Journal of Financial Economics 112 (2014) 213–231

Contents lists available at ScienceDirect

Journal of Financial Economics

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

The price of skill: Performance evaluation by households ☆

Alexi Savov

New York University Stern School of Business, United States

ARTICLE INFO

Article history:
Received 21 May 2012
Received in revised form
22 July 2013
Accepted 4 August 2013
Available online 22 November 2013

JEL classification:
G11
G12
G23

Keywords:
Mutual fund performance
Active funds
Index funds
Fund flows

ABSTRACT

Skilled investors make money off uninformed investors. By acting as intermediaries, they provide a hedge to the uninformed investors themselves. I present a model in which households have imperfect information about expected returns. Non-traded income shocks lead them to rebalance, sometimes at the wrong time. Active funds hedge this risk by trading on superior information. In equilibrium, they pay off when non-traded income disappoints, earning a premium that makes them appear to underperform index funds after fees. Empirical results using aggregate fund flows support the model. A corresponding asset pricing test can account for the apparent underperformance of active funds.

1. Introduction

If you were an uninformed investor, what would you do? Choosing the right asset mix requires good knowledge of expected returns. A lack of such knowledge can lead to poor timing, buying high and selling low. Buy-and-hold avoids this problem, but there are good reasons to trade, like outside income shocks. The alternative is to hire a professional. An informed fund manager corrects the poor timing of her investors by trading against the error in their beliefs. Her fund pays off precisely when her investors discover that error. For those with no outside income, natural buy-and-hold investors, this service has limited value. But for those who buy and sell, it can be a hedge worth a high premium. Whereas buy-and-hold investors are best served by low-cost index funds, others may choose active funds even if their costs are high.

I present a model of portfolio delegation in an economy with time-varying expected returns. Households learn about expected returns from stock returns and economic conditions. Markets are incomplete so this learning is imperfect. For example, expected returns may depend on persistent shocks to non-traded income such as wages, private business, or real estate. A set of active fund managers observes the true expected returns and trades accordingly, generating alpha. For example, they may be able to forecast earnings more accurately than households. In equilibrium, active fund returns are abnormally high when household expectations are off the mark. In particular, they are high when non-traded income disappoints. As a result, active funds attract investors with substantial non-traded income exposure. These investors are willing to pay a premium above and beyond the gross alpha of their funds. This is because the value of active investing depends on the price of non-traded risk, which is not reflected in traded benchmarks. In this

☆ I am grateful to the editor and anonymous referee, my dissertation advisors George Constantinides, Douglas Diamond, Eugene Fama, and Ralph Koijen, and to Zhiguo He, Marcin Kacperczyk, Alan Moreira, and Stijn Van Nieuwerburgh. I thank seminar participants at Chicago Booth, Vanderbilt Owen, Wisconsin School of Business, Stanford GSB, UCLA Anderson, MIT Sloan, Wharton, NYU Stern, HBS, ASU Carey, Wash U Olin, Berkeley Haas, and USC Marshall.

E-mail address: asavov@stern.nyu.edu

0304-405X/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jfineco.2013.11.005

way, the model can account for the apparent underperformance of active funds first noted by Jensen (1968).

The model generates several additional testable implications. First, net active fund returns are countercyclical: they are high when equity risk premiums are high.[1] This result is consistent with the findings of Kosowski (2011) and Glode (2011), among others. In the model, in addition to non-traded income risk, active funds also load on traded risk, paying off when returns come in higher than anticipated. When the price of stock market risk rises, net returns must also rise to compensate investors. To implement this mechanism, I incorporate decreasing returns to scale so that high outflows during market downturns lead to higher gross and net returns.

The model also leads to several novel implications that I examine in the empirical section. I derive these implications by connecting the household stochastic discount factor to aggregate fund flows. Income shocks cause households to revise their beliefs about expected returns and rebalance their portfolios. As a result, aggregate fund flows reveal shocks to non-traded assets.

Consistent with the prediction that active fund investors respond to shocks whereas index fund investors tend to buy and hold, I find that aggregate flows into active funds are strongly related to contemporaneous market returns whereas flows into index funds are not. In the model, both active and index fund investors are equally informed and both learn from stock returns, but active funds endogenously attract investors with substantial non-traded assets who tend to rebalance.

Consistent with the prediction that mutual fund investors have poor timing, I show that fund inflows predict lower market returns between six and 12 months out. The negative predictability is concentrated among active funds and absent among index funds. Thus, active fund investors are not only more likely to rebalance, they are also more likely to do so at the wrong time, buying high and selling low, as predicted.

In the model, active fund managers correct the mistiming of their investors. I check this prediction in the data by using fund flows to predict the returns of a strategy that is long active funds and short the market. I find that high inflows—though they predict low market returns—also predict high active-minus-market returns. The predictability is strongest at the three- to six-month horizon. It is present among active funds and absent among index funds. These results suggest that active fund managers help to undo the poor timing of their investors.

As a final exercise, I run an asset pricing test using a factor based on aggregate fund flows that is constructed from the cross section of equities. In the model, controlling for the market return, fund inflows reveal negative shocks to non-traded income. The predicted premium on a fund-flows factor is therefore negative. Active funds endogenously provide a hedge for non-traded risk and so their returns covary positively with fund flows. In the data, active funds load positively on the fund-flows factor. The estimated premium is large and negative, and it is not subsumed by the standard set of factors. The positive loading and negative premium account for between 25% and over 100% of the apparent underperformance of active funds.

The empirical analysis concludes with a set of robustness checks and extensions, which are presented more fully in the paper's Online Appendix. Results are robust to the exclusion of institutional funds, funds primarily sold through retirement plans, and small-cap and sector funds. The lagged market return negatively predicts the relative returns of active funds versus the market, consistent with countercyclical performance. Most of the empirical results extend to the universe of corporate bond funds, except that it is index bond fund flows that negatively predict bond market returns.

The paper also makes a modeling contribution by presenting a fully tractable framework that incorporates time-varying expected returns, non-traded income risk, investor learning, and portfolio delegation. An extension of the model presented in Appendix B also covers the case of multiple stocks where stock picking, not just market timing, contributes to skill.

Among the key modeling simplifications, I assume an exogenous linear return process to make the learning problem tractable. A general equilibrium version of the model can be constructed in the spirit of Grossman and Stiglitz (1976, 1980). However, it is important that the source of noise in prices, modeled as a reduced-form supply shock in Grossman and Stiglitz (1980), be directly linked to the marginal utility of investors (and not just through their portfolios). That is, the investors and the noise traders must be the same group; it is the noise traders who are willing to pay a high premium for active management. This is the key idea of the paper.

The rest of this paper proceeds with a review of the literature in Section 2, the presentation of the model in Section 3, empirical analysis in Section 4, and concluding sentences in Section 5.

2. Related literature

This paper descends from the mutual fund performance literature started by Jensen (1968) and extended by Carhart (1997) and many others. Elton and Gruber (2013) summarize that most studies find small positive gross alphas that are not high enough to cover fees and expenses. Fama and French (2010) show that a value-weighted portfolio of all equity mutual funds has a statistically significant net alpha of up to -1% per year. In light of the evidence, Gruber (1996) calls the popularity of active funds a puzzle.[2] Indeed, canonical delegation models like Berk and Green (2004) predict net alphas of zero, not less than zero.

This puzzle motivates a set of recent studies. Closest to this paper is Glode (2011). In Glode (2011), managers optimally choose to deliver superior returns in high marginal utility states. Model misspecification as in Roll (1977) can lead to downward-biased performance attribution. Model misspecification also emerges here due to non-traded assets, but managers have no way of allocating

[1] Campbell (1999), among others, shows that equity risk premiums are indeed higher in recessions.

[2] Active mutual funds manage close to $5 trillion.

alpha; it arises endogenously under asymmetric information. Micro-founding alpha in this way makes the model directly testable by linking performance to shocks to perceived expected returns, which are reflected in fund flows. In other words, fund flows can be used to proxy for the missing factor posited by Roll (1977). Active funds provide a natural asset class to implement this proxy.

Another point of departure from Glode (2011) concerns the countercyclical nature of mutual fund performance. In Glode (2011), fund returns covary with the stochastic discount factor (SDF). Here, the covariance between fund returns and the SDF is itself endogenously countercyclical. This means that fund returns are not only positively related to SDF innovations (a return surprise effect), they are also predictably higher conditional on a downturn (an expected return effect). The evidence in Kosowski (2011) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (accepted for publication) supports the view that net returns are predictable.

Two other recent papers that address the underperformance puzzle are Pástor and Stambaugh (2012) and Gennaioli, Shleifer, and Vishny (2012). Pástor and Stambaugh (2012) make the point that learning the full production function of the active fund industry is difficult so that even a long record of underperformance need not cause the industry to shrink. In Gennaioli, Shleifer, and Vishny (2012), investors choose funds on the basis of trust. Both mechanisms are distinct and can be viewed as complementary. Neither makes the additional predictions derived and tested here.

A common thread among existing mutual fund theories is the exogenous specification of alpha. For instance, in Berk and Green (2004) alpha is an additive term in the active fund return. As Sharpe (1991) points out, however, the net supply of gross alpha across the whole market is zero. Existing theories implicitly assume an omitted group of investors with negative alpha. In contrast, this paper takes the point of view of the omitted group itself, which turns out to be a natural clientele for active funds.

Other recent papers on portfolio delegation employ market segmentation or participation constraints. Examples include Vayanos and Woolley (2013), He and Krishnamurthy (2013), and Kaniel and Kondor (2013). These papers focus on the asset pricing implications of delegation. In this paper, asset prices are given and the focus is on active versus passive investing. To explore this margin, I necessarily allow households to trade all assets freely.

Admati and Ross (1985) point out that when active funds trade on superior information, Jensen's alpha becomes unreliable as uncertainty about the manager's portfolio adds to the variance of her returns. In this paper, the manager's actions add to the *co*variance of her returns with household consumption. This covariance is endogenously negative, which biases the usual performance tests towards the appearance of underperformance.

Dybvig and Ross (1985) and Grinblatt and Titman (1989) also critique Jensen's alpha on the basis that delegated returns are inherently nonlinear. Here, they are log-linear. Though I use the term "alpha" for expositional purposes, I judge performance based on the attractiveness of active funds to buy-and-hold investors, not strictly alpha. If a buy-and-hold investor, someone with no exposure to (the systematic component of) non-traded income, strictly prefers an index fund to an active fund, I say that active funds appear to underperform index funds on a buy-and-hold basis.

The mutual fund literature has also focused on conditional performance, specifically the countercyclical nature of mutual fund returns. In Kacperczyk, Van Nieuwerburgh, and Veldkamp (2012), skill is in greater supply during recessions as there is more volatility, which improves gross alpha. However, as there are no household investors, the paper makes no predictions on net alpha and does not address the underperformance puzzle. The intuition of Berk and Green (2004) suggests that net alpha should remain at zero. In this paper, it is the demand for skill that falls as equity risk premiums rise, so that net performance improves, as in the data. The two theories can therefore work side by side.

A large literature explores the cross section of skill. Berk and van Binsbergen (2012) find persistent skill before fees and expenses, as required here. Several studies delve into the nature of skill: Kacperczyk, Sialm, and Zheng (2005), Kacperczyk and Seru (2007), and Cremers and Petajisto (2009) show that managers who have more concentrated portfolios, rely less on public signals, or depart more from a set of indexes, have higher risk-adjusted returns. These findings confirm the intuition that superior information is at the heart of skill.[3]

Wermers (2000) and others find some evidence of stock-picking ability on the part of active mutual funds, and not much evidence of market-timing ability (for an exception, see Jiang, Yao, and Yu, 2007). Kacperczyk, Van Nieuwerburgh, and Veldkamp (accepted for publication) argue that funds optimally alternate between stock picking and market timing over the business cycle. They find evidence of market timing during downturns. Ferson and Schadt (1996) find that high inflows are associated with lower betas, as the model predicts. While the paper focuses on a one-stock model where only market timing is possible, the multiple-stock extension in Appendix B introduces stock-picking. As in Kacperczyk, Van Nieuwerburgh, and Veldkamp (accepted for publication), the relative contributions to performance of stock-picking and market-timing strategies depend on the covariance structure of stock returns and on the level and distribution of risk premiums in the economy.

