Assessing Asset Pricing Models Using Exchange-Traded Fund Flows[*]

K.C. John Wei
Hong Kong Polytechnic University
School of Accounting and Finance
E-mail: john.wei@poyu.edu.hk

Linti Zhang
Hong Kong Polytechnic University
School of Accounting and Finance
E-mail: linti1002.zhang@polyu.edu.hk

Abstract

We test and compare the response of capitals flowing into and out of U.S. equity exchange-traded funds (ETFs) to alphas relative to various asset pricing models to infer how ETF investors evaluate risk. Results from both linear regressions and non-linear horse race tests reveal that the market-adjusted return (MAR) is the best predictor for subsequent ETF flows. By testing the sensitivities of fund flows to components of ETF returns attributable to different factors, we find that investors account for the market risk the most and to some extents the momentum, mispricing-based, and behavior-based factors. In contrast, they largely ignore other common factors when allocating capitals. The superior performance of the MAR is robust to controlling for Morningstar ratings, while it can be partially explained by the preferences for ETF investors for assets with higher prospect and salience theories values.

JEL Classification: G12; G14

Keywords: Equity exchange-traded funds (ETFs); Asset pricing models; Fund flows

[*] John Wei is also a Professor Emeritus at Hong Kong University of Science and Technology. John Wei acknowledges financial support from the Research Grants Council of the Hong Kong Special Administrative Region, China (GRF15507320). All remaining errors are our own.

Corresponding author: K.C. John Wei, School of Accounting and Finance, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: johnwei@ust.hk; Tel: 852-2766-4953; Fax: 852-2330-9845.

Assessing Asset Pricing Models Using Exchange-Traded Fund Flows

Abstract

We test and compare the response of capitals flowing into and out of U.S. equity exchange-traded funds (ETFs) to alphas relative to various asset pricing models to infer how ETF investors evaluate risk. Results from both linear regressions and non-linear horse race tests reveal that the market-adjusted return (MAR) is the best predictor for subsequent ETF flows. By testing the sensitivities of fund flows to components of ETF returns attributable to different factors, we find that investors account for the market risk the most and to some extents the momentum, mispricing-based, and behavior-based factors. In contrast, they largely ignore other common factors when allocating capitals. The superior performance of the MAR is robust to controlling for Morningstar ratings, while it can be partially explained by the preferences for ETF investors for assets with higher prospect and salience theories values.

JEL Classification: G12; G14

Keywords: Equity exchange-traded funds (ETFs); Asset pricing models; Fund flows

1. Introduction

Exchange-traded funds (ETFs) have experienced tremendous growth in the past decade. At the end of 2020, total net assets managed by ETFs has amounted to $7.9 trillion globally. ETFs are a convenient way for retail and institutional investors seeking to gain from the exposures to broad markets, sectors, or factors. The advantages of ETFs include easy diversification, low expense ratio, low trading cost, and tax efficiency. Although ETFs have attracted a lot of attention from media, investors, and regulators and become one of the retail investors' preferred choices for long-term investment, academic research is still at its infant stage. Most academic studies focus on the impacts of equity ETF ownership on the underlying stocks' price efficiency, return volatility, liquidity, and mispricing, while how ETF investors evaluate risk is an under-researched question.

This paper fills this gap by inferring ETF investors' preferences for the return-risk trade-offs from the response of their capital allocation decisions to model-adjusted returns. The failure of the capital asset pricing model (CAPM) to describe cross-sectional stock returns has stimulated extensive literature to develop new models by incorporating non-market risk factors to explain stock market anomalies. Since Fama and French (1993) propose a three-factor model (FF3) that add the size and value factors to the CAPM motivated by the small and value firm premiums, a surge of studies have added other new factors to the FF3 model as new anomalies are discovered. Well-known examples include the four-factor model of Carhart (1997, Car4) by adding the momentum factor; the seven-factor model of Pástor and Stambaugh (2002) by further adding three industry factors (PS7); the five-factor model of Fama and French (2015) by adding the profitability and investment factors (FF5); the four-factor model of Hou, Xue, and Zhang (2015) based on q-theory (QF4); the four-factor model of Stambaugh and Yuan (2017) motivated by 11 prominent stock market anomalies (SY4); and the three-factor model of Daniel, Hirshleifer, and Sun (2020) which captures investors' underreaction in the short run and overconfidence in the long-run (DHS3). Many studies compare empirical validity of various newly developed models by testing the models' capabilities to explain returns on individual stocks or characteristics-sorted portfolios. However, whether the multifactor models provide a better description of equity prices remains a controversial topic. Using passively managed and well-diversified ETFs as alternative test assets provides an out-of-sample comparison of how investors assess asset pricing models when they make capital allocation decisions.

Another advantage of using ETFs as test assets is that one can rely on the net capital flowing into or out of the funds to infer investors' preferences for asset pricing models. In a competitive capital market, investors would fiercely chase positive net present value (NPV) investment opportunities and trade assets to eliminate such opportunities. If fund investors tend to use a particular model to assess historical fund performance, they should buy (sell) funds with positive (negative) alphas relative to the model, resulting in an association between historical fund alphas and future fund flows. Built on this insight, Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) use mutual fund flows to infer which asset pricing model investors use and find that the CAPM-alpha is the best predictor for future capital allocations. We apply the same methodologies of these two studies to domestic equity ETFs in the U.S. market to examine which model is used by ETF investors to evaluate fund performance.

In particular, this paper tests and compares the capabilities of nine candidate asset pricing models to affect ETF investors' capital allocation decisions: including the pure behavior model that investors benchmark funds against the market portfolio and eight prominent factor models. For each model, we calculate weighted-average historical model-adjusted returns or alphas over various horizons to identify *ex-ante* positive NPV investment opportunities, and then examine how investors allocate capital in response to the abnormal returns. Following Berk and van Binsbergen (2016), we first examine the univariate relation between fund flow and alpha relative to each model by regressing signed fund flows on signed alphas. The estimated regression coefficient measures the strength of the fund flow-alpha relation. Confirming the predictability of alphas relative to various models for future fund flows documented by prior studies, we find a strong association between all model-adjusted returns and subsequent capital flows. After accounting for other potential predictors for fund flows, the signs of fund alphas on average correctly predict the signs of fund flows in the next month with the lowest probability of 52.99% for the PS7 model and the highest probability of 54.84% for the MAR. To examine the partial predictive power of the models for subsequent cash inflows or outflows, fund flows are regressed on alphas relative to the nine models simultaneously. We find that the regression coefficient on MAR is higher than those on other models' alphas in terms of both statistical and economic significances. CAPM-alpha and Car4-alpha also appear to have some power to predict fund flows incremental to alphas relative to other models, although the magnitudes of the partial effects are much lower than that of MAR. For all evaluation horizons, model-adjusted alphas along with fund characteristics and lagged fund flows only explain around 3% of total variations in percentage of fund flows. The finding that no model can predict signs of fund flows with probability exceeding 55% and the low adjusted R^2 of the linear panel regressions indicate that a large fraction of ETF investors' capital allocation decisions remains unexplained.

We then compare the relative performance to predict fund flows for each pair of the nine asset-pricing models. Models can be ranked by the beta coefficients of signed fund flows on signed model-adjusted returns, with a larger coefficient indicating better performance. We regress fund flows on the difference between signed alphas relative to each pair of the models to formally test the difference in beta coefficients. For all evaluation horizons, MAR and CAPM-alpha have significantly higher beta coefficients than alphas relative to other models. However, the OLS-based pairwise test cannot reliably distinguish between the MAR and the CAPM in predicting subsequent fund flows. To attenuate possible non-linearity in the fund flow-alpha relation, we conduct pairwise horse race tests based on portfolio-sorting to compare model performance as in Barber, Huang and Odean (2016). We identify cases where the rankings of fund performance diverge between paired models. We then examine these cases further on the magnitude of the difference in coefficients of fund flows on ranking dummies and the proportion of cases that a model predicts future flows better than the other model. In a large majority of the cases, the rankings of fund performance according to the MAR predict cash inflows or outflows more correctly, when the other models disagree with the MAR in fund performance. The CAPM is the second-best performing model that wins the horse race among all models except the MAR. We also notice that the behavioral-based DHS3 model appears to perform relatively better than the other multifactor models.

Taken together, the results from the main analysis reveal that the MAR dominates the other factor models in predicting subsequent fund flows. When investors evaluate the performance of ETFs, returns on the market portfolio seem to be the foremost benchmark in comparison to factors associated with well-known cross-sectional stock anomalies. Meanwhile, the superior performance of the MAR does not imply that investors completely ignore common risk factors when allocating capital. To dip into the extent to which ETF investors ignore common factors in the existing asset pricing models, we decompose fund returns into abnormal returns and returns attributable to each factor and examine the responses of fund flows to each component of returns. In the panel regressions using components of returns to predict fund flows, the coefficients on returns traced to all factors are significantly positive, suggesting that investors fail to fully account for the exposures to the comovement of stock prices. However, the degrees of ignorance differ across factors. The part of return attributed to the market risk has the smallest coefficient, which is significantly lower than the coefficient on alphas, suggesting that investors are most concerned about the market risk. The partial effects of returns attributed to the momentum factor (UMD) and the factors related to short-term underreaction (PEAD) and long-term overconfidence (FIN) are also reliably lower than those of model-adjusted returns. The size-related return has the strongest effect on fund flows, which means that the size premium is the most neglected factor when evaluating the passively managed funds.

We next explore and test two explanations for the superior performance of the MAR over all of the factor models assessed in predicting capitals invested in or withdrawn from ETFs. First, the results can be driven by investors' reliance on Morningstar Ratings, which provide a simple evaluation metric to rate funds. As the third -party financial intermediary ignores factors in more recently developed models and neglects cross-fund differences in factor loadings, the failure of multifactor models may be attributed to the limitations in the rating algorithm. Second, as the ETF market is dominated by retail investors who are more vulnerable to cognitive and behavioral bias, they may evaluate funds according to prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman,1992) and/or salience theory (Bordalo, Gennaioli, and Shleifer, 2012; 2013). In particular, ETFs with higher respect theory value or with more salient payoff can have greater demands and attract more cash inflows. To the extent that MAR is more correlated with the two behavior measures, the MAR would outperform factor-based models in affecting capital allocations. We find that controlling for Morningstar Ratings has a negligible effect on our main findings, while the outperformance of the MAR is partially attenuated by prospect theory value and salience theory value.

The rest of the paper is organized as follows. In section 2, we discuss relevant prior studies. Section 3 describes our sample and detailed research methodologies. In Section 4, we present and discuss the results of main empirical analysis used to test and compare the performance of alternative asset-pricing models to predict fund flows. Section 5 conducts additional analysis and robustness tests, and the final section concludes the paper.

2. Related Literature

2.1. The ETF markets

Since the SPDR ETF (S&P 500 Index Trust) launched in January 1993, equity ETFs have grown rapidly in the past decades, especially for factor-investing ETFs. The U.S. remains the largest with $5.4 trillion and 2,204 funds compared to only $0.8 trillion and 797 funds at the end of 2009 (Investment Company Fact Book). According to BlackRock, the factor-investing industry is currently estimated at $1.9 trillion and is projected to grow to $3.4 trillion by 2022. Lettau and Madhavan (2018) regard "ETF represents one of the most important financial innovations in decades." ETFs are typically a convenient tool for retail and institutional investors who seek to track the return performance of a particular index such as the S&P 500, a specific sector such as hi-tech, or a certain style/factor such as growth/value. Most ETFs are like passive mutual funds, which hold well diversified portfolios. However, ETFs are also different from mutual funds in many fundamental dimensions. ETFs can be freely traded like stocks in the secondary stock exchanges with much lower trading costs, lower bid-ask spreads, and higher liquidity than underlying stocks, while open-ended mutual funds cannot. ETFs are also more transparent in reporting their stock holdings than mutual funds: daily versus quarterly. Gastineau (2010) indicates that many traders use ETFs to hedge (either long or short) market and/or sector exposures to gain from their firm-specific private information.

Given their relative advantages over mutual funds and individual stocks, ETFs have attracted a lot of attention from media, investors, and regulators and have become one of the retail investors' preferred choices for long-term investment. Nevertheless, academic research on ETFs is still at its infant stage. Most studies focus on the impacts of equity ETF ownership on the underlying stocks' volatility (Krause, Ehsani, and Lien, 2014; Ben-David, Franzoni, and Moussawi, 2018), price efficiency (Israeli, Lee, and Sridharan, 2017; Glosten, Nallareddy, and Zou, 2019), return comovement (Da and Shive, 2018), and liquidity (Agarwal, Hanouna, Moussawi, and Stahel, 2018), among others.

From the viewpoint of ETF investors, it is obvious that they care the information on the cross-section of risk and return the most. Only a few studies examine the pricing of individual ETFs in the cross section. For example, in their recent study, Brown, Davies, and Ringgenberg (2021) find that ETF flows contain information about non-fundamental demand shocks which have significant impacts on ETF prices. However, compared to extensive studies about cross-sectional stock prices, how investors evaluate risk and return of individual ETFs is not well-explored by existing literature.

2.2. Literature on asset pricing models

The risk and return relation is a fundamental concept in finance. Based on Markowitz's (1952) portfolio theory, Sharpe (1964) develops the capital asset pricing model (CAPM) which dictates that the expected return on an asset is determined by its market beta. Extensive literature has tested the empirical validity of the CAPM and finds that the model does not perform well in describing asset returns. Prior studies find that stocks with higher market betas do not earn significantly higher average returns (Fama and French, 1992; He and Ng, 1994), which is contrary to what the model predicts. The literature has documented a growing number of stock market anomalies that differences in average returns on stocks sorted by firm characteristics cannot be explained by differences in market betas.

A large number of studies argues that the reason for the failure of the CAPM is that the model omits relevant factors. As a result, various multifactor models have been proposed to better accommodate stock anomalies. Motivated by the present value equation and empirical findings that small firms (Banz, 1981) and stocks with higher P/E ratios returns (Basu, 1977; Rosenberg, Reid, and Lanstein, 1985) earn higher average returns, Fama and French (1993) propose to add the size and value factors. They find that the three-factor model (FF3) can better explain portfolios of stocks sorted by firm size and valuation ratio. Carhart (1997) finds that the FF3 fails to explain momentum in mutual fund performance, while the momentum anomaly can be largely explained by a four-factor model (Car4) that adds the momentum factor.

Over the past two decades, the FF3 and Car4 models have become academic standards. Recently, new asset pricing models have been proposed to better accommodate newly discovered anomalies that cannot be explained by the two models. In particular, Fama and French (2015) proposes to add the profitability and investment factors to the FF3 and demonstrates that the five-factor model (FF5) outperforms the FF3 in explaining the higher average returns of stocks with higher profitability documented by Novy-Marx (2013) and more conservative investment by Titman, Wei, and Xie (2004) and Cooper, Gulen, and Schill (2008). Their follow-up study (Fama and French, 2016) shows that the FF5 performs better than the CAPM, FF3 and Car4 in explaining the volatility puzzle (Ang, Hodrick, Xing, and Zhang, 2006) and net share issuance anomaly

(Loughran and Ritter, 1995; Daniel and Titman, 2006; Pontiff and Woodgate, 2008). Motivated by the q-theory, Hou, Xue, and Zhang (2015) construct a four-factor model (QF4) that consists of the size, profitability, and investment factors from portfolios sorted by size, investment-to-assets, and quarterly profitability, in addition to the market factor. They find that the QF4 provides comparable or better performance in digesting about 40 significant anomalies in the U.S. individual stock market.

More recently, instead of basing upon rational asset pricing theory, Stambaugh and Yuan (2017) propose an empirical approach of averaging stock market anomalies to construct mispricing-based factors. Specifically, they retrieve common information from 11 prominent anomalies separated into two clusters with the greatest co-movements in time-series and cross-sectional anomaly rankings. The first mispricing factor "MGMT" is created from the first group of six anomalies related to managers' decisions, and the second mispricing factor "PERM" is created from the second group of five anomalies related to firm performance. The four-factor model (SY4) which combines the two mispricing factors with the market and size factors proves to have the ability to explain a wide range of anomalies exceeding that of the FF5 and QF4 models.

Another strand of the literature argues that investors' cognitive biases can have implications on asset prices, and many papers propose behavior theories which can better explain many anomalies that remain puzzling to rational pricing models. For instance, Hirshleifer, Lim, and Teoh (2011) argue that investors with limited attention will underreact to new information in the short run, resulting in the post earnings announcement drift (PEAD). Gervais and Odean (2001) develop a model under which biases in investors' learning process result in overconfident traders, who will initially hold on to their biased beliefs of privative information. The reluctance to correct for misbeliefs upon the arrival of new information results in persistent mispricing. Managers who possess superior information about firm intrinsic value will time the issuing or repurchasing of equity (FIN) to take advantage of mispriced stocks. Thus, the net equity issue anomaly is related to investors' overconfidence. Motivated by the above behavior models, Daniel et al. (2020) augment the market factor with two behavior factors (PEAD and FIN) that capture the investors' irrational behavior in the short run and long run. The three-factor behavioral model (DHS3) appears to outperform other rational- and empirical-based models in accommodating a wide range of anomalies.

Despite the extensive work devoted to understanding the determinants of stock prices and returns, the literature still lacks a consensus in which model provides the best fits of the observed stock prices. The majority of empirical tests on the asset pricing models has been focusing on the equity market. Recent studies have extended the literature by testing the performance of various models in the bond and derivative markets. Yet no prior studies have tested the abilities of different models in describing prices and returns on ETFs that investors care the most.

2.3. Literature about fund flows

Extensive literature has examined the determinants of fund flows. Early studies document that capitals flowing into or out of mutual funds are driven by historical fund performance. For instance, Ippolito (1992), Chevalier and Ellison (1997), and Sirri and Tufano (1998) find that equity mutual fund flows with higher average returns in the past attract more cash inflows. Some papers also examine how fund flows respond to abnormal fund returns in relative to particular benchmarks. Gruber (1996) finds that fund alphas relative to a four-factor model strongly predicts future mutual fund flows. Del Guercio and Tkac (2002) compare the performance of raw returns and Jensen's alphas to predict capitals invested in or withdrawn from mutual funds and pension funds. They find that both unadjusted and market risk adjusted returns have incremental power to predict future fund flows. Fung, Hsieh, Naik, and Ramadorai (2008) document that alphas relative to a seven-factor model including market and size factors along with factors in the bond and currency markets is strongly related to future flows into or out of hedge funds. Clifford, Fulkerson, and Jordan (2014) examine the determinants of ETF flows and find that fund flows are driven by historical raw returns on funds, which is likely the result of naïve extrapolation bias.

Recently, the literature has witnessed renewed interests in investigating the fund flow-alpha relation, as comparisons of the sensitivities of fund flows to alphas relative to different models offer an alternative approach to distinguish competing asset pricing models in describing what investors care the most when making investment decisions. In a competitive financial market, investors should chase abnormally profitable investment opportunities and trade to eliminate such opportunities quickly. In equilibrium, prices of financial securities are set so that expected returns are determined by exposures to factors. Prior studies usually test asset pricing models by adjusting observed asset returns for model factors and comparing the magnitudes of alphas relative to models, with smaller alphas indicating better model performance.

