What do fund flows reveal about asset pricing models and investor sophistication?

Narasimhan Jegadeesh and Chandra Sekhar Mangipudi☆

Goizueta Business School

Emory University

December, 2018

☆Narasimhan Jegadeesh is the Dean's Distinguished Professor at the Goizueta Business School, Emory University, Atlanta, GA 30322, ph: 404-727-4821, email: Jegadeesh@emory.edu., Chandra Sekhar Mangipudi is a Doctoral student at Emory University, email: chandra.sekhar.mangipudi@emory.edu . We would like to thank Jonathan Berk, Kent Daniel, Amit Goyal, Cliff Green, Jay Shanken, and the seminar participants at the 2018 SFS Cavalcade North America, Louisiana State University and Virginia Tech for helpful discussions, comments, and suggestions.

What do fund flows reveal about asset pricing models and investor sophistication?

Recent literature uses the relative strength of the relation between fund flows and alphas with respect to various multifactor models to draw inferences about the best asset pricing model and about investor sophistication. This paper analytically shows that such inferences are tenable only under certain assumptions and we test their empirical validity. Our results indicate that any inference about the true asset pricing model based on alpha-flow relations is empirically untenable. The literature uses a multifactor model that includes all factors as the benchmark to assess investor sophistication. We show that the appropriate benchmark excludes some factors when their betas are estimated from the data, but even with this benchmark the rejection of investor sophistication in the literature is empirically tenable.

An extensive literature documents that net fund flows into mutual funds are driven by funds' past performance. For example, Patel, Zeckhauser, and Hendricks (1994) document that equity mutual funds with bigger returns attract more cash inflows and they offer various behavioral explanations for this phenomenon. Other papers that document a positive relation between fund flows and past performance include Ippolito (1992), Chevalier and Ellison (1997), and Sirri and Tufano (1998).

Some papers in the early literature also examine whether abnormal performance (or alphas) measured with respect to some benchmarks better predict fund flows than others. For example, Gruber (1996) compares the mutual fund flow-performance relation for alphas measured with respect to one- and four-factor models, while Del Guercio and Tkac (2002) compares sensitivity of flows to raw returns vis-à-vis alphas from market model in mutual funds and pension funds. Fung et. al. (2008) makes similar comparisons with a different set of factor models for a sample of hedge funds.

While comparison of flow-alpha relations across models was not the primary focus of earlier papers, recent papers in this area have shown a renewed interest in such comparisons using a broader range of asset pricing and factor models. The primary driving force for this resurgence is the argument that these comparisons can potentially help us answer important economic questions that extend beyond a descriptive analysis of mutual fund flows. For example, Barber, Huang and Odean (2016) (hereafter "BHO") compare the relation between fund flows and alphas measured with respect to various models to evaluate mutual fund investors' sophistication. They argue that sophisticated investors should use all common factors to compute alphas and evaluate fund performance regardless of the underlying true asset pricing model. BHO find that fund flows are more highly correlated with market model alphas than with other alphas. Because investors do not seem to be using alphas with respect to a model that includes all common factors, BHO conclude that investors in aggregate are not sophisticated in how they use past returns to assess fund performance.

Berk and van Binsbergen (2016) (hereafter "BvB") argue that such comparisons serve as a new and fundamentally different test of asset pricing models and that the results can determine which asset pricing model is the closest to the true asset pricing model in the economy. Because of the asset pricing model implications, they include several versions of equilibrium consumption-

CAPM as well in their comparisons. Agarwal, Green and Ren (2017) and Blocher and Molyboga (2017) carry out similar tests with samples of hedge funds.

BvB find that fund flows are most highly correlated with alphas computed with a market model in their tests as well. They conclude that therefore the CAPM "is still the best method to use to compute the cost of capital of an investment opportunity." Berk and van Binsbergen (2017) also prescribe that practitioners should use the CAPM to make capital budgeting decisions based on this evidence. The true asset pricing model has been a holy grail of the finance literature and BvB's conclusions potentially have broad implications that go well beyond just the mutual fund literature.

The far reaching inferences drawn in the recent literature based on comparisons of flow-alpha relations stand in contrast with the much more limited inferences drawn in the early literature. A natural question that arises is, under what assumptions can one draw reliable inferences about asset pricing models or about investor sophistication based on these results? Are the inferences about asset pricing models and investor sophistication in the recent literature empirically tenable?

We address these questions in this paper. We analytically show that one can draw reliable inferences about the true asset pricing model based on flow-alpha relations only if certain critically important assumptions are valid, and their validity can only be empirically determined. For example, it is possible that in some situations CAPM may not be true but investors may still optimally use the market model to estimate alphas. Also, in some other situations, it is possible that CAPM may be true but investors may optimally use a multifactor model to estimate alphas. There are also situations where investors may optimally use the market model to estimate alpha when CAPM is true, which would justify inferences about asset pricing model. Therefore, one cannot identify the true asset pricing model solely based on flow-alpha comparison without further tests to determine which of these multiple possibilities are true in the data.

We find similar issues with drawing inferences about investor sophistication as well. Sophisticated investors would use the model that yields the most precise alpha estimates. We show that the optimal model depends on the following factors: the underlying true asset pricing model, the incremental explanatory power of each factor in a multifactor model, the dispersion of factor betas across funds and the potential error in estimating factor betas. Our results indicate that this optimal model need not be the true asset pricing model, nor does it need to use all common factors to estimate betas. Therefore, the optimal model can only be empirically identified and we need the identity of this model to draw reliable inferences about investor sophistication based on flow-alpha relations.

We empirically assess whether inferences about asset pricing models and investor sophistication based on flow-alpha relations are tenable. Our tests estimate the relevant parameters from the data and run simulation experiments under various "true" asset pricing models. These tests enable us to determine the multifactor model that provides the most precise estimator of alphas in the data and assess the tenability of the inferences about asset pricing and investor sophistication in the literature.

1. Fund flows and alphas: Foundation for empirical tests and inferences

This section presents a model that forms the basis for our analysis of the implications of flow-alpha relations for asset pricing models and tests of investor sophistication. Broadly, we use the model to answer the following questions:

(a) How do investors optimally update their priors about the skills of fund managers when they observe fund returns each period?

(b) How are equilibrium fund flows related to the information investors use to update their priors?

(c) What are the implications of the answers to the above questions for interpreting the results of an alpha-fund flow horse race with alphas computed using different multifactor models?

We answer these questions using the Berk and Green (2004) model augmented with a multifactor return generating process and an equilibrium asset pricing model that we describe in the next subsection.

1.1 Return generating process and asset pricing model

The following K-factor model is the true asset pricing model:

$$E[r_i] = r_f + \sum_{k=1}^{K} \beta_{k,i}\gamma_k, \tag{1}$$

where $E[r_i]$ is the expected return on asset i, r_f is the risk-free rate, $\beta_{k,i}$ is the beta of asset i with respect to factor k, and γ_k is the premium for a unit of factor risk. If $K=0$ then all assets have same expected returns and $E[r_i] = E[r_m]$, where r_m is the market return. We refer to the model with $K=0$ as the "no-beta risk premium" (NBRP) model. For the CAPM, $K=1$ and for Fama-French three-factor model, which we refer to as FF3, $K=3$.

Asset returns follow the J-factor model below:

$$r_{i,t} = E[r_i] + \sum_{k=1}^{J} \beta_{k,i}\, f_{k,t} \, + \, \xi_{i,t}, \tag{2}$$

where $f_{k,t}$ is the realization of the common factor k, and $\xi_{i,t}$ asset specific return at time t. Factor realization $f_{k,t}$ is the innovation or the unexpected component of factor k. For instance, if $F_{k,t}$ is the total factor realization of the k^{th} factor then $f_{k,t} = F_{k,t} - E[F_{k,t}]$ and $E[f_{k,t}] = 0$. If this factor is traded and it is prices then $E[F_{k,t}] = \gamma_k$. The return generating process has J common factors and in general $J \geq K$, where K factors are priced and J-K factors are unpriced. Factor returns and asset specific returns are all normally distributed.

For ease of exposition, the analytics section assumes that sample average market beta, which we denote as $\bar{\beta}_{1,p} = 1$, and $\bar{\beta}_{k,p} = 0$ for $k > 1$. This assumption is mathematically true if the average mutual fund mirrors the market index, but to the extent funds deviate from the market index average fund betas would differ from the market average. However, our simulations use average betas that match the corresponding parameters in the sample.

1.2 Fund alphas and optimal signal

This subsection presents a rational expectations model that describes the relation between investors' assessment of fund manager skills and fund flows. We use the Berk and Green (2004) model augmented with an equilibrium asset pricing model and a return generating process described in the last subsection.

The model assumes the following:

(a) All agents in the rational expectations economy are symmetrically informed.

(b) The manager of fund *p* is endowed with stock selection skills that allow them to generate gross returns of ϕ_p^K in excess of the *K*-factor asset pricing benchmark. Investors know the true asset pricing model.

(c) Fund manager skill $\phi_p^K \sim N\left(\phi_0, \frac{1}{\nu}\right)$, where ϕ_0 is average skill and $\frac{1}{\nu}$ is the unconditional variance of skill across funds at time 0, and ν is the precision. ϕ_0 and ν are common knowledge.

(d) The cost of active management is *C(q)* where *q* is the size of the fund, and *C(q)* is a convex function of *q*. Therefore, $q \geq 0, C(q) \geq 0,\ C'(q) > 0$ and $C''(q) > 0$. Also, $C(q) = 0$ and lim $\lim_{q \to \infty} C'(q) = \infty$. Berk and Green (2004) argue that the cost per unit of fund would increase with fund size because of potentially larger price impact when funds trade larger positions and also because as fund size grows managers may run out of ideas and resort to closet indexing for part of their funds.

In addition to these costs, fund *p* charges investors a fee of F_p per unit. The total cost and fees per unit of the fund is $c(q) = {}^{C(q)}/_{q} + \mathrm{F}_p$. $c(q)$ is common knowledge.

(e) Let $R_{p,t}$ and $r_{p,t}$ be fund *p*'s gross and net returns at time *t*, respectively. $R_{p,t} = r_{p,t} + c(q_{t-1})$. Funds' net returns are observable, both to investors in the model economy and to econometricians. Investors can also compute $R_{p,t}$ since they know *q* and *c(q)* but econometricians observe only $r_{p,t}$.

(f) Eqs. (1) and (2) specify the expected returns and the return generating process in this economy, which are both common knowledge. The net return at time *t* is:[1]

$$r_{p,t} = \phi_p^K + \underbrace{r_f + \sum_{k=1}^{K} \beta_{k,p}\gamma_k}_{\text{Expected return Eq. (1)}} + \underbrace{\sum_{k=1}^{J} \beta_{k,p}\, f_{k,t} + \xi_{p,t}}_{\text{Unexpected return Eq. (2)}} - c(q_{t-1}), \tag{3}$$

Investors have a diffuse prior at time 0 about the skills of all funds at time *t=0*, and hence investors expectation of skill is ϕ_0 for all funds. Berk and Green (2004) specify that investors use fund returns in excess of their benchmarks to update their priors about skill.

