Journal of Financial Economics 119 (2016) 1–23

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Assessing asset pricing models using revealed preference ☆

Jonathan B. Berk [a,b,*], Jules H. van Binsbergen [b,c,d,**]

[a] Graduate School of Business, Knight Management Center, Stanford University, Stanford, CA 94305, United States
[b] National Bureau of Economic Research, United States
[c] The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, United States
[d] Tilburg University, Netherlands

ARTICLE INFO

Article history:
Received 25 January 2015
Received in revised form
25 February 2015
Accepted 3 March 2015
Available online 24 August 2015

JEL classification:
G23
G20
G11
J31

Keywords:
Asset Pricing Test
Factor Models
CAPM
Mutual Funds
Flows

ABSTRACT

We propose a new method of testing asset pricing models that relies on quantities rather than just prices or returns. We use the capital flows into and out of mutual funds to infer which risk model investors use. We derive a simple test statistic that allows us to infer, from a set of candidate models, the risk model that is closest to the model that investors use in making their capital allocation decisions. Using our method, we assess the performance of the most commonly used asset pricing models in the literature.

1. Introduction

All neoclassical capital asset pricing models assume that investors compete fiercely with each other to find positive net present value investment opportunities, and in doing so, eliminate them. As a consequence of this competition, equilibrium prices are set so that the expected return of every asset is solely a function of its risk. When a positive net present value (NPV) investment opportunity presents itself in capital markets (that is, an asset is mispriced relative to the model investors are using) investors react by submitting buy or sell orders until the opportunity no longer exists (the mispricing is removed). These buy and sell orders reveal the preferences of investors and therefore they reveal which asset pricing model investors are using. By observing whether or not buy and sell orders occur in reaction to the existence of positive net present value

☆ We are grateful to the referee for his especially clear and insightful comments. We are grateful to John Cochrane, George Constantinides, Peter DeMarzo, Wayne Ferson, Ravi Jagannathan, Valentine Haddad, Lars Hansen, John Heaton, Binying Lui, Tim McQuade, Lubos Pastor, Paul Pfleiderer, Monika Piazzesi, Anamaria Pieschacon, Martin Schneider, Ken Singleton, Rob Stambaugh, and seminar participants at the 2015 AFA meetings, Harvard, the Kellogg Junior Finance Conference, Notre Dame, Princeton University, Stanford GSB, the Stanford Institute for Theoretical Economics (SITE), University of Chicago, University of Washington Summer Finance Conference, Jackson Hole Finance Group, and Washington University in St. Louis for their comments and suggestions.

* Corresponding author at: Graduate School of Business, Knight Management Center, Stanford University, Stanford, CA 94305, United States. Tel.: +1 650 721 1280.

** Corresponding author at: The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, United States. Tel.: +1 215 573 1606.

E-mail addresses: jbberk@stanford.edu (J.B. Berk), julesv@wharton.upenn.edu (J.H. van Binsbergen).

http://dx.doi.org/10.1016/j.jfineco.2015.08.010
0304-405X/

investment opportunities as defined by a particular asset pricing model, one can infer whether investors price risk using that asset pricing model.

There are two criteria that are required to implement this method. First, one needs a mechanism that identifies positive net present value investment opportunities. Second, one needs to be able to observe investor reactions to these opportunities. We demonstrate that we can satisfy both criteria if we implement the method using mutual fund data. Under the assumption that a particular asset pricing model holds, we use the main insight from Berk and Green (2004) to show that positive (negative) abnormal return realizations in a mutual fund investment must be associated with positive net present value buying (selling) opportunities. We then measure investor reactions to these opportunities by observing the subsequent capital flow into (out of) mutual funds.

Using this method, we derive a simple test statistic that allows us to infer, from a set of candidate models, the model that is closest to the asset pricing model investors are actually using. Our test can be implemented by running a simple univariate ordinary least squares (OLS) regression using the t-statistic to assess statistical significance. We illustrate our method by testing the following models: the Capital Asset Pricing Model (CAPM), originally derived by Sharpe (1964), Lintner (1965), Mossin (1966) and Treynor (1961), the reduced form factor models specified by Fama and French (1993) and Carhart (1997) that are motivated by Ross (1976), and the dynamic equilibrium models derived by Merton (1973), Breeden (1979), Campbell and Cochrane (1999), Kreps and Porteus (1978), Epstein and Zin (1991), and Bansal and Yaron (2004).

We find that the CAPM is the closest model to the model that investors use to make their capital allocation decisions. Importantly, the CAPM better explains flows than no model at all, indicating that investors do price risk. Most surprisingly, the CAPM also outperforms a naive model in which investors ignore beta and simply chase any outperformance relative to the market portfolio. Investors' capital allocation decisions reveal that they use the CAPM beta.

Our result, that investors appear to be using the CAPM to make their investment decisions, is very surprising in light of the well documented failure of the CAPM to adequately explain the cross-sectional variation in expected stock returns. Although, ultimately, we leave this as a puzzle to be explained by future research, we do note that the poor performance of the reduced form factor models relative to the CAPM implies that investors do not use the additional factors in those models to measure risk. Much of the flows in and out of mutual funds remain unexplained. To that end the paper leaves as an unanswered question whether the unexplained part of flows results because investors use a superior, yet undiscovered, risk model, or whether investors use other, non-risk-based, criteria to make investment decisions.

It is important to emphasize that implementing our test requires accurate measurement of the variables that determine the Stochastic Discount Factor (SDF). In the case of the CAPM, the SDF is measured using market prices which contain little or no measurement error, and more importantly, can be observed by investors as accurately as by empiricists. Testing the dynamic equilibrium models relies on observing variables such as consumption, which investors can measure precisely (they presumably know their own consumption) but empiricists cannot, particularly over short horizons. Consequently, our tests cannot differentiate whether these models underperform because they rely on variables that are difficult to measure, or because the underlying assumptions of these models are flawed.

Because we implement our method using mutual fund data, one might be tempted to conclude that our tests only reveal the preferences of mutual fund investors, rather than all investors. But this is not the case. When an asset pricing model correctly prices risk, it rules out positive net present value investment opportunities in all markets. Even if no investor in the market with a positive net present value opportunity uses the asset pricing model under consideration, so long as there are investors in other markets that use the asset pricing model, those investors will recognize the positive net present value opportunity and will act to eliminate it. That is, if our test rejects a particular asset pricing model, we are not simply rejecting the hypothesis that mutual fund investors use the model, but rather, we are rejecting the hypothesis that any investor who could invest in mutual funds uses the model.

Of course, the possibility exists that investors are not using a risk model to price assets. In that case our tests only reveal the preferences of mutual fund investors because it is possible, in this world, for investors in other markets to be uninterested in exploiting positive net present value investment opportunities in the mutual fund market. However, mutual fund investors actually represent a very large fraction of all investors. In 2013, 46% of households invested in mutual funds. More importantly, this number rises to 81% for households with income that exceeds $100,000.[1]

The first paper to use mutual fund flows to infer investor preferences is Guercio and Tkac (2002). Although the primary focus of their paper is on contrasting the inferred behavior of retail and institutional investors, that paper documents that flows respond to outperformance relative to the CAPM. The paper does not consider other risk models. Clifford, Fulkerson, Jordan, and Waldman (2013) study the effect of increases in idiosyncratic risk on inflows and outflows separately (rather than the net flow) and show that both inflows and outflows increase when funds take on more idiosyncratic risk (as defined by the Fama-French-Carhart factor specification). Barber, Huang, and Odean (2014) also use fund flows to infer investor risk preferences and find (using a different method) that investors use the CAPM rather than the other reduced-form factor models that have been proposed.[2]

[1] As reported in the 2014 Investment Company Fact Book, Chapter Six, Figures 6.1 and 6.5 (see http://www.icifactbook.org).

[2] The first draft of Barber et al. (2014) paper was posted to the Social Science Research Network (SSRN) five months subsequent to the initial posting of our paper.

2. A new asset pricing test

The core idea that underlies every neoclassical asset pricing model in economics is that prices are set by agents chasing positive net present value investment opportunities. When financial markets are perfectly competitive, these opportunities are competed away so that, in equilibrium, prices are set to ensure that no positive net present value opportunities exist. Prices respond to the arrival of new information by instantaneously adjusting to eliminate any positive net present value opportunities that arise. It is important to appreciate that this price adjustment process is part of all asset pricing models, either explicitly (if the model is dynamic) or implicitly (if the model is static). The output of all these models – a prediction about expected returns – relies on the assumption that this price adjustment process occurs.

The importance of this price adjustment process has long been recognized by financial economists and forms the basis of the event study literature. In that literature, the asset pricing model is assumed to be correctly identified. In that case, because there are no positive net present value opportunities, the price change that results from new information (i.e., the part of the change not explained by the asset pricing model) measures the value of the new information.

Because prices always adjust to eliminate positive net present value investment opportunities, under the correct asset pricing model, expected returns are determined by risk alone. Modern tests of asset pricing theories test this powerful insight using return data. Rejection of an asset pricing theory occurs if positive net present value opportunities are detected, or, equivalently, if investment opportunities can be found that consistently yield returns in excess of the expected return predicted by the asset pricing model. The most important shortcoming in interpreting the results of these tests is that the empiricist is never sure whether a positive net present value investment opportunity that is identified ex post was actually available ex ante.[3]

An alternative testing approach, that does not have this shortcoming, is to identify positive net present value investment opportunities ex ante and test for the existence of an investor response. That is, do investors react to the existence of positive net present value opportunities that result from the revelation of new information? Unfortunately, for most financial assets, investor responses to positive net present value opportunities are difficult to observe. As Milgrom and Stokey (1982) show, the price adjustment process can occur with no transaction volume whatsoever, that is, competition is so fierce that no investor benefits from the opportunity. Consequently, for most financial assets the only observable evidence of this competition is the price change itself. Thus, testing for an investor response is equivalent to standard tests of asset pricing theory that use return data.

The key to designing a test to directly detect investor responses to positive net present value opportunities is to find an asset for which the price is fixed. In this case the market equilibration must occur through volume (quantities). A mutual fund is just such an asset. The price of a mutual fund is always fixed at the price of its underlying assets, or the net asset value (NAV). In addition, fee changes are rare. Consequently, if, as a result of new information, an investment in a mutual fund represents a positive net present value investment opportunity, the only way for investors to eliminate the opportunity is by trading the asset. Because this trade is observable, it can be used to infer investments investors believe to be positive net present value opportunities. One can then compare those investments to the ones the asset pricing model under consideration identifies to be positive net present value and thereby infer whether investors are using the asset pricing model. That is, by observing investors' revealed preferences in their mutual fund investments, we are able to infer information about what (if any) asset pricing model they are using.

2.1. The mutual fund industry

Mutual fund investment represents a large and important sector in U.S. financial markets. In the last 50 years, there has been a secular trend away from direct investing. Individual investors used to make up more than 50% of the market, today they are responsible for barely 20% of the total capital investment in U.S. markets. During that time, there has been a concomitant rise in indirect investment, principally in mutual funds. Mutual funds used to make up less than 5% of the market, today they make up 1/3 of total investment.[4] Today, the number of mutual funds that trade in the U.S. outnumber the number of stocks that trade.

Berk and Green (2004) derive a model of how the market for mutual fund investment equilibrates that is consistent with the observed facts.[5] They start with the observation that the mutual fund industry is like any industry in the economy – at some point it displays decreasing returns to scale.[6] Given the assumption under which all asset pricing models are derived (perfectly competitive financial markets), this observation immediately implies that all mutual funds must have enough assets under management so that they face decreasing returns to scale. When new information arrives that convinces investors that a particular mutual fund represents a positive net present value investment, investors react by investing more capital in the mutual fund. This process continues until enough new capital is invested to eliminate the opportunity. As a consequence, the model is able to explain two robust empirical facts in the mutual fund literature: that mutual fund flows react to past performance while future performance is largely unpredictable.[7] Investors “chase” past performance because it is informative: mutual fund managers that do

[3] For an extensive analysis of this issue, see Harvey, Liu, and Zhu (2014).

[4] See French (2008).

[5] Stambaugh (2014) derives a general equilibrium version of this model based on the model in Pastor and Stambaugh (2012).

[6] Pastor, Stambaugh, and Taylor (2015) provide empirical evidence supporting this assumption.

[7] An extensive literature has documented that capital flows are responsive to past returns (see Chevalier and Ellison, 1997; Sirri and

well (poorly) have too little (much) capital under management. By competing to take advantage of this information, investors eliminate the opportunity to predict future performance.

A key assumption of the Berk and Green (2004) model is that mutual fund managers are skilled and that this skill varies across managers. Berk and van Binsbergen (2015) verify this fact. They demonstrate that such skill exists and is highly persistent. More importantly, for our purposes, they demonstrate that mutual fund flows contain useful information. Not only do investors systematically direct flows to higher skilled managers, but managerial compensation, which is primarily determined by these flows, predicts future performance as far out as ten years[8]. Investors know who the skilled managers are and compensate them accordingly. It is this observation that provides the starting point for our analysis. Because the capital flows into mutual funds are informative, they reveal the asset pricing model investors are using.

2.2. Private information

Most asset pricing models are derived under the assumption that all investors are symmetrically informed. Hence, if one investor faces a positive NPV investment opportunity, all investors face the same opportunity and so it is instantaneously removed by competition. The reality is somewhat different. The evidence in Berk and van Binsbergen (2015) of skill in mutual fund management implies that at least some investors have access to different information or have different abilities to process information. As a result, under the information set of this small set of informed investors, not all positive net present value investment opportunities are instantaneously competed away.

As Grossman (1976) argued, in a world where there are gains to collecting information and information gathering is costly, not everybody can be equally informed in equilibrium. If everybody chooses to collect information, competition between investors ensures that prices reveal the information and so information gathering is unprofitable. Similarly, if nobody collects information, prices are uninformative and so there are large profits to be made collecting information. Thus, in equilibrium, investors must be differentially informed as in, for example, Grossman and Stiglitz (1980). Investors with the lowest information gathering costs collect information so that, *on the margin*, what they spend on information gathering, they make back in trading profits. Presumably, these investors are few in number so that the competition between them is limited, allowing for the existence of prices that do not fully reveal their information. As a result, information gathering is a positive net present value endeavor for a limited number of investors.

The existence of asymmetrically informed investors poses a challenge for empiricists wishing to test asset pricing models derived under the assumption of symmetrically informed investors. Clearly, the empiricist's information set matters. For example, asset pricing models fail under the information set of the most informed investor, because the key assumption that asset markets are competitive is false under that information set. Consequently, the standard in the literature is to assume that the information set of the uninformed investors only contains publicly available information all of which is already impounded in all past and present prices, and to conduct the test under that information set. For now, we will adopt the same strategy but will revisit this assumption in Section 6.2, where we will explicitly consider the possibility that the majority of investors' information sets includes more information than just what is already impounded in past and present prices.