One limitation of using asset prices data is that the detection of abnormal investment opportunities and the evaluation of model performance occur at the same time, and it is unclear whether an *ex-post* positive NPV opportunity exits *ex-ante*. Relying on quantities of funds can resolve the problem by using historical alphas to identify profitable investment opportunities *ex-ante* and examine the how investors allocate capitals in response to the identified opportunities. Stronger associations between fund flows and fund alphas relative to a model suggest better performance of the model. Motivated by this insight, Berk and van Binsbergen (2016) (BvB hereafter) propose a simple test statistic, the beta coefficient of signed (percentage) mutual fund flow on signed fund alpha relative to a model, to quantify the model's performance. They find that the CAPM has a test statistic significantly higher than those of the market-adjusted return, the FF3, the Car4, and several dynamic equilibrium models. Another independent work by Barber, Huang, and Odean (2016) (BHO hereafter) conducts non-linear horse race tests to compare the performance between pairs of various models to predict mutual fund flows. They also find that capital flows are most sensitive to CAPM-alphas.

However, the literature has not reached a consensus about how to interpret the documented differences in sensitivities of fund flows to alphas relative to different models. While BvB (2016) argue that the superior performance of the CAPM indicates that the model is the closest to the true asset pricing model, BHO (2016) take it as evidence that mutual fund investors are unsophisticated investors who fail to account for factors known to be related to cross-sectional equity returns. Jegadeesh and Mangipudi (2021) argue that the results of fund-alpha horse race tests among models depend on how precise empiricists are able to estimate alphas and factor loadings. The model of which alphas and factor loadings can be estimated with the highest precision will win the horse race test. The results of their simulation analysis show that fund-alpha horse race tests as in BvB (2016) and BHO (2016) cannot be used to identify the true asset pricing model or to infer investors' level of sophistication.

Evans and Sun (2021) find that the differences in the fund flow-alpha relations among models are mitigated by the differences in the correlations between model-adjusted returns and Morningstar Ratings, which is a simple evaluation metric popular among fund investors. Specifically, they find that the outperformance of the CAPM over the FF3 significantly weakens after July 2002, when the third-party institution modified its rating algorithm from benchmarking individual funds against the pooled funds to comparing funds within peer funds with similar size, investment style, or industry tilt.

Blocher and Molyboga (2017) and Agarwal, Green, and Ren (2018) conduct the similar fund flow-alpha tests with hedge funds. Both studies find that CAPM-alpha performs better than alphas relative to more sophisticated models in predicting hedge fund flows. However, no prior studies have used ETF fund flows to infer ETF investors' preferences for asset pricing models.

2.4. Contributions to the literature

In this paper, we test and compare the performance of various asset pricing models in the U.S. domestic equity ETFs market by investigating how fund flows respond to model-adjusted returns. The paper extends and contributes to several strands of literature and has practical implications. First, we examine how investors evaluate risks and returns of ETFs, which is an under-researched topic in the literature about the ETF market. Understanding the cross-section of risk-adjusted returns on equity ETFs can benefit investors, especially household retail investors, to make a better investment decision on ETFs that match their needs. Knowledge about ETF investors' risk preferences can also help fund managers to use more suitable benchmarks to gauge fund performance.

Second, using the passively managed funds as test assets, we provide an out-of-sample test on the abilities of prominent factor models to describe what investors care the most when making their investment choices. When comparing the performance of different asset pricing models, most studies typically use characteristic-shorted (long only or long-short) portfolios. However, these portfolios, in particular long-short portfolios, are costly and difficult for any investors, especially retail investors, to replicate or implement. The advantage of using ETFs is that we can apply the methodologies of BvB (2016) and BHO (2016) to assess model performance using quantities data, which enable us to directly test investors' responses to profitable investment opportunities. While the two studies only consider rational-based asset pricing models, we include the recently developed mispricing-based and behavioral-based factor models in the model competition. The outcomes of the horse race tests can reveal whether accounting for investors' psychological biases help better explain their trading behaviors. In addition, the relative performance of models to predict ETF flows can help resolve controversies over the implications of differential model performance in the mutual fund flow literature. If mutual fund investors' preferences for the CAPM indicates that the model is the true asset pricing model and it can be generally used by all investors across markets as in BvB (2016), we expect that the CAPM should also provide the best prediction for fund flows in the ETF market. Meanwhile, the investor sophistication argument of BHO (2016) predicts that the differences in responses of fund flows to alphas would be related to the aggregate level of investors' sophistication in a market. To the extent that the ETF market is dominated by less sophisticated investors, the investor sophistication argument predicts that ETF investors will have a stronger response of fund flows to the models with fewer factors.

Third, the paper extends the previous studies on determinants of fund flows to the ETF market. Existing studies about fund flows have been focusing on the mutual or hedge fund market, and only a few studies examine ETF flows. ETFs have features distinctive from those of mutual funds or hedge funds. Unlike mutual funds and hedge funds, investor can easily short ETFs with low costs. Besides, ETFs are passively managed funds aiming to track an underlying index with absence of skill, while all hedge funds and most of mutual funds are actively managed portfolios aiming to beat the market. Mutual fund investors are expected to trade in pursuit of superior management skills, while most ETFs are regarded as factor/style investing (Barberis and Shleifer, 2003). The differences in investment philosophies of investors in these markets indicate that the empirical finings in the mutual fund or hedge fund market may not be generalized to the ETF market. In light of the rapidly growing popularity of ETFs as an investment option among investors, it is important to understand the determinants of the capital flows into or out of individual ETFs.

3. Data and Research Methodologies

3.1. Data and sample

Daily and monthly prices and trading information of ETFs, including prices, returns, trading volume, and number of shares outstanding, are obtained from the Center for Research in Security Prices (CRSP). Data on fund characteristics, including fund name, fund style, Lipper code, expense ratio, and total net asset value (TNA), are retrieved from CRSP Survivor-Bias-Free U.S. Mutual Fund. ETFs in the two databases are matched by CUSIP. We retain funds with CRSP share code of 73 and ETF flag of "F". CRSP fund returns are net of expenses and fees. Following existing studies,[1] we compute gross monthly returns on ETFs by adding back the annual expense ratio divided by 12 in a given year.[2]

Our sample includes domestic equity ETFs of the U.S. market (with Levels 1 and 2 of CRSP style code = "ED"). To exclude positive and negative leverage and synthetic ETFs, we drop ETFs whose name contains "bond", "hedged", "bear", "2X", "-1X", -2X, or "-3X", or whose Level 4 of CRSP style code is "Hedged (H)" or "Short (S)". Data on ETFs in the U.S. market are available since 1993. To ensure a reasonable number of ETFs in the cross-section, we consider the sample period from January 2000 throughout December 2019, during which at least 50 funds are in the sample each month. The majority of the ETFs are passive investment funds with the objective to closely track an index, while a small number of funds are index-enhanced funds that seek to outperform the overall or a segment of the market. To rule out the possibility that investors evaluate

[1] For instance, Koijen (2014) and Cohen, Cova, and Pástor (2005), among others.
[2] Missing expense ratio of a fund for a year is replaced by the median value of expense ratio of the fund in the future years. Only ETFs that can be matched with expense ratio data in at least one year are included in the sample.

actively managed and passively managed funds in very different ways, we exclude index-enhanced funds from the sample. We also drop fund-months for funds with lower than $5 million total net value at the end of the previous month. Our final sample includes 1,128 U.S. domestic ETFs with 72,504 fund-month observations.

3.2. Methodologies to infer ETF investors' preferences of asset pricing models

The identification of which model is used by ETF investors to evaluate fund performance requires two steps. First, one needs to identify the abnormal performance of ETFs relative to a model. Second, we should be able to examine investors' capital allocation decisions in response to the profitable investment opportunities.

3.3. Measurements of fund flows

Following prior literature, (monthly) fund flow is measured by the percentage growth in net value of total assets under management in a month compared to the previous month, assuming that capitals flowing into or out of a fund occurs at the month end. For an ETF p in a month t, its fund flow is given by

$$F_t^p = \frac{\mathrm{TNA}_t^p}{\mathrm{TNA}_{t-1}^p} - (1 + R_t^p), \tag{1}$$

where TNA_t^p and TNA_{t-1}^p are the fund's total net value of assets at the end of month t and $t-1$, and R_t^p denotes the return on fund p in month t. To reduce the impact of outliers, fund-month observations with monthly fund flow lower than -90% or higher than 1000% are removed.

3.3.1. Measurements of ETF performance relative to competing models

Investors evaluate ETFs in relative to some benchmark to identify profitable investment opportunities. Unsophisticated investors with limited information processing abilities may simply use historical gross fund return to measure fund performance. BHO (2016) argue that sophisticated investors should attend to funds' exposures to all factors which have implications for cross-sectional asset returns, no matter the factors arise from fund risks, investors' behavior biases, or market frictions. For example, prior studies have documented the differences in average returns among stocks with different firm size and valuation ratios and for firms in different industries. An ETF tracking a particular segment of the equity market (like small stocks, value firms, or firms of a specific industry) may have exposures to the size, value, and industry factors different from those of an ETF tracking the overall aggregate market (like the S&P 500 index). Sophisticated investors aware of the pattern of cross-sectional variations in equity returns should strip out the components of fund returns attributable to the differences in factor loadings and industry tilts when assessing fund performance.

This paper considers nine prominent competing models discussed early, in which ETF investors may use to judge if a fund can deliver a positive alpha: (i) the market-adjusted return (MAR), which assumes that investors use the aggregate market portfolio[3] as the benchmark to assess the performance of ETFs; (ii) the standard capital asset pricing model (CAPM); (iii) the three-factor model of Fama and French (1993) (FF3), which augments the CAPM with the size (SMB) and value (HML) factors; (iv) the four-factor model of Carhart (1997) (Car4) that adds the momentum factor (UMD) to the FF3; (v) a seven-factor model that adds three industry factors[4] as in Pástor and Stambaugh (2002) (PS7); (vi) the five-factor model of Fama and French (2015) (FF5), which adds the profitability (RMW) and investment (CMA) factors to the FF3; (vii) the four-factor model of Hou, Xue, and Zhang (2015) motivated by the q-theory; (viii) the four-factor mispricing model of Stambaugh and Yuan (2017) (SY4), which supplements the CAPM with a size factor

[3] We use the CRSP value-weighted portfolio as a proxy for the market portfolio.
[4] Following Pástor, and Stambaugh (2002), we retrieve the first three principal components of the Fama and French 17 industry portfolios as the three industry factors. Details of construction of industry factors are provided in the Online Appendix.

and two mispricing factors MGMT and PERM; and (ix) the three-factor behavior model of Daniel, Hirshleifer and Sun (2020) (DHS3), which augments the market factor with PEAD and FIN.

We use historical cumulative abnormal returns of ETFs relative to different models to assess fund performance. Using the FF3 as an example, for an ETF p in month t, we regress fund monthly returns in excess of the risk-free rate on contemporaneous returns on the three factors in the previous 24 to 60 months (as available) to estimate factor loadings of the ETF

$$R_\tau^p - R_\tau^f = \alpha_t^p + \beta_t^{p,\text{MKT}}\text{MKT}_\tau + \beta_t^{p,\text{SMB}}\text{SMB}_\tau + \beta_t^{p,\text{HML}}\text{HML}_\tau + \epsilon_\tau^p,\quad \text{where } t-60 \le \tau \le t-1, \tag{2}$$

where R_τ^p is the return on ETF p in month τ and R_τ^f is the risk-free rate, MKT_τ, SMB_τ and HML_τ are return on the market, size, and value factors in month τ. $\beta_t^{p,\text{MKT}}$, $\beta_t^{p,\text{SMB}}$, and $\beta_t^{p,\text{HML}}$ represent the fund's exposures to the three factors estimated using information available before month t. Betas of ETFs are estimated in a rolling window of data to capture time-varying factor loadings of funds. The estimated abnormal return (alpha) on fund p relative to the FF3 in month t is the difference between the fund's realized excess return and the components attributable to the fund's exposures to the three risk factors

$$\hat{\alpha}_t^{p,\text{FF3}} = (R_t^p - R_t^f) - \hat{\beta}_t^{p,\text{MKT}}\text{MKT}_t - \hat{\beta}_t^{p,\text{SMB}}\text{SMB}_t - \hat{\beta}_t^{p,\text{HML}}\text{HML}_t. \tag{3}$$

Investors are assumed to evaluate ETFs by its historical abnormal returns, but their evaluation horizon is not clear. On one hand, a short evaluation horizon can provide more recent and relevant information about future fund performance. In a perfectly efficient market, investors should react to arriving new information immediately. On the other hand, a long evaluation horizon can help reduce noises in prices and returns in the short run. Investors need to trade off relevance and reliance when making asset allocation decisions based on past fund performance. Besides, market frictions and investors' limited attention can result in delayed responses to fund performance. To ensure that the empirical results are robust to the assumption of the assessment horizon, we consider a variety of evaluation periods from 3 months to 36 months.

Following BHO (2016), we assume that investors put greater weights on recent fund alphas than on distant fund alphas when allocating capitals. For each valuation horizon from $t - T$ to $t - 1$, the parameter (λ) of the exponential decay model for fund alphas is estimated by maximizing the log likelihood function. In month t for an ETF p, we calculate a weighted average of fund alphas relative to the FF3 during each evaluation horizon T as follows:

$$\text{Alpha}_t^{p,\text{FF3}} = \frac{\sum_{\tau=1}^{T} e^{-\hat{\lambda}(\tau-1)} \hat{\alpha}_{t-\tau}^{p,\text{FF3}}}{\sum_{\tau=1}^{T} e^{-\hat{\lambda}(\tau-1)}}, \tag{4}$$

where $\hat{\alpha}_{t-\tau}^{p,\text{FF3}}$ denotes the estimated abnormal return of the ETF relative to the FF3 in month $t - \tau$ from Eq. (3) and $\hat{\lambda}$ is the estimated rate of decay. Alphas relative to the other eight models can be calculated in the same vein. We then use the following specification to estimate the fund flow sensitivity to fund performance (b):

$$F_t^p = a + b \times \text{Alpha}_t^{p,\text{Model}} + cX_{t-\tau}^p + \theta_t + \varepsilon_t^p. \tag{5}$$

Where $\text{Alpha}_t^{p,\text{Model}}$ is the performance of fund *p* relative a particular asset pricing model at month *t* as described in Eq. (4) and X denotes the vector of lagged control variables that may affect future fund flows. We control for the effects of lagged monthly fund flows from month from $t - T$ to $t - 1$, (natural logarithm of) TNA and fund age in the previous month $t - 1$, total volatility measured by standard deviation of monthly fund returns from month $t - 12$ to $t - 1$, fund categories, and the time fixed effect θ_t. The coefficient *b* represents the relation between fund flows and lagged market excess return on ETFs.

3.3.2. Methods to test the relation between fund alphas and fund flows

ETF investors are assumed to allocate capitals based on the weighted-average cumulative fund alphas computed in Eq. (4). Funds with positive alphas indicate superior performance and should attract net capital inflows. Similarly, funds with negative alphas indicate underperformance and should be penalized with capital withdrawals. A stronger association between future fund flows and historical fund alphas relative to a model suggest greater preferences of investors for the model. We use several methods to test and compare the responses of capital allocation decisions to fund alphas relative to the nine competing models.

First, we investigate sensitivities of fund flows to factor-adjusted returns for each asset pricing model. For a given model, we follow BvB (2016) to regress signed fund flows ($\phi(F_t^P)$) on signed alphas ($\phi\left(\text{Alpha}_t^p\right)$) to estimate the beta coefficient of fund flows on fund alphas relative to the model ($\beta_{F\alpha}$):

$$\beta_{F\alpha} = \frac{\text{Cov}\left[\phi(F_t^P), \phi\left(\text{Alpha}_t^p\right)\right]}{\text{Var}\left[\phi\left(\text{Alpha}_t^p\right)\right]}, \tag{6}$$

$$\text{where } \phi(x) = \begin{cases} \dfrac{x}{|x|}, x \neq 0 \\ 0, \quad x = 0. \end{cases}$$

The beta coefficient of fund flows on model-adjusted returns measures the extent to which investors consider the model's factors when evaluating fund performance. Under the extreme case where all ETF investors use a particular model to identify profitable investment opportunities and allocate capitals to the funds accordingly, signs of fund flows would be perfectly correlated with signs of model-adjusted returns with a beta coefficient of 1. In the other extreme, if investors' capital allocation decisions and fund alphas relative to a model are not correlated, the regression coefficient should be zero. A higher beta coefficient in Eq. (6) indicates stronger responses of fund flows to model-adjusted returns.

Second, we use the multivariate regression approach to test the marginal predictive power of the nine competing models for future fund flows. Specifically, ETF fund flows are regressed on alphas relative to the nine tested models along with a vector of control variables (X_t^p) and time fixed effects (θ_t) as in Eq. (5):

$$\begin{aligned} F_t^p = \lambda_0 &+ \lambda_1 \text{ MAR}_t^p + \lambda_2 \text{ Alpha}_t^{p,\text{CAPM}} + \lambda_3 \text{Alpha}_t^{p,\text{FF3}} + \lambda_4 \text{Alpha}_t^{p,\text{Car4}} \\ &+ \lambda_5 \text{ Alpha}_t^{p,\text{PS7}} + \lambda_6 \text{ Alpha}_t^{p,\text{FF5}} + \lambda_7 \text{ Alpha}_t^{p,\text{QF4}} \\ &+ \lambda_8 \text{ Alpha}_t^{p,\text{SY4}} + \lambda_9 \text{ Alpha}_t^{p,\text{DHS3}} + \lambda_X X_t^p + \theta_t + \varepsilon_t^p. \end{aligned} \tag{7}$$

If a model has incremental power to predict fund flows, the regression coefficient on the alpha relative to the model should be significantly positive. To address heteroscedasticity over time and to attenuate the effects of outliers, we also consider standardizing raw alphas by the cross-sectional standard deviation of alphas in each month and to use percentage ranking of alphas.

Third, we examine the relative performance for each pair of the nine competing models to predict future fund flows. We run the following panel regression to distinguish the abilities of signs of alphas relative to a model m against a model n to predict signs of fund flows ($\phi\left(F_t^p\right)$)

$$\phi\left(F_t^p\right) = \gamma_0 + \gamma_1 \left[\frac{\phi\left(\text{Alpha}_t^{p,m}\right)}{\text{Var}\left[\phi\left(\text{Alpha}_t^{p,m}\right)\right]} - \frac{\phi\left(\text{Alpha}_t^{p,n}\right)}{\text{Var}\left[\phi\left(\text{Alpha}_t^{p,n}\right)\right]}\right] + e_t^p. \tag{8}$$

A statistically and significantly positive (negative) regression coefficient $\gamma_1$suggests that model m is a more (less) preferable model used by investors to rate fund performance than model n.