[1] Funds' gross returns follow the return generating process (2), plus ϕ_p^K. Investors earn net returns in (3) after all costs.

This assumption is adequate for their purposes, but since our objective is to understand the inferences one can draw from fund flow and alpha relations, investors in our model estimate alphas with respect to the particular factor model that enables them to optimally update their priors.

If investors use an η-factor model, then fund p's alpha at time t, say $\hat{\alpha}_{p,\eta}$, is:

$$\hat{\alpha}_{p,\eta,t} = \left(r_{p,t} - r_{f,t}\right) - r_f - \sum_{k=1}^{\eta} \beta_{k,p} F_{k,t}\,, \tag{4}$$

where $F_{k,t}$ is realized factor returns. If $\eta = K$, then Eqs. (1) and (2) imply:

$$\hat{\alpha}_{p,K,t} = \sum_{k=K+1}^{J} \beta_{k,p}\, f_{k,t} \, + \, \xi_{p,t}. \tag{5}$$

We can decompose Eq. (4) as follows:

$$\hat{\alpha}_{p,\eta,t} = \begin{cases} \underbrace{-\sum_{k=K+1}^{\eta} \beta_{k,p} E(F_k)}_{Model\ Misspcification\ error} \; + \; \underbrace{\sum_{k=\eta+1}^{J} \beta_{k,p}\, f_{k,t} \, + \, \xi_{p,t}}_{Estimation\ error} & \text{if } \eta \geq K \\ \underbrace{\sum_{k=\eta+1}^{K} \beta_{k,p} E(F_k)}_{Model\ Misspcification\ error} \; + \; \underbrace{\sum_{k=\eta+1}^{J} \beta_{k,p}\, f_{k,t} \, + \, \xi_{p,t}}_{Estimation\ error} & \text{if } \eta < K \end{cases} \equiv \theta_{\eta,p} + \varepsilon_{\eta,p}. \tag{6}$$

The first part of the equation is model misspecification error, which we denote by $\theta_{\eta,p}$, which is zero if $\eta = K$. If $\eta > K$, we mistakenly attribute premium for risks that are not truly priced in the economy and therefore we add noise to our alpha estimate. For example, if CAPM were the true model but we use FF3 to compute alphas, we mistakenly assume that funds with positive HML or SMB command bigger expected returns than their true expected returns. This misspecification adds to alpha estimation error. The second part of the equation is statistical estimation error which we denote by $\varepsilon_{\eta,p}$.

How does the precision of alpha estimator affect the decision of investors about how they should optimally update their priors about fund manager skill each period and arrive at their investment decisions? The following proposition presents the distribution of investors' posterior each period conditional on using a particular η-factor model to compute alphas. As we state in the corollary to the proposition, investors would choose the η-factor model that yields the most precise posterior.

Proposition 1: Suppose investors use an η-factor model to compute alphas. Let $\phi^K_{p,\eta,t}$ be the mean of investors' time *t* posterior of fund k's skill conditional on the realization of $X_{p,\eta,1}, X_{p,\eta,2}$, ... $X_{p,\eta,t}$, where $X_{p,\eta,t} = \hat{\alpha}_{p,\eta,t} + c(q_{t-1})$, and let $\bar{X}_{p,\eta,t}$ be the mean of these realizations. Investors' posterior of ϕ^K_p is normally distributed with mean $\phi^K_{p,\eta,t}$, where:

$$\phi^K_{p,\eta,t} = \frac{\nu\, \phi_0 + t\vartheta_{\hat{\alpha},\eta}\, \bar{X}_{p,\eta,t}}{\nu + t\vartheta_{\hat{\alpha},\eta}}, \tag{7}$$

and precision $\nu + t\vartheta_{\hat{\alpha},\eta}$, where $\vartheta_{\hat{\alpha},\eta} = \frac{1}{\sigma^2_{\hat{\alpha},\eta}}$. Note that the precisions of $X_{p,\eta,t}$ and $\hat{\alpha}_{p,\eta,t}$ are equal conditional on information available at time *t-1* since $c(q_{t-1})$ is known at that time.

Proof: See Theorem 1 in DeGroot (1970, p. 167).

Berk and Green (2004) also use this theorem to show that ϕ_t and $c(q_t)$ satisfy the recursive relations

$$\begin{aligned} \phi_t &= \phi_{t-1} + \frac{\omega}{\gamma + t\omega} \times \hat{\alpha}_{p,\eta,t}, \\ c(q_t) &= c(q_{t-1}) + \frac{\omega}{\gamma + t\omega} \times \hat{\alpha}_{p,\eta,t} \end{aligned} \tag{8}$$

Equation (8) shows how investors update their priors each period based on the new information they get at time *t* (i.e. in $\hat{\alpha}_{p,\eta,t}$).

In Eqs. (7) and (8) investors update their priors using alphas computed with net returns although skill allows them to generate gross returns. The reason is that $c(q_{t-1})$ is known at time *t-1* and therefore new information at time *t* is entirely captured by alphas.

Our model is the same as Berk and Green (2004) except Berk and Green (2004) assume that the only new information in fund return in each period is the difference between fund returns and benchmark returns. However, we allow for the fact that investors can potentially extract more information from realized returns if they use the information in the return generating process.

Corollary: Investors would use the η-factor model with the smallest variance (or largest precision) to revise their priors about fund skills.

We denote the number of factors in the factor model with the smallest total error as η^*.

1.3 Alphas and Fund Flows

We use the results from Berk and Green (2004) to describe the relation between fund flows and the signals that investors use to update their priors about fund manager skills. In Berk and Green, the mutual fund market is perfectly competitive. Therefore, expected alpha net of fees and costs for investing in any mutual fund equals zero in equilibrium:

$$E_t(r_{p,t+1}) \equiv \phi_p - c(q_t) = 0 \tag{9}$$

Each period, investors form their posteriors using Equation (8) and the information conveyed by $\hat{\alpha}_{p,\eta^*,t}$ to update their priors about ϕ_p. Equation (9) market forces would ensure that the amount of funds that flow into or out of mutual funds result in changes in $c(q_t)$ that exactly offset any revisions in skill assessments (i.e. $\phi_{p,t} - \phi_{p,t-1}$). Berk and Green (2004) then derive the relation between the signal investors use to update their priors in Eq. 8 (i.e. $\hat{\alpha}_{p,\eta^*,t}$) and net flow of funds following updates to their priors. Their results indicate that fund flows positively covary with $\hat{\alpha}_{p,\eta^*,t}$. The exact functional form of this relation is not important for our purposes, and we formally state the positive covariance in the proposition below:

Proposition 2: Let $\Gamma_{p,t}$ be the net inflow of funds into fund *p* at time *t*. The covariance between the signal that investors use to update their priors at time *t* and $\Gamma_{p,t}$ is positive:

$$Cov(\hat{\alpha}_{p,\eta^*,t}, \Gamma_{p,t}) > 0 \tag{10}$$

Proof: See Figure 1 in Berk and Green (2004).

Equation (35) in Berk and Green is the basis for this proposition. This equation is quadratic with positive slope for $\hat{\alpha}_{p,\eta^*,t} > 0$, but the slope is hard to compute analytically for $\hat{\alpha}_{p,\eta^*,t} < 0$, because of their boundary conditions that funds liquidate if investors' posteriors are below a critical value. Therefore, we rely on the numerical results that they present in their Figure 1 and their stated conclusions in the paper to prove this proposition. This proposition basically states that bigger alphas would attract more funds than smaller alphas and it makes economic sense. What is more important for our purposes is that Berk and Green's results directly relate the signals investors use to update their priors to fund flows in equilibrium.[2]

1.4 Alpha-fund flows horse race

To compare alphas-flow relations across models, consider the following univariate cross-sectional regression specification used in the literature:

$$\Gamma_{p,t} = a_{\eta,t} + b_{\eta,t} \times \hat{\alpha}_{p,\eta,t} + \omega_{p,\eta,t}. \tag{11}$$

The estimate $\hat{b}_\eta$ measures the strength of the alpha-flow relation of each estimator. The literature runs a horse race among multifactor models based on $\hat{b}_\eta$. What can we learn about the true asset pricing model or about investor sophistication based on this horse race?

The slope coefficient estimate is:

$$plim\ b_\eta = \frac{Cov(\Gamma_p, \hat{\alpha}_{p,\eta,t})}{\sigma^2_{\hat{\alpha}_{p,\eta,t}}} \tag{12}$$

where $\sigma^2_{\hat{\alpha}_{p,\eta,t}}$ is cross-sectional variance of $\hat{\alpha}_{p,\eta,t}$. As the Corollary to Proposition 1 states, investors optimally use the η^*-factor model to estimate alphas. We can therefore express other estimators as:

$$\hat{\alpha}_{p,\eta,t} = \hat{\alpha}_{p,\eta^*,t} + \zeta_{p,\eta,t}, \tag{13}$$

where $Var(\zeta_{p,\eta,t}) > 0$ for $\eta \neq \eta^*$, and $Cov(\zeta_{p,\eta,t}, \hat{\alpha}_{p,\eta^*,t}) = 0$.

In equilibrium, flow positively covaries with $\hat{\alpha}_{p,\eta^*,t}$ as Berk and Green (2004) show and we restate in Proposition 2 and the covariance is not through the noise in less efficient estimators. Therefore,

$$Cov(\Gamma_p, \hat{\alpha}_{p,\eta,t}) = Cov(\Gamma_p, \hat{\alpha}_{p,\eta^*,t}), \text{ and} \tag{14}$$

[2] Roussanov, Ruan and Wei (2018) present a modified model that incorporates investors search costs and mean reversion in skill. Their model also yields a similar positive correlation.

$$plim\ b_\eta = \frac{Cov(\Gamma_p, \hat{\alpha}_{p,\eta^*,t}\ )}{\sigma^2_{\hat{\alpha}_{p,\eta^*,t}} + Var(\zeta_{p,\eta,t})}$$

Since $Var(\zeta_{p,\eta,t}) > 0$ for $\eta \neq \eta^*$, and b_{η^*} will win this horse race in a rational economy. Therefore, to determine what implications we can draw from this horse race about asset pricing models and investor sophistication, we need to identify η^* based on the parameters in the data.

Note that our analysis does not require that the true relation between alpha and flow be linear. In fact, Berk and Green (2004) show that the equilibrium relation is non-linear. We only use the fact that in equilibrium fund flow positively covaries with the signal investors use to update their priors.

