$ submatrix $\hat{B}'_{t-1}\hat{B}_{t-1}$ of X_{t-1}.

This bias-corrected estimator was originally proposed by Theil & Theil (1971) and Litzenberger & Ramaswamy (1979). Shanken (1992) generalize the EIV-corrected estimator and show that this estimator is consistent when N_t diverges. Chordia et al. (2015) gauge the statistical properties of the EIV-corrected estimator in simulations and show that the negative bias is reduced in comparison to the OLS estimator. Raponi et al. (2017) employ this estimator in a small T environment to test several prominent beta-pricing specifications of Fama-French using individual stocks. They find significant pricing ability of all factors, while the same risk premia often appear insignificantly different from zero when estimated using the traditional approach.

The EIV-corrected estimator subtracts the estimated covariance matrix of the estimator of β_{it} from $\hat{B}'_{t-1}\hat{B}_{t-1}$, to better approximate the true value of $B'_{t-1}B_{t-1}$. However, under a finite T there is the possibility that this correction will overshoot, turning the matrix in parenthesis nearly singular or even not positive definite. This may lead to extreme estimates of Γ_t and nonsensical inference.

To prevent this we apply the following procedure. Following Chordia et al. (2015), we reduce the likelihood of overshooting due to outliers by winsorizing each element of the estimated covariance matrix at the 5% and 95% levels across the cross-section of funds at each time t. Then, we apply the shrinkage procedure of Raponi et al. (2017) using a shrinkage scalar λ ($0 \leq \lambda \leq 1$):

$$\hat{\Gamma}_t^{\text{EIV}} = \left[\hat{X}'_{t-1}\hat{X}_{t-1} - \lambda \sum_{i=1}^{N_t} M'\hat{\Sigma}_{\beta_{it-1}} M \right]^{-1} \hat{X}'_{t-1} R_t. \tag{A.2}$$

When λ is one we obtain the estimator in Eq.(A.1), whereas when λ is zero, we obtain the OLS estimator. The choice of shrinkage parameter λ is dependent on the eigenvalues of the matrix in parenthesis. Starting from $\lambda = 1$, if the minimum eigenvalue of this matrix is negative, we lower λ by an arbitrary small amount set to 0.05. We also apply this shrinkage in case the difference between the EIV-corrected and OLS coefficients is bigger than 100%.

Appendix C: Probability Density Functions

In this appendix several univariate and multivariate probability density functions are given which are used throughout this paper. For univariate densities, we indicate the first moment around the mean by μ, whereas for multivariate densities this is indicated by $\boldsymbol{\mu}$.

C.1 Univariate Distributions

Normal density:

The pdf of a normal distributed random variable Z with mean μ and variance σ^2, that is, $Z \sim \mathcal{N}(\mu, \sigma^2)$ is given by

$$p(Z|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Z-\mu)^2}{2\sigma^2}\right). \tag{A.3}$$

Inverted Gamma-2 density:

The pdf of an inverted Gamma-2 distributed random variable Z with scale parameter $s > 0$ and degrees of freedom $v > 0$, that is, $Z \sim IG2(s, v)$ is given by

$$p(Z|s, v) = c^{-1} Z^{-\frac{v+2}{2}} \exp\left(-\frac{s}{2Z}\right), \tag{A.4}$$

where $c = \Gamma(v/2)\left(\frac{2}{s}\right)^{v/2}$. The mean and variance of Z are given by

$$E[Z] = \frac{s}{v-2} \qquad \text{for } v > 2$$

$$\text{Var}[Z] = \frac{2}{v-4}(E[Z])^2 \qquad \text{for } v > 4.$$

C.2 Multivariate Distributions

Multivariate Normal density:

The pdf of a multivariate normal distributed random variable Z with k-dimensional location parameter $\boldsymbol{\mu}$ and k x k positive definite symmetric covariance matrix Σ, that is, $Z \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ is given by

$$p(Z|\boldsymbol{\mu}, \Sigma) = \left(\frac{1}{\sqrt{2\pi}}\right)^k |\Sigma|^{-\frac{1}{2}} \exp(-\frac{1}{2}(Z - \boldsymbol{\mu})'\Sigma^{-1}(Z - \boldsymbol{\mu})), \tag{A.5}$$

where the mean and variance of Z are $\boldsymbol{\mu}$ and Σ, respectively.