Finally, to attenuate the potential non-linear relation between fund flows and past abnormal returns, we conduct horse race tests based on portfolio sorting between each pair of asset pricing model as in BHO (2016). Specifically, using the CAPM and the FF3 as an example, at the end of each month t, ETFs are independently sorted into 5-by-5 portfolios by weighted-average cumulative historical alphas relative to the CAPM or the FF3. Then in month t, percentage of fund flows are regressed on quintile rankings of CAPM-alphas and FF3-alphas

$$F_t^p = \delta_0 + \sum\sum \delta_{ij} D_{ijt}^p + \delta_X X_t^p + \theta_t + e_t^p. \tag{9}$$

where D_{ijt}^{p} is a dummy variable that equals 1 if the fund's CAPM-alpha falls into the i^{th} quintile and the fund's FF3-alpha falls into the j^{th} quintile in month t, and 0 otherwise. The dummy variable for $i = 3$ and $j = 3$ is excluded. The regression coefficients δ_{ij} for $i \neq j$ capture the differences in performance between the two models to predict fund flows. For a particular pair of ranking $i > j$, i.e., a fund's CAPM-alpha has a higher ranking than its FF3-alpha, we compare coefficient δ_{ij} with coefficient δ_{ji} with the same magnitude but reversing orders of rankings. If investors tend to prefer the CAPM over the FF3 when evaluating fund performance, coefficient δ_{ij} would be significantly greater than coefficient δ_{ji}. If, instead, investors do account for the size and value factors when allocating capitals, they would have a stronger reaction towards the ranking based on FF3-alphas, i.e., $\delta_{ji} > \delta_{ij}$.

We calculate the difference in coefficients between the pair of ranking dummies with the same magnitude but reversed orderings ($\delta_{ij} - \delta_{ji}$). The differences in coefficients are summed up and we test the null hypothesis that $\sum(\delta_{ij} - \delta_{ji}) = 0$. A significantly positive (negative) summed difference constitutes evidence of a greater (less) preference among ETF investors for the CAPM over the FF3. We also calculate the proportion of cases with positive regression coefficients and compute the binominal statistics to test the null hypothesis that $prob(\delta_{ij} > \delta_{ij})$ = 50%. The proportion of positive coefficient differences is the sample probability that CAPM-alphas are more correctly to predict fund flows than do FF3-alphas. If investors attend to the market factor only as in the CAPM while ignore the size and value effects included in the FF3, $prob(\delta_{ij} > \delta_{ij})$ should be significantly higher than 50%, and vice versa if investors do account for the size and value effects when identifying profitable investment opportunities.

4. Main Empirical Results

4.1. Descriptive statistics

Table 1 tabulates the descriptive statistics of returns and fund characteristics of the whole ETF sample and ETFs by investment styles or by industry sectors during the sample period from January 2000 to December 2019. The 1,128 well-diversified ETFs earn an average excess return of 0.71% per month with a standard deviation of 5.52%. On average, funds attract net capital inflows of 1.74% per month. Capitals flows into or out of ETFs have substantial cross-sectional variation: the third quartile fund has a positive fund flow of 3.41%, while the first quartile fund has a capital withdrawal of −1.57%. An average ETF has $2.3 billion of TNA under management with a fund age (defined as the time since the fund inception date) of about 82 months. The passively managed ETFs charge investors an average (annual) expense ratio of 0.42%, much lower than the counterpart for actively managed mutual funds reported by prior studies. The mean value of volatility of monthly ETF returns is 4.78%, comparable with that of the overall equity market index in the U.S.

(Insert Table 1 here)

Morningstar Ratings is a popular performance metric among fund investors and can affect their investment decisions. To better understand the sample, we group ETFs into 18 categories, with 9 size and value/growth investment style boxes and 9 industry categories. The sample includes 30,980 observations of equity-style ETFs and 34,894 observations of industry funds.[5] For style funds categorized by market capitalization and P/E ratio, our sample slightly tilts toward large market cap funds and is evenly distributed among value and growth funds. For funds concentrating on industries, ETFs tracking the Basic Material and Science & Technology industries have relatively a larger number of observations. Looking across fund categories reveals that average fund returns, fund flows, and fund characteristics differ across investment styles and

[5] We find that 6,630 fund-month observations cannot be assigned to a Morningstar rating either because the funds have missing Lipper code, or the funds do not fall into any of the 18 categories.

industry concentrations. Style funds earn higher average excess returns (0.83% per month) and attract greater percentage fund flows (1.89% per month) than sector funds with the corresponding values of 0.65% and 1.67%. Style-based funds have much larger average fund size of over $3.7 billion, charge lower fees, and have lower total volatility then do sector funds.

Within style-based ETFs, the mean values of monthly return volatility for funds tracking large and small firms are 3.77% and 4.97%, respectively, suggesting that returns on small-cap funds tend to be more volatile than returns on large-cap funds. Large-value funds attract the highest average fund inflows of 2.31% per month, despite having the lowest average excess return of 0.71% per month. Within sector funds, ETF investors are most inclined to allocate capitals into Science & Technology funds, which are the most profitable with an average excess return of 1.00% per month. To control for potential effects of fund styles on investors' capital allocation decisions, in the main analysis, we control for fund categories in all regression analysis.

Table 2 reports the mean values of (monthly) fund alphas relative to the nine competing models (Panel A), the correlation matrix of model-adjusted returns (Panel B), and the average and standard deviation of ETFs' loadings on factors in different models (Panel C and Panel D). The aggregate sample of 1,128 ETFs on average earns a negative market-adjusted return of -7 bps per month, which suggests that the passively managed funds tend to underperform the aggregate equity index. Adjusting for the market risk magnifies the average underperformance of ETFs (average α_{CAPM} = -0.22%), as the average fund has a market beta slightly greater than 1 (average β_{MKT} = 1.03).

(Insert Table 2 here)

The mean values of alphas relative to all multifactor models are negative, which suggests that the underperformance of ETFs is a general phenomenon regardless of risk adjustments. Alphas relative to different models are positively correlated, with the correlation coefficient ranging from 0.584 between MAR and PS7-alpha to 0.903 between MAR and CAPM-alpha. The high correlations indicate that model-adjusted returns contain common information about fund performance. Looking across fund categories, we find that models can diverge in terms of rankings of fund performance. For example, small funds appear to underperform large funds according to the CAPM, while the average alpha relative to the QF4 of small funds is slightly higher than that of large funds. Such cases of inconsistent rankings suggest that alphas relative to different models also contain unique information despite their high correlations. As a result, we can use such cases to distinguish model performance, which will be discussed later.

The average ETF in the sample has positive betas on the size factor (0.172) and the first industry factor (0.163) and a negative beta on the investment factor of the QF4 (-0.15), suggesting that funds have mild tilts towards small stocks and firms with aggressive capital investment. The magnitudes of average values of betas on other factors are all lower than 0.150, which indicates that a typical ETF does not have obvious tilts towards stocks sorted by other anomaly variables.

There are considerable variations in factor loadings in the pooled sample. The standard deviation of beta coefficients ranges from 0.276 for β_{UMD} to 0.563 for β_{ROE}. Variations in betas across categories make major contributions to high standard deviations. Funds in different categories appear to have different tilts towards several factors. For the market risk, average small and growth ETFs have slightly higher market betas than large and value ETFs. There are greater variations in market betas among sector funds. ETFs focusing on the Science and Technology industry tends to have the highest market beta (1.328) and Utility funds on average have the lowest sensitivity to market movement of 0.558.

Beta coefficients on the size and value factors generally line up with investment styles. Small-cap funds have an average size beta of 0.762, while large-cap funds have an average negative exposure to the size factor (-0.116). Variations in exposures to the value factor among funds with different focuses on size are relatively small, with small-cap funds having a mild tilt towards value stocks and large-cap funds having a mild tilt towards growth stocks. As expected, the mean value of beta on HML is positive (0.342) for value ETFs and it is on average negative (-0.252) for growth ETFs. Value and growth ETFs have similar average exposures to SMB. ETFs sorted by size or value/growth styles have no obvious differences in loadings on the other factors. One exception is that average value versus growth funds have very different ROE-betas (0.306 versus -0.349), suggesting that value funds behave more like stocks with high profitability and growth funds co-move more with unprofitable firms.

Sector funds seem to have different exposures to several non-market factors. For instance, funds tracking the utility industry on average have a negative exposure to SMB and positive exposures to RMW and CMA, suggesting that Utility funds behave more like firms with large cap, high profitability, and conservative investment. In contrast, a typical Science and Technology fund has a positive size beta and negative profitability and investment betas, indicating that stocks in the industry behave more like small and unprofitable firms that invest aggressively.

It is worth noting that beta coefficients also vary across individual ETFs, although the within-category variations are relatively smaller than the variations for the pooled sample. The large variation in funds' loadings on factors both between and within categories indicates that it is not plausible to assume that all funds or funds in the same category have the same risk exposures when investors identify abnormal investment opportunities.

4.2. Tests on the univariate relation between fund flows and fund alphas

We first use the univariate regression approach to test the predictability of alphas on future fund flows for each tested model. Table 3 reports beta coefficients and associated t-staititics of signed fund flows on signed fund performance measured by historical alphas with various horizons, with and without controlling for fund characteristics and time fixed effects. The t-staititics are double-clustered by funds and months to account for residual serial correlations over time for the same fund and cross-sectional correlations among funds in the same month. It can be proved that the regression coefficient is related to the average probability that the sign of fund performance correctly predicts the sign of fund flows:

$$\beta_{F\alpha} = prob\left[\phi\left(F_t^p\right) = 1 | \ \phi\left(\text{ALPHA}_t^p\right) = 1\right] \\ + prob[\phi\left(F_t^p\right) = -1 \ | \ \phi\left(\text{ALPHA}_t^p\right) = -1] - 1. \tag{10}$$

For the easiness of interpretation, we also report $\frac{\beta_{F\alpha}+1}{2}$ (in percentage) as Prob.

(Insert Table 3 here)

The table shows that for all models and all evaluation horizons, the regression coefficients are positive with highly statistical significance. The results indicate that regardless of the way of risk adjustments, historical alphas of ETFs are a strong predictor for future capitals flows into or out of ETFs. The significantly positive fund flow-performance relation is robust to controlling for fund characteristics, although the strength of predictive power shrinks when other potential predictors for future flows are included.

Comparing across models, for all evaluation horizons, historical market-adjusted fund returns *(*MAR) appear to provoke the strongest response of ETF investors, as evidenced by the highest beta coefficients. Except for the longest evaluation horizon of 36 months, the CAPM is the second best-performing model to predict fund flows. Including non-market risk factors in other multifactor models tend to deteriorate the fund flow-alpha relation, as in most cases the regression coefficients are smaller than the counterparts for MARs and CAPM-alphas. Although underperforming the MAR and the CAPM in predicting capital flows, the DHS3 behavioral model performs relatively better than the rest of the models. The DHS3 model even performs better than the CAPM for the three-year evaluation horizon when fund characteristics are included as controls. Among all models, the PS7 model seems to be the least promising one to describe ETF investors' capital allocating decisions.

Looking across assessment horizons, the magnitudes of beta coefficients are comparable for horizons equal to or shorter than 24 months, while they shrink a lot for the longest horizon of 36 months, which puts some weights on abnormal returns in the distant past. Besides, the differences in predictive powers of alphas relative to different models for fund flows appear to diminish when the evaluation horizon gets longer. The results indicate that ETF investors' reactions are more sensitive to more recent than to more distant information.

Despite the high statistically significant relation between historical alphas, we find the magnitudes of regression coefficients are much lower than 1 and no model can correctly predict signed fund flows with probability higher than 60%. The significant yet small beta coefficients indicate that ETF investors do allocate capitals to take advantage of profitable opportunities, but a large fraction of their investment decisions remains unexplained by past fund performance.

4.3. Multivariate linear regression

Results from the univariate regression suggest that the fund flow-performance relation exists for all the tested models. As shown in Table 2, model-adjusted returns can be highly positively correlated. To examine the marginal predictive power of each model for fund flows over other models, we regress monthly percentage of fund flows on alphas relative to the competing models as shown in Eq. (7). Table 4 reports regression coefficients on alphas relative to different models computed over various horizons, associated t-statistics clustered by funds and months, and adjusted R^2. We consider three specifications for the key independent variables of interests: (i) "Raw" refers to cumulative model-adjusted returns in the original format; (ii) "Std" stands for raw alphas standardized by cross-sectional standard deviation of alphas among ETFs in a given month; and (iii) "Rank" is the percentage ranking of raw alphas of funds compared with other funds in the same month. Regardless of dependent variable specifications, regression coefficients on MARs are positive with the highest statistical significance and the largest magnitude at any evaluation horizons. The Car4 also has statistically significant coefficients for all horizons and for all the three versions of alphas, although the marginal effects of Car-alphas on fund flows are much smaller than the marginal effects of MARs. For instance, on average, one percentage increase in MAR computed over the previous three month leads to a 0.627% increase in fund flows in the subsequent month, which more than triple the counterpart for Car (0.187%). Raw CAPM-alphas do not possess significant predictive power for capitals flows at any conventional significance level, whereas standardized and percentage ranking of CAPM-alphas have significantly positive partial coefficients in most cases. Alphas relative to other models generally have minimal and insignificant marginal predictability for fund flows. We find that for all evaluation horizons, historical fund alphas relative to all models and other fund characteristics jointly only explain around 3% of total variation in fund flows. The low adjusted R^2s suggest that investors use criteria other than historical alphas to make capital allocation decisions.

(Insert Table 4 here)

To sum up, results from the multivariate linear regressions show that MARs have the most statistically and economically significant power to predict fund flows incremental to the effects of alphas relative to other factor models. CAPM-alphas and Car4-alphas also exhibit some incremental ability to predict capital allocation decision, while the effects of alphas relative to other models are subsumed by MARs.

4.4. Pairwise comparison of models

4.4.1. Test differences in responses of signs of fund flows to signs of alphas

Results of regressions to test the univariate relation between fund flows and model-adjusted alphas reveal that investors react to fund performance evaluated under different models with different levels of sensitivities. To formally compare the relative abilities of alternative asset pricing models to predict fund flows, we test the differences in regression coefficients of fund flows to alpha relative to each pair of the nine tested models for various horizons.

The (double-clustered) t-statistics of differences in beta coefficients are presented in Table 5. Panel A reports the results of regressions with the differences in signs of alphas normalized by variance for pairs of models. Except for the longest horizon of 36 months, the MAR and CAPM outperform the other multifactor models with high statistical significance. For all horizons, the MAR performs marginally better than the CAPM, and the outperformance seems to strengthen with extending evaluation horizons. For the longest horizon, the MAR significantly outperforms all other models including the CAPM. The null hypothesis of indifferent performance between the CAPM and the FF5, PS7, or QF4 can be rejected with t-statistics greater than 3, while the difference in performance of the CAPM compared with other models (FF3, Car4, SY4, or DHS3) are not statistically significantly distinguishable from zero. Among the multifactor models, the DHS3 seems to be the most promising model to describe capital allocation decisions. The regression coefficients of signs of fund flows on signs of DHS3-alphas are significantly more positive than its counterparts for other models including non-market factors in most cases. The PS7 model appears to deliver the worst performance among all models.

(Insert Table 5 here)

Panel B of Table 5 presents results of pairwise comparisons between models with control variables. Overall, accounting for differences in fund characteristics make it more difficult to distinguish fund performance, as evidenced by the t-statistics with smaller magnitudes. Yet the relative performance of models remains largely unchanged. The MAR and CAPM are the best-performing models, while including non-market factors does not provide improvement and in some cases even significantly weakens the flow-performance relations.

In short, results of the pairwise comparison of beta coefficients in linear regressions to test how fund flows respond to fund performance reveal that the MAR and CAPM best explain investors' capital allocation decisions. ETF investors either prefer to use the market portfolio as the benchmark or only attend to the market risk of funds when evaluating fund performance. Moreover, they fail to account for other well-known factors that are correlated with cross-sectional equity returns.

4.4.2. Pairwise comparison using nonlinear horse race tests

One issue with the above analysis is that the regression approach assumes a linear fund flow-performance relation. To address this issue, we conduct the portfolio sorting-based horse race test between each pair of asset pricing models. Table 6 tabulates results of horse race tests to compare predictive power of model-adjusted returns computed over the intermediate (18 months) horizons.[6] Let $\text{Alpha}_t^{p,\text{row}}$ and $\text{Alpha}_t^{p,\text{column}}$ denote abnormal fund returns relative to the model in a row and the model in a column. Funds are sorted into quintile portfolios by $\text{Alpha}_t^{p,\text{row}}$ and by $\text{Alpha}_t^{p,\text{column}}$ independently, with i (j) being the quintile ranking based on the row (column)

[6] Results of horse race tests on fund performance evaluated over the previous 3, 6, 12, 24, or 36 months are qualitatively similar.

model. Fund flows are regressed on the ranking dummies and control variables as in Eq. (9).

For each pair of models in a row and in a column, we report the sum of the difference (in percentage) between summed differences in regression coefficients on dummies of rankings with the same magnitudes but reversed orderings. The associated t-statistic of the test on the null hypothesis of zero summed coefficient differences is clustered by funds and months. We also report the proportion of positive differences in paired coefficients and p-values of the binominal tests on the null hypothesis that the differences in coefficients are positive in 50%f of all cases.

(Insert Table 6 here)

Results of the horse race tests among models show that the MAR declares victory over all the other models at all evaluation horizons. For an assessment period of 18 months, the difference in summed coefficients on ranking dummies of the MAR over the CAPM is 8.28% with a t-staistic of 3.11, and the differences over the other models are all higher than 13% with t-staistics greater than 7. The MAR wins the CAPM in 80% of the cases with a binominal p-value of 0.011 and wins over all other models in 100% of the cases with p-values lower than 0.0001. Except for the MAR, the CAPM dominates the other models in predicting fund flows with high statistical significance. Among the multifactor model, the behavioral DHS3 model only loses to the MAR and the CAPM, but DHS3-alphas are significantly better predictors for fund flows than other models. The model incorporating industry factors (PS7) is the worst-performing model which loses to all the races. In sum, results of non-linear horse race tests generally confirm results of linear regressions.

5. Additional Analysis and Robustness Tests

5.1. Decomposition of returns

Results of the main analysis reveal that the MAR has the strongest effect on the way that ETF investors allocate capitals flowing into or out of the funds. ETF investors appear to be unsophisticated agents who do not fully account for the parts of returns attributable to common factors related to cross-sectional stock returns when making capital allocation decisions. However, the results do not indicate that investors completely ignore differences in factor loadings among ETFs when evaluating fund performance. In this section, we examine how investors respond to the components of returns related to various factors.