The literature on fund flows is largely focused at the fund level. Gruber (1996) shows that investors respond to past returns, whereas Zheng (1999) finds that this "smart-money" effect is short-lived. Berk and van Binsbergen (2012) use a measure of skill based on the intuition of Berk and Green (2004) and find a stronger connection with flows. Edelen (1999) shows that more active mutual funds experience substantially higher investor turnover, consistent with this paper.[4] As fund-level flows are largely driven by fund-level performance, the variation that these

[3] In this paper, managers with more private information indeed generate higher returns before fees, but their net returns are actually lower as they can charge a high premium. Index funds, which have the lowest gross alpha (zero), also have the highest net alpha (also zero). This result motivates the paper.

[4] Edelen (1999) calculates that taking into account the liquidity services provided by mutual funds improves their measured performance. This mechanism cannot account for the performance gap between active and index funds as both types provide the same liquidity services. The cost of providing liquidity has also declined as markets have

papers exploit drowns out the variation exploited here, which is based on aggregate shocks. Among the few papers on aggregate flows, Warther (1995) finds a strong relationship between contemporaneous returns and flows, but not much forecastability, perhaps due to a short eight-year sample that ends in 1992. More recently, Ben-Rephael, Kandel, and Wohl (2012) find that net exchanges, a subset of total flows, have surprisingly strong forecasting power.

3. Model

I present an asset pricing model in which active mutual funds underperform traded benchmarks net of fees. Households invest with active funds because fund managers possess superior knowledge of conditional expected returns. From the household point of view, active fund returns are positively correlated with news about risk premiums. When risk premiums are countercyclical as in the data, active funds provide a valuable hedge against non-traded income. Investors pay for this hedge in the form of low net returns. Appendix B features an extension to multiple stocks, incorporating stock-picking skill. I develop testable hypotheses by relating investor beliefs to aggregate fund flows.

3.1. Setup

Active trading is futile in an unpredictable setting (e.g., roulette). Instead, consider an economy beset by persistent fluctuations (e.g., business cycles). Let y_t represent log-income and write

$$\Delta y_{t+1} = \overline{y} + x_t + \varepsilon^y_{t+1} \tag{1}$$

$$x_{t+1} = \rho x_t + \varepsilon^x_{t+1}, \tag{2}$$

with $\rho \in (0,1)$, $\varepsilon^y \sim N(0,\sigma_y^2)$, $\varepsilon^x \sim N(0,\sigma_x^2)$, and $\varepsilon^y \perp \varepsilon^x$ for simplicity. The persistent state variable x controls the conditional mean of income growth Δy. It also controls risk premiums: A single risky asset interpreted as a stock index has an exogenous log-return[5]

$$r_{t+1} = \overline{r} - \phi x_t + \varepsilon^r_{t+1}. \tag{3}$$

Let $\varepsilon^r \sim N(0,\sigma_r^2)$ with $Cov(\varepsilon^r,\varepsilon^y) = \sigma_{ry}$ and $\varepsilon^r \perp \varepsilon^x$ for simplicity. The parameter ϕ controls the variation of expected returns. The natural case to consider is $\phi > 0$ so that risk premiums are countercyclical as in the data. Asset pricing models that deliver this type of specification include Campbell and Cochrane (1999) where the state variable is surplus consumption,[6] and Santos and Veronesi (2006) where it corresponds to the relative share of traded to non-traded income. Exogenous linear returns make the learning problem tractable.

A one-period risk-free bond pays a constant log-return r_f.

Markets are incomplete so that $|Corr(\varepsilon^r,\varepsilon^y)| < 1$. To motivate this assumption, the return shock ε^r can be interpreted as a traded cash flow shock (e.g., dividend news) that is only imperfectly correlated with total income y. For example, non-traded income may be derived from human capital, private business, or real estate.[7]

Market incompleteness implies that stock returns and income growth do not reveal the true state of the economy x, creating private demand for information and an opportunity for active trading. In this way, prices are "noisy" and x serves the role of the stochastic supply by the same label in Grossman and Stiglitz (1980).[8]

Households rely on returns r and income growth Δy to learn about x,[9] forming beliefs based on the innovations representation

$$\widehat{x}_{t+1} = \rho \widehat{x}_t + \widehat{\varepsilon}^x_{t+1} \tag{4}$$

$$\Delta y_{t+1} = \overline{y} + \widehat{x}_t + \widehat{\varepsilon}^y_{t+1} \tag{5}$$

$$r_{t+1} = \overline{r} - \phi \widehat{x}_t + \widehat{\varepsilon}^r_{t+1}. \tag{6}$$

Throughout the paper, hats denote conditional expectations under the household filtration. Since both the signals and the state have linear dynamics, the filtered shocks are computed using the standard Kalman filter. The formulas are in Appendix A. I focus on the steady state of the filter for simplicity. The steady state error dispersion

$$\Sigma_x = \lim_{t\to\infty} \sqrt{E_t[((x_t - \widehat{x}_t)^2]} \tag{7}$$

measures household uncertainty about the true state of the economy and so $\phi\Sigma_x$ measures household uncertainty about expected returns. The question arises, how much are households willing to pay to resolve this uncertainty?

The answer depends on the price of shocks to x, which in turn depends on the price of non-traded income shocks. This price is a free parameter due to market incompleteness. Let m_t be the household log-stochastic discount factor. Without loss of generality, write[10]

$$\Delta m_{t+1} = m_{0,t} + m_{r,t}\widehat{\varepsilon}^r_{t+1} + m_{y,t}\left[\widehat{\varepsilon}^y_{t+1} - \left(\frac{\widehat{\sigma}_{ry}}{\widehat{\sigma}_r^2}\right)\widehat{\varepsilon}^r_{t+1}\right]. \tag{8}$$

The composite shock in the brackets is by construction orthogonal to the return shock. In this way, m_r can be called the price of traded risk, and m_y the shadow price of non-traded risk. Note that both $m_r < 0$ and $m_y < 0$ so that households dislike exposure to these shocks.

(footnote continued)
evolved. To illustrate, the typical active equity fund today holds only 3% cash, down from 10% in 1990.

[5] See Appendix B for a multiple-stock extension.

[6] In Campbell and Cochrane (1999), output follows a random walk but surplus consumption (over which preferences are defined) evolves similarly to y here.

[7] In the U.S. from 1984 to 2012, the correlation between dividends and employee compensation is 31% quarterly and 60% annually and the correlation between dividends and proprietors' income is 39% quarterly and 54% annually. Employee compensation represents 67% of total income and proprietors' income accounts for 9%, versus only 4% for dividend income.

[8] The CARA-normal shocks framework of Grossman and Stiglitz (1980) does not produce closed-form solutions once intermediaries are introduced.

[9] Additional public signals are easily accommodated as long as they do not reveal x perfectly.

[10] Note that $\widehat{\varepsilon}^x$ does not enter separately since it is a linear combination of $\widehat{\varepsilon}^r$ and $\widehat{\varepsilon}^y$ (households learn about x from returns and income).

The prices of the stock and bond impose restrictions on the coefficients of m:

$$1 = \mathrm{E}_t[e^{\Delta m_{t+1} + r_f}] \quad \text{and} \quad 1 = \mathrm{E}_t[e^{\Delta m_{t+1} + r_{t+1}}]. \tag{9}$$

Since Δm and r are log-normal,

$$r_f = -\mathrm{E}_t[\Delta m_{t+1}] - \tfrac{1}{2}\,\mathrm{Var}_t(\Delta m_{t+1}) \tag{10}$$

$$\mathrm{E}_t[r_{t+1}] + \tfrac{1}{2}\,\mathrm{Var}_t(r_{t+1}) - r_f = -Cov_t(\Delta m_{t+1}, r_{t+1}). \tag{11}$$

Substituting for Δm into Eq. (10) gives

$$r_f = -m_{0,t} - \frac{1}{2}\left[m_{r,t}^2 \hat{\sigma}_r^2 + m_{y,t}^2\left(\hat{\sigma}_y - \frac{\hat{\sigma}_{ry}}{\hat{\sigma}_r}\right)^2\right]. \tag{12}$$

To maintain a constant interest rate, the mean and variance of the stochastic discount factor must move against each other so that the forces of intertemporal smoothing and precautionary saving offset. Campbell and Cochrane (1999) use an analogous construction and motivate it by noting that while risk premiums vary substantially, real interest rates are stable.

From Eq. (11), the equity premium pins down the price of traded risk m_r:

$$\bar{r} - \phi\hat{x}_t + \tfrac{1}{2}\hat{\sigma}_r^2 - r_f = -m_{r,t}\hat{\sigma}_r^2. \tag{13}$$

Time-varying expected returns require a time-varying price of risk. Periods of low income growth x are associated with a stronger covariance between stocks and the stochastic discount factor (or simply a more volatile stochastic discount factor), and hence a higher market Sharpe ratio.

Pricing stocks and bonds does not pin down the price of non-traded risk; m_y remains a free parameter. In fact, it can differ across households as markets are incomplete. It turns out that m_y is the key parameter for pricing actively managed mutual funds.

3.2. Active funds

Consider an active fund manager who knows the true expected return.[11] For example, she might derive an information advantage from sophisticated analysis of financial statements and other public sources. The cost of this type of analysis may be prohibitive for an individual household. Given prices, a superior forecast of fundamentals translates directly into a superior forecast of expected returns (Campbell and Shiller, 1988). For this reason, I model the manager's expertise directly in terms of expected returns without any loss of generality.

The active manager opens her fund to households, investing a fraction $\chi_t > 0$ of assets under management in stocks. At the end of each period, her fund delivers a net return

$$e^{r_{t+1}^A} = [e^{r_f} + \chi_t(e^{r_{t+1}} - e^{r_f})]e^{-f_t} \tag{14}$$

$$r_{t+1}^A = r_f + \log[1 + \chi_t(e^{r_{t+1}} - e^{r_f})] - f_t. \tag{15}$$

To investors, fees and expenses f_t represent a drag on returns paid in exchange for good performance. In a model with explicit information production, f_t would reflect the cost of acquiring information, the cost of trading, the competitiveness of the mutual fund industry, the need to compensate managers for risk, and agency costs (see Berk and Green, 2004). Rather than modeling the production of information, I focus on its resale to households and calculate their willingness to pay. It is this willingness to pay that is highlighted as a puzzle in the mutual fund literature.

The portfolio choice χ_t generally depends on the true expected return and so it cannot be made public in real time. Households could potentially infer χ_t with a one-period lag from ex post fund returns. This would curtail (though not eliminate) the information advantage of fund managers, and for this reason I preclude it in the household learning problem. To motivate this assumption, consider adding an unpriced disturbance to a fund's return that reflects unmodeled aspects of the manager's trading strategy. Such a disturbance would have no effect on the main results and is safely left out. In practice, funds invest in many securities using diverse strategies and it is impossible to infer their actions from a single ex post return (see the multiple-stock version of the model in Appendix B).

3.2.1. Negative net alpha

From the point of view of households, active funds represent an additional asset available for trading. As such, they must be priced by the household stochastic discount factor m:

$$1 = \mathrm{E}_t[e^{\Delta m_{t+1} + r_{t+1}^A}] \tag{16}$$

$$1 = \mathrm{E}_t[e^{\Delta m_{t+1}}(e^{r_f} + \chi_t(e^{r_{t+1}} - e^{r_f}))e^{-f_t}]. \tag{17}$$

Suppose that χ_t is log-normal under public information with $\mathrm{E}_t[\chi_t] = 1$ so that households expect their managers to be fully invested in stocks. Then willingness to pay f_t is given by[12]

$$e^{f_t} - 1 = e^{Cov_t(\log\chi_t, \Delta m_{t+1})}[e^{Cov_t(\log\chi_t, r_{t+1})} - 1]. \tag{19}$$

A linearization provides a more intuitive version of this formula:

$$f_t \approx [1 + Cov_t(\log\chi_t, \Delta m_{t+1})]\, Cov_t(\log\chi_t, r_{t+1}). \tag{20}$$

Households' willingness to compensate active managers has two components. The second arises from the ability to anticipate stock returns, $Cov_t(\log\chi_t, r_{t+1})$. I call this component gross alpha.