Inverted Wishart density:

The pdf of an inverted Wishart distributed k x k symmetric positive definite matrix Z with degrees of freedom $v \geq k$ and symmetric positive definite scale matrix S, that is, $Z \sim IW(S, v)$ is given by

$$p(Z|S, v) = c^{-1} \frac{|S|^{v/2}}{|Z|^{(v+k+1)/2}} \exp\left(-\frac{1}{2}\text{tr}(Z^{-1}S)\right), \tag{A.6}$$

where $c = 2^{\frac{vk}{2}} \Gamma_k \left(\frac{v}{2}\right)$. The mean of Z is given by

$$E[Z] = \frac{S}{v - k - 1} \qquad \text{for } v > k + 1.$$

For a nice representation of the variance of Z, we refer to Nydick (2012).

Appendix D: Gibbs Sampler

We provide the Gibbs sampler and the posterior distributions of the parameters in our model in Eqs.(2), (3) and (4), which we repeat here for convenience

$$R_{i\tau,t} = \alpha_{it} + \beta_{it}' F_{\tau,t} + \epsilon_{i\tau,t}, \ \ \epsilon_{i\tau,t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_{it}}), \ \tau = t - \mathcal{T} + 1, ..., t, \ i = 1, ..., N_t \tag{A.7}$$

$$\alpha_{it} = \delta_{0t} + \delta_{1t}' Z_{it-1} + \eta_{it}, \ \ \eta_{it} \sim \mathcal{N}(0, \sigma^2_{\eta t}), \ t = 1, ..., T, \tag{A.8}$$

$$\delta_t = \bar{\delta} + v_t, \ \ v_t \sim \mathcal{N}(0, \Sigma_\delta), \ t = 1, ..., T. \tag{A.9}$$

For the sake of convenience, we can write the model in matrix form as

$$R_{i,t} = X_t B_{it} + \epsilon_{i,t} \tag{A.10}$$

$$\alpha_t = Y_{t-1}\delta_t + \eta_t \tag{A.11}$$

$$\delta_t = \bar{\delta} + v_t, \tag{A.12}$$

where $R_{i,t}$ is a $\mathcal{T} \times 1$ vector with daily returns of fund i during the rolling window ending in month t, X_t is a $\mathcal{T} \times (K{+}1)$ matrix with a column of ones and the corresponding factor returns, $B_{it} = [\alpha_{it}, \beta_{it}]$' contains the parameters of the factor model, and $\epsilon_{i,t}$ is a $\mathcal{T} \times 1$ vector with residuals. In the second equation, α_t is a $N_t \times 1$ vector with factor model alphas in month t, Y_{t-1} is a $N_t \times (L{+}1)$ matrix with a column of ones and the lagged fund-level characteristics for each fund, and η_t is a $N_t \times 1$ vector with residuals.

D.1 Gibbs Sampler

To initialize the Gibbs sampler, we require starting values for the model parameters. For the fund-specific parameters, we estimate Eq.(2) for each fund in each rolling window and use the estimates as starting values for α_{it} and β_{it}. Each month t, we regress the estimates of alpha on a constant and the lagged characteristics using the cross-section of funds. The resulting estimates are used as starting values for δ_t. The starting values for $\sigma^2_{\epsilon_{it}}$ and $\sigma^2_{\eta t}$ are set to the sample residual variance estimator from the corresponding regressions. For the hyperparameters $\bar{\delta}$ and Σ_δ, we set the starting values equal to the sample mean and covariance matrix of the estimates $\{\delta_t\}_{t=1}^T$, respectively. The Gibbs sampler uses an iterative procedure to create Markov chains by simulating from full conditional posteriors. The algorithm below summarizes the steps in the Gibbs sampler.