Specifically, using the FF3 as an example, re-arranging Eq. (3), the total excess return on a fund ($R_t^p - R_t^f$) can be decomposed into the abnormal part ($\hat{\alpha}_t^{p,\mathrm{FF3}}$) and the components attributable to the market ($\hat{\beta}_t^{p,\mathrm{MKT}}\mathrm{MKT}_t$), size ($\hat{\beta}_t^{p,\mathrm{SMB}}\mathrm{SMB}_t$), and value ($\hat{\beta}_t^{p,\mathrm{HML}}\mathrm{HML}_t$) factors. A factor-related component is the product of the fund's beta on a factor and the realization of the factor in that month. A weighted-average cumulative return related to each factor can be computed in a similar way to compute a weighted-average alpha under the exponential decay function as described in Eq. (4). Using the market factor as an example, the component of returns attributable to the market risk (MKTRet_t^p) evaluated over the previous T months is given by

$$\mathrm{MKTRet}_t^p = \frac{\sum_{\tau=1}^{T} e^{-\hat{\lambda}(\tau-1)}\left[\hat{\beta}_{t-\tau}^{p,\mathrm{MKT}}\mathrm{MKT}_{t-\tau}\right]}{\sum_{\tau=1}^{T} e^{-\hat{\lambda}(\tau-1)}}. \tag{11}$$

The components of returns related to other factors in other models can be computed in the same way. For a particular asset pricing model, we can then examine investors' responses to different components of fund returns by regressing fund flows on alphas relative to the model and returns related to factors in the model. For instance, the panel regression to test the relations between fund flows and components of returns decomposed according to the FF3 is given by

$$F_t^p = \varphi_0 + \varphi_\alpha \text{Alpha}_t^p + \varphi_{\text{MKT}} \text{MKTRet}_t^p + \varphi_{\text{SMB}} \text{SMBRet}_t^p \\ + \varphi_{\text{HML}} \text{HMLRet}_t^p + \varphi_X X_t^p + \theta_t + e_t^p, \tag{12}$$

where X_t^p denotes the vector of control variables defined previously and θ_t the time fixed effects; φ_α captures the response of fund flows to model-adjusted alpha, and φ_{MKT}, φ_{SMB}, and φ_{HML} measure the effects of components of returns traced to market, size, and value factors on future capital flows. If ETF investors are aware of a fund's market risk and the cross-sectional differences in returns of stocks with different size and growth potential, they should be able to strip out the parts of returns attributable to the fund's exposures to the three factors when evaluating fund performance. In this case, investors would direct capitals into or out of ETFs according to the alpha only, while factor-related returns should not affect their decision making (i.e., $\varphi_\alpha > 0$, $\varphi_{\text{MKT}} = \varphi_{\text{SMB}} = \varphi_{\text{HML}} = 0$). If investors do not fully account for factors when rating funds, say, they ignore the value factor, then the part of return attributable to HML would go into the estimated alpha. In this case, we would detect a relation to the HML-related return (i.e., $\varphi_{\text{HML}} > 0$), and the regression coefficient on the true alpha will be biased towards 0. Thus, the relative strength of sensitivities of fund flows to the HML-related return and to the alpha ($\varphi_{\text{HML}}/\varphi_\alpha$) is an indicator of the degree of the HML factor is ignored by ETF investors when allocating capitals.

We decompose total fund excess returns into alphas and factor-related returns under the eight factor models of interest and run the panel regression for each model. Panel A of Table 7 reports regression coefficients, double clustered t-statistics, and adjusted R^2 for an evaluation horizon of 18 months. The coefficients on alphas relative to all models are positive and highly statistically significant. The results indicate that fund flows do react to profitable investment opportunities and are in line with the results of regressions to test univariate fund flow-alpha relations. Moreover, all factor-related returns including the market-related factor also reliably and positively affect future fund flows. The finding suggests that investors fail to distinguish the parts of performance traced to the co-movements of stock returns when making investment decisions. Whereas the CAPM-alpha has a significantly stronger effect on future fund flows than alphas relative to other multifactor models, the positive coefficient on MKTRet indicates that ETF investors fail to fully attend to the market risk as well.

(Insert Table 7 here)

We also notice that the coefficients on factor-related returns differ across factors. Panel B of Table 7 presents the ratio of the coefficient on alpha relative to a particular model and the coefficient on the component of return related to each factor of the model, and p-values for the test that the ratio is equal to 1. Regardless of the risk adjustments, the market-related return bears the regression coefficient with the lowest magnitude relative to the regression coefficient on alphas. The ratio $\varphi_{\mathrm{MKT}}/\varphi_{\alpha}$ is lower than one third and is significantly lower than 1 with a p-value of 0 after four decimals for all models.

In all cases, the partial effect of returns related to the size factor on fund flows is greater than it is for alphas, and the differences in regression coefficients are significant at the conventional level (except for the size factor in the QF4 with a p-value for the test of $\varphi_{\mathrm{SMB}}/\varphi_{\mathrm{QF4\text{-}\alpha}} = 1$ being 0.091). The coefficient on the value factor return is higher than the coefficient on alpha as well, but the null hypothesis of equal regression coefficients cannot be rejected at the 5% significance level. The coefficient on the momentum factor return is 54.44% of the coefficient on the Car4-alpha, and the difference in coefficients is highly significant with a p-value close to 0, which indicates that investors to some extent attend to the momentum factor. Results for the profitability and investment factor returns are mixed depending on factor constructions. The profitability factor RMW in the FF5 has significantly stronger partial effects on fund flows than the FF5-alpha, while the coefficient of the profitability factor ROE return in the QF4 is indistinguishable from the coefficient on the QF4-alpha. Investment factors do not affect allocation of capitals more than do abnormal returns. The ratio of $\varphi_{\mathrm{CMA}}/\varphi_{\mathrm{FF5}\text{-}\alpha}$ is insignificantly higher than 1, while the ratio $\varphi_{\mathrm{IA}}/\varphi_{\mathrm{QF4}\text{-}\alpha}$ is significantly lower than 1. The results indicate that investors to some extent consider factors related to firm investment, while large ignore factors related to firm profitability.

ETF investors appear to have weaker reactions to parts of returns traced to mispricing- or behavioral-based factors than to alphas. The mispricing factor related to MGMT has a regression coefficient insignificantly higher than the coefficient on the SY4-alpha. The other three factors, including the PERM, the PEAD and the FIN, all have significantly lower coefficients than abnormal returns. The weaker association between fund flow and components of returns attributable to factors in the SY4 and the DHS3 suggest that to some extent investors are able to account for funds' exposures to factors related to mispricing in stocks and to investor cognitive bias. Regarding the industry factors, investors appear to attend to first industry factor, as the coefficient on IndPCA1 is 60% of the coefficient on the PS7-alpha with the ratio of $\varphi_{\mathrm{IndPCA1}}/\varphi_{\mathrm{PS7}\text{-}\alpha}$ being significantly lower than 1 with a p-value close to zero. In contrast, ETF investors appear to ignore industry tilts in the second and third industry factors, as regression coefficients on IndPCA2 and IndPCA3 are both higher than the coefficient on the abnormal return, although the differences in coefficients are not statistically significant.

In sum, the results of the regressions testing the relation between fund flows and decomposed fund returns reveal that ETF investors attend most to the market risk and least to the size factor. In addition, ETF investors also to some extent consider the momentum, mispricing, and behavioral factors when allocating capitals.

5.2. Effects of Morningstar ratings on fund flows

Choosing among ETFs can be a difficult and complex task in light of the large number of funds available. The preceding analysis reveals that ETF investors are unsophisticated agents who fail to fully account market risks and largely ignore exposures to non-market factors when allocating capitals. When evaluating fund performance to make their investment decisions, instead of searching for all relevant information, unsophisticated investors may simply refer to evaluation metrics provided by third parties or to the disclosure of fund performance provided by fund managers. Introduced in 1985, Morningstar Rating has been a popular and widely used tool among fund investors to assist decision making, and fund companies have been using it for advertising. Before June 2002, in each month, Morningstar ranks individual funds based on risk-adjusted returns in the previous three, five, and ten years relative to the aggregate market. However, on and after July 2002, Morningstar has changed its rating algorithm by grouping funds into 18 categories with similar size and investment style or industry concentration and benchmarking the performance of the fund relative to that of peer funds in the same group. Therefore, when ranking funds, Morningstar only considers the market, size, value, and industry factors, while the other factors proposed in newly developed asset pricing models and within-category variations in exposures to the four factors are ignored. The limitations in the evaluating methods may explain why the MAR outperforms the other factor models in affecting ETF investors' investment decisions, if they rely heavily on the ratings released by the third-party institutions when choosing funds.

We follow the methodology of Morningstar to compute Morningstar Rating (details on the construction of fund ratings are provided in the Online Appendix). Panel A of Table 8 presents mean values of fund characteristics and weighted-average historical alphas over the previous 18 months of ETFs sorted by overall Morningstar Ratings. Funds in the 1, 2, 3, 4 and 5 groups correspond to funds with ratings in the ranges $[1, 1.5)$, $[1.5, 2.5)$, $[2.5, 3.5)$, $[3.5, 4.5)$ and $[4.5, 5]$, respectively. Fund size and age seem not to be strongly correlated with ratings. Funds in the middle range of ratings have the largest net assets under management and have the longest history, and funds with the worst and best performance tend to be small funds with short history. Volatility of returns tends to decrease as rating gets higher, as Morningstar penalizes funds with higher risk when evaluating fund performance. Fund flows in the next month increase monotonically from funds with the lowest rating to funds with the highest rating. The positive flow-rating relation indicates that ETFs rated higher by Morningstar attract more capitals from investors. We also find a strong association between overall ratings and cumulative abnormal fund returns relative to all models, but the strength of the association appears to differ across models. Morningstar ratings have a stronger correlation with the models with fewer factors (the MAR, CAPM, or FF3). The higher correlations with the three models are in line with the fact that the third-party institution evaluates the performance of funds in relative to the aggregate market or to peer funds with similar size and investment style.

(Insert Table 8 here)

With the calculated ratings, we first examine the effects of fund ratings on the relative performance of models to describe investors' trading behavior. Specifically, we regress percentage of fund flows in the next month on the percentage rankings of weighted-average alphas over various evaluation horizons and overall Morningstar Rating while controlling for other potential predictors for fund flows. Panel B of Table 8 reports the results of the panel regressions. Accounting for the effects of ratings hardly affect the positively partial effects of model-adjusted returns on fund flows but it does not exert influences on the relative performance of models. The same as the counterparts when ratings are not included, the coefficients on the MAR, CAPM-alpha and Car-alpha are still significantly positive, and the MAR still has the coefficient with the largest magnitude and statistical significance. The results indicate that the incremental predictive power for fund flows attributed to the MAR, CAPM-alpha, and Car4-alpha, and the most superior performance of the MAR over the other models are not driven by the effects of Morningstar fund ratings. We also document a highly significant positive coefficient on Morningstar rating. Controlling for historical model-adjusted returns on ETFs, an increase of one star in Morningstar Ratings would cause about a 0.3% increase in percentage fund flows in the subsequent month. The strongly positive fund flow-rating relation suggests that ETF investors do refer to Morningstar Ratings when allocating capitals.

5.3. Effects of behavioral theories on fund flows

The analysis above reveals that ETF investors are not sophisticated enough to consider well-known patterns of cross-sectional variation in stock returns when making investment decisions. We also find evidence that the behavioral model performs relatively better than the risk-based models, indicating that ETF investors' trading activities may be more correlated with the comovement in stock returns traced to irrational behavior than those attributed to equity risk. The findings are not surprising given the fact that the ETF market is dominated by retail investors who may have very limited abilities to collect and process relevant information to evaluate risk and return of available investment opportunities properly.

The behavioral finance literature has documented that unsophisticated investors are more likely to commit to psychological and cognitive biases, which drive them to evaluate risk in violation of the expected utility theory and affect the pricing of financial assets in the market. Behavioral finance theories incorporating irrational thinking of agents have been proposed to explain asset pricing anomalies. In particular, Tversky and Kahneman (1992) extend their original prospect theory (Kahneman and Tversky, 1979) and develop cumulative prospect theory, which asserts that investors evaluate assets by calculating distorted probability-weighted average values of gains or losses relative to a reference point. The prospect theory captures agents' mental accounting, loss aversion, and overweighting on extreme events. Barberis, Mukherjee, and Wang (2016) show that in an economy where a subset of the overall population of investors behave according to the prospect theory, preferences for assets with higher prospect theory value (PT) would drive up demands for such assets and result in lower expected returns on the assets. More recently, Bordalo, Gennaioli, and Shleifer (2012, 2013) develop the salience theory that incorporates investors' excess focus on salient payoffs compared to those of alternative options. The salience theory (ST) suggests that in equilibrium, investors with salience thinking would have greater demands and ask for lower expected rate of return on stocks with more salient payoffs.

If ETF investors evaluate risk according to prospect theory or exhibit salience thinking, the PT and ST measures may affect their capital allocating decisions. To the extent that the MAR is more correlated with the two behavior measures, the superior predictability of the MAR for fund flows may be driven by the effects of PT and ST. We follow Tversky and Kahneman (1992) and Barberis et al. (2016) to compute the prospect theory value (PT value) and Bordalo et al. (2013) and Cosemans and Frehen (2021) to compute the salience theory value (ST value) of ETFs. Panel A of Table 9 tabulates average characteristics, future one-month percentage of fund flows and historical weighted-average alphas over the previous 18 months of ETFs, sorted by PT and ST measures. At the bottom, the correlations between these two behavior measures and fund flows and fund alphas. Both the PT and ST measures are positively correlated with fund flows: the percentage of fund flows to the ETFs in the next month increases from 0.01% with the lowest PT value to 3.68% with the highest PT value, and it increases from 0.61% with the lowest salient payoff to 2.71% with the most salient payoff. The two behavioral theory measures are also positively correlated with the fund alphas. The MAR and CAPM-alphas have the highest correlations with both PT and ST values. Alphas relative to the behavior DHS3 model also have relatively strong association with the two behavior measures compared to the other multifactor models.

(Insert Table 9 here)

We next use the multivariate regression to investigate the effects of the two behavioral theory measures on the fund flow-performance relation. Specifically, fund flows are regressed on the historical cumulative alphas relative to the nine tested models, the ST and PT measures, and a battery of control variables and time fixed effects defined previously. We consider three versions of the key variables (alphas, and ST and PT values): the raw value, the value normalized by cross-sectional standard deviation, and the percentage ranking of value. Panel B of Table 9 tabulates estimated coefficients on the two behavioral theory measures and model-adjusted returns computed over various horizons. The coefficients on all versions of PT and ST values are positive and highly statistically significant, and the strong predictive power of the two behavioral theory measures are incremental to alphas measured over all horizons. The economic magnitude of the effect of the PT value is much larger than that of the ST value. If an ETF's ranking of PT value increases by 1%, the fund would attract 0.045% more net capital inflows, while the corresponding value for ST value is only 0.004%. Accounting for the effects of the two behavioral theory measures weakens the positive fund flow-performance relation, especially for the MAR. For instance, the coefficient on the raw cumulative MAR in the previous 3 months is reduced by half from 0.627 (t -stat = 9.10) to 0.316 (t-stat = 4.55) after including the PT and ST measures. However, for all horizons, the raw MAR still has the highest and most significant coefficient compared to raw alphas relative to other models. The reduction in the predictive power of the MAR for fund flows is more pronounced when the percentage ranking of alphas, and PT and ST measures are used. For all horizons, the Car4 model supersedes the dominant position of the MAR as the best predictor for future fund returns. At the 18-month horizon, the ranking of the CAPM-alpha has a slightly greater partial effect on fund flows than the ranking of the MAR, and at the 36-month horizon, the coefficient on the MAR becomes insignificant.

In short, we find evidence that investors' capital allocation decisions are affected by the PT and ST values of ETFs. Accounting for the possibility that ETF investors evaluate risk according to prospect theory and excessively focus on salient payoffs weakens the partial effect of the MAR on future fund flows. The results indicate that the superior performance of the MAR to explain capital flows in or out of ETFs can be partially explained by investor irrational behaviors as described by the two behavior-based models.

6. Conclusion

As a convenient and efficient tool to gain exposures to the broad or specific sectors of the market, ETFs have gained remarkedly growing popularity among investors. How investors evaluate risk-adjusted performance of ETFs in helping them to make investment decisions is still mysterious. We shed new light on the question by inferring ETF investors' preferences of risk from fund capital flows.

The main take way from our study is that market-adjusted returns reliably perform better in predicting future fund flows than all other factor models examined. The strongest association between fund flows and the MAR indicates that ETF investors direct capitals simply based on past performance in relative to the market portfolio and invest capitals in funds that outperform the aggregate market and withdraw capitals from funds that underperform the market, regardless of their tilts towards factors. The superior performance of the MAR is not driven by investors' reliance on alternative metrics provided by third-party institutions and is partially attenuated by investors' preferences for assets with higher values under prospect theory and more salient payoffs under silence theory. Overall, the results seem to support the argument that the ETF market is dominated by unsophisticated agents who fail to account for exposures to factors known to be associated with cross-sectional equity returns. Instead, they seem to chase returns out of naïve extrapolation, and their investment decisions are affected by behavior biases not consistent with a rational framework.

Our finding that the MAR outperforms the CAPM in affecting capital flows stands in contrast to the findings documented in the mutual fund or hedge fund market. The inconsistent findings indicate that BvB's (2016) argument that risk preferences in the mutual fund market can be generalized to investors in the other markets, and that the best performance of the CAPM in predicting mutual fund flows suggests that the model is the true asset pricing model may be questionable. The relative underperformance of the CAPM is in line with the failure of the model in describing prices in the stock market. Similar to studies on mutual fund flows, we also find that the CAPM significantly outperform the newly developed multifactor models in affecting future capital flows into or out of the ETF market, which is contrary to the literature's arguments that adding non-market factors can help improve the CAPM. The in-sample better capabilities of the multifactor models than the CAPM to explain cross-sectional asset returns are not robust out of sample.

Except for the unsophisticated agent explanation, we acknowledge that there are other plausible reasons for the outperformance of the MAR over the existing common factor models that future research can test. First, we find that the PS7 model that includes the three industry factors, whose alphas have high estimation errors (Jegadeesh and Mangipudi, 2021), deliver the worst performance in predicting fund flows. The underperformance of a factor model may be attributable to noise in data that cause biases in estimated factor loadings and model-adjusted returns. Using techniques that can estimate beta coefficients more precisely may attenuate the underperformance of a factor model relative to the MAR. Second, we find that even for the best predictor for fund flows, it can only correctly predict fund inflows or outflows with a probability lower than 60%, and model-adjusted returns along with fund characteristics jointly explain a small proportion of total variation in ETF fund flows. As a large fraction of fund flows remain unexplained, it is likely that investors gauge fund performance according to some latent factors not well captured by the existing asset pricing models. We find that the behavior-based model (DHS3) wins the horse race tests against the other multifactor models. The return decomposition analysis shows that ETF investors appear to account for the momentum and short-term and long-term behavior factors more than for the other risk-based factors. The results suggest that behavior factors may be more promising in affecting ETF fund flows. Future studies may develop new models with factors that better capture commonality in investors' psychological biases in the ETF market to better explain investors' capital allocation decisions.

References

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

Agarwal, Vikas, Hanouna, P., Moussawi, R. and Stahel, C.W., 2018, November. Do ETFs increase the commonality in liquidity of underlying stocks? In 28th Annual Conference on Financial Economics and Accounting, Fifth Annual Conference on Financial Market Regulation.

Ang, A., Hodrick, R.J., Xing, Y. and Zhang, X., 2006. The cross-section of volatility and expected returns. Journal of Finance 61(1), 259-299.

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9(1), 3-18.

Barber, B.M., Huang, X. and Odean, T., 2016. Which factors matter to investors? Evidence from mutual fund flows. Review of Financial Studies 29(10), 2600-2642.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2016. Prospect theory and stock returns: An empirical test. Review of Financial Studies 29(11), 3068-3107.

Barberis, Nicholas, and Andrei Shleifer, 2003. Style investing. Journal of financial Economics 68(2), 161-199.

Basu, S., 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. Journal of Finance 32, 663-682.