1.5 Precision of alpha estimate

Equation (6) decomposes the measurement error in alphas into model misspecification error and statistical estimation error. The model misspecification error depends on the true model which the econometrician does not observe. However, we can analytically determine the variance of the estimation error component. As a starting point, suppose betas are estimated without error. Recall that $\varepsilon_{\eta,p} = \sum_{k=\eta+1}^{J} \beta_{k,p} f_{k,t} + \xi_{p,t}$. Therefore,

$$\sigma^2_{\varepsilon_{\eta,p}} = \sigma^2_{r_p}\left(1 - R^2_{adj,\eta,p}\right) \tag{15}$$

where $\sigma^2_{r_p}$is the variance of fund returns and $R^2_{adj,\eta,p}$ the fraction of fund return variance that is explained by η factors with appropriate adjustment for degrees of freedom. Suppose $\varepsilon_{\eta,p}$ is uncorrelated across funds.[3] Then statistical estimation error is the average of $\sigma^2_{\varepsilon_{\eta,p}}$across funds. Therefore, Eq. (15) indicates that any common factors that increases R^2_{adj}, whether that factor is priced or unpriced, would reduce estimation error.

Measurement errors in betas would also add to statistical estimation error and affect the choice of factors that one would include in computing alphas. For instance, a factor that may

[3] For $\eta < J$, the error terms will be correlated across funds and our simulation allow for such correlations. Also, our simulations allow R^2_{adj} to vary across funds.

marginally increase R^2_{adj} may still not be desirable if the measurement error in beta with respect to that factor increases the alpha estimation error. This issue is particularly important if that factor is correlated with other factors in the regression because the addition of that factor would increase the measurement errors of other factor betas as well.

There are two potential sources of measurement error in betas. Even if true betas were constant, beta estimates using a time-series would contain statistical estimation errors. Additionally, fund betas would vary over time because individual stock betas may be time-varying and active funds typically revise their portfolios over time. Therefore, the difference between the true betas at time *t+1* and the average beta during the estimation period would add to the measurement error in betas.

Eq. (6) indicates that if $\eta < K$, each factor that we omit from K adds to both misspecification error and to estimation error. Therefore, if betas with respect to a priced factor are known without error then inclusion of that factor would reduce measurement error. However, if $\eta > K$, each additional factor would reduce the estimation error, but add to misspecification error. Therefore, whether investors would optimally include these additional factors depend on the relative contribution to these components in the data, which can only be empirically determined.

1.6 CAPM vs. No-beta risk premium model: An illustrative example

This subsection considers an example that illustrates the contribution of $\sigma^2_{\varepsilon_\eta}$ and $\sigma^2_{\theta_\eta}$ to precision of the alpha estimates. Suppose asset returns are generated by the following single factor model:

$$r_{p,t} = E[r_p] + \beta_p \times f_t + \xi_{p,t}. \tag{16}$$

Expected returns are determined by one of the following two models:

i. NBRP model: The expected returns on all stocks are equal, i.e.

$$E[r_p] = E[r_m] \ \forall \ p \tag{17}$$

where $E[r_m]$ is the expected return on the market portfolio.

ii. CAPM:

$$E[r_p] = r_f + \beta_p(E[r_m] - r_f) \tag{18}$$

Consider the following two Estimators of alpha:

- Market adjustment (No-beta risk model):

$$\hat{\alpha}_{p,0} = r_{p,t} - r_{m,t} \tag{19}$$

- Market model adjustment (CAPM):

$$\hat{\alpha}_{p,1} = r_{p,t} - [r_f + \hat{\beta}_p(r_{m,t} - r_f)] \tag{20}$$

where $\hat{\beta}_p$ is computed using market model regression.

The variance of measurement errors of $\hat{\alpha}_{p,0}$ and $\hat{\alpha}_{p,1}$, which include both model misspecification error and statistical estimation error are tabulated below (Appendix 1 presents the derivations):

	Alpha Estimator	
True model:	a. Estimated with Mkt Adj. (Eq. 19)	b. CAPM (Eq. 20)
i. No-beta risk premium model	$\sigma_u^2 \vert r_{m,t} = \sigma_\beta^2 (r_{m,t} - E[r_m])^2 + \sigma_\xi^2 \vert r_{m,t}$	$\sigma_u^2 \vert r_{m,t} = \sigma_\beta^2 E(r_m)^2 + \sigma_{\hat{\beta}-\beta}^2 r_{m,t}^2 + \sigma_\xi^2 \vert r_{m,t}$
ii. CAPM	$\sigma_u^2 \vert r_{m,t} = \sigma_\beta^2\, r_{m,t}^2 + \sigma_\xi^2 \vert r_{m,t}$	$\sigma_u^2 \vert r_{m,t} = \sigma_{\hat{\beta}-\beta}^2 r_{m,t}^2 + \sigma_\xi^2 \vert r_{m,t}$

The variables in the table above are:

Variables	Definition

$\sigma_u^2 \mid r_{m,t}$	Variance of total measurement error conditional on the realization of market return i.e. $\sigma_u^2 \mid r_{m,t} = \sigma_{\hat{\alpha}-\alpha}^2 \mid r_{m,t}$
σ_β^2	Variance of true beta across funds.
$\sigma_{\hat{\beta}-\beta}^2$	Variance of measurement error across funds both due to the standard error of regression estimates and also due to time-variation in beta.
$\sigma_{\hat{\beta}}^2$	Variance of $\hat{\beta}_p$ across funds $= \left(\sigma_\beta^2 + \sigma_{\hat{\beta}-\beta}^2\right)$
$\sigma_\xi^2 \mid r_{m,t}$	Variance of fund specific returns (assumed to be the same for all funds for expositional convenience) conditional on the realization of market return

The results in the above table illustrate the factors that contribute to total measurement error and the inherent trade-offs. For example, the term in cell (i)(b) can be grouped as:

$$\sigma_u^2 \mid r_{m,t} = \underbrace{\sigma_\beta^2 E[r_m]^2}_{Model\ Misspecification\ Error} + \underbrace{\sigma_{\hat{\beta}-\beta}^2 r_{m,t}^2 + \sigma_\xi^2 \mid r_{m,t}}_{Estimation\ Error} \quad (21)$$

The first term in this expression is variance of model misspecification error, which arises because of using market model adjustment in equation (20) when 'no-beta risk' model is true. The last two terms are due to statistical estimation error.

To consider the trade-offs between model misspecification error and estimation error, consider the last row where CAPM is true. The variances of estimation errors in alpha using Equations (19) and (20) are given in the last row of the table, and they both contain the term σ_ξ^2. The variance of alpha estimated with Equation (19) contains the additional term $\sigma_\beta^2 r_{m,t}^2$, which is the cross-sectional variation of true fund beta, and that with Equation (20) contains the term $\sigma_{\hat{\beta}-\beta}^2$, which is the variance of measurement error in beta. If the beta estimates are sufficiently noisy (i.e. big $\sigma_{\hat{\beta}-\beta}^2$) or if differences in betas across funds are small, then the variance of measurement error with Eq. (19) could be smaller than with estimator (20). In this case, we can infer from equation

(14) that the slope coefficient b_η in equation (11) would be bigger for the market adjusted $\hat{\alpha}$ from estimator (19) compared to the market model $\hat{\alpha}$ from estimator (20). In other words, estimate of alpha using Eq. (19) would win out in a horse race of slope coefficients against alpha estimated with Eq. (20) even when CAPM is true (a counterexample to the underlying assumption in BvB), and even if investors were truly sophisticated (a counterexample to the underlying assumption in BHO because investors optimally do not use all factors in the return generating process). Of course, this is only an illustrative example, and we should empirically examine the true parameters to understand what we can learn from the horse races.

2. Simulation Experiment

BVB hypothesize that the winner of the alpha-fund flow horse race is the true asset pricing model, but BHO hypothesize that the winner would be the model that includes all priced and unpriced factors if investors are sophisticated. We can formally state their hypotheses as follows:

Suppose the true asset pricing model is a K-factor asset pricing model and returns are generated by a J-factor model. When we fit regression (11) with alpha computed with each η-factor model where $0 \leq \eta \leq J$, the biggest slope coefficient obtains when η=η^* i.e. when $\hat{\alpha}_{p,\eta^*}$ is computed with respect to an η^*-factor model.

A1. Asset Pricing test hypothesis: The model that yields the biggest correlation is the true asset pricing model, i.e. η^*=K.

A2. Investor Sophistication hypothesis (BHO): The most accurate model is the J-factor model that generates asset returns, i.e. η^*=J.

However we show in Section 1 the winner need not necessarily be a *K* or *J* factor model because the winner broadly depends on the following factors: (i) extent to which various factor models explain fund returns (i.e. model R^2_{adj}), (ii) beta estimation error $\left(\sigma^2_{\hat{\beta}-\beta}\right)$, (iii) variation of betas across funds $\left(\sigma^2_{\beta}\right)$ and (iv) the "true" asset pricing model. Therefore, the winner would depend on the characteristics of the data, and we can only empirically identify η^*.

To do so, we can estimate the first three of the four items we list above from the data but we do not know the "true" asset pricing model. Therefore, we estimate the first three items and use these parameters to generate simulated returns under each asset pricing model. We then run the horse race with regressions (11) in the simulation to determine which factor model would win the race in a rational expectations economy, which in turn would inform us the implications we can draw from the horse race.

2.1 Data and Simulation parameters

We estimate the parameters for the simulation with the sample of funds in the CRSP survivor-bias free mutual fund database. Our sample includes all actively managed domestic equity funds in the January 1990 to June 2017 sample period. Our sample is comprised of all actively managed domestic equity funds. CRSP identifies these funds with objective codes 'EDC' and 'EDY.' When a fund has multiple share classes, we add assets in all share classes to compute its TNA and we compute fund level return as the weighted average of returns of individual share classes with lagged TNA as weights. The sample for month *t* includes all funds with at least $10 million assets under management as of the end of month *t-1*. We follow BHO and exclude funds that had flows smaller than -90% or greater than 1000% in any month from the sample to avoid the effect of outliers. The sample for month *t* includes only funds that have returns data in all months from *t*-61 to *t*-1 to estimate betas.[4]

Table 1 presents the summary statistics for the funds in our sample. The sample is comprised of 1224 funds per month on average. The average monthly fund flow into a fund is 0.25% of its TNA the previous month. Around half of the funds in the sample have either an entry or exit load.

2.2 R^2_{adj} and beta measurement error: A first look

We use the seven factor model from BHO as the *J*-factor model that generates returns. The seven factors are the three Fama-French factors (market $(mkt - r_f)$, SMB and HML), Carhart (1996) momentum factor (UMD), and three industry factors (IND_1, IND_2 and IND_3). Following BHO, we construct the three industry factors as the first three principal components of residuals from regressing Fama-French 17 equal weighted industry portfolios on FFC4 factors.

[4] This sample selection criterion excludes funds from the sample during the first 60 months of their existence. Therefore, our sample is not exposed to potential incubation bias that Evans (2010) and Elton, Gruber and Blake (2001) document.