Algorithm 1 MCMC Gibbs sampler

INPUTS: $\{\{\{R_{i\tau,t}, Z_{it-1}, F_{\tau,t}\}_{i=1}^{N_t}\}_{\tau=t-\mathcal{T}+1}^{t}\}_{t=1}^{T}$ (data)
OUTPUTS: $\{\theta^{(m)}\}_{m=1}^{M}$ (approximate sample from the joint posterior)

1: Set starting values for model parameters:
$\theta^{(0)} = \{\{\{\alpha_{it}^{(0)}, \beta_{it}^{(0)}, \sigma_{\epsilon_{it}}^{2\ (0)}\}_{i=1}^{N_t}, \delta_t^{(0)}, \sigma_{\eta t}^{2\ (0)}\}_{t=1}^{T}, \bar{\delta}^{(0)}, \Sigma_\delta^{(0)}\}$ and set m to 0.

2: Update parameters given current draws $\theta^{(m)}$:

(i) Sample $\alpha_{it}^{(m+1)}, \beta_{it}^{(m+1)} \mid \sigma_{\epsilon_{it}}^{2\ (m)}, \delta_t^{(m)}, \sigma_{\eta t}^{2\ (m)}$, for $i = 1,...,N_t$, $t = 1,...,T$.

(ii) Sample $\sigma_{\epsilon_{it}}^{2\ (m+1)} \mid \alpha_{it}^{(m+1)}, \beta_{it}^{(m+1)}$, for $i = 1,...,N_t$, $t = 1,...,T$.

(iii) Sample $\delta_t^{(m+1)} \mid \{\alpha_{it}^{(m)}\}_{i=1}^{N_t}, \sigma_{\eta t}^{2\ (m)}, \bar{\delta}^{(m)}, \Sigma_\delta^{(m)}$, for $t = 1,...,T$.

(iv) Sample $\sigma_{\eta t}^{2\ (m+1)} \mid \{\alpha_{it}^{(m+1)}\}_{i=1}^{N_t}, \delta_t^{(m+1)}$, for $t = 1,...,T$.

(v) Sample $\bar{\delta}^{(m+1)} \mid \{\delta_t^{(m+1)}\}_{t=1}^{T}, \Sigma_\delta^{(m)}$.

(vi) Sample $\Sigma_\delta^{(m+1)} \mid \{\delta_t^{(m+1)}\}_{t=1}^{T}, \bar{\delta}^{(m+1)}$.

3: Set $m = m + 1$, and go to step 2.

D.2 Joint Posterior Distribution

The joint posterior distribution is proportional to the product of the likelihood function of the data and the prior distributions. Denote $\theta = \{\{\{\alpha_{it}, \beta_{it}, \sigma_{\epsilon_{it}}^2\}_{i=1}^{N_t}, \delta_t, \sigma_{\eta t}^2\}_{t=1}^{T}, \bar{\delta}, \Sigma_\delta\}$ as the model parameters. Denote R as the matrix which stacks the returns in the rolling windows across all funds over the entire sample period. Substituting the prior specifications as in Section 3.A.1, the joint posterior is given by:

$$
\begin{aligned}
p(\theta|R) &\propto p(R|\theta)p(\theta) \\
&\propto \prod_{t=1}^{T}\prod_{i=1}^{N_t} (\sigma_{\epsilon_{it}}^2)^{-\frac{\mathcal{T}}{2}} \exp\left[-\frac{1}{2\sigma_{\epsilon_{it}}^2}(R_i - X_t B_{it})'(R_i - X_t B_{it})\right] \\
&\times \prod_{t=1}^{T} (\sigma_{\eta t}^2)^{-\frac{N_t}{2}} \exp\left[-\frac{1}{2\sigma_{\eta t}^2}(\alpha_t - Y_{t-1}\delta_t)'(\alpha_t - Y_{t-1}\delta_t)\right] \\
&\times \prod_{t=1}^{T}\prod_{i=1}^{N_t} |\Sigma_\beta|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\beta_{it} - \bar{\beta})'\Sigma_\beta^{-1}(\beta_{it} - \bar{\beta})\right] \\
&\times \prod_{t=1}^{T} |\Sigma_\delta|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\delta_t - \bar{\delta})'\Sigma_\delta^{-1}(\delta_t - \bar{\delta})\right] \\
&\times |\Sigma_d|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\bar{\delta} - d)'\Sigma_d^{-1}(\bar{\delta} - d)\right]
\end{aligned}
$$