The first component relates to the timing of active fund returns. From the point of view of households, the portfolio choice of the active manager adds a layer of uncertainty. This uncertainty is necessarily systematic as it is correlated with news about the true state of the

[11] More generally, an active fund manager will attract investors as long as her uncertainty about expected returns is smaller than that of households, $\phi\Sigma_x$.

[12] More generally, if χ is log-normal, willingness to pay also depends on a fund's average risk exposure

$$e^{f_t} - 1 = e^{\mathrm{E}_t[\log\chi_t] + \frac{1}{2}Var_t(\log\chi_t) + Cov_t(\log\chi_t, \Delta m_{t+1})}\left[e^{Cov_t(\log\chi_t, r_{t+1})} - 1\right]. \tag{18}$$

economy. The compensation of the active manager therefore depends on the pricing of this uncertainty as captured by the term $Cov_t(\log \chi_t, \Delta m_{t+1})$. From Eq. (19), households are willing to pay active managers more than their gross alpha whenever

$$Cov_t(\log \chi_t, \Delta m_{t+1}) > 0, \tag{21}$$

that is, when active funds provide a hedge. In this case, they will appear to underperform traded benchmarks. As I show below, this turns out to be the natural case to consider.

The covariance in Eq. (21) is a function of the portfolio policy χ. The particular form of χ generally depends on the manager's preferences and any agency frictions that arise under delegation. However, as long as the manager is concerned with maximizing fee income, χ will be increasing in the expected stock return, and this is what is needed here. Given that, as Admati and Pfleiderer (1990) point out, the precise functional form of the manager's policy is secondary as households can manipulate their fund allocation to achieve a desired overall exposure.[13] With this in mind, I consider a convenient example for χ to convey the intuition of the model:

$$\chi_t = e^{-\eta_t \left((x_t - \widehat{x}_t)/\Sigma_x\right) - \eta_t^2/2}. \tag{22}$$

This policy can be motivated as an approximation to the optimal strategy of a manager with CRRA preferences maximizing end-of-period wealth.[14] Recall from Eq. (3) that expected returns are decreasing in x so χ is indeed increasing in expected returns. Also, χ is log-normal with mean one as required by Eq. (19). The parameter $\eta_t > 0$ controls the aggressiveness with which managers trade on private information. I assume η_t is public (in Pástor and Stambaugh, 2012, it is not).

Fig. 1 illustrates the resulting active fund return. In panel A, a strategy long active funds and short the index pays off whenever returns are high and income is low, or returns are low and income is high.[15] This is because with $\phi > 0$, returns and income have opposing exposures to x. At the same time, the stochastic discount factor, shown in panel B, is high when both income and returns are low. If the price of non-traded shocks m_y is high enough (i.e., if m is steep enough in y), then the high active fund payoff in the low-income-high-return state is a valuable hedge, as shown below.

To formalize the intuition of Fig. 1 and calculate the resulting active fund compensation, substitute Eq. (22) into (19) using (Eqs. (5) and 6):

$$e^{f_t} - 1 = e^{\eta_t[\phi m_{r,t} - (1+\phi\widehat{\sigma}_{ry}/\widehat{\sigma}_r^2) m_{y,t}]\Sigma_x}[e^{\eta_t(\phi\Sigma_x)} - 1]. \tag{23}$$

Gross alpha depends on the amount of private information Σ_x that the active manager possesses, its importance ϕ for expected returns, and the aggressiveness η_t with which fund assets are deployed.

Net alpha, on the other hand, depends on the price of non-traded risk m_y. From Eq. (23), if non-traded risk is not priced, $m_y = 0$, net alpha cannot be negative since the price of traded risk m_r is negative. In this case, high expected returns (low x) are good news.

The opposite is true when expected returns are high during bad times as in the data. In this case, the price of non-traded risk is negative, $m_y < 0$. If $|m_y|$ is large enough and $\widehat{\sigma_{ry}} > 0$ (as in the data) so that

$$|m_{y,t}| > \left(\frac{\phi}{1 + \phi \frac{\widehat{\sigma_{ry}}}{\widehat{\sigma_r^2}}} \right) |m_{r,t}|, \tag{24}$$

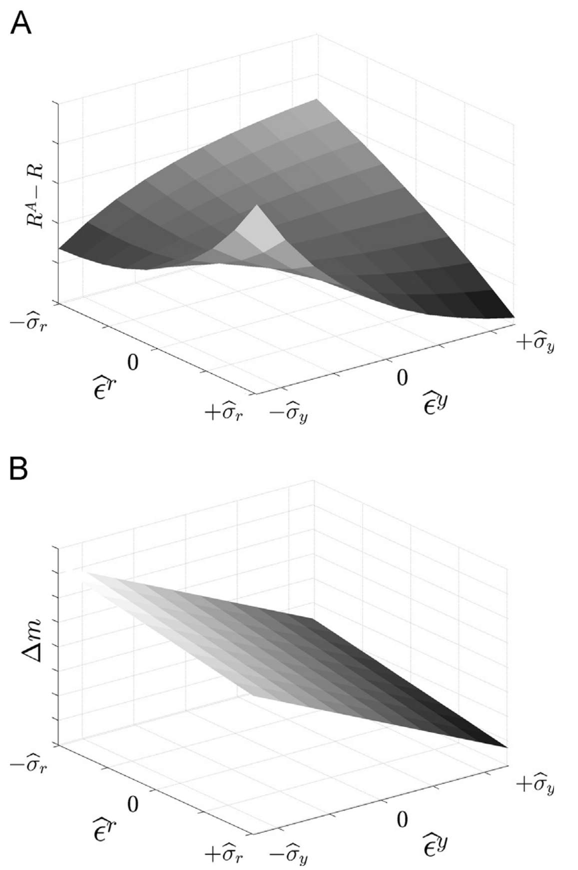

Fig. 1. The active fund return and the stochastic discount factor. This figure illustrates the return of the active fund $R^A = e^{r^A}$ in excess of the index $R = e^r$, and the stochastic discount factor Δm as a function of the filtered return and income shocks $\widehat{\varepsilon}^r$ and $\widehat{\varepsilon}^y$. The active return is given by Eqs. (14) and (22). The stochastic discount factor is given by Eq. (8). (The various parameters are set to $\eta = 1$, $r_f = 0$, $\rho = 0.9$, $\phi = .5$, $\sigma_x = 0.05$, $\sigma_y = 0.1$, $\sigma_r = 0.2$, $\sigma_{ry} = 0$, $m_r = -0.2$, $m_y = -1$.)

[13] More precisely, when only a stock and bond are traded, households can undo any linear transformation of the fund's portfolio policy. The coefficients can be time-varying. In continuous time with no jumps, nonlinear transformations can also be undone.

[14] Suppose the manager's relative risk aversion coefficient is γ_t. Campbell and Viceira (2002) show that the optimal portfolio policy can be approximated as $\chi_t \approx (\mathrm{E}_t[r_{t+1}] + \frac{1}{2}\sigma_r^2 - r_f)/\gamma_t\sigma_r^2 \approx ((\overline{r} + \frac{1}{2}\sigma_r^2 - r_f)/\gamma_t\sigma_r^2 - 1) + e^{(1/\gamma_t\sigma_r^2)(\mathrm{E}_t[r_{t+1}] - \overline{r})}$. This matches the reduced-form expression in Eq. (22) up to a constant.

[15] Merton (1981) shows that trading skill generates a payoff that resembles an option straddle (buy low, sell high). Fig. 1 confirms this result.

net alpha will also be negative. Intuitively, Eq. (24) says that net alpha is negative whenever the net effect of low income growth x is negative: the bad news about lower income growth must more than offset the good news about future investment returns.

Recall that individual households can differ in their shadow prices of non-traded risk since markets are incomplete. Therefore, an investor with no non-traded income exposure strictly prefers an index fund over an active fund with a negative net alpha. Buy-and-hold investors are strictly better off with index funds, whereas investors with substantial non-traded income risk form the natural clientele of active funds.

Eqs. (19) and (23) suggest that the net performance of active funds can be used to infer the price of non-traded risk m_y. In Section 3.3 below, I show how aggregate fund flows can be used to identify the time series of shocks to non-traded income, which offers a potential response to the critique of Roll (1977) and a test of the model.

3.2.2. Countercyclical performance

The evidence suggests that net mutual fund alphas increase during downturns (Moskowitz, 2000; Kosowski, 2011; Glode, 2011). Eq. (23) provides a mechanism: if as x falls the price of traded risk m_r increases faster in magnitude than the price of non-traded risk m_y, willingness to pay for active fund delegation falls. When this is the case, high expected stock returns reflect a particularly high distaste for stock market risk. By betting in the direction of future returns, active fund managers amplify this risk, leading investors to demand higher net returns. Formally, using Eqs. (13), (23), and $\widehat{\sigma}_{ry} > 0$, countercyclical performance requires

$$\frac{\partial m_{y,t}}{\partial \widehat{x}} < \frac{\phi^2/\widehat{\sigma}_r^2}{1+\phi\frac{\widehat{\sigma}_{ry}}{\widehat{\sigma}_r^2}}. \tag{25}$$

This condition is satisfied in the model of Santos and Veronesi (2006) where it is the share of non-traded income that drives risk premiums, which makes the prices of traded and non-traded income risk move in opposite directions. It is also likely to be satisfied in stochastic volatility models where the volatility of stocks rises by more than the volatility of non-traded income during downturns, perhaps because stocks represent a claim with embedded leverage.

In practice, mutual funds rarely change their fees though they sometimes offer waivers and discounts. More commonly, funds suffer large outflows during downturns. Coupled with a decreasing returns-to-scale technology (e.g., Berk and Green, 2004), outflows raise net returns without lowering fees.

To implement this mechanism, suppose that the aggressiveness η with which funds trade is decreasing in assets under management A:

$$\eta_t = \eta(A_t), \quad \eta' < 0. \tag{26}$$

For the purposes of this paper, the diseconomies can be either at the fund level as found by Chen, Hong, Huang, and Kubik (2004) or at the industry level as in Pástor and Stambaugh (2012). With a fixed fee, $f_t = f$, Eq. (23) becomes

$$e^f - 1 = e^{\eta(A_t)[\phi m_{r,t} - (1+\phi\widehat{\sigma}_{ry}/\widehat{\sigma}_r^2)m_{y,t}]\Sigma_x}[e^{\eta(A_t)(\phi\Sigma_x)} - 1]. \tag{27}$$

In a market downturn, as m_r becomes more negative, mutual funds become riskier, which raises their required net returns. Since fees remain fixed, funds see investor outflows, which raises their gross returns to restore equilibrium. The model with decreasing returns to scale predicts that a rise in Sharpe ratios leads to both fund outflows and higher net returns.

3.3. Testing the model with fund flows

A direct test of the model requires identifying shocks to non-traded assets and relating them to market returns and active fund returns. In this section, I show how to use aggregate fund flows to proxy for beliefs and, by extension, income shocks. This connection forms the basis for the empirical tests that follow.

The household stochastic discount factor in Eq. (8) relates expected returns to risk exposures. Eq. (13) shows that high expected returns (low $\widehat{x}$) correspond to a high covariance with the stochastic discount factor (a more negative m_r). Intuitively, when expected returns are high, households keep a larger fraction of their wealth in stocks. The portfolio model in the Online Appendix links portfolio weights and stochastic discount factor exposures in more detail.[16] For present purposes, note that portfolios can adjust via either flows or capital appreciation. The goal here is to show that flows identify shocks to beliefs $\widehat{x}$ holding x fixed, i.e., shocks to household error, which in turn are driven by income shocks.