2.3. Method

To formally derive our testing method, let q_{it} denote assets under management (AUM) of fund i at time t and let θ_i denote a parameter that describes the skill of the manager of fund i. At time t, investors use the time t information set I_t to update their beliefs on θ_i resulting in the distribution function $g_t(\theta_i)$ implying that the expectation of θ_i at time t is

$$\bar{\theta}_{it} \equiv \mathrm{E}\left[\theta_i | I_t\right] = \int \theta_i g_t\left(\theta_i\right) d\theta_i. \tag{1}$$

We assume throughout that $g_t(\cdot)$ is not a degenerate distribution function. Let R^n_{it} denote the excess return (that is, the net return in excess of the risk free rate) earned by investors between time $t-1$ and t. We take as our Null hypothesis that a particular asset pricing model holds. Let R^B_{it} denote the risk adjustment prescribed by this asset pricing model over the same time interval. Note that q_{it}, R^n_{it}, and R^B_{it} are elements of I_t. Let $\alpha_{it}(q)$ denote investors' subjective expectation of the risk-adjusted return they make, under the Null hypothesis, when investing in fund i that has size q between time t and $t+1$, also commonly referred to as the *net alpha*:

$$\alpha_{it}(q) = \bar{\theta}_{it} - h_i(q), \tag{2}$$

where $h_i(q)$ is a strictly increasing function of q, reflecting the fact that, under the assumptions underlying every asset pricing model, all mutual funds must face decreasing returns to scale in equilibrium. Under the Null that the asset pricing model under consideration holds perfectly, in equilibrium, the size of the fund q_{it} adjusts to ensure that there are no positive net present value investment opportunities so $\alpha_{it}(q_{it}) = 0$ and

$$\bar{\theta}_{it} = h_i\left(q_{it}\right). \tag{3}$$

At time $t+1$, the investor observes the manager's return outperformance,

$$\varepsilon_{it+1} \equiv R^n_{it+1} - R^B_{it+1}, \tag{4}$$

which is a signal that is informative about θ_i. The conditional distribution function of ε_{it+1} at time t, $f(\varepsilon_{it+1}|\alpha_{it}(q_{it}))$,

---

(*footnote continued*)
Tufano, 1998) and future investor returns are largely unpredictable (see Carhart, 1997).

[8] Busse and Irvine (2006) document that alpha outperformance is associated with future fund flows and attribute this relationship to investor inferences of managerial skill.

satisfies the following condition in equilibrium:

$$\mathrm{E}[\varepsilon_{it+1}\big|I_t] = \int \varepsilon_{it+1} f\left(\varepsilon_{it+1}|\alpha_{it}(q_{it})\right) d\varepsilon_{it+1} = \alpha_{it}(q_{it}) = 0. \tag{5}$$

Our testing method relies on the insight that, under the Null hypothesis, good news, that is, $\varepsilon_{it} > 0$, implies good news about θ_i and bad news, $\varepsilon_{it} < 0$, implies bad news about θ_i. The following proposition shows that, in expectation, this condition holds generally. That is, on average, a positive (negative) realization of ε_{it} leads to a positive (negative) update on θ_i implying that before the capital response, the fund's alpha will be positive (negative).

Proposition 1. On average, a positive (negative) realization of ε_{it} leads to a positive (negative) update on θ_i:

$$\mathrm{E}[\alpha_{it+1}(q_{it})\varepsilon_{it+1}\big|I_t] > 0.$$

Proof.

$$\begin{aligned}\mathrm{E}[\alpha_{it+1}(q_{it})\varepsilon_{it+1}\big|I_t] &= \mathrm{E}[\mathrm{E}[\alpha_{it+1}(q_{it})\varepsilon_{it+1}\big|\theta_i]\big|I_t] = \mathrm{E}[(\theta_i - h_i(q_{it}))\mathrm{E}[\varepsilon_{it+1}\big|\theta_i]\big|I_t]\\ &= \mathrm{E}[(\theta_i - h_i(q_{it}))(\theta_i - h_i(q_{it}))\big|I_t] > 0. \ \square\end{aligned}$$

Unfortunately this proposition is not directly testable because $\alpha_{it+1}(q_{it})$ is not observable. Instead what we observe are the capital flows that result when investors update their beliefs. Our next objective is to restate the result in Proposition 1 in terms of capital flows.

What Proposition 1 combined with (3) tells us is that positive (negative) news must, on average, lead to an inflow (outflow). However, without further assumptions, we cannot quantify the magnitude of the capital response. The magnitude of the capital response is primarily driven by two factors – the form of the fund's decreasing returns to scale technology and the distribution of investors' priors and posteriors. Neither factor is directly observable so they must be inferred from the flow of funds relation itself. Doing so requires disentangling the two effects. A large flow of funds response can be driven by either a relatively flat decreasing returns to scale technology or a prior that is uninformative. In addition, both factors are likely to vary cross-sectionally. Because the size of a fund is determined endogenously, small funds are likely to differ from large funds in their returns to scale technology. Similarly, the informativeness of returns, and therefore how investors update their priors, is likely to differ across funds. Finally, there is no theoretical reason for the relation between flows and returns to be linear, in Berk and Green (2004), for example, it is quadratic. Furthermore, the empirical evidence suggests that this relation is not linear.

Rather than lose generality by making further assumptions on the technology and how investors update, we can sidestep this issue by focusing only on the direction of the capital response. With that in mind we begin by first defining the function that returns the sign of a real number, taking values 1 for a positive number, -1 for a negative number and zero for zero:

$$\phi(x) \equiv \begin{cases} \dfrac{x}{|x|}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Next, let the flow of capital into mutual fund i at time t be denoted by F_{it}, that is,

$$F_{it+1} \equiv q_{it+1} - q_{it}.$$

The following lemma proves that the sign of the capital inflow and the alpha inferred from the information in ε_{it+1} must be the same.

Lemma 1. The sign of the capital inflow and the alpha inferred from the information in ε_{it+1} must be the same:

$$\phi(F_{it+1}) = \phi(\alpha_{it+1}(q_{it})).$$

Proof.

$$\begin{aligned}\phi(\alpha_{it+1}(q_{it})) &= \phi(\alpha_{it+1}(q_{it}) - \alpha_{it+1}(q_{it+1})) = \phi(h(q_{it+1}) - h(q_{it}))\\ &= \phi(q_{it+1} - q_{it}) = \phi(F_{it+1}),\end{aligned}$$

where the first equality follows from (5) and the second line follows from the fact that $h(q)$ is a strictly increasing function. □

We are now ready to restate Proposition 1 as a testable prediction.

Proposition 2. The regression coefficient of the sign of the capital inflows on the sign of the realized return outperformance is positive, that is,

$$\beta_{F\varepsilon} \equiv \frac{cov(\phi(F_{it+1}), \phi(\varepsilon_{it+1}))}{var(\phi(\varepsilon_{it+1}))} > 0. \tag{6}$$

Proof. See Appendix.

This proposition provides a testable prediction and thus a new method to reject an asset pricing model. Under our method, we define a model as working when investors' revealed preferences indicate that they are using that model to update their inferences of positive net present value investment opportunities. Because flows reveal investor preferences, a measure of whether investors are using a particular asset pricing model is the fraction of decisions for which outperformance (as defined by the model) implies capital inflows and underperformance implies capital outflows. The next lemma shows that $\beta_{F\varepsilon}$ is a simple linear transformation of this measure.

Lemma 2. The regression coefficient of the sign of the capital inflows on the sign of the realized return outperformance can be expressed as follows:

$$\begin{aligned}\beta_{F\varepsilon} &= \Pr\left[\phi(F_{it}) = 1\middle|\phi(\varepsilon_{it}) = 1\right] + \Pr\left[\phi(F_{it}) = -1\middle|\phi(\varepsilon_{it}) = -1\right] - 1\\ &= \Pr\left[\phi(F_{it}) = 1\middle|\phi(\varepsilon_{it}) = 1\right] - \Pr\left[\phi(F_{it}) = 1\middle|\phi(\varepsilon_{it}) = -1\right].\end{aligned}$$

Proof. See Appendix.

To understand the implications of Lemma 2, note that we can use the lemma to express the relation as follows:

$$\frac{\beta_{F\varepsilon} + 1}{2} = \frac{\Pr\left[\phi(F_{it}) = 1\middle|\phi(\varepsilon_{it}) = 1\right] + \Pr\left[\phi(F_{it}) = -1\middle|\phi(\varepsilon_{it}) = -1\right]}{2},$$

that is, by adding one to $\beta_{F\varepsilon}$ and dividing by two we recover the average probability that conditional on outperformance being positive (negative), the sign of the fund flow is positive (negative). If outperformance predicted the direction of fund flows perfectly, both conditional probabilities would be 1 and so $\beta_{F\varepsilon} = 1$.[9] At the other extreme, if there is no relation between outperformance and flows, the conditional probabilities sum to one, implying that $\beta_{F\varepsilon} = 0$. Thus, we would expect the beta estimates to lie between zero and one.

On a practical level, many of the asset pricing models we will consider nest each other. As we will see, we will not be able to reject the Null hypothesis that any of the models we will consider is the true asset pricing model. In that case a natural question to ask is whether a model is "better" than the model it nests. By better we mean the model that comes closest to pricing risk correctly. To formalize this concept, we first assume that a true risk model exists. That is, that the expected return of every asset in the economy is a function only of the risk as measured by that model. Next we consider a set of candidate risk models, indexed by $c \in C$, such that the risk adjustment of each model is given by R_{it}^{c}, so risk-adjusted performance is given by

$$\varepsilon_{it}^{c} = R_{it}^{n} - R_{it}^{c}.$$

Because at most only one element of the set of candidate risk models can be the true risk model, the rest of the models in C do not fully capture risk. We refer to these models as *false risk models*. We will maintain the assumption throughout this paper that if a true risk model exists, any false risk model cannot have additional explanatory power for capital allocation decisions:

$$\Pr\left[\phi(F_{it}) \middle| \phi(\varepsilon_{it}), \phi\left(\varepsilon_{it}^{c}\right)\right] = \Pr\left[\phi(F_{it}) \middle| \phi(\varepsilon_{it})\right]. \tag{7}$$

This assumption is not innocuous. It rules out the possibility that ε_{it}^{c} contains information about managerial ability that is not also contained in ε_{it}.

For a false risk model $c \in C$, let β_{Fc} be the signed flow-performance regression coefficient of that model, that is,

$$\beta_{Fc} \equiv \frac{cov\left(\phi(F_{it}), \phi\left(\varepsilon_{it}^{c}\right)\right)}{var\left(\phi\left(\varepsilon_{it}^{c}\right)\right)}.$$

The next proposition proves that the regression coefficient of the true model (if it exists) must exceed the regression coefficient of a false model.

Proposition 3. The regression coefficient of the sign of the capital inflows on the sign of the realized return outperformance is maximized under the true model, that is, for any false model c,

$$\beta_{F\varepsilon} > \beta_{Fc}.$$

Proof. See Appendix.

[9] Proposition 1 holds only in expectation, implying that even under the true asset pricing model, $\beta_{F\varepsilon}$ need not be 1. Restrictive distributional assumptions are required to ensure that, under the true model, $\beta_{F\varepsilon} = 1$.

We are now ready to formally define what we mean by a model that comes closest to pricing risk. The following definition defines the best model as the model that maximizes the fraction of times outperformance by the candidate model implies outperformance by the true model and the fraction of times underperformance by the candidate model implies underperformance by the true model.

Definition 1. Model c is a better approximation of the true asset pricing model than model d if and only if:

$$\begin{aligned} &\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi\left(\varepsilon_{it}^{c}\right) = 1\right] + \Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi\left(\varepsilon_{it}^{c}\right) = -1\right] \\ &\quad > \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi\left(\varepsilon_{it}^{d}\right) = 1\right] \\ &\qquad + \Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi\left(\varepsilon_{it}^{d}\right) = -1\right]. \end{aligned} \tag{8}$$

With this definition in hand, we now show that the models can be ranked by their regression coefficients.

Proposition 4. Model c is a better approximation of the true asset pricing model than model d if and only if $\beta_{Fc} > \beta_{Fd}$.

Proof. See Appendix.

The next proposition provides an easy method for empirically distinguishing between candidate models.

Proposition 5. Consider an OLS regression of $\phi(F_{it})$ onto $\frac{\phi(\varepsilon_{it}^{c})}{var(\phi(\varepsilon_{it}^{c}))} - \frac{\phi(\varepsilon_{it}^{d})}{var(\phi(\varepsilon_{it}^{d}))}$:

$$\phi(F_{it}) = \gamma_0 + \gamma_1\left(\frac{\phi\left(\varepsilon_{it}^{c}\right)}{var\left(\phi\left(\varepsilon_{it}^{c}\right)\right)} - \frac{\phi\left(\varepsilon_{it}^{d}\right)}{var\left(\phi\left(\varepsilon_{it}^{d}\right)\right)}\right) + \xi_{it}$$

The coefficient of this regression is positive, that is, $\gamma_1 > 0$, if and only if, model c is a better approximation of the true asset pricing model than model d.

Proof. See Appendix.

3. Asset pricing models

The Null hypothesis in this paper is that the particular asset pricing model under consideration holds, implying that capital markets are competitive and investors are rational. Although these assumptions are clearly restrictive, it is important to emphasize that they are not part of our testing method, but instead are imposed on us by the models we test. Conceivably our method could be applied to behavioral models in which case these assumptions would not be required.

Our testing method can be applied to both reduced-form asset pricing models, such as the factor models proposed by Fama and French (1993) and Carhart (1997), as well as to dynamic equilibrium models, such as the consumption CAPM (Breeden, 1979), habit formation models (Campbell and Cochrane, 1999) and long-run risk (LRR) models that use recursive preferences (Epstein and Zin, 1991; Bansal and Yaron, 2004). For the CAPM and factor models, R_{it}^{B} is specified by the beta relationship. We regress the excess returns to investors, R_{it}^{n}, on the risk factors over the life of the fund to get the model's betas. We then use the beta relation to calculate R^B_{it} at each point in time. For example, for the Fama-French-Carhart (FFC) factor specification, the risk adjustment R^B_{it} is then given by

$$R^B_{it} = \beta_i^{mkt}\mathrm{MKT}_t + \beta_i^{sml}\mathrm{SML}_t + \beta_i^{hml}\mathrm{HML}_t + \beta_i^{umd}\mathrm{UMD}_t,$$

where MKT_t, SML_t, HML_t, and UMD_t are the realized excess returns on the four factor portfolios defined in Carhart (1997). Using this risk-adjusted return, we calculate (4) over a T-period horizon ($T>1$) as follows:

$$\varepsilon_{it} = \prod_{s=t-T+1}^{t} (1 + R^n_{is} - R^B_{is}) - 1. \tag{9}$$

The returns of any dynamic equilibrium model must satisfy the following Euler equation in equilibrium:

$$\mathrm{E}_t[M_{t+1}R^n_{it+1}] = 0, \tag{10}$$

where $M_t > 0$ is the stochastic discount factor (SDF) specified by the model. When this condition is violated a positive net present value investment opportunity exists.

The dynamic equilibrium models we consider are all derived under the assumption of a representative investor. Of course, this assumption does not presume that all investors are identical. When investors are not identical, it is possible that they do not share the same SDF. Even so, it is important to appreciate that, in equilibrium, all investors nevertheless agree on the existence of a positive net present value investment opportunity. That is, if (10) is violated, it is violated for every investor's SDF.[10] Because our testing method only relies on the existence of this net present value investment opportunity, it is robust to the existence of investor heterogeneity.

The outperformance measure for fund i at time t is therefore

$$\alpha_{it} = \mathrm{E}_t[M_{t+1}R^n_{it+1}]. \tag{11}$$

Notice that $\alpha_{it} > 0$ is a buying opportunity and so capital should flow into such opportunities. We calculate the outperformance relative to the equilibrium models over a T-period horizon as follows:

$$\varepsilon_{it} = \frac{1}{T}\sum_{s=t-T+1}^{t} M_s R^n_{is}. \tag{12}$$

Notice that in this case, T must be greater than one because when $T=1$, $\phi(\varepsilon_{it})$ is not a function of M_s.