Before we proceed with the simulation, we take a first look at some of the determinants of the accuracy of alpha estimates. One important determinant is the incremental explanatory power of each additional factor. We fit the following time series regression with η factors each month *t* using data for each fund from months t-60 to t-1 and compute average R^2_{adj} for each model:

$$r_{p,\tau} = a_{p,\eta,t} + \sum_{k=1}^{\eta} \beta_{k,p,t}\, F_{k,\tau} + \, e_{p,\eta,\tau}, \qquad \tau = t - 60 \text{ to } t - 1. \tag{22}$$

Table 2 reports the time-series averages. For the market model, we compute R^2_{adj} as $1 - \left(\sum(r_{p,t} - r_{mkt})^2 / \sum(r_{p,t} - \bar{r}_p)^2\right)$.

Market-adjusted returns have the lowest R^2_{adj} of .774. The R^2_{adj} for the single factor market model is bigger at .820. R^2_{adj} increase to .892 for the Fama-French three-factor model, but the increase is fairly gradual as we go from the Fama-French three-factor model to the seven factor model.

Another important component in the measurement error of $\hat{\alpha}$ is the variance of measurement error in betas across funds $\left(\sigma^2_{\hat{\beta}-\beta}\right)$. The term $\sigma^2_{\hat{\beta}-\beta}$ would differ from the time series variance of OLS estimation error in regression (11) for two reasons. First, if the fund-specific returns are correlated across funds, then the average variance of time-series errors will not equal $\sigma^2_{\hat{\beta}-\beta}$. Secondly, as we discussed earlier the OLS estimates are unbiased estimates of mean betas during the estimation periods and any difference between this average and the realized beta in month *t+1* is an additional source of measurement error.

To estimate the magnitude of this error we first estimate the following regressions for each fund for each month:

$$\begin{aligned} (r_{p,\tau} - r_{f,\tau}) &= \alpha^{past}_{p,k,t} + \beta^{past}_{p,k,t}\, F_{k,\tau} + \, e_{p,k,\tau} \qquad & \tau = t - 60 \text{ to } t - 1, \\ (r_{p,\tau} - r_{f,\tau}) &= \alpha^{future}_{p,k,t} + \beta^{future}_{p,k,t}\, F_{k,\tau} + \, e_{p,k,\tau} \qquad & \tau = t \text{ to } t + 11 \end{aligned} \tag{23}$$

where $F_{k,\tau}$ is the factor with respect to which betas are estimated. Suppose betas for a particular fund are constant over time.

$$\hat{\beta}^{past}_{p,k,t} = \beta_{p,k} + u^{past}_{p,k,t}, \text{ and} \tag{24}$$

$$\hat{\beta}_{p,k,t}^{future} = \beta_{p,k} + u_{p,k,t}^{future}$$

where $\beta_{p,k}$is fund p's true beta with respect to factor k.

Consider the following cross-sectional regression for month t:

$$\hat{\beta}_{p,k,t}^{future} = a_t + b_t \times \hat{\beta}_{p,k,t}^{past} + e_{p,t} \tag{25}$$

Since we use non-overlapping sample periods to estimate $\beta_{p,k,t}^{past}$ and $\beta_{p,k,t}^{future}$, $u_{p,k,t}^{past}$ and $u_{p,k,t}^{future}$ are uncorrelated. With a sufficiently large number of funds, the probability limit of the slope coefficient is:

$$\text{plim}\, b_t = \frac{var(\beta_{p,k})}{var(\beta_{p,k}) + var\left(u_{p,k,t}^{past}\right)} \tag{26}$$

Therefore, the slope coefficient of regression (24) is the ratio of the cross-sectional variance of the factor betas divided by the sum of this variance plus the variance of the measurement error. If this slope coefficient is smaller than 0.5 then the variance of true beta is smaller than the variance of measurement error.

We fit regression (24) each month for each of the betas. All betas are estimated using univariate regressions as per equation (23). Table 3 reports the time-series averages of the slope coefficients for each beta. The slope coefficients are all greater than .75 for betas with respect to the three Fama-French factors, but they are less than .5 for UMD and industry factors. Therefore, the variance of measurement error is bigger than the variance of true betas for the latter set of factors.

2.3 Simulation: Experimental design

To understand how the true asset pricing model and estimation error in alphas impact the outcome of the alpha-flow horse race regressions, we simulate a mutual fund economy with parameters that match the actual sample of domestic equity funds described in section 2.1. and we match the entry and exit of funds in the simulation to that in the actual data. In this simulated economy, the fund size evolves over time with flows, net returns generated from managerial skill and net returns from passive factor exposures. And fund size affects net returns through its effect on costs.

The sample of mutual funds and their TNA evolve as follows in the simulation:

a. **Fund origin:** We start the simulation with the number of funds equal to that in the sample on January 1985. At origin, the TNA of all funds is \$10.

b. **Skill:** The average four factor alpha in our actual sample of domestic equity funds, gross of fund fees F_p, is around 5 bps per month. To account for unobservable costs $C(q)/q$, we add 10 bps per month to this estimate to account for average transaction costs[5]. We randomly draw ϕ_p for each fund from a normal distribution with mean equal to 0.15% and standard deviation of 0.2% per month.[6]

c. **Betas**: We randomly generate the seven factor betas for each fund from a normal distribution with means and standard deviations equal to the parameters tabulated in Table 4.[7] Each factor beta is drawn independently and is constant over the entire sample period.

d. **Fund specific return**: We generate monthly fund specific return $\epsilon_{p,t}$ for each fund from a normal distribution with mean zero and standard deviation equal to 2.5%.

e. **Asset pricing model and expected returns**: Steps (b) through (d) describe the return generating process for the funds and this process does not vary with the asset pricing model. However, different common factors that are priced vary across asset pricing models and hence different asset pricing models imply different expected return for each fund. The term $E^{model}(r_p - r_f)$ is the "true" expected excess return and it depends on the model. We conduct simulations under three asset pricing models and expected excess returns under each model are computed as follows:

    - NBRP risk model: $E^{NR}(r_p - r_f) = 0.699\%,$ (27)
    - CAPM: $E^{CAPM}(r_p - r_f) = \beta_{p,m} \times (\overline{mkt - r_f}),$

[5] Elton et. al. (2012) report that the transaction costs are of the same order of magnitude as expense ratios which average to around 10 bps per month.

[6] The monthly cross-sectional variance of $\hat{\alpha}$s in the real data is the variance of true alphas plus the measurement error of alphas. The measurement error variance in $\hat{\alpha}$s is the squared OLS standard errors from the time-series regressions used to estimate alphas. The average standard deviation of the difference across models is roughly 0.2% per month.

[7] As Eq. (26) shows, the standard deviation of true beta distribution in the data is the standard deviation of estimated beta multiplied by the square root of the respective slope coefficients in Table 3.

- Fama-French three factor model (FF3): $E^{FF3}(r_p - r_f) = -0.016\% + \beta_{p,m} \times (\overline{mkt - r_f}) + \beta_{p,smb} \times (\overline{SMB}) + \beta_{p,hml} \times (\overline{HML}),$

The overbars above common factor returns indicate sample means. The constant in the equation for each model is chosen so that the average fund returns equal sample average of market excess returns.

f. **Gross returns:** We generate fund returns using the following seven-factor model:

$$R_{p,t} = \phi_p + E^{model}(r_p) + \beta_{p,m} \times \left(\widetilde{mkt - r_f}\right)_t + \beta_{p,smb} \times \widetilde{SMB}_t + \beta_{p,hml} \times \widetilde{HML}_t + \beta_{p,umd} \times \widetilde{UMD}_t + \beta_{p,ind1} \times \widetilde{IND1}_t + \beta_{p,ind2} \times \widetilde{IND2}_t + \beta_{p,ind3} \times \widetilde{IND3}_t + \epsilon_{p,t} \tag{28}$$

g. **Cost function**: Following Berk, Green (2004), we specify cost per unit size as $c(q_{t-1}) = \delta \times q_{t-1}$, where q represents the Total Net Assets (TNA) and the parameter δ captures decreasing returns to scale. We set δ = 0.2 bps/$ 100 mn, which closely matches the value reported in Table 3 of Pastor, Stambaugh, Taylor (2015).

h. **Net returns**: We compute net returns as $r_{p,t} = R_{p,t} - c(q_{t-1})$.

i. **Fund flow**: For each month, we compute flows using the flowing equation:

$$flow_{p,t} = a + b \times \hat{\alpha}_{p,\eta^*,t} + \psi_{p,t}. \tag{29}$$

We estimate *a* and *b* from the data, and our estimates are a = -0.00225 and b =.2, using $\eta^* = 7$. In the simulation, we draw $\psi_{p,t}$ from a normal distribution with mean zero and standard deviation of 0.09 (9%). All these parameters match the corresponding parameters in the data.

j. **Fund exit and entry:** If the number of funds in the data in month *t* is smaller than the number of funds in month *t-1*, the appropriate number of funds exit the simulation sample as well. We sort funds in the simulated sample based on their TNA at the end of *t-1* and drop the bottom most funds (i.e. least TNA) equal in number to the actual exits for that month. If the number of funds in the data in month *t* is bigger than the number of funds in month *t-1*, the appropriate number of funds enter the sample with TNA of $10 mn.

We repeat the simulation 50 times.

2.4 Simulation: Tests and results

We first examine the relation between alphas and fund flows under various models. We conduct two sets of tests. In the first set of tests, we use true betas and compute $\hat{\alpha}_{p,\eta,t}$ using Eq. (4). In the other set of tests, we examine the effect of beta measurement error on the choice of optimal factor model. For this set of tests, we estimate betas for each month *t* with simulated returns using Regression (22).[8] We then compute alpha for month *t* as follows:

$$\hat{\alpha}_{p,\eta,t} = r_{p,t} - \sum_{k=1}^{\eta} \hat{\beta}_{k,p,t} F_{k,t}\,, \tag{30}$$

where $\hat{\beta}_{k,p,t}$ is the time *t* estimate of beta.

We first examine the components of measurement error in compute $\hat{\alpha}_{p,\eta,t}$. Allowing for measurement error in betas, we can generalize Eq. (6)

$$\hat{\alpha}_{p,\eta,t} = \begin{cases} \underbrace{-\sum_{k=K+1}^{\eta} \beta_{k,p} E(F_k)}_{Model\ Misspcification\ error} + \underbrace{\sum_{k=\eta+1}^{J} \beta_{k,p}\, f_{k,t} + \sum_{k=1}^{\eta} \left(\hat{\beta}_{k,p,t} - \beta_{k,p}\right) F_{k,t} + \xi_{p,t}}_{Estimation\ error} & \text{for } \eta \geq K, \\ \underbrace{\sum_{k=\eta+1}^{K} \beta_{k,p} E(F_k)}_{Model\ Misspcification\ error} + \underbrace{\sum_{k=\eta+1}^{J} \beta_{k,p}\, f_{k,t} + \sum_{k=1}^{\eta} \left(\hat{\beta}_{k,p,t} - \beta_{k,p}\right) F_{k,t} + \xi_{p,t}}_{Estimation\ error} & \text{for } \eta < K \end{cases} \tag{31}$$

The term $\sum_{k=1}^{\eta} (\hat{\beta}_{k,p,t} - \beta_{k,p})\, F_{k,t}$ is the error due to measurement error in betas. Therefore,

$$Var(\hat{\alpha}_{p,\eta}) = Var(\theta_{\eta}) + Var(\varepsilon_{\eta}) + Cov(\theta_{\eta}, \varepsilon_{\eta}), \tag{32}$$

where θ_{η} and ε_{η} denote the misspecification error and estimation error components.