As flows involve trade between investors, and since households are relatively uninformed, consider the stochastic discount factor λ of an informed investor, one who observes the underlying shocks:

$$\Delta\lambda_{t+1} = \lambda_{0,t} + \lambda_{r,t}\varepsilon^r_{t+1} + \lambda_{y,t}\left[\varepsilon^y_{t+1} - \left(\frac{\sigma_{ry}}{\sigma_r^2}\right)\varepsilon^r_{t+1}\right] + \lambda_{x,t}\varepsilon^x_{t+1}. \tag{28}$$

In equilibrium, both informed and uninformed investors must be marginal in the stock market:

$$\overline{r} - \phi x_t + \tfrac{1}{2}\sigma_r^2 - r_f = -\lambda_{r,t}\sigma_r^2 \tag{29}$$

$$\overline{r} - \phi\widehat{x}_t + \tfrac{1}{2}\widehat{\sigma}_r^2 - r_f = -m_{r,t}\widehat{\sigma}_r^2. \tag{30}$$

Thus, we can relate the error in household beliefs $x - \widehat{x}$ to differences in stochastic discount factor exposures:

$$\phi(x_t - \widehat{x}) = \lambda_{r,t}\sigma_r^2 - m_{r,t}\widehat{\sigma}_r^2 + \tfrac{1}{2}\left(\sigma_r^2 - \widehat{\sigma}_r^2\right). \tag{31}$$

Since the volatilities are constant, a shock to household error $x - \widehat{x}$ must be met by a change in relative exposures $\lambda_{r,t}\sigma_r^2 - m_{r,t}\widehat{\sigma}_r^2$. By contrast, a shock to $\widehat{x}$ that leaves $x - \widehat{x}$ fixed, i.e., a true shock to expected returns, leaves the

[16] For example, with CRRA investors with risk aversion γ, the stochastic discount factor is $-\gamma\Delta c_{t+1}$. The exposure m_r is approximated by the household's total wealth share in stocks multiplied by $-\gamma$. (When non-traded income is correlated with returns, an adjustment is made for the non-traded income share of wealth.) See the Online Appendix.

difference between exposures unchanged even as their levels change. We can interpret the former shock as a flow from one group to the other (one group raises its exposure while the other reduces it) and the latter as a compensating price change (both groups bid prices up or down without trading). As in Grossman and Stiglitz (1980), incomplete information precludes optimal risk sharing.

The connection between belief shocks and flows leads to a connection between non-traded income and flows. Consider a positive transitory shock to non-traded income, $\varepsilon^y > 0$. By optimal filtering, households raise their expectations, $\hat{x}$ goes up: households anticipate high income growth and lower stock returns. As the true x is unchanged, $x-\hat{x}$ is lower. By Eq. (31), $\lambda_r\sigma_r^2 - m_r\hat{\sigma}_r^2$ is also lower, so m_r rises relative to λ_r (becomes less negative). Thus, household exposure to stocks declines, leading to outflows. A negative income shock analogously leads to inflows.

A positive return shock $\varepsilon^r > 0$ also leads to inflows by a similar argument. Thus, flows are negatively correlated with non-traded income and positively correlated with stock returns. It is therefore necessary to control for market exposure when constructing a fund-flows factor.

The information structure of the model effectively assumes that fund flows are not publicly observable, at least not in real time. This assumption can be microfounded by adding a layer of unpriced idiosyncratic shocks to household income, left out here for simplicity. These would make the flows of individual households uninformative about aggregate shocks without changing the stochastic discount factor. This is a necessary condition for avoiding the no-trade theorem of Glosten and Milgrom (1985).

Since active funds attract investors with substantial non-traded income, it is their flows that should reveal non-traded shocks while also responding to contemporaneous stock returns. As they are related to household error, they should negatively predict subsequent stock returns. They should also predict relatively high active fund returns, as active managers take the other side. Since income shocks and flows move in opposite directions, the premium on a fund-flows risk factor should be negative (at least after controlling for correlation with the market). Active funds should load positively on this factor in a way that accounts for their apparent underperformance. I test all these predictions in the next section.

4. Empirical results

In this section, I test the predictions of the model using mutual fund flows. After documenting the buy-and-hold underperformance puzzle, I examine the relationship between flows and contemporaneous market returns, flows and future market returns, and flows and future active-minus-market returns. I then construct a flows-based factor that according to the model mimics shocks to non-traded assets and use it to price active and index funds. The section concludes with some robustness checks and extensions.

4.1. Data and preliminary results

I use monthly data on fund flows and returns from Morningstar. I consider U.S. equity open-end mutual funds

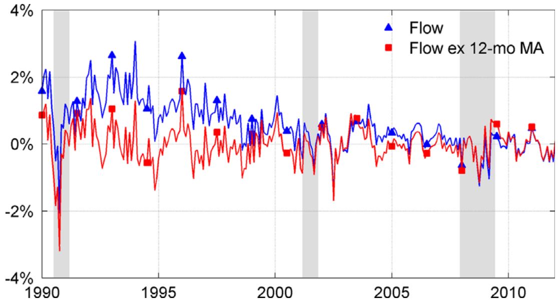

Fig. 2. The time series of fund flows. "Flow" is net cash flow (sales minus redemptions) divided by the previous month's assets. The econometric tests of the paper use the de-trended series, "Flow ex 12-mo MA", obtained by subtracting a 12-month moving average. Data shown are from Morningstar, January 1990 to January 2013, matching the availability of returns. Gray bands indicate NBER recessions.

Table 1
Summary statistics.
Means, standard deviations, autocorrelations, and observation counts for key variables. Using fund-level data on net cash flows and returns, I aggregate all funds in a given category into a single value-weighted fund. "All" is comprised of all U.S. equity mutual funds. "Active" and "Index" are for the subsets of active and index funds, respectively. "Flow" is the ratio of net cash flows (sales minus redemptions) over last month's total assets, expressed in percent. To remove time trends, I subtract a 12-month moving average from each month's flow (see Fig. 2). Returns are in percent per month. Monthly data from Morningstar, January 1990 to January 2013.

		Mean	St. dev.	Autocorr.	N
Flows	All	−0.03	0.53	0.43	277
	Active	−0.03	0.54	0.41	277
	Index	−0.05	1.08	0.32	277
Returns	All	0.74	4.48	0.12	277
	Active	0.74	4.50	0.12	277
	Index	0.79	4.34	0.06	277

from January 1989 to January 2013 (no index funds report monthly flows prior to 1989). The data are reported at the share-class level. I aggregate all share classes on a value-weighted basis into three aggregated "funds": all funds, active funds, and index funds. All analysis takes place at this aggregated level. Whereas fund-level flows are driven by the search for skill in the cross section (Chevalier and Ellison, 1997; Sirri and Tufano, 1998), aggregate flows reflect aggregate shocks, as in the model.

I calculate flows as net cash flows (also called dollar flows) over lagged total net assets. Fig. 2 plots the time series of aggregate flows. A downward trend reveals a slowdown in the growth rate of the mutual fund sector. Since returns are stationary, I remove this trend by subtracting a 12-month moving average. I use de-trended flows in all tests, which limits the final sample to 277 monthly observations between January 1990 and January 2013.

Table 1 shows descriptive statistics for the key variables, flows and returns. Flows are moderately persistent with a half-life of about 25 days. A typical month sees flows on the order of a half a percent of total assets. Index fund flows are more volatile, a result concentrated in the first half of the sample when index funds are relatively

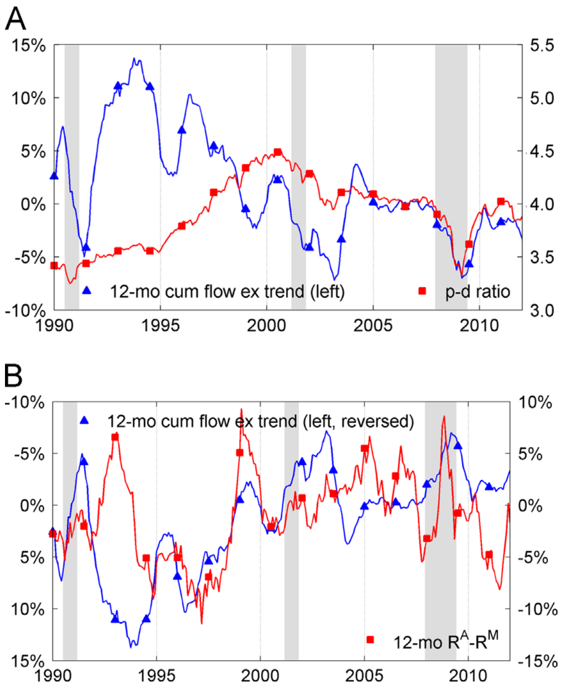

Fig. 3. Cumulative flows and expected returns. "12-mo cum flow ex trend" is the sum of all flows over the past 12 months net of a linear time trend; "*p*–*d* ratio" is the 12-month trailing price–dividend ratio of the S&P 500. The return "12-mo R^A-R^M is the return on a strategy long the aggregate active mutual fund portfolio and short the market portfolio. Data are from Morningstar, January 1990 to January 2013. Gray bands indicate NBER recessions.

small. The average monthly returns of index funds are five basis points higher than active funds before adjusting for risk.

Fig. 3 looks at the relationship between the size of the active fund sector and returns. In the model with decreasing returns to scale, times of high risk premiums are associated with outflows and hence higher net returns.[17] Panel A of Fig. 3 plots 12-month cumulative fund flows net of a linear time trend against the market price–dividend ratio, a proxy for expected returns. The second half of the sample shows a strong positive relationship between cumulative flows and the price–dividend ratio, as predicted. In the first half, when the mutual fund sector is small and growing at a faster pace, this relationship is not evident, though it remains clear that the sector contracts during recessions, when risk premiums are high.

Panel B of Fig. 3 looks at decreasing returns to scale by plotting past cumulative flows (on a reversed axis) against future active fund returns in excess of the market. The two series have a sample correlation of -29% (Newey-West *t*-stat of 1.63). Visually, the co-movement between the two series is strongest in the mid-to-late 1990s and the financial crisis of 2008. Overall, Fig. 3 offers some support to the decreasing returns-to-scale channel, which is standard in the literature (Berk and Green, 2004; Pástor and Stambaugh, 2012). Chen, Hong, Huang, and Kubik (2004) offer direct evidence.

4.2. Main results

Table 2 confirms the well-documented starting premise of the paper: active mutual funds underperform standard benchmarks, whereas index funds do not. The average fund underperforms by 0.07% to 0.10% per month, or 0.84% to 1.20% per year. These numbers are consistent with Fama and French (2010). Active funds tend to overweight small stocks, which lowers their three-factor alphas. Although the CAPM alphas are not statistically significant, the three- and four-factor alphas are significant at the 95% confidence level for both the total and active aggregated funds.

By contrast, index funds have tiny alphas of $\pm 0.01\%$, never more than half a standard error from zero. Index funds load more heavily on large stocks as many explicitly seek to replicate the S&P 500 index. Perhaps surprisingly, index funds also appear to load positively on the value factor and slightly negatively on the momentum factor. Overall, Table 2 demonstrates that active funds underperform traded benchmarks by about as much as their expense ratios (which range between 0.85% and 1% per year over 1990 to 2006 according to French, 2008).

4.2.1. Flows and returns

Turning to the core predictions of the model, Table 3 shows that active fund flows are strongly related to contemporaneous market returns, whereas index fund flows are not. I place the market return on the left side of the regression to run a horse race between active and index fund flows on the right. Overall, a one-standard deviation inflow is associated with a 1.8% higher market return, and this relationship is significant at the 1% level. The middle columns of Table 3 show that the effect is concentrated among active funds, which have a coefficient 4.5 times that of index funds. Whereas the stand-alone coefficient on index fund flows is also highly significant, it drops by a factor of seven and falls within a standard error from zero when active fund flows are included in the regression. By contrast, the coefficient on active fund flows is undiminished. Looking at R^2, market returns account for 16% of the variation in active flows versus only 3% for index flows.

The results of Table 3 are consistent with the predictions of the model, where active funds attract investors who are more likely to rebalance due to outside income. These investors rely on market returns to learn about expected returns. As a result, they respond strongly to current market returns (they also respond to non-traded income shocks).

Table 4 shows results from regressions of future stock market returns on current flows. I include the market price–dividend ratio as an additional forecasting variable in some specifications since households in the model condition on prices. The top panel of Table 4 shows that aggregate flows negatively forecast market returns at the six- and 12-month horizons. (The point estimates are negative but insignificant at the three-month horizon.) A one-standard deviation monthly inflow is associated with

[17] In the absence of decreasing returns to scale, net returns must rise by another mechanism such as fee waivers or greater skill.

Table 2
Mutual fund performance.