To compute these outperformance measures, we must compute the stochastic discount factor for each model at each point in time. For the consumption CAPM, the stochastic discount factor is

$$M_t = \beta\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma},$$

where β is the subjective discount rate and γ is the coefficient of relative risk aversion. The calibrated values we use are given in the top panel of Table 1. We use the national income and product accounts (NIPA) data from the Bureau of Economic Analysis to compute consumption growth of non-durables and services.

[10] In an incomplete market equilibrium, investors may use different SDFs but the projection of each investor's SDF onto the asset space is the same.

Table 1
Parameter calibration.
The table shows the calibrated parameters for the three structural models that we test: power utility over consumption (the consumption CAPM), external habit formation preferences (as in Campbell and Cochrane, 1999) and Epstein Zin preferences as in Bansal and Yaron (2004).

Consumption CAPM				
Subj. disc. factor	Risk aversion			
β	γ			
0.9989	10			
Epstein Zin preferences (LRR)				
Subj. disc. factor	Risk aversion	IES	Weight in bonds	
δ	γ	ψ	w	
0.9989	10	1.5	0%, 70%, 90%	
Habit formation preferences				
Subj. disc. factor	Risk aversion	Mean growth	Habit persistence	Consumption vol
δ	γ	g	ϕ	σ
0.9903	2	0.0020	0.9885	0.0076

For the long-run risk model as proposed by Bansal and Yaron (2004), the stochastic discount factor is given by

$$M_t = \delta^{\theta}\left(\frac{C_t}{C_{t-1}}\right)^{-\theta/\psi}\left(1 + R^a_t\right)^{-(1-\theta)},$$

where R^a_t is the return on aggregate wealth and where θ is given by

$$\theta \equiv \frac{1-\gamma}{1-\frac{1}{\psi}}.$$

The parameter ψ measures the intertemporal elasticity of substitution (IES). To construct the realizations of the stochastic discount factor, we use parameter values for risk aversion and the IES commonly used in the long-run risk literature, as summarized in the middle panel of Table 1. In addition to these parameter values, we need data on the returns to the aggregate wealth portfolio. There are two ways to construct these returns. The first way is to estimate (innovations to) the stochastic volatility of consumption growth as well as (innovations to) expected consumption growth, which combined with the parameters of the long-run risk model lead to proxies for the return on wealth. The second way is to take a stance on the composition of the wealth portfolio, by taking a weighted average of traded assets. In this paper, we take the latter approach and form a weighted average of stock returns (as represented by the Center for Research in Security Pricing (CRSP) value-weighted total market portfolio) and long-term bond returns (the returns on the Fama-Bliss long-term bond portfolio (60–120 months)) to compute the returns on the wealth portfolio. Given the calibration in Table 1, the implied value of θ is large making the SDF very sensitive to the volatility of the wealth portfolio. Because the volatility of the wealth portfolio is sensitive to the relative weighting of stocks and bonds, we calculate the SDF over a range of weights (denoted by w) to assess the robustness with respect to this assumption.[11]

For the Campbell and Cochrane, 1999 habit formation model, the stochastic discount factor is given by

$$M_t = \delta\left(\frac{C_t}{C_{t-1}}\frac{S_t}{S_{t-1}}\right)^{-\gamma},$$

where S_t is the consumption surplus ratio. The dynamics of the log consumption surplus ratio s_t are given by

$$s_t = (1-\phi)\bar{s} + \phi s_{t-1} + \lambda(s_{t-1})(c_t - c_{t-1} - g),$$

where $\bar{s}$ is the steady state habit, ϕ is the persistence of the habit stock, c_t the natural logarithm of consumption at time t, and g is the average consumption growth rate. We set all the parameters of the model to the values proposed in Campbell and Cochrane (1999), but we replace the average consumption growth rate g, as well as the consumption growth rate volatility σ with their sample estimates over the full available sample (1959–2011), as summarized in the bottom panel of Table 1. To construct the consumption surplus ratio data, we need a starting value. As our consumption data start in 1959, which is long before the start of our mutual fund data in 1977, we have a sufficiently long period to initialize the consumption surplus ratio. That is, in 1959, we set the ratio to its steady state value $\bar{s}$ and construct the ratio for the subsequent periods using the available data that we have. Because the annualized value of the persistence coefficient is 0.87, the weight of the 1959 starting value of the consumption surplus ratio in the 1977 realization of the stochastic discount factor is small and equal to 0.015.

4. Results

We use the mutual fund data set in Berk and van Binsbergen (2015). The data set spans the period from January 1977 to March 2011. We remove all funds with less than five years of data leaving 4275 funds.[12] Berk and van Binsbergen (2015) undertook an extensive data project to address several shortcomings in the CRSP database by combining it with Morningstar data, and we refer the reader to the data appendix of that paper for the details.

To implement the tests derived in Propositions 2 and 5 it is necessary to pick an observation horizon. For most of the sample, funds report their AUMs monthly, however, in the early part of the sample many funds report their AUMs only quarterly. In order not to introduce a selection bias by dropping these funds, the shortest horizon we will consider is three months. Furthermore, as pointed out above, we need a horizon length of more than a month to compute the outperformance measure for the dynamic equilibrium models.

If investors react to new information immediately, then flows should immediately respond to performance and the appropriate horizon to measure the effect would be the shortest horizon possible. But in reality, there is evidence that investors do not respond immediately. Mamaysky, Spiegel, and Zhang (2008) show that the net alpha of mutual funds is predictably non zero for horizons shorter than a year, suggesting that capital does not move instantaneously. There is also evidence of investor heterogeneity because some investors appear to update faster than others.[13] For these reasons, we also consider longer horizons (up to four years). The downside of using longer horizons is that longer horizons tend to put less weight on investors who update immediately, and these investors are also the investors more likely to be marginal in setting prices. To ensure that we do not inadvertently introduce autocorrelation in the horizon returns across funds, we drop all observations before the first December observation for a fund so we thereby ensure that the first observation for all funds occurs in December.

The flow of funds is important in our empirical specification because it affects the alpha generating technology as specified by $h(\cdot)$. Consequently, we need to be careful to ensure that we only use the part of capital flows that affects this technology. For example, it does not make sense to include as an inflow of funds, increases in fund sizes that result from inflation because such increases are unlikely to affect the alpha generating process. Similarly, the fund's alpha generating process is unlikely to be affected by changes in size that result from changes in the price level of the market as a whole. Consequently, we will measure the flow of funds over a horizon of length T as

$$q_{it} - q_{it-T}(1 + R^V_{it}), \tag{13}$$

where R^V_{it} is the cumulative return to investors of the appropriate Vanguard benchmark fund as defined in Berk and van Binsbergen (2015) over the horizon from $t-T$ to t. This benchmark fund is constructed by projecting fund i's return onto the space spanned by the set of available Vanguard index funds which can be interpreted as the investor's alternative investment opportunity. Thus, in our empirical specification, we only consider capital flows into and out of funds net of what would have happened had investors not invested or withdrawn capital and had the fund manager adopted a purely passive strategy. That is, under this definition of capital flows, we are assuming that, in making their capital allocation decisions, investors take into account changes in the size of the fund that result from returns due to managerial outperformance. That said, all of our results are robust to using the fund's own return, R^n_{it}, instead of R^V_{it} in (13).

We begin by examining the correlation structure of performance between mutual funds. One would not expect

[11] See Lustig, Van Nieuwerburgh, and Verdelhan (2013) for a discussion on the composition of the wealth portfolio and the importance of including bonds.

[12] We chose to remove these funds to ensure that incubation flows do not influence our results. Changing the criterion to two years does not change our results.

[13] See Berk and Tonks (2007).

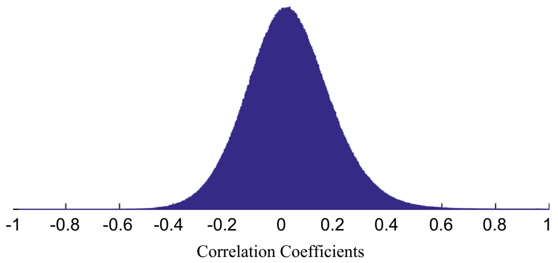

Fig. 1. Correlation between funds. The histogram displays the distribution of the pairwise correlation coefficients between funds of outperformance relative to the Vanguard benchmark.

mutual fund strategies to be highly correlated because otherwise the informational rents would be competed away. It is nevertheless important that we check that this is indeed the case, because otherwise our assumption that $h(\cdot)$ is a function of the size of the fund (rather than the size of the industry) would be subject to question. To examine this correlation, we calculate outperformance relative to the Vanguard benchmark, that is, for each fund we calculate ε_{it} using the Vanguard benchmark. The advantage of computing outperformance this way is that we do not need to take a stand on which risk model best prices risk. Instead, this measure measures outperformance relative to investors' next best alternative investment opportunity – the portfolio of Vanguard index funds that most closely replicates the fund under consideration. We then compute the correlation coefficients of outperformance between every fund in our sample for which the two funds have at least four years of overlapping data. Fig. 1 is a histogram of the results. It is clear from the figure that managers are not using the same strategies – the average correlation between the funds in our sample is 0.03. Furthermore, 43% of funds are negatively correlated and the fraction of funds that have large positive correlation coefficients is tiny (only 0.55% of funds have a correlation coefficient over 50%).

We implement our tests as follows. For each model, c, and each fund, i, we compute monthly outperformance, ε_{it}^c, as we explained in Section 3. That is, for the factor models we generate the outperformance measure for the horizon by using (9) and for the dynamic equilibrium models, we use (12). At the end of this process we have a fund flow and outperformance observation for each fund over each measurement horizon. We then implement the test in Proposition 2 by estimating $\beta_{F\varepsilon}$ for each model, by running a single panel regression. Table 2 reports our results.[14] For ease of interpretation, the table reports $\frac{\beta_{F\varepsilon}+1}{2}$, that is, the average probability that conditional on outperformance being positive (negative), the sign of the fund flow is positive (negative). If flows and outperformance are unrelated, we would expect this measure to equal 50%, that is, $\beta_{F\varepsilon}=0$. The first takeaway from Table 2 is that none of our candidate models can be rejected based on Proposition 2, that is,

Table 2
Flow of funds outperformance relationship (1977–2011).

The table reports estimates of (6) for different asset pricing models. For ease of interpretation, the table reports $(\beta_{F\varepsilon}+1)/2$ in percent, which by Lemma 2 is equivalent to $(\Pr[\phi(F_{it})=1|\phi(\varepsilon_{it})=1]+\Pr[\phi(F_{it})=-1|\phi(\varepsilon_{it})=-1])/2$. Each row corresponds to a different risk model. The first two rows report the results for the market model (CAPM) using the CRSP value-weighted index and the S&P 500 index as the market portfolio. The next three lines report the results of using as the benchmark return, three rules of thumb: (1) the fund's actual return, (2) the fund's return in excess of the risk-free rate, and (3) the fund's return in excess of the return on the market as measured by the CRSP value-weighted index. The next two lines are the Fama-French (FF) and Fama-French-Carhart (FFC) factor specifications. The final four lines report the results for the dynamic equilibrium models: the Consumption CAPM (C-CAPM), the habit model derived by Campbell and Cochrane (1999), and the long-run risk model derived by Bansal and Yaron (2004). For the long-run risk model we consider three different versions, depending on the portfolio weight of bonds in the aggregate wealth portfolio. The maximum number in each column (the best performing model) is shown in bold face.

Model	Horizon					
	3-month	6-month	1-year	2-year	3-year	4-year
			Market models (CAPM)			
CRSP value weighted	**63.63**	**63.49**	**63.38**	64.08	**63.86**	**63.37**
S&P 500	62.52	62.26	61.61	62.20	61.40	60.92
			No model			
Return	58.55	59.77	57.72	59.76	60.83	61.20
Excess return	58.29	59.64	57.57	60.91	61.27	61.69
Return in excess of the market	62.08	61.99	61.19	62.45	62.05	61.76
			Multifactor models			
FF	63.14	62.84	63.05	63.62	63.59	62.43
FFC	63.25	62.92	63.09	63.59	63.46	62.35
			Dynamic equilibrium models			
C-CAPM	58.30	59.52	57.43	60.44	60.67	61.17
Habit	58.29	59.46	57.37	60.45	60.80	61.00
Long-run risk – 0% bonds	57.42	59.54	61.11	**64.54**	63.50	62.17
Long-run risk – 70% bonds	57.07	58.84	59.14	59.95	60.24	59.47
Long-run risk – 90% bonds	57.18	59.27	58.95	58.60	60.21	60.45

[14] The flow of fund data contain very large outliers leading past researchers to winsorize the data. Because we only use the sign of flows, we do not winsorize.

Table 3
Model ranking.
The table shows the ranking of all the models at each time horizon, with the best performing model on top. Factor models are shown in bold ital, dynamic equilibrium models in ital, and roman entries are models that have not been formally derived. The CAPM is coded in both bold ital and ital since it can be interpreted as both a factor model and an equilibrium model. The number following the long-run risk models denotes the percentage of the wealth portfolio invested in bonds.

Horizon (months)					
3	6	12	24	36	48
CAPM	***CAPM***	***CAPM***	*LRR 0*	***CAPM***	***CAPM***
FFC	***FFC***	***FFC***	***CAPM***	***FF***	***FF***
FF	***FF***	***FF***	***FF***	*LRR 0*	***FFC***
CAPM ***SP500***	*CAPM* ***SP500***	*CAPM* ***SP500***	***FFC***	***FFC***	*LRR 0*
Excess market	Excess market	Excess market	Excess market	Excess market	Excess market
Return	Return	*LRR 0*	*CAPM* ***SP500***	*CAPM* ***SP500***	Excess return
C-CAPM	Excess return	*LRR 70*	Excess return	Excess return	Return
Excess return	*LRR 0*	*LRR 90*	*Habit*	Return	*C-CAPM*
Habit	*C-CAPM*	Return	*C-CAPM*	*Habit*	*Habit*
LRR 0	*Habit*	Excess return	*LRR 70*	*C-CAPM*	*CAPM* ***SP500***
LRR 90	*LRR 90*	*C-CAPM*	Return	*LRR 70*	*LRR 90*
LRR 70	*LRR 70*	*Habit*	*LRR 90*	*LRR 90*	*LRR 70*

β_{F_ε} is significantly greater than zero in all cases,[15] implying that regardless of the risk adjustment, a flow-performance relation exists. On the other hand, none of the models perform better than 64%. It appears that a large fraction of flows remain unexplained. Investors appear to be using other criteria to make a non-trivial fraction of their investment decisions.

Which model best approximates the true asset pricing model? Table 3 ranks each model by its β_{Fc}. The best performing model, at all but the 2 year horizon, is the CAPM with the CRSP value weighted index as the market proxy. To assess whether this ranking reflects statistically significant differences, we implement the pairwise linear regression specified in Proposition 5 and report the double-clustered (by fund and time) t-statistics of these regressions in Table 4. No model statistically outperforms the CAPM at any horizon.