Ben-David, I., Franzoni, F. and Moussawi, R., 2018. Do ETFs increase volatility? Journal of Finance 73(6), 2471-2535.

Berk, J.B., and Van Binsbergen, J.H., 2016. Assessing asset pricing models using revealed preference. Journal of Financial Economics 119(1), 1-23.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2012. Salience theory of choice under risk. Quarterly Journal of Economics 127, 1243-1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2013. Salience and asset prices. American Economic Review 103, 623–628.

Brown, D.C., Davies, S. W., and Ringgenberg, M.C., 2021. ETF arbitrage, non-fundamental demand, and return predictability. Review of Finance 25(4), 937-972.

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

Chevalier, J. and Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. Journal of Political Economy 105(6), 1167-1200.

Clifford, C.P., Fulkerson, J.A., and Jordan, B.D., 2014. What drives ETF flows? Financial Review 49(3), 619-642.

Cooper, M.J., Gulen, H. and Schill, M.J., 2008. Asset growth and the cross-section of stock returns. Journal of Finance 63(4), 1609-1651.

Cosemans, Mathijs and Rik Frehen, 2021. Salience theory and stock prices: Empirical evidence. Journal of Financial Economics 140(2), 460-483.

Da, Z. and Shive, S., 2018. Exchange traded funds and asset return correlations. European Financial Management 24(1), 136-168.

Del Guercio, D. and Tkac, P.A., 2002. The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds. Journal of Financial and Quantitative Analysis 37(4), 523-557.

Evans, R.B. and Sun, Y., 2021. Models or stars: The role of asset pricing models and heuristics in investor risk adjustment. Review of Financial Studies 34(1), 67-107.

Fama, Eugene F., and Kenneth R. French, 1992. The cross-section of expected stock returns. Journal of Finance 47(2), 427-465.

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

Fama, Eugene F., and Kenneth R. French, 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, Eugene F., and Kenneth R. French, 2016. Dissecting anomalies with a five-factor model. Review of Financial Studies 29(1), 69-103.

Fung, W., Hsieh, D.A., Naik, N.Y. and Ramadorai, T., 2008. Hedge funds: Performance, risk, and capital formation. Journal of Finance 63(4), 1777-1803.

Gastineau, G.L., 2010. The Exchange-traded Funds Manual (Vol. 186). John Wiley & Sons.

Gervais, S. and Odean, T., 2001. Learning to be overconfident. Review of Financial Studies 14(1), 1-27.

Glosten, L., Nallareddy, S., and Zou, Y., 2021. ETF activity and informational efficiency of underlying securities. Management Science 67(1), 22-47.

Gruber, M.J., 1996. Another puzzle: The growth in actively managed mutual funds. Journal of Finance, 51(3), pp. 783-810

He, J., and Ng, L.K., 1994. Economic forces, fundamental variables, and equity returns. Journal of Business, 599-609.

Hirshleifer, D., Lim, S.S., and Teoh, S.H., 2011. Limited investor attention and stock market misreactions to accounting information. Review of Asset Pricing Studies 1(1), 35-73.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28(3), 650-705.

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

Israeli, D., Lee, C. and Sridharan, S.A., 2017. Is there a dark side to exchange traded funds? An information perspective. Review of Accounting Studies 22(3), 1048-1083.

Jegadeesh, N. and Mangipudi, C.S., 2021. What do fund flows reveal about asset pricing models and investor sophistication? Review of Financial Studies 34(1), 108-148.

Kahneman, Daniel, and Amos *Tversky*, 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 363–391.

Koijen, R.S., 2014. The cross-section of managerial ability, incentives, and risk preferences. Journal of Finance 69(3), 1051-1098.

Krause, T., Ehsani, S. and Lien, D., 2014. Exchange-traded funds, liquidity, and volatility. Applied Financial Economics 24(24), 1617-1630.

Lettau, M. and Madhavan, A., 2018. Exchange-traded funds 101 for economists. Journal of Economic Perspectives 32(1), 135-54.

Loughran, T. and Ritter, J.R., 1995. The new issues puzzle. Journal of Finance, 50(1), pp.23-51.

Markowitz, H., 1952. The utility of wealth. Journal of Political Economy 60(2), 151-158.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics, 108(1), pp.1-28.

Pástor, Ľ. and Robert F. Stambaugh, 2002. Investing in equity mutual funds. Journal of Financial Economics 63(3), 351-380.

Pontiff, J., and Woodgate, A., 2008. Share issuance and cross-sectional returns. Journal of Finance 63, 921–945.

Rosenberg, B., Reid, K. and Lanstein, R., 1985. Persuasive evidence of market inefficiency. Journal of Portfolio Management 11(3), 9-16.

Sharpe, William F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance 19(3), 425-442.

Sirri, E.R. and Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance, 53(5), 1589-1622.

Stambaugh, Robert F., and Yu Yuan, 2017. Mispricing factors. Review of Financial Studies 30(4), 1270-1315.

Tversky, Amos, and Daniel Kahneman, 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5(4), 297-323.

Table 1: Descriptive statistics of fund characteristics

Table 1 presents descriptive statistics of 1,128 U.S. domestic equity exchange-traded funds (ETFs) during the sample period from January 2000 throughout December 2019. #Obs is the number of fund-month observations. Mean is average value. StdDev is standard deviation. p25, p50 and p75 are the first, median, and third quartiles. Excess return is monthly fund return in excess of the risk-free rate. Fund flow is the percentage change in month-end net value of assets under management of ETFs net of fund return. TNA is total net value of assets (in $ million). ExpRatio is annual expense ratio. Age is the number of months since fund inception date. Volatility is standard deviation of monthly fund return in the previous 12 months. Statistics of excess returns, fund flows, expense ratios, and volatility of returns are reported in percentage. Results are reported for the aggregate sample and for 18 narrow categories of ETFs, including nine market-cap and value/growth investment styles groups and nine industry-specific groups. ETFs that do not fall into the 18 categories are put in the "Others" group.

Category	#Obs	Excess Return					Fund Flow					TNA	ExpRatio	Age	Volatility
		Mean	StdDev	p25	p50	p75	Mean	StdDev	p25	p50	p75	Mean	Mean	Mean	Mean
ALL ETFs	72,504	0.71	5.52	−1.88	1.13	3.65	1.74	12.33	−1.57	0.11	3.41	2,296.16	0.42	82.23	4.78
By Investment Styles															
All Style-based ETFs	30,980	0.83	4.63	−1.34	1.28	3.45	1.89	10.55	−0.80	0.25	3.17	3,728.96	0.38	83.65	4.23
By Market Capitalizations															
Large ETF	9,349	0.85	4.12	−1.18	1.28	3.21	2.02	9.93	−0.62	0.48	3.27	7,378.57	0.26	85.80	3.77
Mid ETF	13,676	0.81	4.51	−1.23	1.25	3.37	1.73	10.74	−0.94	0.17	3.11	2,273.71	0.39	79.73	4.12
Small ETF	7,955	0.85	5.34	−1.90	1.34	3.98	2.00	10.91	−0.85	0.17	3.17	1,941.64	0.34	87.86	4.97
By P/E Ratio															
Value ETF	7,460	0.82	4.72	−1.39	1.26	3.35	1.98	10.59	−0.75	0.28	3.03	1,933.16	0.39	86.63	4.30
Blend ETF	15,676	0.82	4.48	−1.23	1.28	3.34	2.02	10.79	−0.73	0.30	3.36	5,147.73	0.32	80.46	4.14
Growth ETF	7,844	0.86	4.81	−1.45	1.30	3.71	1.53	10.01	−1.03	0.13	2.98	2,601.49	0.36	87.19	4.35
By Market Cap and P/E Ratio															
Large-Value ETF	2,188	0.72	4.20	−1.35	1.23	3.10	2.32	10.22	−0.49	0.71	3.14	2,959.18	0.26	83.35	3.73
Large-Blend ETF	4,812	0.89	3.86	−1.04	1.33	3.10	1.94	10.54	−0.69	0.28	3.39	10,075.25	0.25	84.95	3.68
Large-Growth ETF	2,349	0.88	4.53	−1.22	1.25	3.47	1.90	8.20	−0.55	0.68	3.19	5,970.81	0.27	89.82	4.00
Mid-Value ETF	3,130	0.80	4.56	−1.17	1.24	3.22	1.76	9.67	−0.93	0.24	3.17	1,690.96	0.40	80.15	4.21
Mid-Blend ETF	6,916	0.80	4.33	−1.07	1.24	3.25	2.03	11.40	−0.75	0.28	3.28	3,148.26	0.37	78.92	3.99
Mid-Growth ETF	3,630	0.83	4.80	−1.52	1.32	3.68	1.12	10.29	−1.43	0.03	2.74	1,109.97	0.44	80.91	4.28
Small-Value ETF	2,142	0.95	5.41	−1.91	1.33	3.96	1.94	12.12	−0.88	0.08	2.54	1,239.00	0.33	99.46	5.03
Small-Blend ETF	3,948	0.76	5.37	−1.97	1.35	3.92	2.10	9.93	−0.77	0.35	3.47	2,644.46	0.35	77.67	4.95
Small-Growth ETF	1,865	0.91	5.19	−1.72	1.33	4.16	1.86	11.40	−1.04	0.06	2.96	1,260.82	0.32	96.13	4.94
By Sectors															
All Sectors	34,894	0.65	6.20	−2.44	1.04	4.05	1.67	13.99	−2.61	−0.02	3.74	983.99	0.48	83.19	5.36
Basic Materials Funds	8,395	0.23	7.80	−3.98	0.50	4.50	1.42	12.05	−2.61	−0.03	3.42	904.08	0.52	78.41	7.02
Financial Services Funds	3,869	0.58	6.13	−2.39	1.25	3.84	1.97	15.56	−3.11	−0.08	4.06	1,101.10	0.45	84.60	5.13
Health & Biotechnology Funds	3,622	0.93	5.62	−2.15	1.29	4.35	1.55	11.97	−2.36	0.03	3.87	1,200.46	0.46	84.00	5.36
Real Estate Funds	3,439	0.82	5.51	−1.68	1.09	3.75	1.70	11.16	−1.64	0.12	3.33	1,602.51	0.44	81.69	4.35
Science & Technology Funds	4,478	1.00	6.22	−2.10	1.47	4.71	2.36	14.48	−2.15	0.18	4.35	1,117.61	0.51	83.70	5.53
Telecommunication Funds	1,277	0.64	5.72	−2.40	0.75	3.64	1.33	14.23	−3.14	−0.02	4.04	290.44	0.46	91.88	4.83
Utility Funds	1,955	0.56	4.16	−1.85	1.04	3.40	1.58	16.48	−2.43	−0.08	2.86	769.55	0.45	84.90	3.70
Consumer Goods Funds	1,891	0.75	4.72	−1.38	1.03	3.33	1.60	13.81	−2.63	−0.06	3.72	881.20	0.47	88.16	3.84
Consumer Services Funds	2,224	0.89	5.25	−1.88	1.11	4.15	1.81	20.36	−4.76	−0.13	5.45	709.47	0.46	90.46	4.53
Industrials Funds	3,744	0.73	5.44	−2.07	1.04	3.86	1.26	13.81	−2.99	−0.17	3.34	668.33	0.51	81.78	4.91
Others															
All Others	6,630	0.43	5.52	−1.84	0.76	2.88	1.42	10.54	−1.51	0.15	3.39	2,507.13	0.50	70.51	4.34

Table 2: Summary statistics of model-adjusted returns and factor loadings

The table presents descriptive statistics of alphas and factor loadings of ETFs for the aggregate sample and by fund categories. Beta coefficients are estimated by regressing monthly excess fund returns on factors in a model using data in the previous 24 to 60 months (as available). MKT is the market excess return as in the CAPM. SMB and HML are the size and value factors as in the Fama and French (1993) three-factor model (FF3). UMD is the momentum factor as in the Carhart (1997) four-factor model (Car4). RMW and CMA are the profitability and investment factors as in the Fama and French (2015) five-factor model (FF5). ROE and IA are the profitability and investment factors as in the Hou, Xue, and Zhang (2015) four-factor model (QF4). MGMT and PERM are the mispricing factors related to corporate management and firm performance as in the Stambaugh and Yuan (2017) mispricing four-factor model (SY4). FIN and PEAD are the long-term and short-term behavior factors as in the Daniel, Hirshleifer, and Sun (2020) behavioral three-factor model (DHS3). IndPCA1, IndPCA2, and IndPCA3 are the three industry factors retrieved from the industry portfolio returns as in the Pástor and Stambaugh (2002) seven-factor model (PS7). Alpha relative to a model is computed as the difference between excess fund return and the sum of the products of estimated factor loadings and realized factor returns. Panel A reports mean values of (monthly) abnormal returns on ETFs (in percentage) relative to the nine tested models. Panel B tabulates the correlation matrix between pairs of alphas. Panels C and D present means and standard deviations of factor loadings.

Panel A: Average monthly fund alphas

Category	MAR	α_{CAPM}	α_{FF3}	α_{Car4}	α_{PS7}	α_{FF5}	α_{QF4}	α_{SY4}	α_{DHS3}
All ETFs	−0.07	−0.22	−0.20	−0.16	−0.08	−0.18	−0.09	−0.12	−0.14
By Investment Styles									
All Style-based ETFs	0.06	−0.11	−0.05	−0.04	−0.04	−0.06	−0.04	−0.05	−0.08
By Market Capitalizations									
Large ETF	0.05	−0.01	−0.02	−0.02	−0.03	−0.04	−0.05	−0.04	−0.03
Mid ETF	0.06	−0.09	−0.07	−0.06	−0.05	−0.08	−0.04	−0.05	−0.06
Small ETF	0.05	−0.25	−0.04	−0.03	−0.02	−0.06	−0.03	−0.06	−0.18
By P/E Ratio									
Value ETF	0.02	−0.18	−0.04	−0.03	−0.06	−0.09	−0.09	−0.04	−0.17
Blend ETF	0.04	−0.10	−0.03	−0.03	−0.02	−0.05	−0.05	−0.06	−0.08
Growth ETF	0.11	−0.04	−0.07	−0.08	−0.05	−0.05	0.02	−0.05	0.01
By Market cap and P/E ratio									
Large-Value ETF	−0.02	−0.11	−0.03	−0.04	−0.04	−0.06	−0.11	−0.03	−0.14
Large-Blend ETF	0.05	0.00	−0.02	−0.01	−0.03	−0.04	−0.05	−0.05	−0.02
Large-Growth ETF	0.11	0.07	−0.03	−0.03	−0.05	−0.02	0.02	−0.03	0.08
Mid-Value ETF	0.00	−0.18	−0.06	−0.03	−0.07	−0.11	−0.11	−0.05	−0.15
Mid-Blend ETF	0.07	−0.07	−0.04	−0.04	−0.02	−0.06	−0.04	−0.04	−0.04
Mid-Growth ETF	0.10	−0.05	−0.13	−0.13	−0.07	−0.10	0.01	−0.08	0.00
Small-Value ETF	0.09	−0.26	−0.03	−0.01	−0.06	−0.10	−0.07	−0.03	−0.21
Small-Blend ETF	−0.01	−0.30	−0.05	−0.03	−0.01	−0.07	−0.06	−0.09	−0.22
Small-Growth ETF	0.15	−0.13	−0.02	−0.04	0.00	−0.01	0.05	−0.04	−0.06
By Sectors									
All Sectors	−0.10	−0.31	−0.32	−0.24	−0.11	−0.28	−0.13	−0.16	−0.19
Basic Materials Funds	−0.57	−0.88	−1.01	−0.71	−0.16	−1.07	−0.44	−0.56	−0.52
Financial Services Funds	−0.12	−0.45	−0.10	−0.12	−0.17	0.04	−0.33	−0.06	−0.43
Health & Biotechnology Funds	0.27	0.12	0.11	0.04	−0.10	0.21	0.23	0.24	0.25
Real Estate Funds	−0.10	−0.23	−0.18	−0.15	−0.02	−0.10	−0.09	−0.07	−0.15
Science & Technology Funds	0.31	0.05	−0.09	0.00	−0.01	0.11	0.21	0.05	0.16
Telecommunication Funds	−0.07	−0.20	−0.24	−0.16	−0.02	−0.18	0.07	−0.15	−0.12
Utility Funds	−0.20	0.02	−0.17	−0.18	−0.06	−0.20	0.05	−0.10	−0.03
Consumer Goods Funds	0.07	0.11	0.11	0.01	−0.13	0.04	−0.09	−0.17	−0.06
Consumer Services Funds	0.13	−0.10	−0.05	−0.07	−0.19	−0.08	−0.06	−0.05	−0.17
Industrials Funds	−0.05	−0.33	−0.28	−0.23	−0.10	−0.33	−0.19	−0.20	−0.25
Others									
All Others	−0.51	−0.27	−0.28	−0.23	−0.09	−0.23	−0.17	−0.19	−0.22

Table 2: Summary statistics of model-adjusted returns and factor loadings – continued

Panel B: Correlation matrix between alphas

	α_{CAPM}	α_{FF3}	α_{Car4}	α_{PS7}	α_{FF5}	α_{QF4}	α_{SY4}	α_{DHS3}
MAR	0.903	0.791	0.749	0.584	0.733	0.759	0.762	0.811
α_{CAPM}		0.859	0.807	0.624	0.785	0.818	0.819	0.888
α_{FF3}			0.926	0.695	0.920	0.832	0.881	0.799
α_{CAR4}				0.740	0.849	0.831	0.855	0.794
α_{PS7}					0.643	0.655	0.682	0.623
α_{FF5}						0.824	0.841	0.749
α_{QF4}							0.857	0.815
α_{SY4}								0.819