This table compares the after-fee performance of active and index mutual funds. "All" is the return on a value-weighted portfolio of all U.S. equity mutual funds. "Active" is for active funds and "Index" is for index funds. The alphas are in percent per month (e.g., −0.07 is negative seven basis points per month). $Mkt-R_f$ is the excess market return, "SMB" and "HML" are the Fama-French size and book-to-market factors, and *MOM* is the momentum factor. Monthly data from Morningstar, January 1990 to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels, respectively.

	All			Active			Index		
α	−0.07	−0.09**	−0.10**	−0.08	−0.09**	−0.10**	−0.01	−0.01	0.01
	(0.05)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)	(0.02)	(0.02)
$Mkt-R_f$	0.99***	0.98***	0.98***	0.98***	0.98***	0.98***	0.96***	1.00***	0.99***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SMB		0.09***	0.08***		0.10***	0.10***		−0.18***	−0.18***
		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)
HML		0.02*	0.03*		0.02	0.03		0.04***	0.03***
		(0.01)	(0.01)		(0.02)	(0.02)		(0.01)	(0.02)
MOM			0.00			0.01			−0.02***
			(0.01)			(0.01)			(0.01)
R^2	97%	97%	97%	97%	97%	97%	97%	97%	97%
N	277	277	277	277	277	277	277	277	277

Table 3
Fund flows and contemporaneous returns.

This table shows the correlation between fund flows and contemporaneous market returns. Flows are measured as net cash flows over lagged assets, net of a 12-month moving average. "All flow" are the percentage flows of a value-weighted portfolio of all U.S. equity mutual funds. "Active flow" and "Index flow" are for value-weighted portfolios of all active funds and index funds, respectively. R^M is the CRSP value-weighted market return in the same month as the fund flows. Newey-West standard errors with 12 lags. Monthly data from Morningstar, January 1990 to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	R_t^M			
Constant	0.91***	0.91***	0.86***	0.91***
	(0.27)	(0.27)	(0.30)	(0.27)
All flow$_t$	3.38***			
	(0.59)			
All flow$_t$		3.37***		3.28***
		(0.58)		(0.66)
Index flow$_t$			0.76***	0.10
			(0.18)	(0.19)
R^2	16%	16%	3%	16%
N	277	277	277	277

a 1.8% lower return over the next six months (significant at the 10% level) and a 2.5% lower return over the next year (significant at the 5% level). These magnitudes are on the order of half the sample equity premium. Conditioning on a public signal like the price–dividend ratio does not affect the forecasting power of flows, as predicted. Nevertheless, the price–dividend ratio is a strong predictor, which shows that the information contained in fund flows is distinct.

The remaining panels of Table 4 show that the forecasting power of fund flows is concentrated among active funds. The forecasting coefficients for active fund flows are very close to the aggregate numbers, whereas the coefficients on index fund flows are three to 20 times smaller. They are also insignificant in all but one specification, where they come in with a t-stat of 1.78. The last panel of Table 4 shows that in a horse race between active and index funds, active fund flows retain the magnitude of their forecasting coefficients, whereas those of index funds tend to change sign. Active fund flows retain their significance at the 12-month horizon but not at the six-month horizon.

The evidence in Tables 3 and 4 suggests that active fund investors rebalance their portfolios in a systematic way: they buy in response to contemporaneous market returns that are reversed in subsequent months. In the model, households that have this tendency—due to imperfect information and outside income shocks—optimally invest with active funds, even at a high premium as active managers undo their poor timing.

Table 5 brings together the actions of investors and fund managers by showing that high inflows, though they predict lower market returns, also predict high active fund returns relative to the market. The table regresses the future returns of a strategy that is long mutual funds and short the market on current fund flows. The top panel shows that a one-standard deviation inflow is associated with a 0.28% higher active-minus-market return over the next three months (significant at the 1% level) and a 0.46% higher return over the next six months (also significant at the 1% level). The point estimates are large but insignificant at the 12-month horizon. Curiously, the statistical significance of the results in Table 5 is concentrated at a slightly shorter (though overlapping) horizon than in Table 4. However, the point estimates increase similarly with horizon in both tables.

As before, the forecasting power of flows is concentrated among active funds: the coefficients on active fund flows are nearly the same in magnitude and significance to the aggregate numbers, whereas the coefficients on index fund flows are about five times smaller and never much more than one standard error from zero. In a horse race (bottom panel), active fund flows retain their forecasting power whereas index fund flows become even weaker and often flip sign, as predicted.

Table 4
Predicting market returns with fund flows.

Regressions of future market returns on current flows. Flows are net cash flows over lagged assets, net of a 12-month moving average. "All flow" is for a value-weighted portfolio of all U.S. equity mutual funds. "Active flow" and "Index flow" are for value-weighted portfolios of active and index funds, respectively. R^M is the return on the market portfolio from CRSP over a 3-, 6-, and 12-month horizon. p–d is the 12-month trailing market price–dividend ratio. Newey-West standard errors with 12 lags. Monthly data from Morningstar, January 1990 to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	$R^M_{t+1,t+3}$		$R^M_{t+1,t+6}$		$R^M_{t+1,t+12}$	
All fund flows						
Constant	2.53***	20.68***	5.07***	44.77**	10.71***	99.88***
	(0.87)	(9.63)	(1.71)	(17.61)	(3.20)	(28.32)
All flow$_t$	−0.89	−0.89	−3.41*	−3.40*	−4.80**	−4.79**
	(1.22)	(1.25)	(1.93)	(1.85)	(2.41)	(2.09)
$(p-d)_t$		−4.64*		−10.15**		−22.63***
		(2.48)		(4.58)		(7.47)
R^2	0%	3%	2%	9%	2%	18%
Active fund flows						
Constant	2.53***	20.63**	5.08***	44.60**	10.71***	99.03***
	(0.87)	(9.63)	(1.71)	(17.61)	(3.20)	(28.28)
Active flow$_t$	−0.91	−0.89	−3.30*	−3.25*	−4.95**	−4.83**
	(1.25)	(1.28)	(1.91)	(1.86)	(2.43)	(2.13)
$(p-d)_t$		−4.63*		−10.10**		−22.57***
		(2.48)		(4.58)		(7.46)
R^2	0%	3%	2%	9%	2%	18%
Index fund flows						
Constant	2.56***	20.74**	5.11***	46.00***	10.85***	100.19***
	(0.87)	(9.49)	(1.70)	(17.51)	(3.21)	(28.81)
Index flow$_t$	0.08	−0.03	−0.87	−1.11*	−0.21	−0.73
	(0.33)	(0.35)	(0.74)	(0.62)	(1.08)	(0.83)
$(p-d)_t$		−4.65*		−10.45**		−22.84***
		(2.44)		(4.54)		(7.58)
R^2	0%	3%	1%	8%	0%	16%
Active and index fund flows						
Constant	2.54***	20.44**	5.07***	45.19**	10.75***	98.70***
	(0.87)	(9.48)	(1.72)	(17.66)	(3.21)	(28.46)
Active flow$_t$	−1.16	−1.04	−3.09	−2.81	−5.71**	−5.08**
	(1.44)	(1.48)	(1.91)	(1.96)	(2.80)	(2.48)
Index flow$_t$	0.31	0.18	−0.26	−0.54	0.93	0.30
	(0.42)	(0.45)	(0.73)	(0.66)	(1.22)	(0.98)
$(p-d)_t$		−4.58*		−10.26**		−22.48***
		(2.47)		(4.59)		(7.53)
R^2	0%	3%	2%	9%	3%	18%
N (all panels)	274	274	271	271	265	265

4.2.2. *Asset pricing with flows*

Overall, Tables 3–5 suggest that active fund investors are more likely than index fund investors to buy ahead of low market returns, and that active funds offer protection against this bad timing. According to the model, the value of this timely service can raise the compensation of fund managers above their gross alpha. This value depends on the premium for non-traded income risk. The model shows that non-traded income shocks are reflected in the investment flows of mutual fund investors: A positive non-traded income shock conditional on the market return is associated with lower expected returns, and therefore fund outflows. This result suggests that fund flows can serve as an inverse proxy for non-traded income shocks in an asset pricing framework. The premium on an aggregate flow factor should be negative.

To test this prediction, I estimate the premium of a fund-flows factor in the cross section of equities, which is independent of mutual funds. To construct the factor, I value-weight all stocks in CRSP into double-sorted quintiles based on their univariate betas with respect to the market return and the time series of aggregate fund

Table 5
Predicting active-minus-market returns with fund flows.

Regressions of future active fund returns in excess of the market on current flows. Flows are net cash flows over lagged assets, net of a 12-month moving average. "All flow" is for a value-weighted portfolio of all U.S. equity mutual funds. "Active flow" and "Index flow" are for value-weighted portfolios of active and index funds, respectively. $R^A - R^M$ is the return on a strategy long the value-weighted active fund portfolio and short the market over a 3-, 6-, and 12-month horizon. $p-d$ is the 12-month trailing market price–dividend ratio. Newey-West standard errors with 12 lags. Monthly data from Morningstar, January 1990 to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	$(R^A-R^M)_{t+1,t+3}$		$(R^A-R^M)_{t+1,t+6}$		$(R^A-R^M)_{t+1,t+12}$	
All fund flows						
Constant	−0.23	−1.34	−0.48	−2.68	−1.10	−8.40
	(0.18)	(2.33)	(0.37)	(4.59)	(0.77)	(7.96)
All flow$_t$	0.54***	0.54***	0.87***	0.87***	0.90	0.90
	(0.19)	(0.19)	(0.33)	(0.32)	(0.63)	(0.59)
$(p-d)_t$		0.28		0.56		1.87
		(0.60)		(1.17)		(2.02)
R^2	3%	4%	3%	4%	1%	3%
Active fund flows						
Constant	−0.23	−1.32	−0.48	−2.63	−1.10	−8.35
	(0.18)	(2.32)	(0.37)	(4.57)	(0.76)	(7.94)
Active flow$_t$	0.54***	0.54***	0.86**	0.86***	0.89	0.89
	(0.20)	(0.19)	(0.34)	(0.33)	(0.64)	(0.60)
$(p-d)_t$		0.28		0.55		1.85
		(0.59)		(1.17)		(2.02)
R^2	3%	4%	3%	4%	1%	3%
Index fund flows						
Constant	−0.24	−1.47	−0.49	−2.90	−1.12	−8.57
	(0.19)	(2.45)	(0.39)	(4.78)	(0.78)	(8.24)
Index flow$_t$	0.10	0.10	0.17	0.18	0.10	0.14
	(0.09)	(0.08)	(0.17)	(0.16)	(0.27)	(0.27)
$(p-d)_t$		0.31		0.61		1.91
		(0.62)		(1.22)		(2.09)
R^2	0%	1%	0%	1%	0%	2%
Active and index fund flows						
Constant	−0.23	−1.31	−0.48	−2.65	−1.10	−8.30
	(0.18)	(2.28)	(0.37)	(4.51)	(0.76)	(7.95)
Active flow$_t$	0.55**	0.54**	0.86**	0.85**	0.98	0.93
	(0.23)	(0.22)	(0.43)	(0.41)	(0.76)	(0.73)
Index flow$_t$	−0.01	−0.01	−0.01	0.01	−0.10	−0.05
	(0.11)	(0.10)	(0.21)	(0.19)	(0.32)	(0.33)
$(p-d)_t$		0.28		0.56		1.84
		(0.58)		(1.15)		(2.02)
R^2	3%	4%	3%	4%	1%	3%
N (all panels)	274	274	271	271	265	265

flows. The flow factor *FLO* is the return on a strategy long the five quintiles with high flow betas and short the five quintiles with low flow betas (this approach reduces the correlation between *FLO* and the market return).

Table 6 explores the distribution and pricing of *FLO*. Panel A shows that *FLO* has a weak positive 10% correlation with the market (by construction) and a strong 53% correlation with the size factor. The correlation is a weaker −26% with value and an insignificant −8% with momentum. These results suggest that small stocks are more likely to have high returns in months with high fund inflows, making them a potential hedge for non-traded shocks in the context of the model.

Table 6
Aggregate flows as a risk factor.