To assess the relative performance of the models, we begin by first focusing on the behavioral model that investors just react to past returns, the column marked "Ret" in the table. By looking down that column in Table 4, one can see that the factor models all statistically significantly outperform this model at horizons of less than two years. For example, the t-statistic reported in Table 4 that $\beta_{F,CAPM} > \beta_{F,Ret}$ at the 3-month horizon is 4.98, indicating that we can reject the hypothesis that the behavioral model is a better approximation of the true model than the CAPM. Based on these results, we can reject the hypothesis that investors just react to past returns. The next possibility is that investors are risk neutral. Under (7), in an economy with risk-neutral investors, we would find that the excess return best explains flows, so the performance of this model can be assessed by looking at the columns labeled "Ex. ret." Notice that all the risk models nest this model, so to conclude that a risk model better approximates the true model, the risk model must statistically outperform this model. For horizons less than 2 years, the factor models all satisfy this criterion. Unfortunately, none of the dynamic asset pricing models satisfy this criterion. Finally, one might hypothesize that investors benchmark their investments relative to the market portfolio alone, that is, they do not adjust for any risk differences (beta) between their investment and the market. The performance of this model is reported in the column labeled "Ex. mkt." The CAPM statistically significantly outperforms this model at all horizons – investors' actions reveal that they use betas to allocate resources.

Next, we use our method to discriminate between the factor models. Recall that both the Fama-French (FF) and Fama-French-Carhart (FFC) factor specifications nest the CAPM, so to conclude that either factor model better approximates the true model, it must statistically significantly outperform the CAPM. The test of this hypothesis is in the columns labeled "CAPM." Neither factor model statistically outperforms the CAPM at any horizon. Indeed, at all horizons the CAPM actually outperforms both factor models implying that the additional factors add no more explanatory power for flows. Notice that this result does not rely on (7). Because the factor models nest the CAPM, we can conclude that there is no evidence that investors use the factor models.

The relative performance of the dynamic equilibrium models is poor. We can confidently reject the hypothesis that any of these models is a better approximation of the true model than the CAPM. But this result should be interpreted with caution. These models rely on variables like consumption which are notoriously difficult for empiricists to measure, but are observed perfectly by investors themselves.

It is also informative to compare the tests of statistical significance across horizons. The ability to statistically discriminate between the models deteriorates as the horizon increases. This is what one would expect to observe if investors instantaneously moved capital in response to the information in realized returns. Thus, this evidence is consistent with the idea that capital does in fact move quickly to eliminate positive net present value investment opportunities.

The evidence that investors appear to be using the CAPM is puzzling given the inability of the CAPM to correctly

[15] Table 4 reports the double-clustered (by fund and time) t-statistics.

Table 4
Tests of statistical significance.

The first two columns in the table provide the coefficient estimate and double-clustered t-statistic (see Thompson, 2011 and the discussion in Petersen, 2009) of the univariate regression of signed flows on signed outperformance. The rest of the columns provide the statistical significance of the pairwise test, derived in Proposition 5, of whether the models are better approximations of the true asset pricing model. For each model in a column, the table displays the double-clustered t-statistic of the test that the model in the row is a better approximation of the true asset pricing model, that is, that $\beta_{\text{Frow}} > \beta_{\text{Fcolumn}}$. The rows (and columns) are ordered by $\beta_{F\varepsilon}$, with the best performing model on top. The number following the long-run risk models denotes the percentage of the wealth portfolio invested in bonds.

Panel A: 3-Month horizon

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	FFC	FF	CAPM SP500	Ex. mkt	Ret	C-CAPM	Ex. ret	Habit	LRR 0	LRR 90	LRR 70
CAPM	0.273	26.35	0.00	1.15	1.52	4.71	7.28	4.98	5.71	5.77	5.70	6.97	7.13	7.60
FFC	0.265	28.64	−1.15	0.00	0.65	1.69	3.16	4.42	5.07	5.13	5.06	6.34	6.32	6.79
FF	0.263	28.45	−1.52	−0.65	0.00	1.42	2.76	4.35	5.02	5.07	5.00	6.17	6.27	6.69
CAPM SP500	0.250	21.25	−4.71	−1.69	−1.42	0.00	1.25	3.97	4.58	4.62	4.56	5.50	6.01	6.14
Excess market	0.242	22.46	−7.28	−3.16	−2.76	−1.25	0.00	3.40	3.91	3.95	3.90	4.88	5.12	5.45
Return	0.171	10.72	−4.98	−4.42	−4.35	−3.97	−3.40	0.00	1.20	1.18	1.25	1.17	3.60	2.87
C-CAPM	0.166	10.12	−5.71	−5.07	−5.02	−4.58	−3.91	−1.20	0.00	0.12	0.38	0.98	3.54	2.86
Excess return	0.166	10.11	−5.77	−5.13	−5.07	−4.62	−3.95	−1.18	−0.12	0.00	0.03	0.99	4.06	3.16
Habit	0.166	10.00	−5.70	−5.06	−5.00	−4.56	−3.90	−1.25	−0.38	−0.03	0.00	0.97	3.38	2.79
LRR 0	0.148	7.74	−6.97	−6.34	−6.17	−5.50	−4.88	−1.17	−0.98	−0.99	−0.97	0.00	0.30	0.56
LRR 90	0.144	8.32	−7.13	−6.32	−6.27	−6.01	−5.12	−3.60	−3.54	−4.06	−3.38	−0.30	0.00	0.39
LRR 70	0.141	7.93	−7.60	−6.79	−6.69	−6.14	−5.45	−2.87	−2.86	−3.16	−2.79	−0.56	−0.39	0.00

Panel B: 6-Month horizon

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	FFC	FF	CAPM SP500	Ex mkt	Ret	Ex ret	LRR 0	C-CAPM	Habit	LRR 90	LRR 70
CAPM	0.270	21.11	0.00	1.08	1.23	3.24	4.64	2.63	3.17	2.46	3.16	3.18	3.81	4.41
FFC	0.258	21.21	−1.08	0.00	0.35	0.95	1.47	2.21	2.64	2.12	2.68	2.71	2.90	3.57
FF	0.257	22.40	−1.23	−0.35	0.00	0.79	1.38	2.09	2.49	2.03	2.51	2.54	2.89	3.51
CAPM SP500	0.245	14.21	−3.24	−0.95	−0.79	0.00	0.50	1.78	2.09	1.67	2.14	2.18	2.41	2.83
Excess market	0.240	16.03	−4.64	−1.47	−1.38	−0.50	0.00	1.47	1.73	1.44	1.78	1.81	2.08	2.53
Return	0.195	8.44	−2.63	−2.21	−2.09	−1.78	−1.47	0.00	0.32	0.12	0.65	0.76	0.61	0.87
Excess return	0.193	8.26	−3.17	−2.64	−2.49	−2.09	−1.73	−0.32	0.00	0.06	0.93	1.19	0.54	0.94
LRR 0	0.191	7.21	−2.46	−2.12	−2.03	−1.67	−1.44	−0.12	−0.06	0.00	0.01	0.04	0.16	0.54
C-CAPM	0.190	8.03	−3.16	−2.68	−2.51	−2.14	−1.78	−0.65	−0.93	−0.01	0.00	1.07	0.32	0.72
Habit	0.189	7.93	−3.18	−2.71	−2.54	−2.18	−1.81	−0.76	−1.19	−0.04	−1.07	0.00	0.24	0.66
LRR 90	0.185	7.54	−3.81	−2.90	−2.89	−2.41	−2.08	−0.61	−0.54	−0.16	−0.32	−0.24	0.00	0.71
LRR 70	0.177	6.70	−4.41	−3.57	−3.51	−2.83	−2.53	−0.87	−0.94	−0.54	−0.72	−0.66	−0.71	0.00

Panel C: 1-Year horizon

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	FFC	FF	CAPM SP500	Ex mkt	LRR 0	LRR 70	LRR 90	Ret	Ex ret	C-CAPM	Habit
CAPM	0.268	13.54	0.00	0.44	0.47	3.89	6.42	0.74	3.40	2.75	2.25	2.98	2.88	2.90
FFC	0.262	14.30	−0.44	0.00	0.18	1.63	2.39	0.62	2.59	2.32	2.17	2.79	2.77	2.79
FF	0.261	14.55	−0.47	−0.18	0.00	1.47	2.25	0.61	2.64	2.26	2.11	2.67	2.67	2.69
CAPM SP500	0.232	8.31	−3.89	−1.63	−1.47	0.00	0.54	0.18	1.74	1.63	1.69	2.15	2.12	2.14
Excess market	0.224	10.38	−6.42	−2.39	−2.25	−0.54	0.00	0.03	1.43	1.14	1.26	1.60	1.59	1.61
LRR 0	0.222	6.51	−0.74	−0.62	−0.61	−0.18	−0.03	0.00	0.87	0.70	0.99	1.19	1.22	1.24
LRR 70	0.183	5.93	−3.40	−2.59	−2.64	−1.74	−1.43	−0.87	0.00	0.13	0.63	0.92	0.94	0.97
LRR 90	0.179	4.54	−2.75	−2.32	−2.26	−1.63	−1.14	−0.70	−0.13	0.00	0.98	1.60	1.62	1.70
Return	0.154	4.10	−2.25	−2.17	−2.11	−1.69	−1.26	−0.99	−0.63	−0.98	0.00	0.17	0.38	0.46
Excess return	0.151	4.00	−2.98	−2.79	−2.67	−2.15	−1.60	−1.19	−0.92	−1.60	−0.17	0.00	0.57	0.81
C-CAPM	0.149	4.04	−2.88	−2.77	−2.67	−2.12	−1.59	−1.22	−0.94	−1.62	−0.38	−0.57	0.00	1.29
Habit	0.147	3.98	−2.90	−2.79	−2.69	−2.14	−1.61	−1.24	−0.97	−1.70	−0.46	−0.81	−1.29	0.00

Panel D: 2-Year horizon

Model	$\beta_{F\varepsilon}$	Univ t-stat	LRR 0	CAPM	FF	FFC	Ex mkt	CAPM SP500	Ex ret	Habit	C-CAPM	LRR 70	Ret	LRR 90
LRR 0	0.291	7.96	0.00	0.11	0.20	0.21	0.50	0.57	0.77	0.86	0.87	1.49	0.99	1.55
CAPM	0.282	12.80	−0.11	0.00	0.80	0.97	5.73	3.81	1.45	1.50	1.50	2.57	1.42	2.54
FF	0.272	16.17	−0.20	−0.80	0.00	0.13	1.86	1.57	1.37	1.46	1.46	2.14	1.37	2.65
FFC	0.272	16.46	−0.21	−0.97	−0.13	0.00	2.06	1.72	1.31	1.39	1.40	2.19	1.33	2.51
Excess market	0.249	10.89	−0.50	−5.73	−1.86	−2.06	0.00	0.36	0.70	0.83	0.83	1.47	0.89	1.67
CAPM SP500	0.244	8.16	−0.57	−3.81	−1.57	−1.72	−0.36	0.00	0.60	0.74	0.74	1.19	0.84	1.53
Excess return	0.218	7.09	−0.77	−1.45	−1.37	−1.31	−0.70	−0.60	0.00	1.28	1.34	0.37	1.22	1.57
Habit	0.209	6.62	−0.86	−1.50	−1.46	−1.39	−0.83	−0.74	−1.28	0.00	0.24	0.19	1.00	1.17
C-CAPM	0.209	6.66	−0.87	−1.50	−1.46	−1.40	−0.83	−0.74	−1.34	−0.24	0.00	0.18	0.98	1.17
LRR 70	0.199	6.47	−1.49	−2.57	−2.14	−2.19	−1.47	−1.19	−0.37	−0.19	−0.18	0.00	0.07	0.62
Return	0.195	5.99	−0.99	−1.42	−1.37	−1.33	−0.89	−0.84	−1.22	−1.00	−0.98	−0.07	0.00	0.59

Table 4 (*continued*)

Panel D: 2-Year horizon

Model	$\beta_{F\varepsilon}$	Univ *t*-stat	*LRR 0*	***CAPM***	***FF***	***FFC***	Ex mkt	*CAPM* ***SP500***	Ex ret	*Habit*	*C-CAPM*	*LRR 70*	Ret	*LRR 90*
LRR 90	0.172	5.72	−1.55	−2.54	−2.65	−2.51	−1.67	−1.53	−1.57	−1.17	−1.17	−0.62	−0.59	0.00

Panel E: 3-Year horizon

Model	$\beta_{F\varepsilon}$	Univ *t*-stat	***CAPM***	***FF***	*LRR 0*	***FFC***	Ex mkt	*CAPM* ***SP500***	Ex ret	Ret	*Habit*	*C-CAPM*	*LRR 70*	*LRR 90*
CAPM	0.277	13.86	0.00	0.51	0.06	1.04	4.90	3.53	1.24	1.11	1.51	1.61	1.67	1.74
FF	0.272	14.39	−0.51	0.00	0.02	0.43	2.54	2.41	1.21	1.09	1.53	1.64	1.43	1.80
LRR 0	0.270	7.72	−0.06	−0.02	0.00	0.01	0.29	0.44	0.40	0.46	0.49	0.52	0.76	0.65
FFC	0.269	14.42	−1.04	−0.43	−0.01	0.00	2.67	2.55	1.07	0.98	1.34	1.43	1.42	1.67
Excess market	0.241	9.93	−4.90	−2.54	−0.29	−2.67	0.00	0.84	0.37	0.46	0.61	0.69	0.78	0.83
CAPM ***SP500***	0.228	8.05	−3.53	−2.41	−0.44	−2.55	−0.84	0.00	0.05	0.19	0.25	0.31	0.50	0.47
Excess return	0.225	6.91	−1.24	−1.21	−0.40	−1.07	−0.37	−0.05	0.00	0.51	1.09	1.39	0.39	0.70
Return	0.217	5.85	−1.11	−1.09	−0.46	−0.98	−0.46	−0.19	−0.51	0.00	0.02	0.15	0.19	0.33
Habit	0.216	6.91	−1.51	−1.53	−0.49	−1.34	−0.61	−0.25	−1.09	−0.02	0.00	1.58	0.21	0.41
C-CAPM	0.213	6.90	−1.61	−1.64	−0.52	−1.43	−0.69	−0.31	−1.39	−0.15	−1.58	0.00	0.16	0.32
LRR 70	0.205	5.01	−1.67	−1.43	−0.76	−1.42	−0.78	−0.50	−0.39	−0.19	−0.21	−0.16	0.00	0.01
LRR 90	0.204	7.29	−1.74	−1.80	−0.65	−1.67	−0.83	−0.47	−0.70	−0.33	−0.41	−0.32	−0.01	0.00

Panel F: 4-Year horizon

Model	$\beta_{F\varepsilon}$	Univ *t*-stat	***CAPM***	***FF***	***FFC***	*LRR 0*	Ex mkt	Ex ret	Ret	*C-CAPM*	*Habit*	*CAPM* ***SP500***	*LRR 90*	*LRR 70*
CAPM	0.267	13.02	0.00	1.81	1.95	0.14	4.76	0.79	0.90	1.05	1.15	3.93	1.38	1.84
FF	0.249	11.77	−1.81	0.00	0.37	0.03	1.11	0.38	0.57	0.66	0.76	1.62	0.98	1.28
FFC	0.247	11.61	−1.95	−0.37	0.00	0.02	0.96	0.32	0.50	0.58	0.68	1.58	0.92	1.24
LRR 0	0.243	5.02	−0.14	−0.03	−0.02	0.00	0.05	0.05	0.11	0.12	0.14	0.16	0.21	0.38
Excess market	0.235	9.70	−4.76	−1.11	−0.96	−0.05	0.00	0.04	0.24	0.28	0.36	1.26	0.58	1.05
Excess return	0.234	7.20	−0.79	−0.38	−0.32	−0.05	−0.04	0.00	0.52	0.94	1.44	0.32	0.67	0.80
Return	0.224	6.37	−0.90	−0.57	−0.50	−0.11	−0.24	−0.52	0.00	0.03	0.21	0.11	0.38	0.60
C-CAPM	0.223	6.87	−1.05	−0.66	−0.58	−0.12	−0.28	−0.94	−0.03	0.00	1.25	0.11	0.39	0.60
Habit	0.220	6.89	−1.15	−0.76	−0.68	−0.14	−0.36	−1.44	−0.21	−1.25	0.00	0.04	0.30	0.55
CAPM ***SP500***	0.218	7.30	−3.93	−1.62	−1.58	−0.16	−1.26	−0.32	−0.11	−0.11	−0.04	0.00	0.19	0.60
LRR 90	0.209	8.03	−1.38	−0.98	−0.92	−0.21	−0.58	−0.67	−0.38	−0.39	−0.30	−0.19	0.00	0.36
LRR 70	0.189	3.89	−1.84	−1.28	−1.24	−0.38	−1.05	−0.80	−0.60	−0.60	−0.55	−0.60	−0.36	0.00

account for cross-sectional differences in average returns. Although providing a complete explanation of this puzzling finding is beyond the scope of this paper, in the next section we will consider a few possible explanations. We will leave the question of which, if any, explanation resolves this puzzle to future research.