Panel C: Mean values of factor loadings

Category	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}	β_{RWM}	β_{CMA}	β_{ROE}	β_{IA}	β_{MGMT}	β_{PERM}	β_{FIN}	β_{PEAD}	$\beta_{IndPCA1}$	$\beta_{IndPCA2}$	$\beta_{IndPCA3}$
All ETFs	1.026	0.172	0.009	−0.059	0.007	−0.050	−0.098	−0.151	−0.126	−0.051	−0.113	−0.110	0.163	−0.029	−0.007
By Investment Styles															
All Style-based ETFs	1.054	0.249	0.053	−0.008	0.059	−0.011	0.001	−0.036	0.010	−0.012	−0.042	−0.032	0.021	0.083	0.007
By Market capitalizations															
Large ETF	0.955	−0.116	−0.019	−0.003	0.054	−0.002	0.004	0.037	0.005	0.012	0.040	−0.004	−0.009	−0.003	0.004
Mid ETF	1.042	0.201	0.000	−0.015	0.049	−0.006	−0.046	−0.040	−0.045	0.001	−0.058	−0.039	0.050	0.136	0.008
Small ETF	1.191	0.762	0.229	0.000	0.082	−0.031	0.079	−0.115	0.111	−0.062	−0.109	−0.054	0.006	0.091	0.009
By P/E Ratio															
Value ETF	1.042	0.217	0.342	−0.072	0.115	0.115	0.306	−0.122	0.196	−0.142	0.082	−0.135	0.019	0.069	−0.004
Blend ETF	1.035	0.234	0.068	−0.004	0.068	0.007	0.032	−0.021	0.033	−0.010	−0.026	−0.036	0.016	0.062	0.008
Growth ETF	1.104	0.310	−0.252	0.047	−0.013	−0.168	−0.349	0.018	−0.212	0.107	−0.190	0.073	0.033	0.137	0.016
By Market cap and P/E ratio															
Large-Value ETF	0.935	−0.145	0.289	−0.021	0.028	0.114	0.296	−0.033	0.189	−0.128	0.125	−0.082	−0.008	−0.088	−0.004
Large-Blend ETF	0.936	−0.130	−0.003	−0.008	0.071	0.018	0.035	0.042	0.034	0.009	0.062	−0.004	−0.019	−0.001	0.003
Large-Growth ETF	1.015	−0.059	−0.338	0.024	0.042	−0.153	−0.330	0.092	−0.226	0.147	−0.084	0.069	0.012	0.074	0.017
Mid-Value ETF	1.016	0.127	0.292	−0.101	0.140	0.141	0.286	−0.143	0.150	−0.135	0.100	−0.167	0.040	0.130	−0.009
Mid-Blend ETF	1.023	0.180	0.021	−0.007	0.060	0.020	−0.004	−0.014	−0.013	0.008	−0.037	−0.037	0.047	0.099	0.005
Mid-Growth ETF	1.101	0.302	−0.291	0.043	−0.052	−0.183	−0.412	−0.001	−0.273	0.104	−0.234	0.067	0.065	0.213	0.027
Small-Value ETF	1.191	0.718	0.470	−0.083	0.167	0.078	0.345	−0.185	0.270	−0.169	0.011	−0.145	0.018	0.142	0.004
Small-Blend ETF	1.176	0.772	0.238	0.005	0.078	−0.029	0.091	−0.112	0.114	−0.063	−0.114	−0.073	0.004	0.073	0.019
Small-Growth ETF	1.224	0.791	−0.068	0.085	−0.008	-0.160	−0.251	−0.039	−0.076	0.062	−0.238	0.091	−0.005	0.068	−0.008
By Sectors															
All Sectors	1.075	0.160	−0.032	0.101	−0.054	−0.090	−0.195	−0.261	−0.246	−0.090	−0.189	−0.171	0.267	−0.101	−0.009
Basic Materials Funds	1.164	0.185	−0.016	−0.252	0.097	−0.235	−0.428	−0.576	−0.746	−0.168	−0.485	−0.399	1.196	−0.291	−0.139
Financial Services Funds	1.143	0.063	0.637	−0.056	−0.297	−0.222	0.394	−0.155	0.306	−0.397	0.110	−0.230	−0.200	0.016	0.021
Health & Biotech Funds	0.948	0.310	−0.410	0.008	−0.584	0.152	−0.402	−0.186	−0.336	−0.055	−0.401	0.037	−0.205	−0.569	0.048
Real Estate Funds	0.955	0.046	0.112	−0.104	0.154	−0.084	−0.042	−0.318	−0.223	−0.128	0.044	−0.301	0.130	−0.240	0.051
Science & Tech Funds	1.328	0.330	−0.476	−0.091	−0.345	−0.462	−0.674	−0.260	−0.410	0.047	−0.393	0.011	−0.019	0.076	0.116
Telecommunication Funds	1.101	−0.071	−0.282	−0.148	−0.129	0.212	−0.180	−0.290	−0.045	−0.010	−0.175	−0.095	−0.039	−0.208	0.087
Utility Funds	0.558	−0.214	−0.098	0.086	0.188	0.327	−0.048	−0.002	−0.043	0.150	0.002	0.075	0.225	−0.772	−0.005
Consumer Goods Funds	0.725	−0.125	0.015	0.047	0.259	0.370	0.242	0.129	0.279	0.144	0.267	−0.002	−0.188	−0.061	−0.145
Consumer Services Funds	1.115	0.366	0.024	−0.078	0.252	0.002	−0.025	−0.085	0.053	0.003	0.068	−0.145	−0.137	0.809	−0.068
Industrials Funds	1.153	0.257	0.073	−0.096	0.087	−0.027	−0.056	−0.110	−0.054	−0.085	−0.086	−0.155	0.186	0.399	0.061
Others															
All Others	0.632	−0.122	0.020	−0.074	0.082	−0.021	−0.049	−0.117	−0.130	−0.027	−0.045	−0.153	0.279	−0.171	−0.067

Table 2: Summary statistics of model-adjusted returns and factor loadings – continued

Panel D: Standard deviations of factor loadings

Category	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}	β_{RWM}	β_{CMA}	β_{ROE}	β_{IA}	β_{MGMT}	β_{PERM}	β_{FIN}	β_{PEAD}	$\beta_{IndPCA1}$	$\beta_{IndPCA2}$	$\beta_{IndPCA3}$
All ETFs	0.340	0.435	0.478	0.276	0.434	0.527	0.563	0.433	0.478	0.298	0.379	0.361	0.433	0.520	0.219
By Investment Styles															
All Style-based ETFs	0.184	0.385	0.306	0.150	0.227	0.282	0.350	0.260	0.274	0.173	0.205	0.222	0.142	0.270	0.088
By Market Capitalizations															
Large ETF	0.121	0.148	0.275	0.099	0.173	0.239	0.314	0.167	0.242	0.140	0.152	0.137	0.109	0.208	0.070
Mid ETF	0.172	0.236	0.295	0.169	0.253	0.322	0.375	0.291	0.277	0.185	0.206	0.223	0.162	0.303	0.095
Small ETF	0.184	0.196	0.289	0.163	0.235	0.256	0.328	0.271	0.278	0.178	0.227	0.288	0.131	0.250	0.094
By P/E Ratio															
Value ETF	0.205	0.397	0.232	0.146	0.210	0.286	0.266	0.299	0.236	0.173	0.187	0.239	0.137	0.287	0.086
Blend ETF	0.171	0.387	0.204	0.135	0.205	0.245	0.241	0.223	0.220	0.139	0.174	0.200	0.130	0.273	0.078
Growth ETF	0.181	0.363	0.248	0.157	0.264	0.278	0.297	0.268	0.254	0.146	0.185	0.200	0.167	0.240	0.107
By Market Cap and P/E Ratio															
Large-Value ETF	0.123	0.180	0.156	0.076	0.162	0.192	0.194	0.170	0.187	0.096	0.108	0.130	0.115	0.213	0.058
Large-Blend ETF	0.103	0.124	0.146	0.099	0.154	0.205	0.200	0.143	0.168	0.095	0.120	0.122	0.102	0.201	0.058
Large-Growth ETF	0.137	0.147	0.211	0.113	0.212	0.265	0.288	0.186	0.236	0.120	0.168	0.133	0.115	0.189	0.097
Mid-Value ETF	0.191	0.219	0.231	0.165	0.242	0.358	0.309	0.320	0.262	0.192	0.192	0.248	0.165	0.326	0.085
Mid-Blend ETF	0.158	0.234	0.178	0.143	0.218	0.268	0.257	0.244	0.203	0.153	0.154	0.189	0.132	0.300	0.082
Mid-Growth ETF	0.168	0.218	0.252	0.189	0.286	0.303	0.309	0.323	0.255	0.163	0.172	0.204	0.205	0.272	0.119
Small-Value ETF	0.206	0.224	0.249	0.158	0.174	0.239	0.256	0.346	0.224	0.200	0.221	0.295	0.105	0.224	0.108
Small-Blend ETF	0.170	0.184	0.213	0.157	0.231	0.242	0.247	0.238	0.273	0.145	0.215	0.275	0.146	0.286	0.088
Small-Growth ETF	0.179	0.178	0.185	0.129	0.267	0.242	0.253	0.212	0.222	0.125	0.177	0.253	0.125	0.180	0.089
By Sectors															
All Sectors	0.383	0.463	0.582	0.333	0.552	0.650	0.681	0.520	0.562	0.366	0.452	0.436	0.534	0.653	0.274
Basic Materials Funds	0.465	0.475	0.706	0.511	0.668	0.822	0.939	0.673	0.664	0.464	0.523	0.563	0.528	0.730	0.368
Financial Services Funds	0.296	0.407	0.445	0.249	0.391	0.524	0.519	0.426	0.393	0.291	0.370	0.376	0.294	0.400	0.157
Health & Biotechnology Funds	0.313	0.528	0.470	0.206	0.523	0.492	0.527	0.490	0.521	0.296	0.446	0.378	0.263	0.553	0.236
Real Estate Funds	0.400	0.450	0.523	0.251	0.467	0.651	0.506	0.467	0.316	0.374	0.280	0.414	0.256	0.543	0.204
Science & Technology Funds	0.241	0.389	0.332	0.248	0.395	0.494	0.475	0.437	0.420	0.243	0.252	0.281	0.275	0.455	0.220
Telecommunication Funds	0.316	0.513	0.384	0.282	0.559	0.502	0.549	0.421	0.337	0.322	0.396	0.419	0.360	0.508	0.262
Utility Funds	0.306	0.253	0.406	0.216	0.376	0.481	0.586	0.428	0.439	0.257	0.345	0.301	0.336	0.446	0.263
Consumer Goods Funds	0.307	0.422	0.341	0.196	0.373	0.446	0.360	0.349	0.302	0.235	0.245	0.225	0.193	0.357	0.250
Consumer Services Funds	0.148	0.364	0.275	0.230	0.352	0.477	0.384	0.475	0.310	0.274	0.211	0.390	0.219	0.377	0.163
Industrials Funds	0.200	0.402	0.390	0.232	0.388	0.497	0.531	0.314	0.439	0.226	0.319	0.317	0.349	0.520	0.190
Others															
All Others	0.444	0.354	0.516	0.358	0.444	0.664	0.669	0.499	0.633	0.340	0.509	0.393	0.636	0.541	0.296

Table 3: Regression to test univariate relation between fund flows and fund alphas

This table presents the estimated beta coefficients of fund flows on historical cumulative alphas ($\beta_{F\alpha}$) of ETFs relative to the nine tested models and associated t-statistics (in parenthesis). t-statistics are double-clustered by funds and months. Fund alphas are computed over various horizons from 3 months to 36 months. We also report the average probability that signed alphas correctly predict signed fund flows in the next month (Prob), which is equal to ($\beta_{F\alpha}$ + 1) / 2. MAR is market-adjusted return. CAPM is the standard capital asset pricing model. FF3 is the three-factor model of Fama and French (1993). Car4 is the four-factor model of Carhart (1997). PS7 is the seven-factor model which adds three industry factors to the Car4. FF5 is the five-factor model of Fama and French (2015). QF4 is the q-theroy based four-factor factor model of Hou, Xue and Zhang (2015). SY4 is the mispricing-based four-factor model of Stambaugh and Yuan (2017). DHS3 is the behavior-based three-factor model of Daniel, Hirshleifer and Sun (2020). Panel A reports results of univariate regressions with the model-adjusted returns as the single independent variable. Panel B reports results of regressions controlling for a bunch of fund characteristics including (log) fund size, (log) fund age, lagged expense ratio, lagged fund flows, volatility, fund category dummies and month dummies. The sample includes 1,128 U.S. domestic equity ETFs from January 2000 to December 2019.

Panel A: Regressions without controls

	Horizon																	
Model	3 months			6 months			12 months			18 months			24 months			36 months		
	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$
MAR	58.31	0.166	(45.05)	58.99	0.180	(48.87)	59.10	0.182	(49.53)	58.96	0.179	(48.74)	59.07	0.181	(49.34)	57.09	0.142	(38.32)
CAPM	57.88	0.158	(42.57)	58.50	0.170	(45.98)	58.64	0.173	(46.79)	58.37	0.167	(45.33)	58.56	0.171	(46.29)	56.36	0.127	(33.82)
FF3	56.27	0.125	(33.74)	56.86	0.137	(37.00)	57.03	0.141	(37.91)	56.89	0.138	(37.16)	57.00	0.140	(37.72)	56.29	0.126	(33.61)
Car4	56.60	0.132	(35.57)	57.07	0.141	(38.18)	57.27	0.145	(39.28)	57.06	0.141	(38.12)	57.21	0.144	(38.96)	56.43	0.129	(34.48)
PS7	55.26	0.105	(28.30)	55.82	0.116	(31.35)	55.88	0.118	(31.70)	55.86	0.117	(31.58)	55.96	0.119	(32.13)	55.20	0.104	(27.96)
FF5	55.88	0.118	(31.64)	56.36	0.127	(34.23)	56.60	0.132	(35.52)	56.42	0.128	(34.61)	56.60	0.132	(35.53)	55.81	0.116	(31.01)
QF4	56.19	0.124	(33.35)	56.73	0.135	(36.34)	56.77	0.135	(36.54)	56.74	0.135	(36.41)	56.61	0.132	(35.70)	55.41	0.108	(29.09)
SY4	56.22	0.124	(33.52)	56.77	0.135	(36.54)	56.94	0.139	(37.44)	56.91	0.138	(37.33)	56.94	0.139	(37.44)	55.91	0.118	(31.69)
DHS3	57.10	0.142	(38.29)	57.68	0.154	(41.53)	57.84	0.157	(42.42)	57.60	0.152	(41.13)	57.69	0.154	(41.55)	56.11	0.122	(32.67)

Panel B: Regressions with controls

	Horizon																	
Model	3 months			6 months			12 months			18 months			24 months			36 months		
	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$	Prob	$\beta_{F\alpha}$	$t(\beta_{F\alpha})$
MAR	54.84	0.097	(28.65)	55.30	0.106	(31.16)	55.33	0.107	(31.33)	55.30	0.106	(31.18)	55.31	0.106	(31.17)	53.63	0.073	(21.57)
CAPM	54.57	0.091	(27.15)	54.94	0.099	(29.21)	55.03	0.101	(29.71)	54.92	0.098	(29.15)	54.94	0.099	(29.23)	52.96	0.059	(17.49)
FF3	53.48	0.070	(20.87)	53.80	0.076	(22.72)	53.86	0.077	(23.04)	53.90	0.078	(23.33)	53.85	0.077	(22.97)	53.11	0.062	(18.45)
Car4	53.73	0.075	(22.35)	53.97	0.079	(23.60)	54.09	0.082	(24.30)	54.00	0.080	(23.92)	54.04	0.081	(24.06)	53.19	0.064	(18.95)
PS7	52.99	0.060	(18.09)	53.31	0.066	(19.94)	53.35	0.067	(20.19)	53.35	0.067	(20.22)	53.41	0.068	(20.57)	52.80	0.056	(16.93)
FF5	53.28	0.066	(19.66)	53.51	0.070	(21.04)	53.67	0.073	(22.00)	53.57	0.071	(21.44)	53.67	0.073	(22.03)	52.98	0.060	(17.80)
QF4	53.49	0.070	(20.87)	53.91	0.078	(23.35)	53.92	0.078	(23.41)	53.90	0.078	(23.27)	53.82	0.076	(22.82)	52.89	0.058	(17.46)
SY4	53.55	0.071	(21.37)	53.85	0.077	(22.99)	53.97	0.079	(23.69)	53.99	0.080	(23.83)	53.94	0.079	(23.48)	53.15	0.063	(18.85)
DHS3	54.22	0.084	(25.28)	54.64	0.093	(27.58)	54.72	0.094	(28.12)	54.58	0.092	(27.35)	54.59	0.092	(27.37)	53.31	0.066	(19.81)

Table 4: Multivariate linear regressions to test partial effects of performances on flows

This table presents results of the multivariate panel regression to test the marginal predictive power of weighted-average cumulative abnormal fund returns relative to the nine tested asset pricing models for percentage fund flow in the subsequent month. Three versions of fund alphas are considered: (i) the column "Raw" represents alphas in the original form; (ii) the column "Std" represents alphas normalized by standard deviation of alphas across funds in a month; (iii) the column "Rank" is the percentage ranking of alphas compared across funds in a month. As controls, (log) fund size, (log) fund age, lagged expense ratio, lagged fund flows, volatility of returns, category dummies and time dummies are included as independent variables along with the alphas. Alphas are computed over various evaluation horizon from 3 months to 3 years. We report estimated regression coefficients on alphas, double-clustered t-statistics (in parenthesis) and adjusted R^2.

Horizon	3 months			6 months			12 months		
	Raw	Std	Rank	Raw	Std	Rank	Raw	Std	Rank
Alpha(MAR)	0.627	0.013	0.043	0.756	0.013	0.044	0.782	0.013	0.043
t-stat	(9.10)	(9.81)	(12.07)	(9.16)	(10.27)	(12.55)	(8.98)	(10.25)	(12.82)
Alpha(CAPM)	0.025	0.002	0.016	0.108	0.004	0.018	0.113	0.004	0.018
t-stat	(0.27)	(1.33)	(3.20)	(0.92)	(2.35)	(3.68)	(0.91)	(2.27)	(3.64)
Alpha(FF3)	−0.307	−0.005	−0.010	−0.367	−0.005	−0.008	−0.353	−0.004	−0.010
t-stat	(−2.69)	(−2.05)	(−1.62)	(−2.78)	(−2.12)	(−1.37)	(−2.54)	(−1.93)	(−1.55)
Alpha(Car4)	0.187	0.004	0.016	0.230	0.005	0.016	0.250	0.005	0.017
t-stat	(2.44)	(2.25)	(3.22)	(2.76)	(3.11)	(3.37)	(2.80)	(3.12)	(3.59)
Alpha(PS7)	0.046	0.000	0.001	0.098	0.001	0.005	0.101	0.001	0.005
t-stat	(1.01)	(0.03)	(0.26)	(1.86)	(0.83)	(1.85)	(1.82)	(0.88)	(1.70)
Alpha(FF5)	0.064	0.000	0.003	0.067	−0.000	0.003	0.046	−0.001	0.003
t-stat	(0.86)	(0.02)	(0.61)	(0.76)	(−0.14)	(0.73)	(0.51)	(−0.38)	(0.59)
Alpha(QF4)	−0.053	−0.001	0.007	−0.073	−0.000	0.007	−0.084	−0.001	0.007
t-stat	(−0.74)	(−0.48)	(2.03)	(−0.86)	(−0.34)	(2.06)	(−0.92)	(−0.63)	(1.96)
Alpha(SY4)	−0.004	0.000	0.000	−0.004	−0.000	−0.002	−0.024	−0.000	−0.002
t-stat	(−0.05)	(0.25)	(0.01)	(−0.05)	(−0.25)	(−0.58)	(−0.24)	(−0.26)	(−0.53)
Alpha(DHS3)	0.174	0.002	−0.002	0.142	0.001	−0.004	0.176	0.001	−0.002
t-stat	(2.41)	(1.51)	(−0.61)	(1.58)	(0.47)	(−0.97)	(1.86)	(0.86)	(−0.50)
Adj-R^2	0.032	0.034	0.041	0.033	0.035	0.044	0.033	0.035	0.044
Horizon	**18 months**			**24 months**			**36 months**		
	Raw	Std	Rank	Raw	Std	Rank	Raw	Std	Rank
Alpha(MAR)	0.740	0.013	0.044	0.798	0.013	0.043	0.672	0.009	0.033
t-stat	(9.59)	(10.52)	(12.96)	(9.15)	(10.19)	(13.03)	(6.56)	(7.46)	(10.52)
Alpha(CAPM)	0.059	0.004	0.017	0.090	0.004	0.017	0.133	0.003	0.014
t-stat	(0.55)	(1.91)	(3.48)	(0.74)	(2.05)	(3.59)	(0.93)	(1.58)	(3.05)
Alpha(FF3)	−0.318	−0.004	−0.012	−0.351	-0.004	−0.009	−0.289	−0.003	−0.003
t-stat	(−2.51)	(−1.93)	(−1.92)	(−2.53)	(−1.89)	(−1.51)	(−1.85)	(−1.23)	(−0.53)
Alpha(Car4)	0.237	0.005	0.017	0.256	0.004	0.017	0.275	0.003	0.012
t-stat	(2.91)	(3.01)	(3.50)	(2.86)	(2.95)	(3.53)	(2.58)	(2.19)	(2.58)
Alpha(PS7)	0.095	0.001	0.004	0.093	0.001	0.005	0.077	0.001	0.007
t-stat	(1.88)	(0.89)	(1.48)	(1.65)	(0.80)	(1.70)	(1.21)	(1.43)	(2.82)
Alpha(FF5)	0.031	−0.001	0.003	0.054	−0.000	0.003	0.028	−0.000	0.003
t-stat	(0.37)	(−0.45)	(0.66)	(0.59)	(−0.21)	(0.76)	(0.28)	(−0.22)	(0.77)
Alpha(QF4)	−0.069	−0.001	0.007	−0.092	−0.001	0.006	−0.111	−0.002	0.002
t-stat	(−0.86)	(−0.51)	(1.99)	(−1.02)	(−0.69)	(1.63)	(−1.05)	(−1.25)	(0.46)
Alpha(SY4)	−0.008	0.000	0.001	−0.013	−0.000	−0.002	0.024	0.000	−0.001
t-stat	(−0.09)	(0.15)	(0.29)	(−0.13)	(−0.15)	(−0.45)	(0.20)	(0.19)	(−0.20)
Alpha(DHS3)	0.152	0.001	−0.002	0.180	0.002	−0.002	0.110	0.001	−0.005
t-stat	(1.87)	(0.96)	(−0.46)	(1.90)	(1.04)	(−0.47)	(0.97)	(0.48)	(−1.12)
Adj-R^2	0.034	0.036	0.044	0.033	0.035	0.043	0.024	0.025	0.031