Pricing of the aggregate fund flows risk factor *FLO*. *FLO* is constructed from the returns of 25 double-sorted portfolios. Each stock in CRSP is assigned full-sample univariate betas with respect to the market return and the time series of aggregate fund flows (net of a 12-month moving average). Each month, stocks are sorted into value-weighted portfolios based on the quintile of their market and flow betas. *FLO* is a portfolio that is long the five high-flow beta quintiles and short the five low-flow beta quintile portfolios. Panel A displays correlations between *FLO* and standard risk factors. Panel B runs time-series regressions of *FLO* on the other factors. All returns are in percent per month (e.g., −0.83 is minus 83 basis points per month). Data are monthly from Morningstar, January 1990 to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	Panel A: Pairwise factor correlations with FLO			
	$Mkt-R_f$	*SMB*	*HML*	*MOM*
FLO	0.10	0.53	−0.26	−0.08
	Panel B: Time-series regressions of FLO			
	(1)	(2)	(3)	(4)
Constant	−0.83*	−0.92**	−0.94**	−0.76**
	(0.43)	(0.43)	(0.37)	(0.37)
$Mkt-R_f$		0.16*	−0.08	−0.16*
		(0.10)	(0.09)	(0.09)
SMB			1.07***	1.09***
			(0.12)	(0.12)
HML			−0.23*	−0.30***
			(0.12)	(0.12)
MOM				−0.21***
				(0.07)
R^2	0%	1%	29%	31%
N	276	276	276	276

Panel B of Table 6 considers the pricing of *FLO*. *FLO* has a sample average return of −0.83% per month (*t*-stat 1.93), suggesting a negative premium, as expected. Importantly, this premium is not subsumed by the other factors. Controlling for the correlation between *FLO* and the market, the flow premium grows in magnitude to −0.92% and is significant at the 5% level. Adding the size and value factors strengthens it further to −0.94% (5% significance). Momentum tapers the premium to −0.76% without affecting its statistical significance. As *FLO* has by far its largest loading on size, a negative average return despite a positive size premium bolsters the evidence for a negative flow premium.

The next step is to check whether the negative flow premium can help price mutual funds. Since *FLO* was obtained from equity returns, it offers an independent test. For this exercise, I construct the returns of a portfolio

Table 7
Asset pricing with flows.

Results from an asset pricing test using the aggregate fund flows risk factor *FLO*. *FLO* is constructed from the returns of 25 value-weighted portfolios, double-sorted according to their univariate betas with respect to the market and aggregate fund flows. *FLO* is a portfolio that is long the five high-flow beta quintiles and short the five low-flow beta quintile portfolios. The table presents alphas and betas from time-series regressions of the return on an active-minus-index long-short portfolio strategy R^A-R^I on *FLO* and standard risk factors. All returns are in percent per month (e.g., −0.07 is minus seven basis points per month). Data are monthly from Morningstar, January 1990 to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	R^A-R^I					
Alpha	−0.07	0.02	−0.09*	−0.06	−0.11**	−0.08
	(0.08)	(0.07)	(0.05)	(0.05)	(0.05)	(0.05)
FLO		0.09***		0.03***		0.03***
		(0.01)		(0.01)		(0.01)
$Mkt-R_f$	0.03**	0.02	−0.02*	−0.02	−0.01	−0.01
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
SMB			0.28***	0.25***	0.28***	0.24***
			(0.02)	(0.02)	(0.02)	(0.02)
HML			−0.01	−0.01	−0.01	0.00
			(0.02)	(0.02)	(0.02)	(0.02)
MOM					0.03**	0.03***
					(0.01)	(0.01)
R^2	1%	3%	53%	55%	54%	56%
N	276	276	276	276	276	276

long active funds and short index funds. I run time-series regressions of this long-short portfolio on the risk factors.

The results of the pricing test are in Table 7. The active-minus-index fund portfolio has a highly significant positive beta of 0.09 with respect to *FLO* in an augmented CAPM specification. In other words, active funds outperform index funds in months in which mutual fund inflows are high, even after controlling for the market return. The *FLO* beta falls to 0.03 when the size factor is added, but it remains significant at the 1% level. The positive *FLO* beta arises in the model: active funds perform relatively well when non-traded income disappoints and expected returns are revised higher, leading to inflows. As predicted, this correlation remains even conditional on the market return.

The intercepts in Table 7 show that including *FLO* improves the measured performance of active funds. In the augmented CAPM specification, the alpha flips from −0.07% to 0.02% per month. The three- and four-factor alphas increase from −0.09% and −0.11% to −0.06% and −0.08%. In all cases, the alphas are not statistically significant once *FLO* is included. Thus, depending on the specification, the fund-flows factor accounts for between 25% and over 100% of the apparent underperformance of active funds.

It is possible that this result understates the effect of non-traded shocks if the cross section of equities does not fully reveal these shocks. Nevertheless, the results suggest that fund flows are picking up important shocks experienced by household investors.

Table 8
Fund flows and contemporaneous returns, robustness.

This table shows the correlation between fund flows and contemporaneous market returns under four robustness specifications. For benchmark results and methodology, see Table 3, column 4. “No retirement” excludes funds flagged as “available for retirement plan” in Morningstar; “No institutional” excludes funds flagged as institutional; and “No small-cap & sector” excludes funds flagged as small-cap, defined as primarily investing in stocks with market capitalization below $1 billion, or as sector funds; “Bond funds” is for corporate bond mutual funds. Monthly data from Morningstar, January 1990 (January 1992 for bond funds) to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	R_t^M			
	No retirement	No institutional	No small-cap & sector	Bond funds
Constant	0.91***	0.92***	0.92***	0.61***
	(0.27)	(0.27)	(0.27)	(0.08)
Active flow$_t$	3.20***	3.31***	3.28***	0.51***
	(0.64)	(0.61)	(0.74)	(0.11)
Index flow$_t$	0.15	0.07	0.14	0.06
	(0.81)	(0.16)	(0.21)	(0.07)
R^2	16%	17%	14%	14%
N	277	277	277	253

4.3. Robustness

Tables 8–10 examine the robustness of the main results of the paper with respect to variation in the underlying sample of funds. A full set of tables is available in the Online Appendix that accompanies the paper. Tables 8–10 also present results for bond funds, which are discussed separately in Section 4.4.

The first robustness check excludes funds flagged as “available for retirement” in Morningstar. Index funds receive substantial flows through retirement plans on a pre-determined basis, which could potentially account for their distinct flow dynamics.[18]

The second robustness check excludes funds flagged as institutional. The goal here is to determine whether the documented differences in the flow dynamics of index and active funds are correlated with differences in institutional versus retail clienteles.[19]

The third robustness check excludes small-cap funds (defined by Morningstar as funds that primarily invest in stocks with market capitalization under $1 billion), and sector funds.[20] Many index funds explicitly target the S&P 500, a large-cap index. Removing small-cap and sector

[18] At the same time, the model predicts that investors with stable income flows, natural buy-and-hold-investors, are optimally attracted to index funds, so this observation can be interpreted as a prediction.

[19] A priori, one might expect institutional clients to favor either index or active funds. For example, they might be more informed, which would suggest index funds, or since many are in reality wealthy individuals, they might have significant non-traded assets (e.g., private business), which would lead them towards active funds.

[20] An alternative is to only include large-cap funds, which Morningstar defines as primarily investing in stocks with market capitalization over $8 billion. Results are similar to those excluding small-cap stocks. The relationship between flows and contemporaneous returns is unaffected.

Table 9
Predicting market returns with fund flows, robustness.

Regressions of future market returns on current flows under four robustness specifications. For benchmark results and methodology, see Table 4. “No retirement funds” excludes funds flagged as “available for retirement plan” in Morningstar; “No institutional funds” excludes funds flagged as institutional; and “No small-cap & sector funds” excludes funds flagged as small-cap (defined as primarily investing in stocks with market capitalization below $1 billion) or as sector funds; “Bond funds” is for corporate bond mutual funds. Monthly data from Morningstar, January 1990 (January 1992 for bond funds) to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	$R^M_{t+1,t+3}$		$R^M_{t+1,t+6}$		$R^M_{t+1,t+12}$	
			No retirement funds			
Constant	2.54***	20.37**	5.07***	45.16**	10.75***	98.63***
	(0.87)	(9.59)	(1.72)	(17.65)	(3.21)	(28.41)
Active flow$_t$	−1.21	−1.09	−3.08	−2.80	−5.73**	−5.11**
	(1.41)	(1.44)	(1.92)	(1.96)	(2.80)	(2.49)
Index flow$_t$	0.38	0.26	−0.27	−0.54	0.98	0.37
	(0.38)	(0.40)	(0.73)	(0.67)	(1.22)	(0.98)
$(p-d)_t$		−4.56*		−10.25**		−22.46***
		(2.47)		(4.59)		(7.51)
R^2	1%	3%	2%	9%	3%	18%
			No institutional funds			
Constant	2.54***	20.46**	5.06***	45.23**	10.76***	98.10***
	(0.87)	(9.48)	(1.72)	(17.63)	(3.19)	(28.37)
Active flow$_t$	−0.96	−0.85	−2.87	−2.62	−5.70**	−5.15**
	(1.43)	(1.47)	(1.91)	(1.97)	(2.45)	(2.20)
Index flow$_t$	0.26	0.14	−0.20	−0.47	1.33	0.75
	(0.48)	(0.49)	(0.73)	(0.68)	(1.13)	(0.93)
$(p-d)_t$		−4.58*		−10.27**		−22.33***
		(2.45)		(4.59)		(7.50)
R^2	0%	3%	2%	9%	3%	18%
			No small-cap & sector funds			
Constant	2.53***	20.42**	5.05***	45.11***	10.74***	98.76***
	(0.87)	(9.51)	(1.72)	(17.46)	(3.20)	(28.35)
Active flow$_t$	−1.27	−1.12	−3.46*	−3.12	−5.68*	−4.93*
	(1.62)	(1.67)	(2.08)	(2.13)	(2.97)	(2.72)
Index flow$_t$	0.31	0.18	−0.23	−0.52	0.86	0.22
	(0.44)	(0.45)	(0.74)	(0.66)	(1.24)	(1.00)
$(p-d)_t$		−4.57*		−10.24**		−22.50***
		(2.45)		(4.54)		(7.49)
R^2	0%	3%	2%	9%	2%	18%
			Bond funds			
Constant	1.78***	−0.65	3.59***	−1.52	7.20***	−5.16
	(0.31)	(1.22)	(0.59)	(2.52)	(1.08)	(5.07)
(Active flow)$_t$	0.20	0.22	0.61*	0.62	0.37	0.34
	(0.23)	(0.25)	(0.37)	(0.39)	(0.49)	(0.47)
(Index flow)$_t$	−0.17**	−0.12	−0.49***	−0.37**	−0.83***	−0.54***
	(0.08)	(0.09)	(0.15)	(0.15)	(0.24)	(0.20)
Yield$_t$		0.40**		0.84**		2.02**
		(0.20)		(0.41)		(0.79)
R^2	1%	4%	4%	11%	4%	20%

Table 10
Predicting active-minus-market returns with fund flows, robustness.

Regressions of future active fund returns in excess of the market on current flows under four robustness specifications. For benchmark results and methodology, see Table 5, panel 4. “No retirement funds” excludes funds flagged as “available for retirement plan” in Morningstar; “No institutional funds” excludes funds flagged as institutional; and “No small-cap & sector funds” excludes funds flagged as small-cap (defined as primarily investing in stocks with market capitalization below $1 billion) or as sector funds; “Bond funds” is for corporate bond mutual funds. Monthly data from Morningstar, January 1990 to January 2013. Three, two, and one stars denote significance at the 1%, 5%, and 10% levels.