To estimate the variances in Eq (32), we first compute the values of $\theta_{\eta}, \epsilon_{\eta}$ for different η- and K-factor models in our simulated sample based on the analytical expressions in Eq (31). Using

[8] Since we generate excess returns in simulations, Eq. (22) does not use risk-free rate.

these values, we compute the monthly cross-sectional variances of θ, ϵ as well as their covariance. We then average these values across time to get the required estimates.

Table 5 presents the components of alpha estimation error variance for each asset pricing model and η-factor model. Consider the results when true betas are known. Estimation error variance decreases monotonically as we increase η from zero to seven for all asset pricing models. For example, under the CAPM, estimation error variance is 875 for $\eta = 0$, which reduces to reduces to 625 for $\eta = 0$.

The model misspecification error variance increases monotonically as we move away from the true asset pricing model. However, model misspecification error variance is an order of magnitude smaller than estimation error variance. For instance, the smallest estimation error variance is 625 and in comparison the largest model misspecification error variance is 1.75, which is about 2.8% of 625.

The total error variance also monotonically declines as we increase the number of factors, which is similar to the pattern we see for the estimation error. Model misspecification error is so small in all instances that it hardly moves the needle. Therefore, if we can observe betas without error then the *J*-factor model is the optimal model, as long as each factor increases the R^2_{adj}.

When betas are not known and we estimate betas using data for 60 months, estimation error variance exhibits a U-shaped pattern. It decreases as we go from $\eta = 0$ to FF3 and then increases monotonically as we add more factors. For example, under the CAPM, estimation error variance decreases from 874.5 for $\eta = 0$ to 702.4 for $\eta = 3$, but then increases to 760.4 for $\eta = 7$. As we saw in Table 3, beta measurement error is relatively large for UMD and the three factor industry factors. Consequently, accounting for these factors to compute alpha increases estimation error variance. As before, model misspecification error is so small that it does not make difference when we compare total estimation error across models.

We next fit Regression (11) each month and estimate the coefficients and standard errors using the Fama-MacBeth approach. Table 6 reports the results. As we showed analytically, the magnitude of the slope coefficients across models would be negatively related to the precision of alpha estimates, and we see this pattern in Table 6. Without beta measurement error, the slope coefficient increases monotonically as we add factors. For example, under the CAPM, the slope increases from 13.93 for $\eta = 0$ to 19.97 for $\eta = 7$. The average regression R^2 also increases from .29 to .39.

The ordering of the slope coefficients across models when we estimate betas from the data is exactly the opposite of the ordering of estimation error variance in Table 5. For all asset pricing models, we find the biggest slope coefficients for $\eta = 3$. The slope coefficients for $\eta = 4$ and 7 are comparatively smaller, but the magnitude of the difference is not large. For example, with CAPM, the slope coefficient is 17.81 for $\eta = 3$ and 17.55 for $\eta = 4$, and it only reduces to 17.32 for $\eta = 7$.

The slope coefficients are almost identical under different true asset pricing models both when we know the true beta and when we estimate beta from the data. These results indicate that the relation between alpha and flow is not particularly sensitive to the true asset pricing model. Therefore, one cannot identify the true asset pricing model using the alpha-flow horse race.

3. Binary variable regression

Our analysis in the last section uses a linear regression for the alpha-fund flow horse race. However, Berk and Green (2004) show that the equilibrium relation between alpha and fund flows nonlinear. Because of the non-linearity, BvB transform flows and alpha estimates to binary variables and run the horse race with these transformed variables. Specifically, the transformed binary variables are defined as follows:

$$Q_x = \begin{cases} 1 \text{ if } x \geq 0 \\ -1 \text{ if } x < 0 \end{cases} \tag{33}$$

where x is any random variable. BvB run the following OLS regression:

$$Q_{\Gamma_p} = A_\eta + B_\eta \times Q_{\hat{\alpha}_{p,\eta}} + o_{p,\eta}, \tag{34}$$

and compare $\hat{B}_\eta$. To relate our analysis based on Regression (11) to that based on Regression (34), we first establish the following propositions:

Proposition 3: Let $\hat{\alpha}_{p,\eta_1}$ and $\hat{\alpha}_{p,\eta_2}$ be the alphas computed with respect to η_1- and η_2-factor models using Equation (4). $\hat{b}_{\eta_1}$ and $\hat{b}_{\eta_2}$ are the corresponding Regression (11) slope coefficients and $\hat{B}_{\eta_1}$ and $\hat{B}_{\eta_2}$ are the corresponding Regression (34) slope coefficients. Under the augmented Berk and Green model,

If $\hat{b}_{\eta_1} > \hat{b}_{\eta_2}$ then $\hat{B}_{\eta_1} > \hat{B}_{\eta_2}$, when the number of funds in the sample is sufficiently large.

Proof: See Appendix 2.

Corollary: The ordering of the slope coefficients of Regressions (11) and (34) are identical.

Proposition 3 and its corollary show that our analysis of the horse race based on Regression (11) applies exactly to that of the horse race based on Regression (34).

4. Results in Perspective

BvB, BHO, Agarwal, Green and Ren (2017) and Blocher and Molyboga (2017) report that single factor alpha is most highly correlated with fund flows into mutual funds and hedge funds among alphas computed with respect to many multifactor models. BvB and some other papers conclude that these results indicate that the CAPM is the true asset pricing model. However, BHO conclude that these results indicate that investors lack the sophistication to use the most precise model to estimate alphas for their investment decision. What are the assumptions that are necessary to draw these inferences? Are these assumptions satisfied in the data?

Our analysis shows that any inference about the true asset pricing model is tenable only if inclusion of any of the unpriced factors to compute alphas in Eq. (4) increases alpha estimate variance due to model misspecification error more than it reduces statistical estimation error. BvB effectively make such an assumption when they assume "if a true risk model exists, any false risk model cannot have additional explanatory power." BvB note that this assumption "rules out the possibility that ε_{it}^{c} contains information about managerial ability that is not also contained in ε_{it}" where their notations ε_{it} and ε_{it}^{c} denote alpha estimation errors with the true asset pricing model (i.e. $\hat{\alpha}_{p,K}$ estimated using the K-factor model) and with any other multifactor model $\left(\text{i.e. } \hat{\alpha}_{p,\eta}\ \forall\, \eta \neq K\right)$, respectively.

Is this assumption empirically tenable? Our simulation shows for the parameters in the data, the precision of alpha estimate is insensitive to the true asset pricing model. For example, if CAPM were the true model but we estimate alphas using the seven-factor model, the increase in model misspecification error is an order of magnitude smaller than the decrease in estimation error compared with the error in $\hat{\alpha}_{p,K}$. In fact, the winner of the horse race does not depend on the true asset pricing model both if we know the true betas and if we estimated betas from the data. Therefore, any inference about the true asset pricing model based on alpha-fund flow horse race is empirically untenable.

Regarding inferences about investor sophistication, an important question is, what is the appropriate benchmark that sophisticated investors would use? BHO hypothesize that sophisticated investors would use the *J*-factor model, a model that includes all priced and unpriced common factors. Our analysis shows that this hypothesis ignores the potential contribution of model misspecification error and the effect of measurement error in factor betas.

Our simulation results indicate that *J*-factor model indeed wins out horse race regardless of the true asset pricing model if betas are known. However, when we estimate betas with 60 months of data, alphas are estimated more precisely with the three factor model than with the seven factor model. Therefore, in this case the appropriate benchmark for assessing investor sophistication is the three factor model rather than the seven factor model. The evidence in BHO that market model alphas win the horse race indicates that investors use this alpha to inform their investment decision rather than the most precise three factor alpha. Therefore, their conclusion that investors are not sophisticated enough to use the most precise estimate of alpha to inform their mutual fund investment decisions is empirically tenable.

5. Conclusion

Investors reveal their preferences for mutual funds through investments in or withdrawals from them. Since non-satiated investors prefer more abnormal returns to less, investors' fund flows reveal their views on abnormal returns that they can earn from their investments. Because flows reveal investors' perceptions, the recent literature has proposed that a comparison of relations between fund flows and alphas measured with respect to a number of models can be used to identify the best asset pricing model and also to assess investor sophistication.

We show analytically that the empirical tenability of any inferences we draw based on such flow-alpha horse race critically depend on the sources of measurement error in alphas estimated under various models. For instance, we show that we can draw reliable inferences about asset pricing models only if the dominant source of error in alphas is due to the misspecification of the true asset pricing model. However, we find that the true asset pricing model has no effect on the ordering of the flow-alpha relations in our simulations with parameters estimated form the data. These findings indicate that asset pricing model misspecification error is a trivial of alpha estimation error in the data. Therefore, any inference about the true asset pricing model based on the flow-alpha horse race is empirically untenable.

References

Agarwal, V., Green, T. C., & Ren, H. (2017). Alpha or beta in the eye of the beholder: What drives hedge fund flows?. *Journal of Financial Economics*, Forthcoming.

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

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of political economy*, *112*(6), 1269-1295.

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

Berk, J. B., & van Binsbergen, J. H. (2017). How Do Investors Compute the Discount Rate? They Use the CAPM. *Financial Analysts Journal*, *73*(2), 25-32.

Blocher, J., & Molyboga, M. (2017). The revealed preference of sophisticated investors. *European Financial Management*, *23*(5), 839-872.

Busse, J. A., Goyal, A., & Wahal, S. (2010). Performance and persistence in institutional investment management. *The Journal of Finance*, *65*(2), 765-790.

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

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

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

DeGroot, Morris H. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.

Del Guercio, D., & 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.

Elton, E. J., Gruber, M. J., & Blake, C. R. (2001). A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases. *The Journal of Finance*, *56*(6), 2415-2430.

Elton, E. J., Gruber, M. J., & Blake, C. R. (2012). Does mutual fund size matter? The relationship between size and performance. *The Review of Asset Pricing Studies*, *2*(1), 31-55.

Evans, R. B. (2010). Mutual fund incubation. *The Journal of Finance*, *65*(4), 1581-1611.

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

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

Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988. *The Journal of Finance, 48*(1), 93-130.

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

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics*, *116*(1), 23-45.

Patel, J., Zeckhauser, R., & Hendricks, D. (1994). Investment flows and performance: Evidence from mutual funds, cross-border investments and new issues. *Japan, Europe and the international financial markets: Analytical and empirical perspectives*, 51-72.