$$
\times |\Sigma_\delta|^{-\frac{\psi_\delta+k+1}{2}} \exp\left[-\frac{1}{2}\text{tr}(\Sigma_\delta^{-1}[\psi_\delta S_\delta])\right]
$$
$$
\times \prod_{t=1}^{T}\prod_{i=1}^{N_t} \sigma_{\epsilon_{it}}^{(v_\epsilon+2)} \exp\left(-\frac{s_\epsilon}{2\sigma_{\epsilon_{it}}^2}\right)
$$
$$
\times \prod_{t=1}^{T} \sigma_{\eta t}^{(v_\eta+2)} \exp\left(-\frac{s_\eta}{2\sigma_{\eta t}^2}\right), \quad \text{(A.13)}
$$

where k denotes the dimension of Σ_δ.

D.3 Conditional Posterior Distributions

In order to implement the MCMC Gibbs sampler, we need to derive the full conditional posterior distributions for each block of model parameters. The conditional densities can be obtained from the joint posterior density by gathering all the terms that depend on the parameters of interest and ignore all the remaining parameters. We then obtain the conditional density for each block of parameters by rearranging the remaining terms into the kernel of a known distribution.[21] We partition the model parameters θ into the following blocks:

- $\theta^{(1)}$: Fund-specific factor model parameters (α_{it}, β_{it})
- $\theta^{(2)}$: Fund-specific idiosyncratic variance ($\sigma^2_{\epsilon_{it}}$)
- $\theta^{(3)}$: Characteristic loadings (δ_{0t}, δ_{1t})
- $\theta^{(4)}$: Cross-sectional idiosyncratic variance ($\sigma^2_{\eta t}$)
- $\theta^{(5)}$: Time series mean of characteristic loadings ($\bar{\delta}$)
- $\theta^{(6)}$: Time series covariance matrix of characteristic loadings (Σ_δ)

The conditional posteriors for all parameters have convenient functional forms which trace back to a known distribution. Therefore, we use the Gibbs sampler to iteratively draw from the full conditional distributions of $\theta^{(1)}$, $\theta^{(2)}$, $\theta^{(3)}$, $\theta^{(4)}$, $\theta^{(5)}$, and $\theta^{(6)}$. Using the joint posterior distribution in Eq.(A.13), we will derive the full conditional posteriors for each block.

[21] Appendix D provides the probability density functions (pdfs) of the distributions used in our analysis.

D.3.1 Conditional Posterior $[\alpha_{it}\ \beta_{it}]$'

In order to derive the full conditional posterior of $[\alpha_{it}\ \beta_{it}]$', we require the following mathematical relations. First, we require

$$B^{-1} = B^{-\frac{1}{2}\prime} B^{-\frac{1}{2}}, \tag{A.14}$$

which is denoted the Cholesky decomposition, i.e., the decomposition of a positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose. We also require the decomposition rule:

$$(y - X\beta)'(y - X\beta) = (y - X\hat{\beta})'(y - X\hat{\beta}) + (\beta - \hat{\beta})'X'X(\beta - \hat{\beta}), \tag{A.15}$$

where $\hat{\beta} = (X'X)^{-1}X'y$ is the OLS estimator of β.