5. Implications

The empirical finding that the CAPM does a poor job explaining cross-sectional variation in expected returns raises a number of possibilities about the relation between risk and return. The first possibility, and the one most often considered in the existing literature, is that this finding does not invalidate the neoclassical paradigm that requires expected returns to be a function solely of risk. Instead, it merely indicates that the CAPM is not the correct model of risk, and, more importantly, a better model of risk exists. As a consequence researchers have proposed more general risk models that better explain the cross section of expected returns.

The second possibility is that the poor performance of the CAPM is a consequence of the fact that there is no relation between risk and return. That is, that expected returns are determined by non-risk based effects. The final possibility is that risk only partially explains expected returns, and that other, non-risk based factors, also explain expected returns. The results in this paper shed new light on the relative likelihood of these possibilities.

The fact that we find that the factor models all statistically significantly outperform our “no model” benchmarks implies that the second possibility is unlikely. So long as (7) holds, if there is no relation between risk and expected return, there is no reason for the CAPM to best explain investors' capital allocation decisions. That leaves the question of whether the failure of the CAPM to explain the cross section of expected stock returns results because a better model of risk exists, or because factors other than risk also explain expected returns.

Based on the evidence using return data, one might be tempted to conclude, after properly taking into account the data mining bias discussed in Harvey, Liu, and Zhu (2014), that if multi-factor models do a superior job explaining the cross section, they necessarily explain risk better. But this conclusion is premature. To see why, consider the following analogy. Rather than look for an alternative theory, early astronomers reacted to the inability of the Ptolemaic theory to explain the motion of the planets by “fixing” each observational inconsistency by adding an additional epicycle to the theory. By the time Copernicus proposed the correct theory that the Earth revolved around the Sun, the Ptolemaic theory had been fixed so many times it *better* explained the motion of the planets than the Copernican system.[16] Similarly, although the extensions to the CAPM better explain the cross section of asset returns, it is hard to know, using traditional tests, whether these extensions represent true progress towards measuring risk or simply the asset pricing equivalent of an epicycle.

Our results shed light on this question. By our measures, factor models do no better explaining investor behavior than the CAPM even though they nest the CAPM. This fact reduces the likelihood that the reason these models better explain the cross section of expected returns is because they are better risk models. This is a key advantage of our testing method. It can differentiate between whether current extensions to the CAPM just improve the model's fit to existing data or whether they represent progress towards a better model of risk. The extensions of the CAPM model were proposed to better fit returns, not flows. As such, flows provide a new set of moments that those models can be confronted with. Consequently, if the extension of the original model better explains mutual fund flows, this suggests that the extension does indeed represent progress towards a superior risk model. Conversely, if the extended model cannot better explain flows, then we should worry that the extension is the modern equivalent of an epicycle, an arbitrary fix designed simply to ensure that the model better explains the cross section of returns.

Our method can also shed light on the third possibility, that expected returns might be a function of both risk and non-risk-based factors. To conclude that a better risk model exists, one has to show that the part of the variation in asset returns not explained by the CAPM can be explained by variation in risk. This is what the flow of funds data allow us to do. If variation in asset returns that is not explained by the CAPM attracts flows, then one can conclude that this variation is not compensation for risk. Thus, our method allows us to infer something existing tests of factor models cannot do. It allows us to determine whether or not a new factor that explains returns measures risk. What our results imply is that the factors that have been proposed do not measure additional risk not measured by the CAPM. What these factors actually do measure is clearly an important question for future research.

6. Tests of the robustness of our results

In this section we consider other possible alternative explanations for our results. First we look at the possibility that mutual fund fee changes might be part of the market equilibrating mechanism. Then we test the hypothesis that investors' information sets contain more than what is in past and present prices. Finally, we cut the data sample along two dimensions and examine whether our results change in the subsamples. Specifically, we examine whether our results change if we start the analysis in 1995 rather than 1977 and if we restrict attention to large return observations. In both cases we show that our results are unchanged in these subsamples.

6.1. Fee changes

As argued in the introduction, capital flows are not the only mechanism that could equilibrate the mutual fund market. An alternative mechanism is that fund managers adjust their fees to ensure that the fund's alpha is zero. In fact, fee changes are rare, occurring in less than 4% of our observations, making it unlikely that fee changes play any role in equilibrating the mutual fund market. Nevertheless, in this section we will run a robustness check to make sure that fee changes do not play a role in explaining our results.

The fees mutual funds charge are stable because they are specified in the fund's prospectus, so theoretically, a change to the fund's fee requires a change to the fund's prospectus, a relatively costly endeavor. However, the fee in the prospectus actually specifies the maximum fee the fund is allowed to charge because funds are allowed to (and do) rebate some of their fees to investors. Thus, funds can change their fees by giving or discontinuing rebates. To rule out these rebates as a possible explanation of our results, we repeat the above analysis by assuming that fee changes are the primary way mutual fund markets equilibrate.

We define a positive (negative) fee change as an increase (decrease) in the percentage fee charged from the beginning to the end of the horizon. For each fund, in periods that we observe a fee change, we assume the fee change is equilibrating the market and so the flow variable takes the sign of the fee change. In periods without a fee change, we continue to use the sign of the flows. That is, define F^*_{it} as

$$F^*_{it} \equiv \begin{cases} \Delta_{it}, & \Delta_{it} \neq 0 \\ F_{it}, & \Delta_{it} = 0, \end{cases}$$

where Δ_{it} is the fee change experienced by fund i at time t.

Table 5 reports the results of estimating $(\beta_{F^*\varepsilon} + 1)/2$, that is, the average conditional probability using the flow variable that includes fee changes. The results are qualitatively unchanged – the CAPM outperforms all the other models – and quantitatively very similar. More importantly, including fee changes in this way reduces the explanatory power of all the models (the point estimates in Table 5 are lower than in Table 2) so there is no evidence that fee changes play an important role in equilibrating the market for mutual funds.

6.2. Other information sets

Conceivably, the poor performance of some of the models reported in the last section could result because the assumption that the information set for most investors does not include any more information than past and present prices is incorrect. If this assumption is false and the information set of most investors includes information in addition to what is communicated by prices, what appears to us as a positive NPV investment might actually be zero NPV when viewed from the perspective of the actual information available at the time.

If information is indeed the explanation and if investors are right in their decision to allocate or withdraw money,

[16] Copernicus wrongly assumed that the planets followed circular orbits when in fact their orbits are ellipses.

Table 5
Effect of fee changes.

The table shows the effect of assuming that the market equilibrates through fee changes if they occur. That is, we use the sign of the fee change instead of the sign of the flow whenever we have a non zero fee change observation. In period when there is no fee change, we use the sign of the flow as before. The table reports $(\beta_{F^*\varepsilon} + 1)/2$ in percent. The first two rows report the results for the market model (CAPM) using the CRSP value-weighted index and the S&P 500 index as the market portfolio. The next three lines report the results of using as the benchmark return, three rules of thumb: (1) the fund's actual return, (2) the fund's return in excess of the risk-free rate, and (3) the fund's return in excess of the return on the market as measured by the CRSP value-weighted index. The next two lines are the Fama-French (FF) and Fama-French-Carhart (FFC) factor specifications. The final four lines report the results for the dynamic equilibrium models: the Consumption CAPM (C-CAPM), the habit model derived by Campbell and Cochrane (1999), and the long run risk model derived by Bansal and Yaron (2004). For the long-run risk model we consider three different versions, depending on the portfolio weight of bonds in the aggregate wealth portfolio. The maximum number in each column (the best performing model) is shown in bold face.

Model	Horizon					
	3-month	6-month	1-year	2-year	3-year	4-year
			Market models (CAPM)			
CRSP value weighted	**60.70**	**59.46**	**57.39**	**55.50**	**54.02**	**53.62**
S&P 500	59.85	58.37	56.55	54.85	52.75	52.48
			No model			
Return	57.66	57.41	53.68	49.51	50.22	50.93
Excess return	57.42	57.37	54.12	50.95	50.51	50.93
Return in excess of the market	59.45	58.39	56.07	54.76	53.28	52.74
			Multifactor models			
FF	60.21	58.88	57.04	54.91	53.16	52.76
FFC	60.30	58.89	57.14	54.74	53.36	52.81
			Dynamic equilibrium models			
C-CAPM	57.45	57.25	53.96	50.76	50.25	50.51
Habit	57.42	57.21	53.89	50.74	50.15	50.50
Long-run risk 0	50.59	50.64	50.30	49.66	49.63	49.66
Long-run risk 70	55.87	56.05	53.40	50.42	48.64	48.46
Long-run risk 90	56.27	56.79	54.49	52.10	50.98	51.92

Table 6
Out of sample persistence.

The table shows by how much the top alpha/bottom flow tercile outperforms the bottom alpha/top flow tercile, where outperformance is the estimated alpha under the given model. At time τ, we use all the information until that point in time to calculate the fund's information ratio (the estimated alpha divided by its standard error). We also calculate the fund's capital flow over the number of years equal to the specified horizon. We then sort firms into nine flow-performance terciles based on the information ratio and measured capital flow and then measure outperformance over the specified future measurement horizon. At the end of the measurement horizon, we then sort again and repeat the process as many times as the data allow. By the end of the process we have a time series of monthly outperformance measurements for each of the nine portfolios. We then subtract outperformance of the bottom information ratio /top flow from the top information ratio/bottom flow and the table reports the mean (in basis points (b.p.) per month) and t-statistic of this time series.

Model	Horizon (years)		
	2	3	4
CAPM (b.p./month)	0.00	1.57	−1.66
t-statistic	0.00	0.25	−0.26
Fama-French (b.p./month)	18.06	19.65	20.84
t-statistic	3.32	3.62	3.83
Fama-French-Carhart (b.p./month)	13.50	14.83	16.23
t-statistic	2.81	3.08	3.38
C-CAPM (b.p./month)	9.90	9.90	7.21
t-statistic	1.18	1.18	0.86
Habit (b.p./month)	10.22	9.86	7.96
t-statistic	1.21	1.17	0.95
Long-run risk – 0% bonds (%/month)	−13.60	−13.49	−13.58
t-statistic	−1.20	−1.19	−1.20
Long-run risk – 70% bonds (b.p./month)	−9.81	−19.43	−20.67
t-statistic	−0.65	−1.28	−1.36
Long-run risk – 90% bonds (b.p./month)	1.28	3.37	−5.32
t-statistic	0.16	0.43	−0.69

the alpha must be zero even when the flow has the opposite sign to the outperformance. We test this Null hypothesis by double sorting firms into terciles based on their past alpha as well as their past flows. Going forward, over a specified measurement horizon, h, we test to see whether funds in the highest alpha tercile and the lowest flow tercile outperform funds in the lowest alpha tercile and the highest flow tercile.[17] Put differently, we investigate whether previously outperforming funds that nevertheless experience an outflow of funds outperform previously underperforming funds that experience an inflow. Under the Null that the asset pricing model under consideration holds, these two portfolios should perform equally well going forward (both should have a zero net alpha in the measurement horizon).

The main difficulty with implementing this test is uncertainty in the estimate of the fund's betas for the factor models. When estimation error in the sorting period is positively correlated to the error in the measurement horizon, as would occur if we would estimate the betas only once over the full sample, a researcher could falsely conclude that evidence of persistence exists when there is no persistence. To avoid this bias we do not use information from the sorting period to estimate the betas in the measurement horizon. This means that we require a measurement horizon of sufficient length to produce reliable beta

[17] The sorts we do are unconditional sorts, meaning that we independently sort on flows and alpha. The advantage of this is that our results are not influenced by the ordering of our sorts. The downside is that the nine "portfolios" do not have the same number of funds in them.

Table 7

Flow of funds outperformance relationship (1995–2011).

The table reports estimates of (6) for different asset pricing models. For ease of interpretation, the table reports $(\beta_{F_\varepsilon}+1)/2$ in percent, which by Lemma 2 is equivalent to $(\Pr[\phi(F_{it})=1|\phi(\varepsilon_{it})=1]+\Pr[\phi(F_{it})=-1|\phi(\varepsilon_{it})=-1])/2$. Each row corresponds to a different risk model. The first two rows report the results for the market model (CAPM) using the CRSP value-weighted index and the S&P 500 index as the market portfolio. The next three lines report the results of using as the benchmark return, three rules of thumb: (1) the fund's actual return, (2) the fund's return in excess of the risk-free rate, and (3) the fund's return in excess of the return on the market as measured by the CRSP value-weighted index. The next two lines are the Fama-French (FF) and Fama-French-Carhart (FFC) factor specifications. The final four lines report the results for the dynamic equilibrium models: the Consumption CAPM (C-CAPM), the habit model derived by Campbell and Cochrane (1999), and the long-run risk model derived by Bansal and Yaron (2004). For the long-run risk model we consider three different versions, depending on the portfolio weight of bonds in the aggregate wealth portfolio. The maximum number in each column (the best performing model) is shown in bold face.

Model	Horizon		
	3 month	6 month	1 year
Market models (CAPM)			
CRSP value weighted	**62.74**	**62.68**	62.70
S&P 500	61.44	61.23	60.77
No model			
Return	57.94	59.48	57.45
Excess return	57.67	59.27	57.44
Return in excess of the market	61.18	61.31	60.33
Multifactor Models			
FF	62.42	62.20	**62.80**
FFC	62.57	62.35	62.71
Dynamic equilibrium models			
C-CAPM	57.73	59.13	57.29
Habit	57.75	59.10	57.22
Long-run risk 0	56.69	58.70	60.36
Long-run risk 70	56.34	57.90	58.56
Long-run risk 90	56.44	58.80	58.84

estimates, so the shortest measurement horizon we consider is two years.

At time τ, we use all the information until that point in time to calculate the fund's information ratio, that is, we estimate the fund's alpha using all of its return data up to time τ and divide this by the standard error of the estimate. We then calculate the fund's capital flow over the prior h years. We sort firms into nine flow performance terciles based on the estimated information ratio and measured capital flow. We require a fund to have at least three years of historical data to be included in the sort. Because we need at least six months to estimate the fund's betas in the measurement horizon, we drop all funds with less than six observations in the measurement horizon. To remove the obvious selection bias, we estimate the betas over the full measurement horizon, but then calculate ε_{it} by dropping the first six observations, that is, we only use $\{\varepsilon_{i,\tau+6},\ldots,\varepsilon_{i,\tau+h}\}$ when we measure future performance. At the end of the measurement horizon we then sort again and repeat the process as many times as the data allow. By the end of the process we have a time series of monthly outperformance measurements for each of the nine portfolios. We then subtract the outperformance of the bottom information ratio/top flow portfolio from the top information ratio/bottom flow portfolio. Table 6 reports the mean and t-statistic of this time series for horizons $h=2,3$, and 4 years.