Table 5: Pairwise comparison of models using regressions

The table presents t-statistics of regression coefficient from the panel regression to test the relative performances of models in predicting fund flows over various horizons from 3 months to 36 months. For each pair of models, the table shows the double-clustered t-statistics of the coefficient (γ_1) in the following panel regression

$$\phi(F_t^p) = \gamma_0 + \gamma_1 \left[\frac{\phi(\text{Alpha}_t^{p,\text{row}})}{\text{Var}[\phi(\text{Alpha}_t^{p,\text{row}})]} - \frac{\phi(\text{Alpha}_t^{p,\text{column}})}{\text{Var}[\phi(\text{Alpha}_t^{p,\text{column}})]} \right] + e_t^p$$

where F_t^p is the percentage fund flow. $\text{Alpha}_t^{p,\text{row}}$ and $\text{Alpha}_t^{p,\text{column}}$ are historical cumulative alphas relative to the model in a row and the model in a column. Panel A reports results of regression with the difference in variance normalized signs of alphas as the single independent variables. Panel B presents results of panel regressions controlling for fund characteristics, lagged fund flows, and category and time fixed effects.

Panel A: t-statistics of γ_1 from univariate regressions

	3 months								18 months							
	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
MAR	1.65	7.79	6.54	11.63	9.25	8.11	7.96	4.64	2.27	7.94	7.30	11.86	9.71	8.51	7.84	5.20
CAPM		6.12	4.87	9.95	7.58	6.44	6.29	2.98		5.65	5.01	9.56	7.42	6.22	5.55	2.92
FF3			−1.26	3.82	1.46	0.31	0.17	−3.15			−0.65	3.91	1.77	0.56	−0.11	−2.73
Car4				5.08	2.72	1.56	1.42	−1.90				4.56	2.42	1.21	0.54	−2.09
PS7					−2.36	−3.52	−3.65	−6.97					−2.14	−3.35	−4.01	−6.64
FF5						−1.16	−1.29	−4.61						−1.21	−1.88	−4.50
QF4							−0.14	−3.46							−0.67	−3.30
SY4								−3.32								−2.63

	6 months								24 months							
	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
MAR	1.88	8.13	7.34	12.11	10.05	8.64	8.47	5.00	1.96	7.93	7.10	11.88	9.44	9.40	8.15	5.28
CAPM		6.23	5.44	10.20	8.14	6.74	6.57	3.11		5.94	5.12	9.88	7.46	7.41	6.17	3.31
FF3			−0.80	3.96	1.92	0.50	0.34	−3.12			−0.83	3.93	1.51	1.46	0.22	−2.64
Car4				4.76	2.71	1.30	1.13	−2.33				4.76	2.34	2.29	1.05	−1.81
PS7					−2.04	−3.46	−3.63	−7.09					−2.42	−2.48	−3.72	−6.58
FF5						−1.42	−1.58	−5.04						−0.06	−1.30	−4.15
QF4							−0.17	−3.63							−1.24	−4.10
SY4								−3.46								−2.86

	12 months								36 months							
MAR	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
CAPM	1.76	7.93	7.02	12.31	9.58	8.95	8.29	4.83	2.79	3.04	2.53	7.20	4.87	6.40	4.50	3.73
FF3		6.14	5.24	10.51	7.79	7.15	6.50	3.05		0.24	−0.27	4.34	2.05	3.55	1.68	0.92
Car4			−0.91	4.36	1.65	1.00	0.35	−3.10			−0.51	4.12	1.82	3.32	1.44	0.68
PS7				5.28	2.56	1.92	1.27	−2.19				4.64	2.34	3.84	1.96	1.20
FF5					−2.71	−3.36	−4.01	−7.46					−2.29	−0.80	−2.68	−3.43
QF4						−0.65	−1.30	−4.75						1.49	−0.38	−1.14
SY4							−0.65	−4.11							−1.88	−2.64
								−3.45								−0.76

Table 5: Pairwise comparison of models using linear regressions – continued

Panel B: t-statistics of γ_1 from regressions with controls

	3 months								18 months							
	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
MAR	1.16	5.91	4.80	8.03	6.77	5.87	5.60	2.67	1.68	6.09	5.66	8.50	7.51	6.09	5.71	3.16
CAPM		4.74	3.63	6.86	5.60	4.69	4.43	1.50		4.40	3.97	6.80	5.82	4.40	4.02	1.48
FF3			−1.11	2.11	0.86	−0.05	−0.31	−3.24			−0.43	2.40	1.42	−0.01	−0.39	−2.93
Car4				3.22	1.97	1.06	0.80	−2.14				2.83	1.85	0.42	0.05	−2.50
PS7					−1.25	−2.17	−2.43	−5.36					−0.98	−2.41	−2.79	−5.34
FF5						−0.91	−1.18	−4.11						−1.43	−1.81	−4.35
QF4							−0.26	−3.20							−0.38	−2.93
SY4								−2.93								−2.55
								6 months								24 months
	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
MAR	1.57	6.50	5.80	8.66	7.80	6.07	6.30	2.88	1.57	6.32	5.50	8.23	7.08	6.48	5.94	3.13
CAPM		4.91	4.22	7.06	6.21	4.48	4.71	1.31		4.73	3.91	6.63	5.49	4.89	4.35	1.55
FF3			−0.70	2.14	1.30	−0.45	−0.21	−3.62			−0.83	1.89	0.75	0.14	−0.39	−3.20
Car4				2.85	2.00	0.25	0.49	−2.92				2.72	1.58	0.97	0.44	−2.37
PS7					−0.85	−2.60	−2.36	−5.77					−1.13	−1.75	−2.28	−5.10
FF5						−1.75	−1.51	−4.92						−0.62	−1.15	−3.95
QF4							0.24	−3.18							−0.53	−3.35
SY4								−3.41								−2.81
								12 months								36 months
MAR	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
CAPM	1.30	6.36	5.41	8.61	7.21	6.15	5.92	2.66	2.85	2.24	1.91	3.56	2.80	3.18	2.07	1.36
FF3		5.04	4.10	7.27	5.89	4.82	4.60	1.35		−0.62	−0.95	0.67	−0.07	0.29	−0.80	−1.50
Car4			−0.95	2.22	0.84	−0.23	−0.46	−3.71			−0.34	1.30	0.55	0.91	−0.18	−0.89
PS7				3.18	1.80	0.72	0.50	−2.76				1.64	0.89	1.25	0.16	−0.55
FF5					−1.38	−2.46	−2.68	−5.94					−0.75	−0.39	−1.48	−2.19
QF4						−1.08	−1.30	−4.55						0.36	−0.73	−1.43
SY4							−0.22	−3.49							−1.10	−1.81
								−3.26								−0.71

Table 6: Non-linear pairwise horse races between models

This table presents results of non-linear horse races between each pair of the nine competing asset pricing models. In each month, ETFs are sorted into quintiles by alphas relative to a model in the row and a model in the column. Fund Flows are regressed on dummies of intersections of quintile rankings and control variables

$$F_t^p = \delta_0 + \sum\sum \delta_{ij} D_{ijt}^p + \delta_X X_t^p + \theta_t + e_t^p$$

where D_{ijt}^p are dummies taking the value of 1 if a fund's alpha relative to the row-model falls into the i^{th} quintile and the fund's alpha relative to the column-model falls into the j^{th} quintile in the month t while taking the value of 0 if otherwise. The dummy for $i = 3$ and $j = 3$ is excluded. As controls, we include (log) fund size, (log) fund age, lagged expense ratio, volatility, category dummies and time dummies. We report the sum of differences between the coefficient on the pair of ranking dummies with the same magnitudes and reversed orderings, and the t-statistics to test the null hypothesis that the summed differences in coefficients are equal to 0, i.e., $\sum(\delta_{ij} - \delta_{ji}) = 0$. t-statistics are double clustered by funds and months. The table also presents the proportion of cases with positive coefficient difference $P(\delta_{ij} > \delta_{ji})$, and the p-value of the binominal test on the hypothesis that $P(\delta_{ij} > \delta_{ji})$ = 0.

Models	Statistics	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
	Sum of coefficient diff	8.278	15.133	13.273	18.441	16.715	14.026	13.656	13.777
MAR	*t*-stat	(3.11)	(8.08)	(7.36)	(11.64)	(9.90)	(7.64)	(7.61)	(6.31)
	Proportion of coeff diff > 0	0.800	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Binominal *p*-value	(0.011)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Sum of coefficient diff		14.470	11.156	16.874	15.302	11.601	11.409	13.031
CAPM	*t*-stat		(6.03)	(5.34)	(10.21)	(7.91)	(5.38)	(5.45)	(3.84)
	Proportion of coeff diff > 0		1.000	1.000	1.000	1.000	1.000	0.900	0.900
	Binominal *p*-value		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
	Sum of coefficient diff			-2.328	13.136	13.945	0.821	1.603	-3.635
FF3	*t*-stat			(-0.57)	(6.79)	(3.75)	(0.34)	(0.60)	(-1.82)
	Proportion of coeff diff > 0			0.500	0.900	1.000	0.600	0.600	0.500
	Binominal *p*-value			(0.377)	(0.001)	(0.000)	(0.172)	(0.172)	0.377)
	Sum of coefficient diff				14.850	8.694	1.177	1.865	-2.399
Car4	*t*-stat				(7.12)	(3.29)	(0.54)	(0.77)	(-1.26)
	Proportion of coeff diff > 0				1.000	0.900	0.500	0.500	0.300
	Binominal *p*-value				(0.000)	(0.001)	(0.377)	(0.377)	(0.828)
	Sum of coefficient diff					-5.886	-9.885	-11.264	-0.755
PS7	*t*-stat					(-3.52)	(-5.82)	(-6.28)	(-6.67)
	Proportion of coeff diff > 0					0.100	0.100	0.100	0.000
	Binominal *p*-value					(0.989)	0.989)	(0.989)	(0.999)
	Sum of coefficient diff						-6.775	-6.869	-7.691
FF5	*t*-stat						(-3.04)	(-3.09)	(-4.27)
	Proportion of coeff diff > 0						0.000	0.200	0.200
	Binominal *p*-value						(0.999)	(0.945)	(0.945)
	Sum of coefficient diff							-0.582	-4.205
QF4	*t*-stat							(-0.25)	(-1.98)
	Proportion of coeff diff > 0							0.500	0.300
	Binominal *p*-value							(0.377)	(0.828)
	Sum of coefficient diff								-4.103
SY4	*t*-stat								(-1.92)
	Proportion of coeff diff > 0								0.400
	Binominal *p*-value								(0.623)

Table 7: Regressions to Test Responses of Fund Flows to Components of Returns

The table presents results of the panel regressions to test the responses of fund flows to fund returns decomposed into alphas and factor-related parts. For a factor, the factor-related return is the product of a fund's beta coefficient on that factor estimated in the previous 24 to 60 months (as available) and the realized return on the factor in a month. Alpha relative to a model is the difference between excess return on a fund and summed returns attributable to all factors in that model. Using the exponential-decay function, we calculate weighted-average historical alphas (Alpha) and factor-related returns over various evaluation horizons from 3 months to 36 months. For each model, fund returns are regressed on Alpha and cumulative returns traced to each factor in the model. As controls, we include (log) fund size, (log) fund age, lagged expense ratio, lagged fund flows, volatility, category dummies and time dummies. Panel A presents results of the panel regressions for each evaluation period. *t*-statistics of regression coefficients are double-clustered by funds and months and are reported in parenthesis. Panel B reports the ratio of regression coefficient on each factor-related return to regression coefficient on alpha relative to a model. *p*-value for the test on the null hypothesis that the ratio is equal to 1 are reported in parenthesis.

Panel A: Responses of fund flows to return components

Model	Alpha	MKT	SMB	HML	UMD	RMW	CMA	ROE	IA	MGMT	PERM	FIN	PEAD	IndPCA1	IndPCA2	IndPCA3
CAPM	0.826	0.248														
	(20.60)	(11.57)														
FF3	0.816	0.214	1.127	0.789												
	(19.66)	(9.73)	(10.60)	(9.84)												
Car4	0.851	0.199	1.184	0.726	0.551											
	(20.82)	(9.09)	(11.14)	(8.91)	(5.78)											
FF5	0.800	0.214	1.116	0.702		0.899	0.768									
	(19.27)	(9.53)	(10.40)	(8.61)		(7.65)	(9.43)									
QF4	0.800	0.203	1.098					0.789	0.543							
	(19.37)	(9.11)	(10.26)					(8.71)	(5.83)							
SY4	0.827	0.206	1.234							0.724	0.428					
	(20.31)	(9.22)	(10.93)							(9.43)	(4.25)					
DHS3	0.849	0.227										0.510	0.847			
	(21.12)	(10.11)										(6.12)	(3.48)			
PS7	0.899	0.195	1.257	0.588	0.895									0.607	1.043	0.8136
	(21.14)	(8.89)	(11.45)	(7.88)	(6.18)									(12.15)	(8.04)	(5.68)

Panel B: Comparison of regression coefficients on factor-related components and on alphas

Model	MKT	SMB	HML	UMD	RMW	CMA	ROE	IA	MGMT	PERM	FIN	PEAD	IndPCA1	IndPCA2	IndPCA3
CAPM	0.3064														
	(0.0000)														
FF3	0.2752	1.2679	1.1800												
	(0.0000)	(0.0330)	(0.0619)												
Car4	0.2585	1.3176	1.0733	0.5444											
	(0.0000)	(0.0091)	(0.4729)	(0.0000)											
FF5	0.2650	1.3959	1.0243		1.3559	1.1020									
	(0.0000)	(0.0011)	(0.7907)		(0.0140)	(0.4033)									
QF4	0.2351	1.2044					0.9964	0.4781							
	(0.0000)	(0.0909)					(0.9736)	(0.0000)							
SY4	0.2536	1.3701							1.1423	0.7738					
	(0.0000)	(0.0031)							(0.2149)	(0.0420)					
DHS3	0.2896										0.5093	0.6771			
	(0.0000)										(0.0000)	(0.0317)			
PS7	0.2289	1.3618	0.7906	0.8803									0.6751	1.0218	1.1617
	(0.0000)	(0.0017)	(0.0132)	(0.3232)									(0.0000)	(0.8918)	(0.3518)

Table 8: Effects of Morningstar rating on flow-alpha relations

We follow the Morningstar evaluation algorithm to compute the overall fund ratings based on fund risk-adjusted returns in the previous3, 5, and 10 years in comparison to peer funds. Panel A reports average fund characteristics and performances of ETFs sorted by Morningstar ratings, and the variables' correlations with fund ratings. The group 1, 2, 3, 4 and 5 corresponds to ETFs with overall ratings falling into the range [1, 1.5), [1.5, 2.5), [2.5, 3.5), [3.5, 4.5) and [4.5, 5]. %Obs is the percentage of the number of observations for a group out of the total number of observations for the aggregate sample. Size is the month-end net asset value (in $million). Age is the number of months since fund inception date. Volatility is the standard deviation of monthly fund returns over the previous one year. Flow is the percentage of fund flow in the subsequent month. Alpha is the weighted-average historical alpha relative to a particular model computed over the previous 18 months. Panel B reports results of panel regressions with fund flow as the independent variable and percentage ranking of historical weight-average alphas relative to the nine competing models and Morningstar Ratings as dependent variables. Alphas are computed over various horizons. Double-clustered t-statistics of regression coefficients are reported in parenthesis.