Following Bayes' theorem, we can write

$$\begin{aligned}
p(B_{it}|R) &\propto p(R|B_{it})p(B_{it}) \\
&\propto \exp\left[-\frac{1}{2\sigma^2_{\epsilon_{it}}}(R_{i,t} - X_tB_{it})'(R_{i,t} - X_tB_{it})\right] \times \exp\left[-\frac{1}{2}(B_{it} - \bar{B}_{it})'\Sigma^{-1}_{B_{it}}(B_{it} - \bar{B}_{it})\right], \\
&\text{where } \bar{B}_{it} = \begin{bmatrix} \delta_{0t} + \delta'_{1t}Z_{it-1} \\ \bar{\beta} \end{bmatrix} \text{ and } \Sigma_{B_{it}} = \begin{bmatrix} \sigma^2_{\eta t} & 0_{1\times K} \\ 0_{K\times 1} & \Sigma_\beta \end{bmatrix} \\
&= \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2_{\epsilon_{it}}}(R_{i,t} - X_tB_{it})'(R_{i,t} - X_tB_{it}) + (\Sigma^{-\frac{1}{2}}_{B_{it}}\bar{B}_{it} - \Sigma^{-\frac{1}{2}}_{B_{it}}B_{it})'(\Sigma^{-\frac{1}{2}}_{B_{it}}\bar{B}_{it} - \Sigma^{-\frac{1}{2}}_{B_{it}}B_{it})\right)\right] \\
&\text{using Eq.(A.14)} \\
&= \exp\left[-\frac{1}{2}(w_{it} - V_{it}B_{it})'(w_{it} - V_{it}B_{it})\right], \\
&\text{where } w_{it} = \begin{bmatrix} \sigma^{-1}_{\epsilon_{it}}R_{i,t} & \Sigma^{-\frac{1}{2}}_{B_{it}}\bar{B}_{it} \end{bmatrix}' \text{ and } V_{it} = \begin{bmatrix} \sigma^{-1}_{\epsilon_{it}}X_t & \Sigma^{-\frac{1}{2}}_{B_{it}} \end{bmatrix}' \\
&\propto \exp\left[(B_{it} - \hat{B}_{it})'V'_{it}V_{it}(B_{it} - \hat{B}_{it})\right], \text{ using Eq.(A.15)}, \\
&\text{where } \hat{B}_{it} = (V'_{it}V_{it})^{-1}V'_{it}w_{it},
\end{aligned} \tag{A.16}$$

which is the kernel of a normal distribution with mean $\hat{B}_{it}$ and variance $(V'_{it}V_{it})^{-1}$, that is, the full conditional posterior[22] of $B_{it} = [\alpha_{it}\ \beta_{it}]$' follows a multivariate normal distribution and is given by

$$\begin{bmatrix} \alpha_{it} \\ \beta_{it} \end{bmatrix} |\theta^{-(\alpha_{it},\beta_{it})}, R \sim \mathcal{N}\left(\tilde{B}_{it}, \left[\sigma_{\epsilon_{it}}^{-2} X_t' X_t + \Sigma_{B_{it}}^{-1}\right]^{-1}\right), \tag{A.17}$$

with

$$\tilde{B}_{it} = \left[\sigma_{\epsilon_{it}}^{-2} X_t' X_t + \Sigma_{B_{it}}^{-1}\right]^{-1} \left[\sigma_{\epsilon_{it}}^{-2} X_t' R_t + \Sigma_{B_{it}}^{-1} \bar{B}_{it}\right]. \tag{A.18}$$

We derive the full conditional posteriors of δ_t and $\bar{\delta}$ in the same way, leading to the following results.

D.3.2 Conditional Posterior δ_t

The full conditional posterior of δ_t follows a multivariate normal distribution and is given by

$$\delta_t|\theta^{-(\delta_t)}, R \sim \mathcal{N}\left(\tilde{\delta}_t, \left[\sigma_{\eta t}^{-2} Y_{t-1}' Y_{t-1} + \Sigma_{\delta}^{-1}\right]^{-1}\right), \tag{A.19}$$

with

$$\tilde{\delta}_t = \left[\sigma_{\eta t}^{-2} Y_{t-1}' Y_{t-1} + \Sigma_{\delta}^{-1}\right]^{-1} \left[\sigma_{\eta t}^{-2} Y_{t-1}' \alpha_t + \Sigma_{\delta}^{-1} \bar{\delta}\right]. \tag{A.20}$$