The main takeaway from the results reported in Table 6 is that outperformance relative to the CAPM shows no evidence of persistence while outperformance relative to the other factor models is highly persistent and economically large. Consequently, we can confidently reject the Null hypothesis that the poor performance of the factor models relative to the CAPM is attributable to investors using a richer information set than the set containing just the information in past returns.

We find no evidence of predictability for the dynamic equilibrium models. In this case the likelihood that investors have better information is higher because they observe their own consumption. So the lack of predictability is consistent with the possibility that the poor performance of these models is due to the fact that the empiricist measures consumption with error.[18]

6.3. Post-1994 sample

Initial readers of this paper have pointed out that the particular factors that we test in our factor models were discovered in the early nineties and so part of the poor performance of these models might be attributable to the fact that investors did not know about the existence of these factors in the early part of our data set. Under the hypothesis put forward by these readers, once the factors were discovered, people started using them, and so the appropriate time period to compare the CAPM to these factor models is the post-1995 period. Such a view raises interesting questions about the role of economic research. Rather than just trying to discover what asset pricing model people use, under this view, economic researchers also have a role teaching people what model they should be using. To see if there is any support for this hypothesis in the data, we rerun our tests in the sample that excludes data prior to 1995.

Because the time series of this subsample covers just 16 years, we repeat the analysis using horizons of a year or less.[19] Tables 7 and 8 report the results. They are quantitatively very similar to the full sample, and qualitatively the same. At every horizon the performance of the factor models and the CAPM are statistically indistinguishable. At the 3- and 6-month horizons, the CAPM actually outperforms both factor models. It also still significantly outperforms the "no model" benchmarks. In addition, the dynamic equilibrium models continue to perform poorly. In summary, there is no detectable evidence that the discovery of the value, size, and momentum factors had any influence on how investors measure risk.

[18] Note that the outperformance point estimate for the long-run risk model when the wealth portfolio consists entirely of stocks is four orders of magnitude higher than all other models, despite the fact that it is still statistically indistinguishable from zero. As we have already pointed out, given the volatility of stocks, the SDF of this model is extremely volatile leading to highly volatile estimates of outperformance for this model.

[19] Because of the loss in data, at longer horizons the double-clustered standard errors are so large that there is little power to differentiate between models.

Table 8
Tests of statistical significance (post-1994 subsample).

The first two columns in the table provide the coefficient estimate and double-clustered t-statistic (see Thompson, 2011 and the discussion in Petersen, 2009) of the univariate regression of signed flows on signed outperformance. The rest of the columns provide the statistical significance of the pairwise test, derived in Proposition 5, of whether the models are better approximations of the true asset pricing model. For each model in a column, the table displays the double-clustered t-statistic of the test that the model in the row is a better approximation of the true asset pricing model, that is, that $\beta_{\text{Frow}} > \beta_{\text{Fcolumn}}$. The rows (and columns) are ordered by β_{F_ε}, with the best performing model on top and on the left. The number following the long-run risk models denotes the percentage of the wealth portfolio invested in bonds.

Panel A: 3-Month horizon

Model	β_{F_ε}	Univ t-stat	*CAPM*	*FFC*	*FF*	*CAPM* ***SP500***	Ex. mkt	Ret	*Habit*	*C-CAPM*	Ex. ret	*LRR 0*	*LRR 90*	*LRR 70*
CAPM	0.255	21.48	0.00	0.43	0.81	4.94	6.32	4.33	4.81	4.85	4.94	5.89	6.13	6.49
FFC	0.251	24.38	−0.43	0.00	0.71	2.23	3.16	3.95	4.39	4.43	4.53	5.56	5.57	5.96
FF	0.248	23.81	−0.81	−0.71	0.00	1.92	2.76	3.87	4.32	4.36	4.45	5.39	5.51	5.86
CAPM ***SP500***	0.229	16.94	−4.94	−2.23	−1.92	0.00	0.64	3.19	3.56	3.60	3.68	4.44	4.92	5.01
Excess Market	0.224	18.39	−6.32	−3.16	−2.76	−0.64	0.00	2.88	3.21	3.24	3.32	4.10	4.39	4.64
Return	0.159	8.86	−4.33	−3.95	−3.87	−3.19	−2.88	0.00	0.94	1.02	1.19	1.17	3.82	2.86
Habit	0.155	8.36	−4.81	−4.39	−4.32	−3.56	−3.21	−0.94	0.00	0.78	0.81	1.05	3.68	2.90
C-CAPM	0.155	8.41	−4.85	−4.43	−4.36	−3.60	−3.24	−1.02	−0.78	0.00	0.69	1.03	3.72	2.89
Excess Return	0.153	8.32	−4.94	−4.53	−4.45	−3.68	−3.32	−1.19	−0.81	−0.69	0.00	0.99	4.19	3.14
LRR 0	0.134	6.32	−5.89	−5.56	−5.39	−4.44	−4.10	−1.17	−1.05	−1.03	−0.99	0.00	0.28	0.49
LRR 90	0.129	6.61	−6.13	−5.57	−5.51	−4.92	−4.39	−3.82	−3.68	−3.72	−4.19	−0.28	0.00	0.31
LRR 70	0.127	6.32	−6.49	−5.96	−5.86	−5.01	−4.64	−2.86	−2.90	−2.89	−3.14	−0.49	−0.31	0.00

Panel B: 6-Month horizon

Model	β_{F_ε}	Univ t-stat	*CAPM*	*FFC*	*FF*	Ex mkt	*CAPM* ***SP500***	Ret	Ex ret	*C-CAPM*	*Habit*	*LRR 90*	*LRR 0*	*LRR 70*
CAPM	0.254	17.08	0.00	0.52	0.73	3.57	3.44	2.12	2.54	2.56	2.57	3.11	2.06	3.98
FFC	0.247	18.17	−0.52	0.00	0.52	1.40	1.34	1.80	2.16	2.21	2.23	2.43	1.87	3.36
FF	0.244	18.90	−0.73	−0.52	0.00	1.20	1.09	1.65	1.97	2.00	2.02	2.35	1.77	3.23
Excess Market	0.226	13.00	−3.57	−1.40	−1.20	0.00	0.13	1.10	1.34	1.39	1.41	1.68	1.28	2.37
CAPM ***SP500***	0.225	11.13	−3.44	−1.34	−1.09	−0.13	0.00	1.16	1.39	1.46	1.48	1.70	1.29	2.36
Return	0.190	7.24	−2.12	−1.80	−1.65	−1.10	−1.16	0.00	0.58	0.94	0.99	0.84	0.35	1.38
Excess Return	0.185	6.98	−2.54	−2.16	−1.97	−1.34	−1.39	−0.58	0.00	0.93	1.03	0.63	0.27	1.43
C-CAPM	0.183	6.78	−2.56	−2.21	−2.00	−1.39	−1.46	−0.94	−0.93	0.00	0.66	0.39	0.20	1.17
Habit	0.182	6.72	−2.57	−2.23	−2.02	−1.41	−1.48	−0.99	−1.03	−0.66	0.00	0.34	0.19	1.12
LRR 90	0.176	6.32	−3.11	−2.43	−2.35	−1.68	−1.70	−0.84	−0.63	−0.39	−0.34	0.00	0.05	1.27
LRR 0	0.174	5.99	−2.06	−1.87	−1.77	−1.28	−1.29	−0.35	−0.27	−0.20	−0.19	−0.05	0.00	0.55
LRR 70	0.158	5.35	−3.98	−3.36	−3.23	−2.37	−2.36	−1.38	−1.43	−1.17	−1.12	−1.27	−0.55	0.00

Panel C: 1-Year horizon

Model	β_{F_ε}	Univ t-stat	*FF*	*FFC*	*CAPM*	*CAPM* ***SP500***	*LRR 0*	Ex mkt	*LRR 90*	*LRR 70*	Ret	Ex ret	*C-CAPM*	*Habit*
FF	0.256	12.65	0.00	0.33	0.12	1.72	0.65	2.59	1.80	2.49	1.94	2.29	2.32	2.33
FFC	0.254	12.21	−0.33	0.00	0.02	1.77	0.63	2.54	1.80	2.37	1.96	2.35	2.37	2.38
CAPM	0.254	11.16	−0.12	−0.02	0.00	3.40	0.65	6.32	2.00	2.84	1.98	2.43	2.39	2.40
CAPM ***SP500***	0.215	6.47	−1.72	−1.77	−3.40	0.00	0.13	0.48	1.01	1.33	1.35	1.57	1.58	1.60
LRR 0	0.207	6.76	−0.65	−0.63	−0.65	−0.13	0.00	0.01	0.41	0.70	0.77	0.86	0.89	0.92
Excess Market	0.207	8.26	−2.59	−2.54	−6.32	−0.48	−0.01	0.00	0.63	1.06	0.97	1.14	1.15	1.17
LRR 90	0.177	3.78	−1.80	−1.80	−2.00	−1.01	−0.41	−0.63	0.00	0.16	1.16	1.54	1.58	1.67
LRR 70	0.171	4.98	−2.49	−2.37	−2.84	−1.33	−0.70	−1.06	−0.16	0.00	0.45	0.58	0.63	0.66
Return	0.149	3.54	−1.94	−1.96	−1.98	−1.35	−0.77	−0.97	−1.16	−0.45	0.00	0.02	0.21	0.30
Excess Return	0.149	3.46	−2.29	−2.35	−2.43	−1.57	−0.86	−1.14	−1.54	−0.58	−0.02	0.00	0.53	0.77
C-CAPM	0.146	3.51	−2.32	−2.37	−2.39	−1.58	−0.89	−1.15	−1.58	−0.63	−0.21	−0.53	0.00	1.35
Habit	0.144	3.45	−2.33	−2.38	−2.40	−1.60	−0.92	−1.17	−1.67	−0.66	−0.30	−0.77	−1.35	0.00

6.4. *Restricting the sample to large returns*

One important advantage of our method, which uses only the signs of flows and returns, is that it is robust to outliers. However, this also comes with the important potential limitation that we ignore the information contained in the magnitude of the outperformance and the flow of fund response. It is conceivable that investors might react differently to large and small return outperformance. For example, a small abnormal return might lead investors to update their priors of managerial performance only marginally. Assuming that investors face some cost to transact, it might not be profitable for investors to react to this information by adjusting their investment in the mutual fund. To examine the importance of this hypothesis, we rerun our tests in a subsample that does not include small return realizations.

We focus on deviations from the market return, and begin by dropping all return observations that deviate from the market return by less than 0.1 standard deviation (of the panel of deviations from the market return). The first

Table 9
Flow of funds outperformance relationship for observations with extreme returns.

The table re-estimates the regression in Table 2 in different subsamples at the 3 month horizon. In each case observations that do not deviate by a predetermined amount from the market return over the same horizon are discarded. The first column reports results when we discard all observations with returns that did not deviate from the market return by more than 0.1 of a standard deviation. The other columns do the same thing but for progressively larger windows – 0.25, 0.5, 0.75, and 1 standard deviation. For ease of interpretation, the table reports $(\beta_{F_\epsilon}+1)/2$ in percent. The first row of the table reports the fraction of observations that are dropped. Each row after that corresponds to a different risk model. The second and third rows report the results for the market model (CAPM) using the CRSP value-weighted index and the S&P 500 index as the market portfolio. The next three lines report the results of using as the benchmark return, three rules of thumb: (1) the fund's actual return, (2) the fund's return in excess of the risk-free rate, and (3) the fund's return in excess of the return on the market as measured by the CRSP value-weighted index. The next two lines are the Fama-French (FF) and Fama-French-Carhart (FFC) factor specifications. The final four lines report the results for the dynamic equilibrium models: the Consumption CAPM (C-CAPM), the habit model, and the long-run risk model. The maximum number in each column (the best performing model) is shown in bold face.

Model	Drop window (in units of standard deviation)				
	0.1	0.25	0.5	0.75	1
Fraction of data discarded (%)	13.90	32.54	55.95	70.78	80.01
Market models (CAPM)					
CRSP value weighted	**64.97**	**67.14**	**70.53**	73.10	**75.38**
S&P 500	64.08	66.75	70.44	**73.11**	75.38
No model					
Return	59.20	60.53	63.12	65.79	68.25
Excess return	59.02	60.42	63.15	66.04	68.54
Return in excess of the market	63.85	66.31	69.96	72.72	75.06
Multifactor models					
FF	64.41	66.58	69.90	72.48	74.78
FFC	64.40	66.44	69.68	72.19	74.58
Dynamic equilibrium models					
C-CAPM	59.00	60.41	63.19	66.07	68.54
Habit	58.99	60.37	63.17	66.04	68.51
Long-run risk – 0% bonds	58.03	59.13	61.49	63.96	66.03
Long-run risk – 70% bonds	57.80	59.28	62.23	65.07	67.34
Long-run risk – 90% bonds	57.86	59.36	62.01	64.62	66.82

column of Table 9 reports the results of our earlier tests at the 3-month horizon in this subsample.[20] The other columns in the table increase the window of dropped observations: 0.25, 0.5, 0.75, and 1 standard deviation. Table 10 reports the statistical significance in these subsamples of the test derived in Proposition 5. The results are again quantitatively similar to the main sample and qualitatively identical. The CAPM is statistically significantly better at explaining flows than the “no model” benchmarks, and none of the factor models statistically outperform the CAPM.

It might seem reasonable to infer from the results in Table 9 that transaction costs do explain the overall poor performance of all the models in explaining flows. But caution is in order here. Although the CAPM does explain 75% of flow observations at the 1-standard deviation window, in this sample, 80% of the data are discarded. It seems hard to believe that transaction costs are so high that only the 20% most extreme observations contain enough information to be worth transacting on.

7. Conclusion

The field of asset pricing is primarily concerned with the question of how to compute the cost of capital for investment opportunities. Because the net present value of a long-dated investment opportunity is very sensitive to assumptions regarding the cost of capital, computing this cost of capital correctly is of first order importance. Since the initial development of the Capital Asset Pricing Model, a large number of potential return anomalies relative to that model have been uncovered. These anomalies have motivated researchers to develop improved models that “explain” each anomaly as a risk factor. As a consequence, in many (if not most) research studies these factors and their exposures are included as part of the cost of capital calculation. In this paper we examine the validity of this approach to calculating the cost of capital.

The main contribution of this paper is a new way of testing the validity of an asset pricing model. Instead of following the common practice in the literature which relies on moment conditions related to returns, we use mutual fund capital flow data. Our study is motivated by revealed preference theory: if the asset pricing model under consideration correctly prices risk, then investors must be using it, and must be allocating their money based on that risk model. Consistent with this theory, we find that investors' capital flows in and out of mutual funds does reliably distinguish between asset pricing models. We find that the CAPM outperforms all extensions to the original model, which implies, given our current level of knowledge, that it is still the best method to use to compute the cost of capital of an investment opportunity. This observation is consistent with actual experience. Despite the empirical shortcomings of the CAPM, Graham and Harvey (2001) find that it is the dominant model used by corporations to make investment decisions.

The results in the paper raise a number of puzzles. First, and foremost, there is the apparent inconsistency that the CAPM does a poor job explaining cross-sectional variation in expected returns even though investors appear to use the CAPM beta to measure risk. Explaining this puzzling fact is an important area for future research.