Russanov, N., Ruan, H., Wei, Y. (2018). Marketing Mutual Funds. Working Paper

Sirri, E. R., & Tufano, P. (1998). Costly Search and Mutual Fund Flows, *Journal of Finance, LIII (5), October*, 1589-622.

Table 1: Summary statistics

This table presents the summary statistics for the sample of funds included in the sample. The number of fund-month observations is 404,042. The table first computes the respective statistics across funds each month and reports the averages over the entire sample period. The sample period is from January, 1990 to June, 2017.

	Mean	Std. Dev.	Median
Number of funds each month	1224		
Flow (%)	0.25	10.8	-0.42
TNA ($ mn)	1120.4	4507.4	223.6
Age (months)	376.8	306.6	299.2
Expense Ratio (%)	1.22	0.45	1.19
Load Dummy	0.49	0.50	0
Ret. Volatility (t-1,t-12)	4.7	2.3	4.2

Table 2: Factor model R^2

This table fits the following regression:

$$\left(r_{p,t} - r_{f,t}\right) = \alpha_{p,\eta} + \sum_{k=1}^{\eta} \beta_{k,p}\, F_{k,t} \; + \; e_{p,\eta,t}.$$

Where $r_{p,t}, r_{f,t}$ and $F_{k,t}$ are fund return, risk-free rate and realization of factor k in month t, respectively. For each month t, the regression is fitted from $t-60$ to $t-1$. The table reports the cross-sectional averages of time-series means of adjusted R^2 of the OLS regressions under each model. For market-adjusted and benchmark-adjusted returns we compute this metric as $1 - \left(\sum(r_{it} - r_{mkt})^2 / \sum(r_{it} - \bar{r}_i)^2\right)$, $1 - \left(\sum\left(r_{it} - r_{b/m}\right)^2 / \sum(r_{it} - \bar{r}_i)^2\right)$ using full sample of returns for each. Benchmark is the fund benchmark identified by Cremers and Petajisto (2009). The sample period is January, 1990 to June, 2017.

Model	Adj. R^2
Market Adj. Return	0.774
Benchmark Adj. Return	0.870
CAPM	0.820
FF3	0.892
FFC4	0.901
FFC4 + 3 IND	0.910

Table 3: Measurement Errors in betas

This table reports the slope coefficients from the following cross-sectional regressions:

$$\hat{\beta}_{p,k,t}^{future} = a_t + b_t \times \hat{\beta}_{p,k,t}^{past} + e_{p,t},$$

where for each fund f, $\hat{\beta}_{p,k,t}^{future}$ and $\hat{\beta}_{p,k,t}^{past}$ are estimated using time-series regressions with data from t to t+11, and t-1 to t-60, respectively. All betas are estimated with univariate time-series regressions. The above regression is fitted each month for betas with respect to each factor and the table reports time-series averages of the slope coefficients. Standard errors from the second stage of Fama-MacBeth regressions are adjusted for serial correlation using Newey-West correction with lag length of 11 months. Sample period for these regressions is Jan-1990 to Jul-2016. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Betas	Average b_t	Std. Err.
Market	0.821***	0.07
SMB	0.876***	0.03
HML	0.765***	0.05
UMD	0.409***	0.08
IND1	0.356***	0.09
IND2	0.362***	0.09
IND3	0.090	0.10

Table 4: Simulation Parameters

This table shows the parameters used in generating simulated returns and flows during 1990-2017. We generate net returns each month using the following seven-factor model:

$$r_{p,t} = \phi_p^K - \delta \times q_{t-1} + E^{model}(r_p) + \beta_{p,m} \times (\widetilde{mkt - rf})_t + \beta_{p,smb} \times \widetilde{SMB}_t + \beta_{p,hml} \times \widetilde{HML}_t$$
$$+ \quad \beta_{p,umd} \times \widetilde{UMD}_t + \beta_{p,ind1} \times \widetilde{IND1}_t * + \beta_{p,ind2} \times \widetilde{IND2}_t + \beta_{p,ind3} \times \widetilde{IND3}_t$$
$$+ \epsilon_{p,t}$$

where ϕ_p^K is fund manager skill, q_{t-1} is fund's Total Net Assets at the end of $t-1$ and $\delta \times q_{t-1}$ is the cost per unit for active fund management. The variables under $tilde$ are demeaned realizations of the following factors: market, SMB, HML, UMD, and three industry factors and βs are the corresponding factor sensitivities. We generate monthly flow as:

$$flow_{p,t} = a + b \times \hat{\alpha}_{p,\eta^*,t} + \psi_{p,t},$$

where $\hat{\alpha}_{p,\eta^*,t}$ is computed using the fund's realized return and the seven factor returns as $r_{p,t} - \sum_{k=1}^{7} \beta_{p,k} \times F_{k,t}$. All randomly drawn parameters are generated from a normal distribution with means and standard deviations shown in the table.

Panel A: Randomly drawn parameters		
Parameter	Mean	Standard Deviation
ϕ_p^K	0.15%	0.2%
β_{mkt}	1	0.154
β_{smb}	0.25	0.328
β_{hml}	0	0.262
β_{umd}	0	0.096
β_{IND1}	0	0.036
β_{IND2}	0	0.036
β_{IND3}	0	0.024
ϵ	0	0.025 (2.5%)
ψ	0	0.09 (9%)

Panel B: Fixed parameters	
Parameter	Value
δ	0.2 bps/$ 100 mn
$q_{t=0}$	$10 mn
a	-0.00225
b	0.2
η^*	7 factor model

Table 5: Measurement error components in simulated sample

This table shows the empirical estimates of various components in the variance decomposition of measurement error in $\hat{\alpha}$ from Eq (32). Each month, in the simulated sample, we compute the model misspecification error (θ) and statistical estimation error (ϵ) for various combinations of true asset pricing models (K=0, 1, 3) and estimation models (η=0, 1, 3, 4, 7) using the analytical expressions from Eq. (31). From these values, we compute the monthly cross-sectional variances and covariance and then average them across time and across 50 simulation samples and report the values scaled by 10^{-6}. Columns (1), (2), (3) of each Panel in the table below show the variances of estimation error, misspecification error and the covariance of the two respectively. Panel A shows the variance estimates when we use true betas of the funds to compute $\hat{\alpha}_\eta$ in which case the beta measurement error part drops in the estimation error component. And panel B shows the variance estimates where we use 60 month rolling window estimates of $\hat{\beta}$ to compute $\hat{\alpha}_\eta$ in which case the measurement error in betas shows up as part of column (1). The expressions we use to compute θ, ϵ are:

$$\theta = -\sum_{k=K+1}^{\eta} \beta_{k,p} E(F_k) \text{ if } \eta \geq K,\ \theta = \sum_{k=\eta+1}^{K} \beta_{k,p} E(F_k) \text{ if } \eta < K \text{ and } \epsilon = \sum_{k=\eta+1}^{J} \beta_{k,p}\, f_{k,t} + \sum_{k=1}^{\eta} \left(\hat{\beta}_{k,p,t} - \beta_{k,p}\right) F_{k,t} + \xi_{p,t}.$$

Betas used to estimate alphas are:		Panel A: True Betas				Panel B: $\hat{\beta}$s from 60 month rolling regressions			
		σ_ϵ^2 (1)	σ_θ^2 (2)	$Cov(\theta,\epsilon)$ (3)	(1)+(2)+(3)	σ_ϵ^2 (1)	σ_θ^2 (2)	$Cov(\theta,\epsilon)$ (3)	(1)+(2)+(3)
True asset pricing Model (K):	Alpha Estimated Using (η):								
No-beta risk premium model (*K=0*)	Mkt adj. ret.	875.2	0	0	875.2	874.5	0	0	874.5
	Market model	832.3	1.151	0.005	833.5	833.3	1.148	0.005	834.5
	FF3	661.4	1.494	0.003	662.9	702.4	1.49	0.005	703.9
	FFC4	640.2	1.743	-0.001	642.0	717.0	1.739	-0.001	718.7
	FFC4+3 IND	625.0	1.755	-0.001	626.7	760.9	1.751	0.007	762.7
CAPM (*K=1*)	Mkt adj. ret.	875.2	1.151	-0.110	876.3	874.5	1.148	-0.103	875.6
	Market model	832.3	0	0	832.3	833.3	0	0	833.3
	FF3	661.4	0.350	-0.002	661.8	702.4	0.35	-0.006	702.8
	FFC4	640.2	0.602	-0.005	640.8	717.0	0.601	-0.010	717.6

	FFC4+3 IND	625.0	0.614	-0.004	625.6	760.9	0.613	-0.008	761.6
FF3 (K=3)	Mkt adj. ret.	875.2	1.494	-0.052	876.7	874.5	1.49	-0.039	876.0
	Market model	832.3	0.350	0.058	832.7	833.3	0.35	-0.004	833.7
	FF3	661.4	0	0	661.4	702.4	0	0	702.4
	FFC4	640.2	0.252	-0.004	640.5	717.0	0.252	-0.005	717.2
	FFC4+3 IND	625.0	0.264	-0.002	625.3	760.9	0.263	-0.002	761.2

Table 6: Flow-Performance relation in simulated sample

This table presents univariate flow-performance regression results in the simulated sample. Columns (1), (2) and (3) in each of panels A and B report the results with true expected returns generated under No-beta risk premium (NBRP), CAPM, and FF3 models. The alphas which are the independent variables are computed with respect to the models indicated in the first column. Panel A shows the results using true betas to compute these alphas while panel B shows the results with $\hat{\beta}$s estimated using the prior month returns. Monthly flow is simulated in the sample as $flow_{p.t} = -0.00225 + 0.2 * \hat{\alpha}_{p,\eta^*=7,t} + \psi_{p,t}$ which is the dependent variable. The table presents the average value of slope coefficients multiplied by 100 with flows as the dependent variable and alphas as independent variables across 50 simulated samples.