D.3.3 Conditional Posterior $\bar{\delta}$

The full conditional posterior of $\bar{\delta}$ follows a multivariate normal distribution and is given by

$$\bar{\delta}|\theta^{-(\bar{\delta})}, R \sim \mathcal{N}\left(\tilde{\bar{\delta}}, \left[T\Sigma_{\delta}^{-1} + \Sigma_{d}^{-1}\right]^{-1}\right), \tag{A.21}$$

with

$$\tilde{\bar{\delta}} = \left[T\Sigma_{\delta}^{-1} + \Sigma_{d}^{-1}\right]^{-1} \left[\Sigma_{\delta}^{-1} \sum_{t=1}^{T} \delta_t + \Sigma_{d}^{-1} d\right]. \tag{A.22}$$

[22] Let $\theta^{-(x)}$ denote the model parameters excluding the set x.

D.3.4 Conditional Posterior Σ_δ

In order to derive the full conditional posterior of Σ_δ, we require the following mathematical relations. First, we require

$$A = \text{tr}(A) \tag{A.23}$$

$$\text{tr}(ABC) = \text{tr}(CAB) = \text{tr}(BCA) \tag{A.24}$$

$$\text{tr}(A) + \text{tr}(B) = \text{tr}(A + B), \tag{A.25}$$

where tr denotes the trace of a matrix. The above relations state that the trace of a scalar is equivalent to the scalar itself, that the trace is invariant under cyclic permutations, and that the sum of traces equals the trace of the sum.

Following Bayes' theorem, we can write

$$\begin{aligned}
p(\Sigma_\delta|R) &\propto p(\Sigma_\delta)p(\delta_t) \\
&\propto |\Sigma_\delta|^{-\frac{\psi_\delta+k+1}{2}} \exp\left[-\frac{1}{2}\text{tr}(\Sigma_\delta^{-1}[\psi_\delta S_\delta])\right] \times \prod_{t=1}^{T} |\Sigma_\delta|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\delta_t - \bar{\delta})'\Sigma_\delta^{-1}(\delta_t - \bar{\delta})\right) \\
&= |\Sigma_\delta|^{-\frac{\psi_\delta+T+k+1}{2}} \exp\left[-\frac{1}{2}\left(\text{tr}(\Sigma_\delta^{-1}[\psi_\delta S_\delta]) + \text{tr}\left(\sum_{t=1}^{T}(\delta_t - \bar{\delta})'\Sigma_\delta^{-1}(\delta_t - \bar{\delta})\right)\right)\right], \\
&\text{using Eq.(A.23)} \\
&= |\Sigma_\delta|^{-\frac{\psi_\delta+T+k+1}{2}} \exp\left[-\frac{1}{2}\left(\text{tr}(\Sigma_\delta^{-1}[\psi_\delta S_\delta]) + \text{tr}\left(\sum_{t=1}^{T}\Sigma_\delta^{-1}(\delta_t - \bar{\delta})(\delta_t - \bar{\delta})'\right)\right)\right], \\
&\text{using Eq.(A.24)} \\
&= |\Sigma_\delta|^{-\frac{\psi_\delta+T+k+1}{2}} \exp\left[-\frac{1}{2}\text{tr}\left(\Sigma_\delta^{-1}\left[\sum_{t=1}^{T}(\delta_t - \bar{\delta})(\delta_t - \bar{\delta})' + \psi_\delta S_\delta\right]\right)\right], \\
&\text{using Eq.(A.25)},
\end{aligned} \tag{A.26}$$

which is the kernel of an inverted Wishart distribution, that is, the full conditional posterior distribution of Σ_δ is given by

$$\Sigma_\delta|\theta^{(-\Sigma_\delta)}, R \sim IW\left(\sum_{t=1}^{T}(\delta_t - \bar{\delta})(\delta_t - \bar{\delta})' + \psi_\delta S_\delta, \psi_\delta + T\right). \tag{A.27}$$