A second puzzle that bears investigating is the growth in the last 20 years of value and growth mutual funds. There are a number of possibilities. First, investors might see these funds as a convenient way to characterize CAPM beta risk. Why investors would use these criteria rather than beta itself is unclear. If this explanation is correct, the answer is most likely related to the same reason the CAPM does such a poor job in the cross section. Another possibility is that value and growth funds are not riskier and so offer investors a convenient way to invest in positive net present value strategies. But this explanation begs the question of why competition between these funds has not eliminated such opportunities. It is quite likely that by separately investigating what drives

[20] Results for the one-year horizon are reported in the Internet Appendix to this paper. We choose to report the short-horizon results because as before, the results for longer horizons have little statistical power to differentiate between models.

Table 10
Tests of statistical significance in the extreme return sample at the 3-month horizon.

The table re-estimates the regression in Table 2 in different subsamples at the 3-month horizon. In each case observations that do not deviate by a predetermined amount from the market return over the same horizon are discarded. The first two columns in the table provide the coefficient estimate and double-clustered *t*-statistic of the univariate regression of signed flows on signed outperformance. The rest of the columns provide the statistical significance of the pairwise test, derived in Proposition 5, of whether the models are better approximations of the true asset pricing model. For each model in a column, the table displays the double-clustered *t*-statistic of the test that the model in the row is a better approximation of the true asset pricing model, that is, that $\beta_{F\text{row}} > \beta_{F\text{column}}$. The rows (and columns) are ordered by $\beta_{F\varepsilon}$, with the best performing model on top. The number following the long-run risk models denotes the percentage of the wealth portfolio invested in bonds.

Panel A: 0.1 Standard deviation window

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	FF	FFC	CAPM SP500	Ex. mkt	Ret	Ex. ret	C-CAPM	Habit	LRR 0	LRR 90	LRR 70
CAPM	0.299	26.06	0.00	1.67	1.67	4.13	5.81	5.57	6.31	6.28	6.28	7.65	7.64	8.10
FF	0.288	29.52	−1.67	0.00	0.05	0.78	1.44	4.86	5.52	5.50	5.50	6.80	6.71	7.11
FFC	0.288	28.78	−1.67	−0.05	0.00	0.76	1.44	4.82	5.47	5.45	5.44	6.81	6.65	7.09
CAPM SP500	0.282	23.04	−4.13	−0.78	−0.76	0.00	0.72	4.92	5.53	5.53	5.52	6.51	6.98	7.11
Excess Market	0.277	22.59	−5.81	−1.44	−1.44	−0.72	0.00	4.34	4.90	4.87	4.87	6.03	6.06	6.42
Return	0.184	11.24	−5.57	−4.86	−4.82	−4.92	−4.34	0.00	0.77	0.91	0.96	1.21	3.58	2.79
Excess Return	0.180	10.68	−6.31	−5.52	−5.47	−5.53	−4.90	−0.77	0.00	0.21	0.33	1.12	4.13	3.26
C-CAPM	0.180	10.71	−6.28	−5.50	−5.45	−5.53	−4.87	−0.91	−0.21	0.00	0.47	1.08	3.55	2.90
Habit	0.180	10.60	−6.28	−5.50	−5.44	−5.52	−4.87	−0.96	−0.33	−0.47	0.00	1.07	3.41	2.83
LRR 0	0.161	8.26	−7.65	−6.80	−6.81	−6.51	−6.03	−1.21	−1.12	−1.08	−1.07	0.00	0.21	0.35
LRR 90	0.157	8.83	−7.64	−6.71	−6.65	−6.98	−6.06	−3.58	−4.13	−3.55	−3.41	−0.21	0.00	0.19
LRR 70	0.156	8.57	−8.10	−7.11	−7.09	−7.11	−6.42	−2.79	−3.26	−2.90	−2.83	−0.35	−0.19	0.00

Panel B: 0.25 Standard deviation window

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	CAPM SP500	FF	FFC	Ex. mkt	Ret	Ex. ret	C-CAPM	Habit	LRR 90	LRR 70	LRR 0
CAPM	0.343	24.64	0.00	2.15	1.63	1.99	5.07	6.22	6.89	6.87	6.90	8.05	8.49	8.32
CAPM SP500	0.335	23.85	−2.15	0.00	0.42	0.75	1.66	6.27	6.93	6.93	6.96	8.15	8.42	7.96
FF	0.331	29.30	−1.63	−0.42	0.00	0.72	0.71	5.44	6.06	6.04	6.07	7.04	7.48	7.52
FFC	0.329	27.94	−1.99	−0.75	−0.72	0.00	0.35	5.37	5.95	5.92	5.95	6.93	7.39	7.39
Excess Market	0.326	21.89	−5.07	−1.66	−0.71	−0.35	0.00	5.17	5.70	5.68	5.70	6.72	7.12	7.05
Return	0.211	11.76	−6.22	−6.27	−5.44	−5.37	−5.17	0.00	0.54	0.65	0.79	3.35	2.63	1.43
Excess Return	0.208	11.40	−6.89	−6.93	−6.06	−5.95	−5.70	−0.54	0.00	0.15	0.44	3.60	3.04	1.43
C-CAPM	0.208	11.43	−6.87	−6.93	−6.04	−5.92	−5.68	−0.65	−0.15	0.00	0.83	3.39	2.89	1.40
Habit	0.207	11.34	−6.90	−6.96	−6.07	−5.95	−5.70	−0.79	−0.44	−0.83	0.00	3.21	2.79	1.37
LRR 90	0.187	9.55	−8.05	−8.15	−7.04	−6.93	−6.72	−3.35	−3.60	−3.39	−3.21	0.00	0.26	0.26
LRR 70	0.186	9.48	−8.49	−8.42	−7.48	−7.39	−7.12	−2.63	−3.04	−2.89	−2.79	−0.26	0.00	0.22
LRR 0	0.183	8.80	−8.32	−7.96	−7.52	−7.39	−7.05	−1.43	−1.43	−1.40	−1.37	−0.26	−0.22	0.00

Panel C: 0.5 Standard deviation window

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM	CAPM SP500	Ex. mkt	FF	FFC	C-CAPM	Ex. ret	Habit	Ret	LRR 70	LRR 90	LRR 0
CAPM	0.411	21.79	0.00	0.63	3.82	1.69	2.23	7.02	7.08	7.06	6.64	8.20	8.12	8.50
CAPM SP500	0.409	22.12	−0.63	0.00	2.07	1.33	1.87	7.46	7.50	7.51	7.05	8.58	8.61	8.60
Excess Market	0.399	20.34	−3.82	−2.07	0.00	0.16	0.73	6.12	6.18	6.14	5.81	7.19	7.11	7.62
FF	0.398	26.43	−1.69	−1.33	−0.16	0.00	1.21	6.35	6.39	6.37	5.97	7.44	7.38	7.88
FFC	0.394	25.21	−2.23	−1.87	−0.73	−1.21	0.00	6.18	6.21	6.20	5.84	7.25	7.18	7.64
C-CAPM	0.264	11.52	−7.02	−7.46	−6.12	−6.35	−6.18	0.00	0.50	0.42	0.36	2.24	3.72	1.81
Excess Return	0.263	11.39	−7.08	−7.50	−6.18	−6.39	−6.21	−0.50	0.00	−0.22	0.15	2.28	3.78	1.78
Habit	0.263	11.50	−7.06	−7.51	−6.14	−6.37	−6.20	−0.42	0.22	0.00	0.25	2.20	3.66	1.79
Return	0.262	11.59	−6.64	−7.05	−5.81	−5.97	−5.84	−0.36	−0.15	−0.25	0.00	1.74	2.82	1.61
LRR 70	0.245	10.11	−8.20	−8.58	−7.19	−7.44	−7.25	−2.24	−2.28	−2.20	−1.74	0.00	0.71	1.07
LRR 90	0.240	9.87	−8.12	−8.61	−7.11	−7.38	−7.18	−3.72	−3.78	−3.66	−2.82	−0.71	0.00	0.58
LRR 0	0.230	9.35	−8.50	−8.60	−7.62	−7.88	−7.64	−1.81	−1.78	−1.79	−1.61	−1.07	−0.58	0.00

Panel D: 0.75 Standard deviation window

Model	$\beta_{F\varepsilon}$	Univ t-stat	CAPM SP500	CAPM	Ex. mkt	FF	FFC	C-CAPM	Habit	Ex. ret	Ret	LRR 70	LRR 90	LRR 0
CAPM SP500	0.462	20.24	0.00	0.15	1.77	1.49	2.17	6.83	6.90	6.82	6.76	7.54	7.99	8.04
CAPM	0.462	19.79	−0.15	0.00	2.83	1.47	2.09	6.42	6.47	6.44	6.36	7.19	7.55	7.82
Excess Market	0.455	18.78	−1.77	−2.83	0.00	0.58	1.23	5.69	5.73	5.71	5.65	6.45	6.74	7.12
FF	0.450	24.31	−1.49	−1.47	−0.58	0.00	1.53	6.01	6.06	6.00	5.91	6.83	7.17	7.45
FFC	0.444	23.29	−2.17	−2.09	−1.23	−1.53	0.00	5.82	5.87	5.79	5.76	6.56	6.95	7.07
C-CAPM	0.321	11.43	−6.83	−6.42	−5.69	−6.01	−5.82	0.00	0.47	0.31	1.64	2.05	4.34	2.45
Habit	0.321	11.44	−6.90	−6.47	−5.73	−6.06	−5.87	−0.47	0.00	0.04	1.42	2.01	4.25	2.43
Excess Return	0.321	11.31	−6.82	−6.44	−5.71	−6.00	−5.79	−0.31	−0.04	0.00	1.26	2.13	4.65	2.44
Return	0.316	11.32	−6.76	−6.36	−5.65	−5.91	−5.76	−1.64	−1.42	−1.26	0.00	1.26	2.80	1.94

Table 10 (*continued*)

Model	$\beta_{F\varepsilon}$	Univ *t*-stat	*CAPM* ***SP500***	***CAPM***	Ex. mkt	***FF***	***FFC***	*C-CAPM*	*Habit*	Ex. ret	Ret	*LRR 70*	*LRR 90*	*LRR 0*
Panel D: 0.75 Standard deviation window														
LRR 70	0.301	10.22	−7.54	−7.19	−6.45	−6.83	−6.56	−2.05	−2.01	−2.13	−1.26	0.00	1.21	1.89
LRR 90	0.292	10.02	−7.99	−7.55	−6.74	−7.17	−6.95	−4.34	−4.25	−4.65	−2.80	−1.21	0.00	0.78
LRR 0	0.279	9.52	−8.04	−7.82	−7.12	−7.45	−7.07	−2.45	−2.43	−2.44	−1.94	−1.89	−0.78	0.00
Panel E: 1 Standard deviation window														
Model	$\beta_{F\varepsilon}$	Univ *t*-stat	***CAPM***	*CAPM* ***SP500***	Ex. mkt	***FF***	***FFC***	Ex. ret	*C-CAPM*	*Habit*	Ret	*LRR 70*	*LRR 90*	*LRR 0*
CAPM	0.508	17.88	0.00	0.02	2.02	1.26	1.64	5.74	5.74	5.78	5.80	6.26	6.94	7.00
CAPM ***SP500***	0.508	18.14	−0.02	0.00	1.33	1.24	1.66	6.00	6.01	6.05	6.08	6.47	7.24	7.16
Excess Market	0.501	17.23	−2.02	−1.33	0.00	0.58	0.96	5.03	5.03	5.06	5.08	5.62	6.15	6.38
FF	0.496	22.08	−1.26	−1.24	−0.58	0.00	1.10	5.42	5.48	5.54	5.42	6.13	6.72	6.84
FFC	0.492	21.57	−1.64	−1.66	−0.96	−1.10	0.00	5.25	5.32	5.38	5.30	5.89	6.53	6.51
Excess Return	0.371	10.83	−5.74	−6.00	−5.03	−5.42	−5.25	0.00	−0.07	0.24	1.31	2.49	6.32	3.19
C-CAPM	0.371	10.95	−5.74	−6.01	−5.03	−5.48	−5.32	0.07	0.00	0.50	1.56	2.41	5.63	3.15
Habit	0.370	10.98	−5.78	−6.05	−5.06	−5.54	−5.38	−0.24	−0.50	0.00	1.33	2.35	5.19	3.11
Return	0.365	10.73	−5.80	−6.08	−5.08	−5.42	−5.30	−1.31	−1.56	−1.33	0.00	1.43	3.31	2.42
LRR 70	0.347	9.86	−6.26	−6.47	−5.62	−6.13	−5.89	−2.49	−2.41	−2.35	−1.43	0.00	1.18	2.47
LRR 90	0.336	9.80	−6.94	−7.24	−6.15	−6.72	−6.53	−6.32	−5.63	−5.19	−3.31	−1.18	0.00	0.98
LRR 0	0.321	9.13	−7.00	−7.16	−6.38	−6.84	−6.51	−3.19	−3.15	−3.11	−2.42	−2.47	−0.98	0.00

flows into and out of these funds, new light can be shed on what motivates investors to invest in these funds.

Finally, there is the question of what drives the fraction of flows that are unrelated to CAPM beta risk. A thorough investigation of what exactly drives these flows is likely to be highly informative about how risk is incorporated into asset prices.

Perhaps the most important implication of our paper is that it highlights the usefulness and power of mutual fund data when addressing general asset pricing questions. Mutual fund data provide insights into questions that stock market data cannot. Because the market for mutual funds equilibrates through capital flows instead of prices, we can directly observe investors' investment decisions. That allows us to infer their risk preferences from their actions. The observability of these choices and what this implies for investor preferences has remained largely unexplored in the literature.