Panel A: Fund characteristics and performances by Morningstar Rating

						Alpha								
Rating	%Obs	Size	Age	Volatility	Flow	MAR	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
1	7.84	435	59.77	6.20	0.95	−0.95	−1.09	−1.05	−0.91	−0.66	−0.97	−0.73	−0.76	−0.85
2	21.66	1,167	71.57	5.02	1.36	−0.30	−0.49	−0.42	−0.36	−0.21	−0.40	−0.26	−0.29	−0.35
3	36.27	3,179	91.88	4.57	1.41	−0.03	−0.18	−0.16	−0.13	−0.08	−0.15	−0.08	−0.10	−0.12
4	23.53	2,846	93.99	4.48	1.93	0.15	0.02	0.03	0.06	0.11	0.03	0.09	0.05	0.06
5	10.70	1,819	63.28	4.71	3.68	0.39	0.27	0.27	0.32	0.30	0.25	0.28	0.28	0.27
Correlation		0.037	0.041	−0.179	0.053	0.164	0.176	0.167	0.157	0.125	0.145	0.126	0.136	0.150

Panel B: Multivariate regression to test flow-alpha and flow-rating relations

	Horizon					
	3 months	6 months	12 months	18 months	24 months	36 months
Alpha(MAR)	0.043	0.044	0.043	0.044	0.043	0.032
t-stat	(11.94)	(12.42)	(12.70)	(12.84)	(12.88)	(10.13)
Alpha(CAPM)	0.015	0.017	0.017	0.016	0.017	0.013
t-stat	(2.99)	(3.53)	(3.49)	(3.31)	(3.47)	(2.93)
Alpha(FF3)	−0.010	−0.008	−0.009	−0.012	−0.009	−0.003
t-stat	(−1.57)	(−1.32)	(−1.53)	(−1.84)	(−1.48)	(−0.49)
Alpha(Car4)	0.016	0.015	0.017	0.017	0.016	0.012
t-stat	(3.15)	(3.31)	(3.52)	(3.42)	(3.48)	(2.52)
Alpha(PS7)	0.000	0.005	0.004	0.004	0.004	0.007
t-stat	(0.15)	(0.71)	(1.58)	(1.33)	(1.57)	(2.73)
Alpha(FF5)	0.002	0.003	0.002	0.002	0.003	0.003
t-stat	(0.52)	(0.68)	(0.57)	(0.55)	(0.72)	(0.81)
Alpha(QF4)	0.007	0.007	0.007	0.007	0.005	0.002
t-stat	(2.02)	(2.06)	(1.95)	(2.02)	(1.62)	0.52)
ALPHA(SY4)	0.000	−0.002	−0.002	0.001	−0.002	−0.001
t-stat	(0.02)	(−0.60)	(−0.54)	(0.26)	(−0.48)	(−0.29)
Alpha(DHS3)	−0.002	−0.004	−0.002	−0.001	−0.002	−0.004
t-stat	(−0.50)	(−0.90)	(−0.42)	(−0.37)	(−0.39)	(−0.99)
Rating	0.003	0.002	0.002	0.003	0.002	0.001
t-stat	(6.54)	(4.96)	(4.46)	(5.46)	(4.24)	(2.28)
Adj-R^2	0.042	0.044	0.045	0.045	0.044	0.032

Table 9: Effects of behavior models on flow-alpha relations

In each month, ETFs are sorted by prospect theory (PT) and salience theory (ST) values. Panel A reports mean values of fund size (in $ million), fund age, volatility of fund returns, percentage of fund flow in the next month (Flow), and weighted-average alphas over the previous 18 months (Alpha) relative to a model, and their correlations with PT and ST values. Mean values of volatility, fund flow, and alphas are reported in percentage. Panel B reports results of panel regressions to test responses of fund flows to model-adjusted returns and the PT and ST values with control variables. Three versions of key variables (Alphas, PT and ST) are measured by (i) original variables (Raw), (ii) original variables normalized by cross-sectional standard deviation in each month (Std), and (iii) percentage ranking of variables. Double-clustered t-statistics are reported in parenthesis.

Panel A: Fund characteristics and performances by PT and ST values

Groups	Size	Age	Volatility	Flow	Alpha								
					MAR	CAPM	FF3	Car4	PS7	FF5	QF4	SY4	DHS3
Low PT	1,178	75.24	6.08	0.01	−1.66	−1.70	−1.50	−1.38	−0.94	−1.43	−1.26	−1.25	−1.41
2	1,220	78.87	4.74	0.82	−0.29	−0.45	−0.31	−0.25	−0.17	−0.29	−0.21	−0.22	−0.33
3	1,591	80.95	4.44	1.67	0.11	−0.06	0.01	0.03	0.03	0.02	0.05	0.03	−0.01
4	2,915	84.72	4.29	2.42	0.43	0.25	0.21	0.23	0.19	0.20	0.27	0.22	0.26
High PT	4,577	92.01	4.40	3.68	1.01	0.79	0.59	0.61	0.54	0.56	0.67	0.62	0.75
Correlation with PT	0.117	0.107	−0.308	0.106	0.465	0.445	0.373	0.351	0.265	0.327	0.333	0.340	0.395
Low ST	1,613	81.66	5.41	0.61	−0.98	−1.13	−0.98	−0.91	−0.62	−0.93	−0.81	−0.82	−0.95
2	2,523	83.08	4.44	1.55	−0.16	−0.28	−0.22	−0.19	−0.11	−0.21	−0.14	−0.17	−0.21
3	2,806	82.52	4.28	1.69	−0.01	−0.11	−0.10	−0.06	−0.02	−0.09	−0.04	−0.07	−0.06
4	2,696	83.32	4.41	2.06	0.14	0.01	−0.02	0.01	0.04	−0.01	0.05	0.03	0.05
High ST	1,880	81.40	5.40	2.71	0.63	0.38	0.34	0.40	0.37	0.33	0.47	0.44	0.45
Correlation with ST	0.008	0.0014	0.000	0.056	0.274	0.263	0.228	0.225	0.171	0.201	0.213	0.222	0.248

Panel B: Panel regression to test fund-alpha relations controlling for PT and ST

Horizon	3 months			6 months			12 months			18 months			24 months			36 months		
	Raw	Std	Rank	Raw	Std	Rank	Raw	Std	Rank	Raw	Std	Rank	Raw	Std	Rank	Raw	Std	Rank
Alpha(MAR)	0.316	0.006	0.012	0.323	0.005	0.011	0.309	0.005	0.010	0.328	0.005	0.009	0.311	0.005	0.010	0.135	0.002	0.005
t-stat	(4.55)	(4.15)	(3.34)	(3.91)	(3.94)	(3.09)	(3.54)	(3.86)	(2.98)	(4.21)	(4.06)	(2.46)	(3.57)	(3.73)	(2.98)	(1.35)	(1.28)	(1.51)
Alpha(CAPM)	0.069	0.004	0.013	0.171	0.005	0.011	0.193	0.005	0.010	0.125	0.005	0.011	0.170	0.005	0.010	0.202	0.004	0.003
t-stat	(0.73)	(1.94)	(2.61)	(1.48)	(2.83)	(2.23)	(1.58)	(2.79)	(2.13)	(1.18)	(2.45)	(2.31)	(1.40)	(2.59)	(2.17)	(1.41)	(2.04)	(0.78)
Alpha(FF3)	-0.280	-0.002	0.001	-0.356	-0.003	0.001	-0.346	-0.003	-0.001	-0.301	-0.002	-0.003	-0.341	-0.003	-0.001	-0.260	-0.002	-0.003
t-stat	(-2.46)	(-0.96)	(0.13)	(-2.72)	(-1.28)	(0.24)	(-2.51)	(-1.22)	(-0.20)	(-2.38)	(-1.03)	(-0.49)	(-2.48)	(-1.14)	(-0.13)	(-1.68)	(-1.01)	(-0.53)
Alpha(Car4)	0.187	0.003	0.013	0.223	0.004	0.013	0.229	0.004	0.014	0.225	0.004	0.015	0.242	0.004	0.014	0.240	0.003	0.011
t-stat	(2.33)	(1.77)	(2.65)	(2.48)	(2.65)	(2.75)	(2.37)	(2.73)	(2.98)	(2.58)	(2.69)	(3.10)	(2.50)	(2.58)	(2.92)	(2.11)	(1.90)	(2.44)
Alpha(PS7)	-0.185	-0.004	-0.017	-0.193	-0.004	-0.016	-0.190	-0.004	-0.016	-0.188	-0.004	-0.016	-0.201	-0.004	-0.016	-0.132	-0.002	-0.010
t-stat	(-4.08)	(-4.82)	(-6.72)	(-3.60)	(-4.41)	(-6.22)	(-3.35)	(-4.19)	(-6.26)	(-3.66)	(-4.31)	(-5.97)	(-3.52)	(-4.33)	(-6.28)	(-1.98)	(-2.61)	(-4.31)
Alpha(FF5)	-0.010	-0.003	-0.013	0.010	-0.003	-0.011	-0.004	-0.003	-0.011	-0.029	-0.003	-0.011	0.005	-0.003	-0.011	0.001	-0.002	-0.004
t-stat	(-0.13)	(-1.89)	(-2.99)	(0.11)	(-1.87)	(-2.71)	(-0.04)	(-2.05)	(-2.64)	(-0.35)	(-2.24)	(-2.50)	(0.06)	(-1.82)	(-2.55)	(0.01)	(-1.00)	(-1.06)
Alpha(QF4)	-0.059	-0.001	0.003	-0.084	-0.001	0.003	-0.100	-0.002	0.003	-0.081	-0.001	0.003	-0.099	-0.002	0.001	-0.100	-0.002	-0.001
t-stat	(-0.81)	(-0.99)	(0.94)	(-0.96)	(-0.90)	(0.76)	(-1.07)	(-1.09)	(0.75)	(-0.98)	(-1.05)	(0.83)	(-1.07)	(-1.16)	(0.38)	(-0.92)	(-1.29)	(-0.16)
ALPHA(SY4)	-0.054	0.001	-0.002	-0.051	0.000	-0.002	-0.063	0.000	-0.001	-0.060	0.001	0.000	-0.068	0.000	-0.000	-0.037	0.001	0.003
t-stat	(-0.71)	(0.42)	(-0.51)	(-0.56)	(0.20)	(-0.43)	(-0.64)	(0.20)	(-0.24)	(-0.70)	(0.40)	(0.13)	(-0.68)	(0.25)	(-0.10)	(-0.31)	(0.62)	(0.66)
Alpha(DHS3)	0.015	0.000	-0.011	-0.044	-0.001	-0.010	-0.028	-0.001	-0.008	-0.022	-0.001	-0.009	-0.021	-0.000	=0.008	-0.051	-0.000	-0.005
t-stat	(0.20)	(0.24)	(-2.83)	(-0.49)	(-0.67)	(-2.47)	(-0.30)	(-0.39)	(-2.01)	(-0.26)	(-0.36)	(-2.30)	(-0.22)	(-0.17)	(-1.96)	(-0.44)	(-0.23)	(-1.10)
PT	1.352	0.009	0.045	1.360	0.009	0.045	1.375	0.009	0.045	1.334	0.008	0.045	1.376	0.009	0.045	1.494	0.010	0.047
t-stat	(10.96)	(11.94)	(21.04)	(10.98)	(11.51)	(20.91)	(11.08)	(11.40)	(20.82)	(10.66)	(10.77)	(20.57)	(11.11)	(11.59)	(20.89)	(12.29)	(14.10)	(22.27)
ST	0.169	0.000	0.004	0.168	0.000	0.004	0.166	0.000	0.004	0.158	-0.000	0.004	0.170	0.000	0.004	0.183	0.000	0.004
t-stat	(1.70)	(0.29)	(2.27)	(1.68)	(0.22)	(2.23)	(1.66)	(0.18)	(2.23)	(1.58)	(-0.07)	(2.16)	(1.70)	(0.21)	(2.25)	(1.83)	(0.60)	(2.27)
Adj-R^2	0.027	0.029	0.032	0.027	0.029	0.032	0.027	0.029	0.032	0.027	0.029	0.032	0.027	0.029	0.032	0.026	0.027	0.031

Online Appendix for
"Assessing Asset Pricing Models Using Exchange-Traded Fund Flows"

Appendix A1: Construction of industry factors using the principal component analysis (PCA)

The procedures to construct the three industry factors are proceed as follows. First, for an industry portfolio I in one of the Fama and French 17 industries in month t, monthly equal-weighted portfolio returns are regressed on market, size, value, and momentum factors (i.e., the Car4 model) in the previous 120 months

$$R_\tau^I = \alpha_t^I + \beta_t^{I,\mathrm{MKT}}\mathrm{MKT}_\tau + \beta_t^{I,\mathrm{SMB}}\mathrm{SMB}_\tau + \beta_t^{I,\mathrm{HML}}\mathrm{HML}_\tau + \beta_t^{I,\mathrm{UMD}}\mathrm{UMD}_\tau + \varepsilon_\tau^I, \quad \text{where } 1 \le I \le 17, t-120 \le \tau \le t-1, \tag{A1}$$

Where R_τ^I is the return in industry I; β_t^Is denote the industry portfolio's factor loadings estimated using information available before month t; and ε_τ^I represents the residual portfolio return in month τ. For month t, with the estimated factor loadings and realizations of excess returns on industry portfolio I and returns on pricing factors, the residual return on industry portfolio I is estimated as

$$\hat{\varepsilon}_t^I = \left(R_t^I - R_t^f\right) - \left(\beta_t^{I,\mathrm{MKT}}\mathrm{MKT}_t + \beta_t^{I,\mathrm{SMB}}\mathrm{SMB}_t + \beta_t^{I,\mathrm{HML}}\mathrm{HML}_t + \beta_t^{I,\mathrm{UMD}}\mathrm{UMD}_t\right). \tag{A2}$$

We then compute the covariance matrix of the 17-by-120 matrix of the 17 industries' residual returns in the previous 120 months and obtain the eigenvectors and associated eigenvalues of the covariance matrix. Let $[w_t^{k1}, w_t^{k2}, w_t^{k3}]$ denotes the eigenvectors associated with the three largest eigenvalues estimated based on information available before time t. Let E_t denotes the 17-by-1 vector of estimated residual returns on the 17 industry portfolios in month t. The return on the first industry factor in month t is the average residual returns on the 17 industry portfolios weighted by the first eigenvector as follows:

$$\mathrm{IndPCA1}_t = \left(w_t^{k1}\right)' E_t. \tag{A3}$$

Returns on the second and third industry factors can be calculated in the similar way, with weightings on industries replaced by the second and third eigenvectors estimated from PCA analysis. The three industry factors can be considered as the returns on long-short strategies of industry portfolios that can best explain the variations in industry portfolio returns in a month that are left unexplained by the Carhart four-factor model (Car4).

Appendix A2: Method to construct Morningstar Ratings

The Morningstar Ratings is a measure of the risk-adjusted performance of an ETF in comparison to peer funds in the same category. Morningstar ratings range from 1 for the worst-performing funds to 5 for the best-performing funds. According to the documentation of Morningstar, we rate fund performance in the following steps.

First, based on the Lipper objective codes, we assign ETFs to one of the 18 narrow Morningstar categories for U.S. domestic equity ETFs, including 9 equity style boxes by size (small, mid, and big) and by P/E ratio (value, blend, and growth) and 9 sector-specific categories. Funds with missing Lipper codes and funds that do not fall into any of the 18 categories are aggregated into the "Others" group.

Second, based on the expected utility theory, for a fund p over an evaluation horizon in the previous T months, we compute the Morningstar Risk-Adjusted Return (MSAR) as follows:

$$\mathrm{MRAR}(T)_t^p = \left[\frac{1}{T}\sum_{\tau=1}^{T}\left(1+\tilde{R}_\tau^p\right)^{-\gamma}\right]^{-\frac{12}{\gamma}} - 1, \tag{A4}$$

where $\tilde{R}_\tau^p = \left(1+R_\tau^p\right)/\left(1+R_\tau^f\right) - 1$ is the geometric excess return on fund p in month τ, and γ = 2 denotes the degree of risk aversion. MRARs are computed over the previous three, five, and ten years.

Third, for one of the three evaluation horizons, ETFs are ranked by the MRAR within peer funds in descending order. Before June 2002, Morningstar pooled all domestic equity funds into one ranking universe when evaluating fund performance. On and after July 2002, the institution has changed its rating algorithm by ranking ETFs within funds with the same fund category. For an evaluation horizon T, an ETF p is given a star rating $\mathrm{Star}(T)_t^p$ based on the following rule: funds with ranking in the range $[0\%, 10\%)$, $[10\%, 32{,}5\%)$, $[32.5\%, 67.5\%)$, $[67.5\%, 90\%)$, and $[90\%, 100\%)$ receive 5, 4, 3, 2 and 1 star(s), respectively.

Finally, we compute an overall Morningstar Rating as a weighted average of fund stars evaluated over the previous three, five, and ten years depending on fund age. For an ETF with age shorter than 60 months, the overall rating is $100\% \times \mathrm{Star}(3ys)_t^p$; for an ETF with age ranging between 60 and 120 months, the overall rating is $60\% \times \mathrm{Star}(5ys)_t^p + 40\% \times \mathrm{Star}(3ys)_t^p$; for an ETF with age longer than 120 months, the overall rating is $50\% \times \mathrm{Star}(10ys)_t^p + 30\% \times \mathrm{Star}(5ys)_t^p + 20\% \times \mathrm{Star}(3ys)_t^p$.

Appendix A3: Construction of prospect theory (PT) and salience theory (ST) values

We compute the prospect theory value for individual ETFs using the method similar to that of Barberis, Mukherjee, and Wang (2016). Specifically, we assume that investors mentally represent gains and losses by the daily fund returns in excess of the market return on the same day. For an ETF p in month t, suppose that daily market excess returns on the fund are negative in m days and are non-negative in $n = N_t - m$ days, where $N_t = m + n$ is the number of trading days.

Daily market excess returns on the funds are ranked in ascending order from the most negative $\tilde{r}_{-mt}^{p}$ to the most positive $\tilde{r}_{nt}^{p}$. The prospect theory value of ETF p in month t, PT_t^p, is the probability-weighted average value of the daily gains and losses:

$$\mathrm{PT}_t^p = \sum_{d=-m}^{-1} v(\tilde{r}_{dt}^p)\left[\omega^-\left(\frac{d+m+1}{N_t}\right) - \omega^-\left(\frac{d+m}{N_t}\right)\right] + \sum_{d=1}^{n} v(\tilde{r}_{dt}^p)\left[\omega^+\left(\frac{n-d+1}{N_t}\right) - \omega^+\left(\frac{n-d}{N_t}\right)\right], \tag{A5}$$

where $v(\cdot)$ is the value function

$$v(\tilde{r}_{dt}^p) = \begin{cases} (\tilde{r}_{dt}^p)^\alpha, & \tilde{r}_{dt}^p \geq 0, \\ -\lambda(-\tilde{r}_{dt}^p)^\alpha, & \tilde{r}_{dt}^p < 0, \end{cases} \tag{A6}$$

and $\omega^+(\cdot)$ and $\omega^+(\cdot)$ are probability weighting functions for the gain and loss domains

$$\omega^+(p) = \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{\frac{1}{\gamma}}} \text{ and } \omega^-(p) = \frac{p^\varphi}{[p^\varphi + (1-p)^\varphi]^{\frac{1}{\varphi}}}. \tag{A7}$$

We use the values of parameters of the value function ($\alpha = 0.88$; $\lambda = 2.25$) and those of the weighting function ($\gamma = 0.61$; $\varphi = 0.69$) estimated by Tversky and Kahneman (1992).

The salience value theory of ETFs is computed following Bordalo, Gennaioli, and Shleifer (2012; 2013) and Cosemans and Frehen (2021). For ETF p on day d in month t, the salience measure is the scaled distance between the fund's daily return r_{dt}^p and the equal-weighted market portfolio return $\bar{r}_{dt}^M$

$$\sigma(r_{dt}^p, \bar{r}_{dt}^M) = \frac{|r_{dt}^p - \bar{r}_{dt}^M|}{|r_{dt}^p| + |\bar{r}_{dt}^M| + \theta}. \tag{A8}$$

Daily returns on an ETF in a month are ranked by the salience measures and assigned a salience ranking k_{dt}^p ranging from 1 to N_t. The salience weighting of the fund's return on day d is

$$\omega_{dt}^{p\mathrm{ST}} = \delta^{k_{dt}^p} \Big/ \sum_d \left(\frac{\delta^{k_{dt}^p}}{N_t}\right), \qquad \delta \in (0,1]. \tag{A9}$$

We use parameter values estimated by Bordalo et al. (2012) ($\theta = 0.1$ and $\delta = 0.7$) to compute salience weights. Finally, an ETF's salience value in a month is computed as the covariance between daily fund returns and their salience weightings with equal probability each day ($1/N_t$): $\mathrm{ST}_t^p = \mathrm{Cov}(\omega_{dt}^{p\mathrm{ST}}, r_{dt}^p)$.