	Panel A: True betas used to estimate alphas						Panel B: 60 month rolling window $\hat{\beta}$s used to estimate alphas					
	(1)		(2)		(3)		(1)		(2)		(3)	
True asset pricing model (K):	NBRP model		CAPM		FF3		NBRP model		CAPM		FF3	
	Coef/SE	R^2	Coef/SE	R^2	Coef/SE	R^2	Coef/SE	R^2	Coef/SE	R^2	Coef/SE	R^2
Alpha Estimated Using (η):												
Market Adjusted Ret	13.949***	0.294	13.935***	0.293	13.929***	0.293	14.517***	0.339	14.499***	0.339	14.494***	0.339
	(0.445)		(0.442)		(0.440)		(0.624)		(0.623)		(0.622)	
Market model	14.691***	0.305	14.686***	0.305	14.679***	0.305	15.258***	0.352	15.252***	0.351	15.249***	0.351
	(0.462)		(0.460)		(0.458)		(0.674)		(0.673)		(0.673)	
FF3	18.753***	0.370	18.753***	0.369	18.753***	0.369	17.793***	0.392	17.790***	0.392	17.791***	0.391
	(0.541)		(0.539)		(0.535)		(0.694)		(0.694)		(0.693)	
FFC4	19.454***	0.380	19.456***	0.380	19.456***	0.380	17.810***	0.392	17.807***	0.392	17.808***	0.392
	(0.589)		(0.588)		(0.584)		(0.688)		(0.688)		(0.687)	
FFC4+1 IND	19.698***	0.384	19.700***	0.384	19.701***	0.384	17.557***	0.388	17.553***	0.388	17.553***	0.388
	(0.592)		(0.591)		(0.587)		(0.674)		(0.674)		(0.673)	
FFC4+2 IND	19.939***	0.388	19.941***	0.387	19.942***	0.387	17.326***	0.385	17.323***	0.384	17.323***	0.384
	(0.610)		(0.609)		(0.605)		(0.663)		(0.662)		(0.661)	
FFC4+3 IND	19.970***	0.388	19.973***	0.388	19.973***	0.388	16.963***	0.379	16.959***	0.378	16.958***	0.378
	(0.603)		(0.601)		(0.597)		(0.670)		(0.670)		(0.669)	

Coefficient Difference Test						
FFC4 - (FFC4+3 IND)	-0.516***	-0.517***	-0.517***	0.847***	0.849***	0.849***
	(0.082)	(0.082)	(0.083)	(0.171)	(0.171)	(0.171)
FFC4 - True Asset Pricing Model	5.505***	4.770***	0.703***	3.294***	2.556***	0.017
	(0.302)	(0.278)	(0.116)	(0.335)	(0.315)	(0.144)

Appendix 1:

This appendix derives the results presented in section 1.4 of the paper. For expositional convenience, we set the risk-free rate to zero.

Let returns be generated by a single factor model as shown in equation (16). The true model of expected returns is either a no-beta risk premium model in equation (17) or CAPM in equation (18). $\hat{\alpha}$ is estimated using either a market adjustment as shown in equation (19) or a market model adjustment as shown in equation (20).

In the cross-section of funds, the following hold true:

$$\begin{aligned} cov(\beta,\xi) &= 0 \\ cov(\hat{\beta},\xi) &= 0 \\ cov(\hat{\beta}-\beta,\xi) &= 0, \end{aligned} \qquad (A.1.1)$$

where β represents true beta of a fund, $\hat{\beta}$ represents the estimated beta of the fund, $\hat{\beta}-\beta$ is the measurement error in estimated beta, and ξ represents the fund specific returns.

We also have, by definition:

$$cov(\hat{\beta}-\beta,\beta) = 0 \qquad (A.1.2)$$

From the two models of expected returns and two estimators, we have the following four cases.

Case 1: Market adjustment when the no-beta risk model is true

From equations (16), (17), (19):

$$\begin{aligned} \hat{\alpha}_{p,0} = r_{p,t} - r_{m,t} &= \alpha_p + E[r_m] + \beta_p \times f_t + \xi_{p,t} - r_{m,t} \\ &= \alpha_p + E[r_m] + \beta_p \times (r_{m,t} - E[r_m]) + \xi_{p,t} - r_{m,t} \\ &= \alpha_p + u_t \text{ where } u_t = (\beta_p - 1) \times (r_{m,t} - E[r_m]) + \xi_{p,t} \end{aligned}$$

Therefore, the cross-sectional variance of u_t after using the results in $(A.1.1)$ will be:

$$\begin{aligned} \sigma_u^2 \Big| r_{m,t} &= \left(r_{m,t} - E[r_m]\right)^2 \times var\left(\beta_p - 1 \middle| r_{m,t}\right) + \sigma^2_{\xi_{p,t}} \Big| r_{m,t} \\ &= \left(r_{m,t} - E[r_m]\right)^2 \times \sigma^2_{\beta_p} | r_{m,t} + \sigma^2_{\xi_{p,t}} | r_{m,t} \end{aligned}$$

Since the true betas and the fund specific returns are drawn from identical distributions across funds, we can drop the subscript p to arrive at:

$$\sigma_u^2|r_{m,t} = \left(r_{m,t} - E[r_m]\right)^2 \times \sigma_\beta^2|r_{m,t} + \sigma_\xi^2|r_{m,t} \qquad (A.1.3)$$

Case 2: Market model adjustment (i.e. CAPM) when the no-beta risk model is true

From equation (20):

$$\begin{aligned}
\hat{\alpha}_{p,1} &= r_{p,t} - \hat{\beta}_p \times r_{m,t} - \left(1 - \hat{\beta}_p\right) \times r_f \\
&= \alpha_p + E[r_m] + \beta_p \times f_t + \xi_{p,t} - \hat{\beta}_p r_{m,t} - \left(1 - \hat{\beta}_p\right) \times r_f \text{ from equations (13), (14)} \\
&= \alpha_p + E[r_m] + \beta_p \times \left(r_{m,t} - E[r_m]\right) + \xi_{p,t} - \hat{\beta}_p r_{m,t} - \left(1 - \hat{\beta}_p\right) \times r_f \\
&= \alpha_p + u_t \text{ where } u_t = \left(1 - \beta_p\right) \times E(r_m) - \left(\hat{\beta}_p - \beta_p\right) \times r_{m,t} - \left(1 - \hat{\beta}_p\right) \times r_f + \xi_{p,t}
\end{aligned}$$

Using $(A.1.1), (A.1.2)$, and the following two results

$$\begin{gathered}
Cov\left(1 - \beta_p, 1 - \hat{\beta}_p\right) = var\left(\beta_p\right) \\
Cov\left(\hat{\beta}_p - \beta_p, 1 - \hat{\beta}_p\right) = -var\left(\hat{\beta}_p - \beta_p\right),
\end{gathered}$$

the cross-sectional variance of u_t will be:

$$\begin{gathered}
\sigma_u^2\big|r_{m,t} = E(r_m) \times \left(E(r_m) - r_f\right) \times \sigma_\beta^2\big|r_{m,t} + r_{m,t} \times (r_{m,t} - r_f) \\
\times \sigma_{\hat{\beta}-\beta}^2 \Big|r_{m,t} + r_f^2 \times \sigma_{\hat{\beta}}^2 \Big| r_{m,t} + \sigma_\xi^2|r_{m,t}
\end{gathered} \qquad (A.1.4)$$

When the risk-free rate is set to zero:

$$\sigma_u^2\big|r_{m,t} = E(r_m)^2 \times \sigma_\beta^2\big|r_{m,t} + r_{m,t}^2 \times \sigma_{\hat{\beta}-\beta}^2|r_{m,t} + \sigma_\xi^2|r_{m,t} \qquad (A.1.5)$$

Case 3: Market adjustment when CAPM is true

From equations (16), (18), (19):

$$\begin{aligned}
\hat{\alpha}_{p,0} &= \alpha + r_f + \beta_p \times \left(E[r_m] - r_f\right) + \beta_p \times f_t + \xi_{p,t} - r_{m,t} \\
&= \alpha + r_f + \beta_p \times \left(E[r_m] - r_f\right) + \beta_p \times \left(r_{m,t} - E[r_m]\right) + \xi_{p,t} - r_{m,t} \\
&= \alpha + u_t \text{ where } u_t = -\left(1 - \beta_p\right) \times (r_{m,t} - r_f) + \xi_{p,t}
\end{aligned}$$

Using $(A.1.1)$, the cross-sectional variance of u_t is:

$$\sigma_u^2 \left| r_{m,t} = \left(r_{m,t} - r_f\right)^2 \times \sigma_{\beta_p}^2 \right| r_{m,t} + \sigma_{\xi_{p,t}}^2 | r_{m,t} \qquad (A.1.6)$$

Dropping the subscript p since betas and fund specific returns are drawn from identical distributions across funds and with risk free rate set to zero, we get:

$$\sigma_u^2 | r_{m,t} = r_{m,t}^2 \times \sigma_\beta^2 | r_{m,t} + \sigma_\xi^2 | r_{m,t} \qquad (A.1.7)$$

Case 4: Market model adjustment when CAPM is true

From (16), (18), (20):

$$\begin{aligned}\hat{\alpha}_{p,1} &= \alpha + r_f + \beta_p \times \left(E[r_m] - r_f\right) + \beta_p \times f_t + \xi_{p,t} - \left[r_f + \hat{\beta}_p\left(r_{m,t} - r_f\right)\right] \\ &= \alpha + r_f + \beta_p \times \left(E[r_m] - r_f\right) + \beta_p \times \left(r_{m,t} - E[r_m]\right) + \xi_{p,t} - \left[r_f + \hat{\beta}_p\left(r_{m,t} - r_f\right)\right] \\ &= \alpha + u_t \text{ where } u_t = -\left(\hat{\beta}_p - \beta_p\right) \times (r_{m,t} - r_f) + \xi_{p,t}\end{aligned}$$

Using $(A.1.1)$, the cross-sectional variance of u_t is:

$$\sigma_u^2 \left| r_{m,t} = \left(r_{m,t} - r_f\right)^2 \times \sigma_{\hat{\beta}_p - \beta_p}^2 \right| r_{m,t} + \sigma_{\xi_{p,t}}^2 | r_{m,t} \qquad (A.1.8)$$

After dropping subscript p and setting risk free rate to zero, we get:

$$\sigma_u^2 \left| r_{m,t} = r_{m,t}^2 \times \sigma_{\hat{\beta} - \beta}^2 \right| r_{m,t} + \sigma_\xi^2 | r_{m,t} \qquad (A.1.9)$$

Appendix 2:

This appendix proves Proposition 3.

Proof of Proposition 3:

Denote

$$\hat{\alpha}_{p,\eta} = \hat{\alpha}_{p,\eta^*} + v_{p,\eta}, \tag{A.2.1}$$

where $\hat{\alpha}_{p,\eta^*}$ is the alpha estimated using the most optimal η^*-factor model. Under the rational expectations equilibrium of Berk, Green (2004), flow positively covaries with $\hat{\alpha}_{p,\eta^*,t}$ and is uncorrelated with the noise term $v_{p,\eta}$.