	$(R^A-R^M)_{t+1,t+3}$		$(R^A-R^M)_{t+1,t+6}$		$(R^A-R^M)_{t+1,t+12}$	
			No retirement funds			
Constant	−0.23	−1.33	−0.48	−2.70	−1.11	−8.39
	(0.18)	(2.26)	(0.37)	(4.47)	(0.76)	(7.86)
Active flow$_t$	0.55**	0.55**	0.87**	0.86**	1.03	0.98
	(0.24)	(0.22)	(0.43)	(0.41)	(0.75)	(0.72)
Index flow$_t$	−0.03	−0.02	−0.01	0.00	−0.14	−0.08
	(0.11)	(0.10)	(0.21)	(0.19)	(0.31)	(0.32)
$(p-d)_t$		0.28		0.57		1.86
		(0.58)		(1.14)		(2.00)
R^2	3%	4%	3%	4%	2%	4%
			No institutional funds			
Constant	−0.24	−1.19	−0.49	−2.51	−1.12	−7.99
	(0.17)	(2.26)	(0.36)	(4.46)	(0.73)	(7.83)
Active flow$_t$	0.53**	0.52**	0.83**	0.82**	0.85	0.81
	(0.22)	(0.21)	(0.40)	(0.39)	(0.70)	(0.68)
Index flow$_t$	−0.06	−0.05	−0.07	−0.06	−0.19	−0.14
	(0.11)	(0.10)	(0.20)	(0.20)	(0.31)	(0.33)
$(p-d)_t$		0.24		0.52		1.76
		(0.58)		(1.14)		(2.00)
R^2	3%	4%	3%	4%	1%	3%
			No small-cap & sector funds			
Constant	−0.25*	−0.99	−0.53*	−2.18	−1.19*	−7.05
	(0.15)	(1.85)	(0.30)	(3.63)	(0.61)	(6.41)
Active flow$_t$	0.41*	0.41**	0.64*	0.63*	0.90	0.85
	(0.22)	(0.21)	(0.39)	(0.37)	(0.73)	(0.73)
Index flow$_t$	−0.01	−0.01	−0.01	0.00	−0.16	−0.12
	(0.09)	(0.09)	(0.17)	(0.16)	(0.28)	(0.29)
$(p-d)_t$		0.19		0.42		1.50
		(0.47)		(0.92)		(1.61)
R^2	2%	2%	2%	3%	1%	4%
			Bond funds			
Constant	−0.23	1.29*	−0.49*	2.01	−0.99*	3.39
	(0.15)	(0.74)	(0.28)	(1.32)	(0.51)	(2.61)
(Active flow)$_t$	0.14	0.13	0.00	−0.01	0.25	0.26
	(0.12)	(0.12)	(0.10)	(0.11)	(0.22)	(0.21)
(Index flow)$_t$	0.07	0.03	0.16**	0.10	0.27**	0.17
	(0.05)	(0.05)	(0.07)	(0.08)	(0.12)	(0.14)
Yield$_t$		−0.25**		−0.41*		−0.72*
		(0.12)		(0.21)		(0.43)
R^2	2%	7%	2%	9%	4%	14%

funds checks whether differences in the portfolios of index and active funds are related to the main results of the paper.[21]

Table 8 reproduces the contemporaneous return regressions in Table 3, last column. In each robustness specification, active fund flows are strongly correlated with contemporaneous market returns, whereas index fund flows are not. The coefficients are very close to the benchmark results and the statistical significance is identical.

Table 9 reproduces the market return forecasting regressions in Table 4, last panel. In all cases, as in the benchmark results, active flows are significant negative predictors of the market return at the 12-month horizon. Statistical significance is reduced to the 10% level when small-cap and sector funds are excluded but in this case, the six-month forecast gains in significance. As expected, index fund flows have no forecasting power.

Table 10 reproduces the active-minus-market return forecasting regressions in Table 5, last panel. Active fund flows are strong positive predictors of the excess return of active funds versus the market at the three- and six-month horizons across all robustness specifications. The coefficients are generally significant at the 5% level, falling to 10% in some specifications when small-cap and sector funds are excluded. Once again, index fund flows have no forecasting power. The magnitudes are close to the benchmark estimates.

The Online Appendix also contains results for forecasting models that condition on the lagged market return. The results from the market-forecasting regression become a bit weaker, whereas the active-minus-market forecasting regression becomes stronger. The lagged market return is a significant negative predictor of the active-minus-market return, consistent with countercyclical performance. The premium on *FLO* is somewhat larger following a market downturn, consistent with a higher volatility of non-traded income or risk aversion. In a conditional asset pricing model, *FLO* has a bigger impact on mutual fund performance, but statistical power is reduced.

4.4. Bond funds

In a final extension of the empirical analysis, I examine the universe of corporate bond funds. The key challenge in mapping bond funds into the framework of the paper is the nature of bonds as an intermediate asset class in terms of risk. In the model, risky and risk-free asset flows are negatively correlated and so they predict returns in opposite directions. Nevertheless, at nearly $1 trillion in assets, bond funds represent an important asset class.

The Online Appendix documents the existence of an underperformance puzzle among bond funds, though it is about one-third smaller than among equity funds (active bond funds underperform index bond funds by four to seven basis points per month versus six to ten for equity funds). These results are consistent with the literature (Blake, Elton, and Gruber, 1993).

The last column of Table 8 shows that the flows of active bond funds are strongly correlated with contemporaneous market returns, as with equity funds. The coefficients are quite a bit smaller but the statistical significance is just as high due to smaller standard errors.

The last panel of Table 9 shows that in contrast to equity funds, it is index bond fund flows that negatively predict market returns. The forecasting coefficients are small but the statistical significance is high. These results suggest that index bond fund investors are not attaining the high buy-and-hold returns of their funds. Non-traded income shocks generate this bad timing in the model, but the matching between funds and investors mitigates it.

In the active-minus-market forecasting regressions presented in Table 10, neither active nor index bond fund flows are good predictors. In the full set of tables in the Online Appendix, active bond fund flows come in significantly in specifications without index funds, as they do in the benchmark results.

Interestingly, the flows into bond and equity index funds are highly correlated at 40% versus only 12% for bond and equity active funds. This suggests a possible explanation for the predictability results: Index fund investors appear to treat bond and equity funds similarly, going in and out of both at the same time, whereas active fund investors do not. This could happen if active fund investors perceive bonds to be closer to the model's risk-free asset, perhaps because they are less risk averse. A related possibility is that if expected bond returns are stable, bond fund flows might be primarily driven by developments in other markets.

The Online Appendix also contains analogs to the asset pricing results in Tables 6 and 7. For consistency, I examine the pricing power of the equity-based fund-flows factor *FLO* in the context of bond funds. As with equities, *FLO* retains its robust negative premium against a set of bond-market factors that includes the returns of the aggregate bond market, a high-yield strategy, a term-premium strategy, and a Treasury index. Interestingly, *FLO* loads strongly on the high-yield index.

Like active equity funds, active bond funds also load strongly on *FLO*. Adding *FLO* to an asset pricing model narrows the performance gap between active and index bond funds by between one and five basis points per month. Since the gap is narrower to begin with, the effect is somewhat smaller and less significant than for equity funds.

5. Conclusion

Roll (1977) points out that non-traded assets create an omitted variables problem in asset pricing. If econometricians lack full information, then likely so do households. By hiring an active portfolio manager, they can hedge the risk of being wrong. This risk must exist widely and it must be at least in part systematic in order for active trading to be profitable in the first place. This makes it expensive to

(*footnote continued*)
The coefficient estimates for predicting the market return and the active-minus-market return come in significant at the 10% level.

[21] Within the model, an active fund bias towards small stocks could be due to greater uncertainty about their expected returns or a hedging demand by investors with substantial non-traded assets.

hedge; expensive enough to make active funds appear to underperform passive index funds net of fees.

The data suggest that active fund investors rebalance their portfolios in a way that is systematically related to market returns, whereas index fund investors do not. Active fund investors are more likely to buy (sell) ahead of low (high) market returns. Active fund managers help to undo this effect and, depending on the specification, the premium for this service can account for the apparent underperformance of active mutual funds.

In this paper, I study mutual fund performance from the household point of view. Consider the portfolio problem of a household that just got richer, perhaps thanks to a pay raise, a booming business, or house price appreciation: Is now a good time to buy stocks? What is the current expected return? And what if the future is not so bright after all? For some investors, like those with no outside income, these questions never arise and they are better off buying and holding low-cost index funds. For everyone else, there are active funds.

Appendix A. The Kalman filter

The state-space system can be summarized as

$$x_{t+1} = \rho x_t + \varepsilon^x_{t+1} \tag{32}$$

$$\Delta y_{t+1} = \overline{y} + x_t + \varepsilon^y_{t+1} \tag{33}$$

$$r_{t+1} = \overline{r} - \phi x_t + \varepsilon^r_{t+1}. \tag{34}$$

Households update their beliefs by running the population regression

$$x_{t+1} - \rho \widehat{x}_t = L_t \begin{bmatrix} \Delta y_{t+1} - (\overline{y} + \widehat{x}_t) \\ r_{t+1} - (\overline{r} - \phi \widehat{x}_t) \end{bmatrix} + \upsilon_{t+1}. \tag{35}$$

The orthogonality condition gives

$$L_t = \mathrm{E}_t \left[(x_{t+1} - \rho \widehat{x}_t) \begin{bmatrix} \Delta y_{t+1} - (\overline{y} + \widehat{x}_t) \\ r_{t+1} - (\overline{r} - \phi \widehat{x}_t) \end{bmatrix}' \right] \tag{36}$$

$$\cdot \mathrm{E}_t \left[\begin{bmatrix} \Delta y_{t+1} - (\overline{y} + \widehat{x}_t) \\ r_{t+1} - (\overline{r} - \phi \widehat{x}_t) \end{bmatrix} \begin{bmatrix} \Delta y_{t+1} - (\overline{y} + \widehat{x}_t) \\ r_{t+1} - (\overline{r} - \phi \widehat{x}_t) \end{bmatrix}' \right]^{-1}. \tag{37}$$

Let $\Sigma^2_{x,t} = \mathrm{E}_t[(x_t - \widehat{x}_t)^2]$ represent the variance of the error in household beliefs about x. Evaluating,

$$L_t = \rho \Sigma^2_{x,t} [1 \;\; -\phi] \begin{bmatrix} \Sigma^2_{x,t} + \sigma^2_y & -\phi \Sigma^2_{x,t} + \sigma_{ry} \\ -\phi \Sigma^2_{x,t} + \sigma_{ry} & \phi^2 \Sigma^2_{x,t} + \sigma^2_r \end{bmatrix}^{-1}. \tag{38}$$

The error variance evolves according to

$$\Sigma^2_{x,t+1} = \mathrm{E}_{t+1}[(x_{t+1} - \widehat{x}_{t+1})^2] \tag{39}$$

$$\Sigma^2_{x,t+1} = \mathrm{E}_{t+1} \left[\left(\rho (x_t - \widehat{x}_t) + \varepsilon^x_{t+1} - L_t \begin{bmatrix} (x_t - \widehat{x}_t) + \varepsilon^y_{t+1} \\ -\phi (x_t - \widehat{x}_t) + \varepsilon^r_{t+1} \end{bmatrix} \right)^2 \right] \tag{40}$$

$$\Sigma^2_{x,t+1} = \left(\rho - L_t \begin{bmatrix} 1 \\ -\phi \end{bmatrix} \right)^2 \Sigma^2_{x,t} + L_t \begin{bmatrix} \sigma^2_y & \sigma_{ry} \\ \sigma_{ry} & \sigma^2_r \end{bmatrix} L_t' + \sigma^2_x \tag{41}$$

with initial condition $\Sigma^2_{x,0}$. The steady state is given by the solution to

$$\Sigma^2_x = \left(\rho - L \begin{bmatrix} 1 \\ -\phi \end{bmatrix} \right)^2 \Sigma^2_x + L \begin{bmatrix} \sigma^2_y & \sigma_{ry} \\ \sigma_{ry} & \sigma^2_r \end{bmatrix} L' + \sigma^2_x \tag{42}$$

$$L = \rho \Sigma^2_x [1 \;\; -\phi] \begin{bmatrix} \Sigma^2_x + \sigma^2_y & -\phi \Sigma^2_x + \sigma_{ry} \\ -\phi \Sigma^2_x + \sigma_{ry} & \phi^2 \Sigma^2_x + \sigma^2_r \end{bmatrix}^{-1}. \tag{43}$$

Appendix B. Multiple stocks

I consider an extension of the model with multiple risky assets to allow for stock picking in addition to market timing. There are N stocks with returns stacked inside $\mathbf{r}_t = [r^1_t \;\; r^2_t \;\; \cdots \;\; r^N_t]'$ given by

$$\mathbf{r}_{t+1} = \overline{r} - \phi \mathbf{x}_t + \varepsilon^r_{t+1} \tag{44}$$

$$\mathbf{x}_{t+1} = \rho \mathbf{x}_t + \varepsilon^x_{t+1}, \tag{45}$$

where $\overline{r}$ is an $N \times 1$ vector, and ϕ and ρ are scalars. Let $\varepsilon^r \sim \mathrm{N}(0, \sigma_r \sigma_r')$ and $\varepsilon^x \sim \mathrm{N}(0, \sigma_x \sigma_x')$ with $\varepsilon^r \perp \varepsilon^x$ as before. The risk-free rate is r_f. Total log-output y is given by

$$\Delta y_{t+1} = \overline{y} + \mathbf{1}' \mathbf{x}_t + \varepsilon^y_{t+1}, \tag{46}$$

with $\varepsilon^y \sim \mathrm{N}(0, \sigma^2_y)$. For simplicity, suppose $\varepsilon^y \perp \varepsilon^x$. Aggregate growth is composed of growth in N sectors, only partly reflected in the N risky assets. Markets remain incomplete and returns are not fully revealing.