Appendix A

Proof of Proposition 2. The denominator of (6) is positive so we need to show that the numerator is positive as well. Conditioning on the information set at each point in time gives the following expression for the numerator:

$$cov(\phi(F_{it+1}), \phi(\varepsilon_{it+1})) = \mathrm{E}\big[\,\mathrm{E}[\phi(F_{it+1})\phi(\varepsilon_{it+1}))\big|I_t]\big] - \mathrm{E}\big[\,\mathrm{E}[\phi(F_{it+1})\big|I_t]\big]\mathrm{E}\big[\,\mathrm{E}[\phi(\varepsilon_{it+1})\big|I_t]\big]. \tag{14}$$

Taking each term separately,

$$\begin{aligned}\mathrm{E}[\phi(F_{it+1})\phi(\varepsilon_{it+1})\big|I_t] &= \mathrm{E}[\phi(\varepsilon_{it+1})\phi(\alpha_{it+1}(q_{it}))\big|I_t]\\ &= \mathrm{E}[\phi(\varepsilon_{it+1})\phi(\alpha_{it+1}(q_{it}))\big|\theta_i > \bar{\theta}_{it}, I_t]\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\\ &\quad + \mathrm{E}[\phi(\varepsilon_{it+1})\phi(\alpha_{it+1}(q_{it}))\big|\theta_i \le \bar{\theta}_{it}, I_t]\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\\ &= \mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i > \bar{\theta}_{it}, I_t]\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\\ &\quad - \mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i \le \bar{\theta}_{it}, I_t]\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t],\end{aligned}$$

where the first equality follows from Lemma 1 and the last equality follows from (2) and (3) because when $\theta_i > \bar{\theta}_{it}$, then $\alpha_{it+1}(q_{it}) > 0$ and similarly for $\theta_i \le \bar{\theta}_{it}$. Using similar logic,

$$\begin{aligned}\mathrm{E}[\phi(F_{it+1})\big|I_t] &= \mathrm{E}[\phi(\alpha_{it+1}(q_{it}))\big|\theta_i > \bar{\theta}_{it}, I_t]\Pr[\theta_i > \bar{\theta}_{it}\big|I_t] + \mathrm{E}[\phi(\alpha_{it+1}(q_{it}))\big|\theta_i \le \bar{\theta}_{it}, I_t]\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\\ &= \Pr[\theta_i > \bar{\theta}_{it}\big|I_t] - \Pr[\theta_i \le \bar{\theta}_{it}\big|I_t],\end{aligned}$$

and

$$\mathrm{E}[\phi(\varepsilon_{it+1})\big|I_t] = \mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i > \bar{\theta}_{it}, I_t]\Pr[\theta_i > \bar{\theta}_{it}\big|I_t] + \mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i \le \bar{\theta}_{it}, I_t]\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t].$$

Using these three expressions we have

$$
\begin{aligned}
&\mathrm{E}\Big[\mathrm{E}[\phi(F_{it+1})\phi(\varepsilon_{it+1}))\big|I_t]\Big] - \mathrm{E}\Big[\mathrm{E}[\phi(F_{it+1})\big|I_t]\Big]\mathrm{E}\Big[\mathrm{E}[\phi(\varepsilon_{it+1})\big|I_t]\Big] \\
&\quad = \mathrm{E}\Big[\mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i > \bar{\theta}_{it}, I_t]\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\Big] - \mathrm{E}\Big[\mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i \le \bar{\theta}_{it}, I_t]\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\Big] \\
&\qquad - \mathrm{E}\Big[\mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i > \bar{\theta}_{it}, I_t]\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\Big]\Big(\mathrm{E}\Big[\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\Big] - \mathrm{E}\Big[\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\Big]\Big) \\
&\qquad - \mathrm{E}\Big[\mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i \le \bar{\theta}_{it}, I_t]\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\Big]\Big(\mathrm{E}\Big[\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\Big] - \mathrm{E}\Big[\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\Big]\Big) \\
&\quad = \mathrm{E}\Big[\mathrm{E}[\phi(\varepsilon_{it+1})\big|\theta_i > \bar{\theta}_{it}, I_t]\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\Big]\Big(1 - \mathrm{E}\Big[\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\Big] + \mathrm{E}\Big[\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\Big]\Big) \\
&\qquad\quad + \mathrm{E}\Big[\mathrm{E}[-\phi(\varepsilon_{it+1})\big|\theta_i \le \bar{\theta}_{it}, I_t]\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\Big]\Big(1 + \mathrm{E}\Big[\Pr[\theta_i > \bar{\theta}_{it}\big|I_t]\Big] - \mathrm{E}\Big[\Pr[\theta_i \le \bar{\theta}_{it}\big|I_t]\Big]\Big) \\
&\qquad > 0
\end{aligned}
$$

because every term in the last equation is positive. Substituting the above expression into (14) completes the proof.□

Proof of Lemma 2. First, by using Bayes' law and by rearranging terms we have

$$
\begin{aligned}
\Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = -1\Big] &= \frac{\Pr\Big[\phi(\varepsilon_{it}) = -1\big|\phi(F_{it}) = 1\Big]\Pr\Big[\phi(F_{it}) = 1\Big]}{\Pr\Big[\phi(\varepsilon_{it}) = -1\Big]} \\
&= \frac{\Big(1 - \Pr\Big[\phi(\varepsilon_{it}) = 1\big|\phi(F_{it}) = 1\Big]\Big)\Pr\Big[\phi(F_{it}) = 1\Big]}{1 - \Pr\Big[\phi(\varepsilon_{it}) = 1\Big]} \\
&= \frac{\Pr\Big[\phi(F_{it}) = 1\Big] - \Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = 1\Big]\Pr\Big[\phi(\varepsilon_{it}) = 1\Big]}{1 - \Pr\Big[\phi(\varepsilon_{it}) = 1\Big]}.
\end{aligned}
$$

Hence,

$$
\Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = 1\Big] - \Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = -1\Big] = \frac{\Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = 1\Big] - \Pr\Big[\phi(F_{it}) = 1\Big]}{1 - \Pr\Big[\phi(\varepsilon_{it}) = 1\Big]}. \tag{15}
$$

Now note that without loss of generality, we can rescale the sign variables to take values of zero and one by dividing by two and adding one. Because rescaling both the left- and right-hand-side variables does not change the slope coefficient in a linear regression, we can simply write out the OLS regression coefficient as if the variables are rescaled:

$$
\begin{aligned}
\beta_{F\varepsilon} &= \frac{cov\Big(\phi(F_{it}), \phi(\varepsilon_{it})\Big)}{var\big(\phi(\varepsilon_{it})\big)} \\
&= \frac{\Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = 1\Big]\Pr\Big[\phi(\varepsilon_{it}) = 1\Big] - \Pr\Big[\phi(F_{it}) = 1\Big]\Pr\Big[\phi(\varepsilon_{it}) = 1\Big]}{\Pr\Big[\phi(\varepsilon_{it}) = 1\Big]\Big(1 - \Pr\Big[\phi(\varepsilon_{it}) = 1\Big]\Big)} \\
&= \frac{\Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = 1\Big] - \Pr\Big[\phi(F_{it}) = 1\Big]}{1 - \Pr\Big[\phi(\varepsilon_{it}) = 1\Big]},
\end{aligned}
$$

which is (15).□

Proof of Proposition 3. From Lemma 2, all we need to prove is that

$$
\Pr\Big[\phi(F_{it}) = 1\big|\phi(\varepsilon_{it}) = 1\Big] + \Pr\Big[\phi(F_{it}) = -1\big|\phi(\varepsilon_{it}) = -1\Big] > \Pr\Big[\phi(F_{it}) = 1\big|\phi\big(\varepsilon_{it}^{c}\big) = 1\Big] + \Pr\Big[\phi(F_{it}) = -1\big|\phi\big(\varepsilon_{it}^{c}\big) = -1\Big].
$$

Taking each term separately,

$$\begin{aligned}
&\Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] \\
&\quad = \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1,\, \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] \\
&\qquad + \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1,\, \phi(\varepsilon_{it}) = -1\right]\Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] \\
&\quad = \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] \\
&\qquad + \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] \\
&\quad = \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] \\
&\qquad + \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\left(1 - \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right]\right) \\
&\quad < \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] \\
&\qquad + \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right]\left(1 - \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right]\right) \\
&\quad = \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right],
\end{aligned} \tag{16}$$

where the second equality follows from (7) and the inequality follows from Lemma 2 and $\beta_{F\varepsilon} > 0$ (from Proposition 2). Similarly,

$$\begin{aligned}
&\Pr\left[\phi(F_{it}) = -1 \middle| \phi(\varepsilon_{it}^{c}) = -1\right] \\
&\quad = \Pr\left[\phi(F_{it}) = -1 \middle| \phi(\varepsilon_{it}^{c}) = -1,\, \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = -1\right] \\
&\qquad + \Pr\left[\phi(F_{it}) = -1 \middle| \phi(\varepsilon_{it}^{c}) = -1,\, \phi(\varepsilon_{it}) = -1\right]\Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^{c}) = -1\right] \\
&\quad = \Pr\left[\phi(F_{it}) = -1 \middle| \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = -1\right] \\
&\qquad + \Pr\left[\phi(F_{it}) = -1 \middle| \phi(\varepsilon_{it}) = -1\right]\Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^{c}) = -1\right] \\
&\quad < \Pr\left[\phi(F_{it}) = -1 \middle| \phi(\varepsilon_{it}) = -1\right]
\end{aligned}$$

which completes the proof.□

Statement of Lemma 3

Lemma 3. Condition (8) is equivalent to

$$\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] - \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = -1\right] > \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{d}) = 1\right] - \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{d}) = -1\right]$$

which is also equivalent to

$$\frac{cov\left(\phi(\varepsilon_{it}), \phi(\varepsilon_{it}^{c})\right)}{var\left(\phi(\varepsilon_{it}^{c})\right)} > \frac{cov\left(\phi(\varepsilon_{it}), \phi(\varepsilon_{it}^{d})\right)}{var\left(\phi(\varepsilon_{it}^{d})\right)}$$

Proof. The proof follows identical logic as the proof of Lemma 2.□

Proof of Proposition 4. First define

$$\begin{aligned}
\pi_c &= \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = 1\right] - \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{c}) = -1\right] \\
\pi_d &= \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{d}) = 1\right] - \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^{d}) = -1\right].
\end{aligned}$$

Using Lemma 2 and (7), β_{Fc} can be rewritten in terms of π_c:

$$
\begin{aligned}
\beta_{Fc} &= \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}^c) = 1\right] - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}^c) = -1\right] \\
&= \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^c) = 1\right] \\
&\quad + \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^c) = 1\right] \\
&\quad - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right]\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^c) = -1\right] \\
&\quad - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^c) = -1\right] \\
&= \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right]\left(\Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^c) = 1\right] - \Pr\left[\phi(\varepsilon_{it}) = 1 \middle| \phi(\varepsilon_{it}^c) = -1\right]\right) \\
&\quad + \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\left(\Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^c) = 1\right] - \Pr\left[\phi(\varepsilon_{it}) = -1 \middle| \phi(\varepsilon_{it}^c) = -1\right]\right) \\
&= \pi_c\left(\Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right] - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\right)
\end{aligned}
$$

Note that, from Proposition 2 and Lemma 2, the term in parentheses is positive, that is,

$$
\Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right] - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right] > 0. \tag{17}
$$

Assume that model c is a better approximation of the true asset pricing model than model d, that is,

$$
\frac{cov\left(\phi(\varepsilon_{it}), \phi\left(\varepsilon_{it}^c\right)\right)}{var\left(\phi\left(\varepsilon_{it}^c\right)\right)} > \frac{cov\left(\phi(\varepsilon_{it}), \phi\left(\varepsilon_{it}^d\right)\right)}{var\left(\phi\left(\varepsilon_{it}^d\right)\right)}.
$$

By Lemma 3, this relation implies that

$$
\pi_c > \pi_d,
$$

which means that

$$
\begin{aligned}
&\pi_c\left(\Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right] - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\right) \\
&\quad > \pi_d\left(\Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right] - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\right),
\end{aligned}
$$

so

$$
\beta_{Fc} > \beta_{Fd}.
$$

Let us now prove the reverse. Assume that $\beta_{Fc} > \beta_{Fd}$. This means that

$$
\begin{aligned}
&\pi_c\left(\Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right] - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\right) \\
&\quad > \pi_d\left(\Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = 1\right] - \Pr\left[\phi(F_{it}) = 1 \middle| \phi(\varepsilon_{it}) = -1\right]\right),
\end{aligned}
$$

which by (17) implies that $\pi_c > \pi_d$.□

Proof of Proposition 5.

$$
\begin{aligned}
\gamma_1 &= \frac{cov\left(\phi(F_{it}), \dfrac{\phi\left(\varepsilon_{it}^c\right)}{var\left(\phi\left(\varepsilon_{it}^c\right)\right)} - \dfrac{\phi\left(\varepsilon_{it}^d\right)}{var\left(\phi\left(\varepsilon_{it}^d\right)\right)}\right)}{var\left(\dfrac{\phi\left(\varepsilon_{it}^c\right)}{var\left(\phi\left(\varepsilon_{it}^c\right)\right)} - \dfrac{\phi\left(\varepsilon_{it}^d\right)}{var\left(\phi\left(\varepsilon_{it}^d\right)\right)}\right)} \\
&= \frac{\beta_{Fc} - \beta_{Fd}}{var\left(\dfrac{\phi\left(\varepsilon_{it}^c\right)}{var\left(\phi\left(\varepsilon_{it}^c\right)\right)} - \dfrac{\phi\left(\varepsilon_{it}^d\right)}{var\left(\phi\left(\varepsilon_{it}^d\right)\right)}\right)}.
\end{aligned}
$$

By Proposition 4, $\beta_{Fc} > \beta_{Fd}$ if and only if model c is better than model d. It then follows immediately that $\gamma_1 > 0$ because the strict inequality $\beta_{Fc} > \beta_{Fd}$ rules out the possibility that the denominator is zero.□

Appendix B. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jfineco.2015.08.010.

References

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.
Barber, B.M., Huang, X., Odean, T., 2014. What risk factors matter to investors? Evidence from mutual fund flows. Available at SSRN: ⟨http://ssrn.com/abstract=2408231⟩.
Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. Journal of Political Economy 112, 1269–1295.
Berk, J.B., Tonks, I., 2007. Return persistence and fund flows in the worst performing mutual funds. Unpublished Working Paper NBER No. 13042.
Berk, J.B., van Binsbergen, J.H., 2015. Measuring skill in the mutual fund industry. Journal of Financial Economics, forthcoming.
Breeden, D.T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7, 265–296.
Busse, J.A., Irvine, P.J., 2006. Bayesian alphas and mutual fund persistence. The Journal of Finance 61, 2251–2288.
Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.
Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.
Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. Journal of Political Economy 106, 1167–1200.
Clifford, C.P., Fulkerson, J.A., Jordan, B.D., Waldman, S., 2013. Risk and fund flows. Available at SSRN: http://ssrn.com/abstract=1752362.
Epstein, L.G., Zin, S.E., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: an empirical analysis. Journal of Political Economy 99, 263–286.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
French, K.R., 2008. The cost of active investing. Journal of Finance 63, 1537–1573.
Graham, J.R., Harvey, C.R., 2001. The theory and practice of corporate finance: evidence from the field. Journal of Financial Economics 60, 187–243.
Grossman, S., 1976. On the efficiency of competitive stock markets where trades have diverse information. Journal of Finance 31, 573–585.
Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. The American Economic Review 70, 393–408.
Guercio, D.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, 523–557.
Harvey, C.R., Liu, Y., Zhu, H., 2015. and the cross-section of expected returns. Review of Financial Studies, forthcoming.
Kreps, D.M., Porteus, E.L., 1978. Temporal resolution of uncertainty and dynamic choice theory. Econometrica 46, 185–200.
Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. The Review of Economics and Statistics 47, 13–37.
Lustig, H., Van Nieuwerburgh, S., Verdelhan, A., 2013. The wealth-consumption ratio. Review of Asset Pricing Studies 3, 38–94.
Mamaysky, H., Spiegel, M., Zhang, H., 2008. Estimating the dynamics of mutual fund alphas and betas. Review of Financial Studies 21, 233–264.
Merton, R.C., 1973. Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3, 373–413.
Milgrom, P., Stokey, N., 1982. Information, trade, and common knowledge. Journal of Economic Theory 26, 17–27.
Mossin, J., 1966. Equilibrium in a capital asset market. Econometrica 34, 768–783.
Pastor, L., Stambaugh, R.F., 2012. On the size of the active management industry. Journal of Political Economy 120, 740–781.
Pastor, L., Stambaugh, R.F., Taylor, L.A., 2015. Scale and skill in active management. Journal of Financial Economics 116, 23–45.
Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Review of Financial Studies 22, 435–480.
Ross, S.A., 1976. The arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341–360.
Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance 19, 425–442.
Sirri, E.R., Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance 53, 1589–1622.
Stambaugh, R.F., 2014. Investment noise and trends. Journal of Finance 69, 1415–1453.
Thompson, S.B., 2011. Simple formulas for standard errors that cluster by both firm and time. Journal of Financial Economics 99, 1–10.
Treynor, J., 1961. Toward a theory of the market value of risky assets.