Under the model of Berk, Green (2004), $\hat{\alpha}_{p,\eta1}, \hat{\alpha}_{p,\eta2}$ are normally distributed with mean zero and are therefore symmetric around zero. Therefore:

$$\begin{gathered}\Pr\left(Q_{p,\eta} = -1\right) = \Pr\left(Q_{p,\eta} = 1\right) = .5 \text{ for } \eta = \eta_1, \eta_2, \\ E\left(Q_{p,\eta_1}\right) = E\left(Q_{p,\eta_2}\right) = 0 \text{ , and} \\ Var\left(Q_{p,\eta_1}\right) = Var\left(Q_{p,\eta_2}\right) = 1.\end{gathered} \tag{A.2.2}$$

It also follows from the definition in $(A.2.1)$ that:

$$E\left(v_{p,\eta}\right) = 0 \tag{A.2.3}$$

Consider the following OLS regressions from (11) and (34):

$$\begin{gathered}\Gamma_p = a_\eta + b_\eta \hat{\alpha}_{p,\eta} + \omega_{p,\eta} \\ Q_{\Gamma_p} = A_\eta + B_\eta Q_{p,\eta} + o_{p,\eta}\end{gathered}$$

From Regression (11), after using $Cov\left(\Gamma_p, v_{p,\eta}\right) = 0$, we get:

$$b_\eta = \frac{cov(\Gamma_p, \hat{\alpha}_{p,\eta})}{var\left(\hat{\alpha}_{p,\eta}\right)} = \frac{cov(\Gamma_p, \hat{\alpha}_{p,\eta^*})}{var\left(\hat{\alpha}_{p,\eta^*}\right) + var(v_{p,\eta})} \tag{A.2.4}$$

Given that $\hat{b}_{\eta_1} > \hat{b}_{\eta_2}$. Therefore, from $(A.2.4)$ we get:

$$var(v_{p,\eta 1}) < var(v_{p,\eta 2}) \qquad (A.2.5)$$

From Regression (34), after using the result in $(A.2.2)$, we get:

$$B_\eta = \frac{Cov\left(Q_{\Gamma_p}, Q_{p,\eta}\right)}{Var(Q_{p,\eta})} = Cov\left(Q_{\Gamma_p}, Q_{p,\eta}\right) \qquad (A.2.6)$$

To evaluate this covariance term, we use the law of total covariance which states:

$$cov(X,Y) = E\big(cov(X,Y|Z)\big) + cov(E(X|Z), E(Y|Z)) \qquad (A.2.7)$$

Using $(A.2.7)$, we can write:

$$\begin{aligned} &Cov\left(Q_{\Gamma_p}, Q_{p,\eta}\right) \\ &\quad = E\left(cov\left(Q_{\Gamma_p}, Q_{p,\eta}|Q_{p,\eta^*}\right)\right) + cov\left(E\left(Q_{\Gamma_p}|Q_{p,\eta^*}\right), E\left(Q_{p,\eta}|Q_{p,\eta^*}\right)\right) \end{aligned} \qquad (A.2.8)$$

Since Γ_p is independent of the noise part of $\hat{\alpha}_{p,\eta}$, the conditional covariance $cov\left(Q_{\Gamma_p}, Q_{p,\eta}|Q_{p,\eta^*}\right)$ will be zero on average. Hence the first term on the RHS of $(A.2.8)$ will be zero. Expanding the second term in $(A.2.8)$, we get:

$$\begin{aligned} &Cov\left(Q_{\Gamma_p}, Q_{p,\eta}\right) \\ &\quad = E\left[E\left(Q_{\Gamma_p}|Q_{p,\eta^*}\right) \times E\left(Q_{p,\eta}|Q_{p,\eta^*}\right)\right] - E\left[E\left(Q_{\Gamma_p}|Q_{p,\eta^*}\right)\right] \\ &\quad \times E\left[E\left(Q_{p,\eta}|Q_{p,\eta^*}\right)\right] \end{aligned} \qquad (A.2.9)$$

The two terms in $(A.2.9)$ can further be expanded as:

$E\left[E\left(Q_{\Gamma_p}|Q_{p,\eta^*}\right) \times E\left(Q_{p,\eta}|Q_{p,\eta^*}\right)\right] = E\left(Q_{\Gamma_p}|Q_{p,\eta^*} = 1\right) \times E\left(Q_{p,\eta}|Q_{p,\eta^*} = 1\right) \times \Pr\left(Q_{p,\eta^*} = 1\right) + E\left(Q_{\Gamma_p}|Q_{p,\eta^*} = -1\right) \times E\left(Q_{p,\eta}|Q_{p,\eta^*} = -1\right) \times \Pr\left(Q_{p,\eta^*} = -1\right)$, and

$E\left[E\left(Q_{\Gamma_p}|Q_{p,\eta^*}\right)\right] \times E\left[E\left(Q_{p,\eta}|Q_{p,\eta^*}\right)\right] = \left\{E\left(Q_{\Gamma_p}|Q_{p,\eta^*} = 1\right) \times \Pr\left(Q_{p,\eta^*} = 1\right) + E\left(Q_{\Gamma_p}|Q_{p,\eta^*} = -1\right) \times \Pr\left(Q_{p,\eta^*} = -1\right)\right\} \times \left\{E\left(Q_{p,\eta}|Q_{p,\eta^*} = 1\right) \times \Pr\left(Q_{p,\eta^*} = 1\right) + E\left(Q_{p,\eta}|Q_{p,\eta^*} = -1\right) \times \Pr\left(Q_{p,\eta^*} = -1\right)\right\}$

Substituting these into $(A.2.9)$, using $1 - \Pr\left(Q_{p,\eta^*} = 1\right) = \Pr\left(Q_{p,\eta^*} = -1\right)$, and rearranging the terms yields:

$$\begin{aligned} Cov\left(Q_{\Gamma_p}, Q_{p,\eta}\right) &= \left\{E\left(Q_{\Gamma_p}|Q_{p,\eta^*} = 1\right) - E\left(Q_{\Gamma_p}|Q_{p,\eta^*} = -1\right)\right\} \\ &\times \left\{E\left(Q_{p,\eta}|Q_{p,\eta^*} = 1\right) - E\left(Q_{p,\eta}|Q_{p,\eta^*} = -1\right)\right\} \times \Pr\left(Q_{p,\eta^*} = 1\right) \\ &\times \Pr\left(Q_{p,\eta^*} = -1\right) \end{aligned} \tag{A.2.10}$$

From $(A.2.6)$ and $(A.2.10)$ we can see that comparing coefficients $B_{\eta 1}, B_{\eta 2}$ reduces to comparing $\left\{E\left(Q_{p,\eta}|Q_{p,\eta^*} = 1\right) - E\left(Q_{p,\eta}|Q_{p,\eta^*} = -1\right)\right\}$ for $\eta = \eta_1$ & η_2, since η^* is same across the two models.

By definition, this term can be expressed as:

$$\begin{aligned} E\left(Q_{p,\eta}|Q_{p,\eta^*} = 1\right) - E\left(Q_{p,\eta}|Q_{p,\eta^*} = -1\right) &= \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta} \geq 0|\hat{\alpha}_{p,\eta^*} \geq 0\right) \\ &- \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta} < 0|\hat{\alpha}_{p,\eta^*} \geq 0\right) \\ &- \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta} \geq 0|\hat{\alpha}_{p,\eta^*} < 0\right) \\ &+ \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta} < 0|\hat{\alpha}_{p,\eta^*} < 0\right) \end{aligned} \tag{A.2.11}$$

Where the conditional probabilities are defined as:

$$\begin{aligned} \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta} \geq 0|\hat{\alpha}_{p,\eta^*} \geq 0\right) &= \int_0^\infty \Pr\left(\hat{\alpha}_{p,\eta^*} \geq -v_{p,\eta}|\hat{\alpha}_{p,\eta^*}\right) \times f\left(\hat{\alpha}_{p,\eta^*}|\hat{\alpha}_{p,\eta^*} \geq 0\right) \times d\hat{\alpha}_{p,\eta^*} \end{aligned} \tag{A.2.12}$$

with $v_{p,\eta}|\hat{\alpha}_{p,\eta^*}$ distributed as Normal with mean zero.

We get similar expressions for the remaining three terms on the RHS of equation $(A.2.11)$.

When $X \sim N(0, \sigma^2)$, the following definitions apply:

$$\Pr(X \le a) = F(a) = \frac{1}{2} \times \left[1 + \operatorname{erf}\left(\frac{a-\mu}{\sigma\sqrt{2}}\right)\right] = \frac{1}{2} \times \left[1 + \operatorname{erf}\left(\frac{a}{\sigma\sqrt{2}}\right)\right]$$
$$\Pr(X \ge a) = 1 - F(a) = \frac{1}{2} \times \left[1 - \operatorname{erf}\left(\frac{a-\mu}{\sigma\sqrt{2}}\right)\right] = \frac{1}{2} \times \left[1 - \operatorname{erf}\left(\frac{a}{\sigma\sqrt{2}}\right)\right] \qquad (A.2.13)$$

Where $\operatorname{erf}(x)$ is the error function given by:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \times \int_0^x e^{-t^2} dt$$

This is an odd function with $\operatorname{erf}(-x) = -\operatorname{erf}(x)$ and is monotonically increasing in its argument x. From these two properties and the definitions in $(A.2.13)$, we can infer the following:

$$\begin{aligned} \Pr(X \ge a) \text{ is } &\begin{cases} \text{decreasing with } \sigma \text{ if } a < 0 \\ \text{increasing with } \sigma \text{ if } a > 0 \end{cases} \\ \Pr(X \le a) \text{ is } &\begin{cases} \text{increasing with } \sigma \text{ if } a < 0 \\ \text{decreasing with } \sigma \text{ if } a > 0 \end{cases} \end{aligned} \qquad (A.2.14)$$

From $(A.2.5)$, we have $\sigma_{v_{p,\eta 1}} < \sigma_{v_{p,\eta 2}}$. Therefore, from $(A.2.12)$ and $(A.2.14)$, we can see that:

$$\begin{aligned} \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 1} \ge 0 | \hat{\alpha}_{p,\eta^*} \ge 0\right) &> \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 2} \ge 0 | \hat{\alpha}_{p,\eta^*} \ge 0\right), \\ \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 1} < 0 | \hat{\alpha}_{p,\eta^*} < 0\right) &> \Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 2} < 0 | \hat{\alpha}_{p,\eta^*} < 0\right), \\ -\Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 1} < 0 | \hat{\alpha}_{p,\eta^*} \ge 0\right) &> -\Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 2} < 0 | \hat{\alpha}_{p,\eta^*} \ge 0\right), \\ -\Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 1} \ge 0 | \hat{\alpha}_{p,\eta^*} < 0\right) &> -\Pr\left(\hat{\alpha}_{p,\eta^*} + v_{p,\eta 2} \ge 0 | \hat{\alpha}_{p,\eta^*} < 0\right) \end{aligned} \qquad (A.2.15)$$

Substituting $(A.2.15)$ into $(A.2.11)$ gives:

$$E\left(Q_{p,\eta 1} | Q_{p,\eta^*} = 1\right) - E\left(Q_{p,\eta 1} | Q_{p,\eta^*} = -1\right) > E\left(Q_{p,\eta 2} | Q_{p,\eta^*} = 1\right) - E\left(Q_{p,\eta 2} | Q_{p,\eta^*} = -1\right) \qquad (A.2.16)$$

Finally, substituting this into $(A.2.10)$ and using the definition of B_η from $(A.2.6)$, we get $B_{\eta 1} > B_{\eta 2}$. Q.E.D.