The linearity of the signals (returns and income) and the state imply the optimality of the Kalman filter. The household information filtration has the form

$$\widehat{\mathbf{x}}_{t+1} = \rho \widehat{\mathbf{x}}_t + \widehat{\varepsilon}^x_{t+1} \tag{47}$$

$$\Delta y_{t+1} = \overline{y} + \mathbf{1}' \widehat{\mathbf{x}}_t + \widehat{\varepsilon}^y_{t+1} \tag{48}$$

$$\mathbf{r}_{t+1} = \overline{r} - \phi \widehat{\mathbf{x}}_t + \widehat{\varepsilon}^r_{t+1}. \tag{49}$$

Let the steady state error covariance matrix be $\Sigma_x' \Sigma_x = \lim_{t \to \infty} \mathrm{E}_t[(\mathbf{x}_t - \widehat{\mathbf{x}}_t)(\mathbf{x}_t - \widehat{\mathbf{x}}_t)']$. The stochastic discount factor has the form

$$\Delta m_{t+1} = m_{0,t} + m'_{r,t} \widehat{\varepsilon}^r_{t+1} + m_{y,t} [\widehat{\varepsilon}^y_{t+1} - (\widehat{\sigma}_y \widehat{\sigma}_r')(\widehat{\sigma}_r \widehat{\sigma}_r')^{-1} \widehat{\varepsilon}^r_{t+1}]. \tag{50}$$

The vector m_r contains the risk prices of the N traded assets and m_y is the price of non-traded income risk. No-arbitrage requires

$$1 = \mathrm{E}_t[e^{\Delta m_{t+1} + r_f}] \quad \text{and} \quad \mathbf{1} = \mathrm{E}_t[e^{\Delta m_{t+1} + \mathbf{r}_{t+1}}]. \tag{51}$$

The risk-free rate fixes the level of the stochastic discount factor:

$$r_f = -m_{0,t} - \tfrac{1}{2} \left[m'_{r,t} \left(\widehat{\sigma}_r \widehat{\sigma}_r' \right) m_{r,t} + m^2_{y,t} \widehat{\sigma}^2_y \left(1 - \widehat{\sigma}_r' (\widehat{\sigma}_r \widehat{\sigma}_r')^{-1} \widehat{\sigma}_r \right) \right]. \tag{52}$$

The expected returns of the N stocks pin down the prices of the traded shocks:

$$\overline{r} - \phi \widehat{\mathbf{x}}_t + \tfrac{1}{2} \, \mathbf{diag}\left(\widehat{\sigma}_r \widehat{\sigma}_r' \right) - r_f = -\left(\widehat{\sigma}_r \widehat{\sigma}_r' \right) m_{r,t}. \tag{53}$$

As in the one-stock case, the price of non-traded risk m_y remains a free parameter.

Consider an active fund manager who knows the true state vector $\mathbf{x}$ and invests accordingly. For example, consider the portfolio weight vector $\chi_t = [\chi_t^1 \ \chi_t^2 \ \cdots \ \chi_t^N]'$ given by

$$\chi_t = \frac{1}{N} e^{-\eta_t (\Sigma_x' \Sigma_x)^{-1/2} (\mathbf{x}_t - \widehat{\mathbf{x}}_t) - \eta_t^2/2}. \tag{54}$$

Households expect their managers to be fully invested in the equally weighted market portfolio. Managers use their private information to overweight undervalued stocks. This strategy combines market timing and stock picking. Stock picking tends to dominate when households have an accurate estimate of the average expected return but not of any one stock's expected return. This situation arises if the volatility of aggregate income is low but idiosyncratic (or sector) volatility is high. This result is similar to Kacperczyk, Van Nieuwerburgh, and Veldkamp (2012) where managers rationally stock-pick in booms and market-time during downturns when aggregate volatility is high.

Pricing the active mutual fund gives

$$e^{f_t} - 1 = \sum_{i=1}^{N} e^{Cov_t(\log \chi_t^i, \Delta m_{t+1})} [e^{Cov_t(\log \chi_t^i, r_{t+1}^i)} - 1]. \tag{55}$$

A fund that trades in N assets is priced as a portfolio of funds that each trade in one asset. Evaluating,

$$e^{f_t} - 1 = \frac{1}{N} \mathbf{1}' e^{\eta_t (\Sigma_x' \Sigma_x)^{1/2} [\phi m_{r,t} - (\mathbf{1} + \phi(\widehat{\sigma}_r \widehat{\sigma}_r')^{-1} \widehat{\sigma}_r \widehat{\sigma}_y) m_{y,t}]} \left[\mathbf{1}' e^{\eta_t \phi \, \mathbf{diag}(\Sigma_x' \Sigma_x)^{1/2}} - 1 \right]. \tag{56}$$

Looking across stocks, higher expected returns (low m_r^i) contribute to higher net returns as in the time series. At the same time, stocks with more uncertain expected returns (high $(\Sigma_x' \Sigma_x)^{ii}$) contribute to lower net returns.

Overall, the multiple-stock extension of the model incorporates stock picking in addition to market timing. When aggregate uncertainty is high, market timing becomes more profitable but required net returns are high so fund compensation is relatively low. By contrast, stock picking is most profitable when there is lots of dispersion in expected returns across stocks.

Appendix C. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jfineco.2013.11.005.

References

Admati, A.R., Pfleiderer, P., 1990. Direct and indirect sale of information. Econometrica 58, 901–928.

Admati, A.R., Ross, S.A., 1985. Measuring investment performance in a rational expectations equilibrium model. J. Bus. 58, 1–26.

Ben-Rephael, A., Kandel, S., Wohl, A., 2012. Measuring investor sentiment with mutual fund flows. J. Financ. Econ. 104, 363–382.

Berk, J.B., van Binsbergen, J.H., 2012. Measuring managerial skill in the mutual fund industry. NBER Working Paper No. 18184.

Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. J. Polit. Econ. 112, 1269–1295.

Blake, C.R., Elton, E.J., Gruber, M.J., 1993. The performance of bond mutual funds. J. Bus. 66, 371–403.

Campbell, J., Viceira, L., 2002. Strategic Asset Allocation: Portfolio Choice for Long-term Investors. Oxford University Press, USA.

Campbell, J.Y., 1999. Asset prices, consumption, and the business cycle. In: Taylor, J., Woodford, M. (Eds.), Handbook of Macroeconomics, vol. 1, North Holland, Amsterdam, pp. 1231–1303. (Chapter 19).

Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. J. Polit. Econ. 107, 205–251.

Campbell, J.Y., Shiller, R.J., 1988. The dividend–price ratio and expectations of future dividends and discount factors. Rev. Financ. Stud. 1, 195–228.

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

Chen, J., Hong, H., Huang, M., Kubik, J., 2004. Does fund size erode mutual fund performance? The role of liquidity and organization. Am. Econ. Rev. 94, 1276–1302.

Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. J. Polit. Econ. 105, 1167–1200.

Cremers, K.M., Petajisto, A., 2009. How active is your fund manager? A new measure that predicts performance. Rev. Financ. Stud. 22, 3329–3365.

Dybvig, P.H., Ross, S.A., 1985. Differential information and performance measurement using a security market line. J. Finance 40, 383–399.

Edelen, R., 1999. Investor flows and the assessed performance of open-end mutual funds. J. Financ. Econ. 53, 439–466.

Elton, E.J., Gruber, M.J., 2013. Mutual funds. In: Constantinides, G.M., Harris, M., Stulz, R.M. (Eds.), Handbook of the Economics of Finance, vol. 2B. , North Holland, Amsterdam, pp. 1011–1061. (Chapter 15).

Fama, E.F., French, K.R., 2010. Luck versus skill in the cross-section of mutual fund returns. J. Finance 65, 1915–1947.

Ferson, W.E., Schadt, R.W., 1996. Measuring fund strategy and performance in changing economic conditions. J. Finance 51, 425–461.

French, K.R., 2008. Presidential address: the cost of active investing. J. Finance 63, 1537–1573.

Gennaioli, N., Shleifer, A., Vishny, R., 2012. Money doctors. NBER Working Paper No. 18174.

Glode, V., 2011. Why mutual funds underperform. J. Financ. Econ. 99, 546–559.

Glosten, L.R., Milgrom, P.R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. J. Financ. Econ. 14, 71–100.

Grinblatt, M., Titman, S., 1989. Portfolio performance evaluation: old issues and new insights. Rev. Financ. Stud. 2, 393–421.

Grossman, S.J., Stiglitz, J.E., 1976. Information and competitive price systems. Am. Econ. Rev. 66, 246–253.

Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. Am. Econ. Rev. 70, 393–408.

Gruber, M.J., 1996. Another puzzle: the growth in actively managed mutual funds. J. Finance 51, 783–810.

He, Z., Krishnamurthy, A., 2013. Intermediary asset pricing. Am. Econ. Rev. 103, 732–770.

Jensen, M.C., 1968. The performance of mutual funds in the period 1945–1964. J. Finance 23, 389–416.

Jiang, G., Yao, T., Yu, T., 2007. Do mutual funds time the market? Evidence from portfolio holdings. J. Financ. Econ. 86, 724–758.

Kacperczyk, M., Seru, A., 2007. Fund manager use of public information: new evidence on managerial skills. J. Finance 62, 485–528.

Kacperczyk, M., Sialm, C., Zheng, L., 2005. On the industry concentration of actively managed equity mutual funds. J. Finance 60, 1983–2011.

Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L., 2012. Rational attention allocation over the business cycle. NBER Working Paper No. 15450.

Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L. Time-varying fund manager skill. J. Finance, http://dx.doi.org/10.1111/jofi.12084, accepted for publication.

Kaniel, R., Kondor, P., 2013. The delegated Lucas tree. Rev. Financ. Stud. 26, 929–984.

Kosowski, R., 2011. Do mutual funds perform when it matters most to investors? U.S. mutual fund performance and risk in recessions and expansions. Q. J. Finance 1, 607–664.

Merton, R.C., 1981. On market timing and investment performance. I. An equilibrium theory of value for market forecasts. J. Bus. 54, 363–406.

Moskowitz, T.J., 2000. Mutual fund performance: an empirical decomposition into stock- picking talent, style, transactions costs, and expenses: discussion. J. Finance 55, 1695–1703.

Pástor, L., Stambaugh, R.F., 2012. On the size of the active management industry. J. Polit. Econ. 120, 740–781.

Roll, R., 1977. A critique of the asset pricing theory's tests. Part I: on past and potential testability of the theory. J. Financ. Econ. 4, 129–176.

Santos, T., Veronesi, P., 2006. Labor income and predictable stock returns. Rev. Financ. Stud. 19, 1–44.

Sharpe, W.F., 1991. The arithmetic of active management. Financ. Analysts J. 47, 7–9.
Sirri, E.R., Tufano, P., 1998. Costly search and mutual fund flows. J. Finance 53, 1589–1622.
Vayanos, D., Woolley, P., 2013. An institutional theory of momentum and reversal. Rev. Financ. Stud. 26, 1087–1145.
Warther, V.A., 1995. Aggregate mutual fund flows and security returns. J. Financ. Econ. 39, 209–235.
Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transactions costs, and expenses. J. Finance 55, 1655–1695.
Zheng, L., 1999. Is money smart? A study of mutual fund investors' fund selection ability. J. Finance 54, 901–933.