D.3.5 Conditional Posterior $\sigma^2_{\epsilon_{it}}$

Following Bayes' theorem, we can write

$$\begin{aligned} p(\sigma^2_{\epsilon_{it}}|R) &\propto p(R|\sigma^2_{\epsilon_{it}})p(\sigma^2_{\epsilon_{it}}) \\ &\propto \sigma_{\epsilon_{it}}^{-(\mathcal{T}+v_\epsilon+2)} \exp\left[-\frac{1}{2\sigma^2_{\epsilon_{it}}}\left((R_{i,t}-X_tB_{it})'(R_{i,t}-X_tB_{it})+s_\epsilon\right)\right], \end{aligned} \tag{A.28}$$

which is the kernel of an inverted Gamma-2 distribution, that is, the full conditional posterior distribution of $\sigma^2_{\epsilon_{it}}$ is given by

$$\sigma^2_{\epsilon_{it}}|\theta^{(-\sigma^2_{\epsilon_{it}})}, R \sim IG2\left((R_{i,t}-X_tB_{it})'(R_{i,t}-X_tB_{it})+s_\epsilon, \mathcal{T}+v_\epsilon\right). \tag{A.29}$$

D.3.6 Conditional Posterior $\sigma^2_{\eta t}$

The full conditional posterior of $\sigma^2_{\eta t}$ follows an Inverted Gamma-2 distribution and is given by

$$\sigma^2_{\eta t}|\theta^{(-\sigma^2_{\eta t})}, R \sim IG2\left((\alpha_t - Y_{t-1}\delta_t)'(\alpha_t - Y_{t-1}\delta_t)+s_\eta, N_t+v_\eta\right). \tag{A.30}$$

Appendix E: Construction of Benchmark Portfolios of DGTW Characteristic Selectivity Measure

We construct the benchmark portfolios in the spirit of Daniel, Grinblatt, Titman, and Wermers (DGTW; 1997). We proceed as follows. We conduct a three-dimensional sort along the size, value and momentum dimensions. All stocks having accounting data from Compustat, as well stock return and market capitalization data are sorted into quintile portfolios by their market capitalization. The breakpoints for the sort on size is based on stocks listed on the NYSE, while the analysis includes all stocks listed on the NYSE, NYSE MKT (formerly known as AMEX), and NASDAQ. Each size quintile is further subdivided into quintiles based on the book-to-market ratio of stocks, resulting into 25 fractile portfolios. Then, each fractile portfolio is further subdivided into quintiles based on the past twelve-month cumulative return excluding the most recent month. This 5x5x5 sorting procedure leads to 125 portfolios, each with a unique combination of size, value, and momentum rankings.[23]

Next, to adjust for profitability and investment effects, we also conduct a 3x3x3x2x2 sorting procedure along the size, value, momentum, profitability, and investment dimensions. Similar to the previous sorting procedure, we first sort stocks into three portfolios based on market capitalization, using breakpoints derived from NYSE listed stocks. Then we further subdivide the portfolios, first into three value portfolios and then into three momentum portfolios, resulting into 27 portfolios. Then, we subdivide each fractile portfolio into two portfolios based on the operating profitability of stocks. We conclude the sorting procedure by further dividing the fractile portfolios into two portfolios based on the investment variable. Thus, we end up with 108 portfolios, each with a distinct combination of size, value, momentum, profitability, and investment rankings.

The 3x3x3x2x2 (5x5x5) sorting procedure is repeated at the end of each quarter, such that the 108 (125) portfolios are reconstituted each quarter-end, and are held for the subsequent quarter. We calculate the value-weighted average return for each portfolio during the months in the holding period. We store quarter-end rankings for all stocks, such that we can track the rankings of all portfolio holdings, which are valid for the subsequent quarter.

[23] For example, a stock sorted in size portfolio five, value portfolio five, and momentum five, is a large, value and high momentum stock.