Why mutual funds "underperform" ☆

Vincent Glode *

The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA

A B S T R A C T

I propose a parsimonious model that reproduces the negative risk-adjusted performance of actively managed equity mutual funds. In the model, a fund manager can generate state-dependent active returns at a disutility. Negative expected performance and mutual fund investing simultaneously arise in equilibrium because the active return the fund manager generates covaries positively with a component of the pricing kernel that the performance measure omits, consistent with recent empirical evidence. Using data on U.S. funds, I also document new empirical evidence consistent with the model's cross-sectional implications.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Jensen (1968), Malkiel (1995), and Fama and French (2010), among others, show that actively managed U.S. equity mutual funds significantly "underperform" passive investment strategies, net of fees. Yet, despite the apparent inferiority to passive investment strategies, more than two trillion dollars were invested in these funds by the end of 2008.[1] This paper shows that investing in actively managed funds expected to perform poorly unconditionally can be rational if these funds tend to perform abnormally well when the economy is doing poorly, as Moskowitz (2000), Kosowski (2006), and Staal (2006) find.

I derive a partial equilibrium model of optimal fee setting and active management by a skilled fund manager. The model builds on an insight from Berk and Green (2004) that the fund manager owns the bargaining power in his relationship with investors. However, unlike

---

[1] See the 2009 Investment Company Fact Book at http://www.icifactbook.org/.

Berk and Green (2004), I allow the fund manager in my model to generate active returns that depend on the state of the economy. I investigate how this ability might influence the fee the fund manager will charge and the performance an econometrician will attribute to him. The model shows that mutual fund investing and negative expected fund performance can simultaneously arise in a setting with skilled fund managers facing rational investors.

The intuition behind my model is that a fund manager who can generate state-specific active returns, at a given disutility or cost, will be better off doing so for states in which investors are willing to pay more for these returns. Thus, the fund manager will optimally focus his effort toward realizing good performance during periods where investors' marginal utility of consumption is high (i.e., in bad states of the economy), and will generate active returns that covary positively with the pricing kernel. Investors will be willing to pay for this (partial) insurance against pricing kernel variations. The fee the fund manager is able to charge in equilibrium will equal the certainty equivalent of the value he adds through active management. As originally anticipated by Moskowitz (2000), I show that a *misspecified* performance measure, i.e., one based on an unbiased but imperfect proxy for the pricing kernel, will underestimate the value created by active management when active returns are positively correlated with the true pricing kernel. Consequently, the skilled fund manager in my model will wrongly appear to underperform passive investment strategies net of fees.

That misspecification in the performance measure leads to the measurement of abnormal performance should not come as a surprise (see, e.g., Berk, 1995). What is both unique and non-trivial about the result derived here is the demonstration that a fairly general type of misspecification should lead to the measurement of *negative unconditional performance* in equilibrium when the fund manager implements an investment strategy that *insures investors against bad states of the economy*. Negative expected performance and mutual fund investing can simultaneously arise in equilibrium because the active return the fund manager generates covaries positively with a component of the pricing kernel the performance measure omits. This paper should not, however, be regarded as claiming that negative performance, per se, is desired by investors or that, in reality, all fund managers are skilled. It instead demonstrates that well-documented facts often considered as anomalous can be reproduced in a model with rational agents.

I calibrate the model to the U.S. economy and reproduce quantitatively the measured underperformance of U.S. funds. I also use data on 3,147 funds over the 1980–2005 period and find new empirical evidence consistent with the model's cross-sectional predictions. Relative to other funds, funds with poor unconditional performance tend to charge high fees and generate highly countercyclical risk-adjusted returns. This finding might explain the survival of several funds with poor unconditional performance and suggest the existence of a recession-related misspecification in popular performance measures.

Ideally, a fund's risk-adjusted performance would be measured by the fund's realized excess return, net of fees,

minus a risk premium for the covariance between the fund's return and a pricing kernel. In practice, the most popular measure of mutual fund performance is the intercept (alpha) from a regression of a fund's excess returns, net of fees, on the excess returns of passive investment strategies. The linear combination of these passive excess returns proxies for the empirically unobservable pricing kernel. Gruber (1996) argues that, since these passive excess returns are associated with zero-cost portfolios, alpha should be zero for a random portfolio. When he finds that the average alpha for actively managed U.S. equity funds is negative and smaller in absolute value than the average fee these funds charge, Gruber concludes that fund managers add value, on average, but charge investors more than the value they add.[2] According to this argument, the widely documented negative alphas indicate that investing in actively managed mutual funds destroys value and is irrational from an investor's standpoint. Yet, according to the 2009 Investment Company Fact Book, only 13% of the assets invested in U.S. domestic equity funds by the end of 2008 were invested in passively managed funds.

My model rationalizes mutual fund investing despite the negative alphas. In equilibrium, a skilled fund manager will choose an active management policy that maximizes his expected utility while satisfying an investor's participation constraint. This policy will, however, result in the measurement of a negative alpha unless the performance measure the econometrician uses allows for a perfect specification of the pricing kernel. But as Roll (1977), Berk (1995), and Fama (1998) argue, we should not expect perfect specification to occur in empirical practice. Hence, my paper might shed some light on why, on average, actively managed U.S. equity mutual funds underperform passive investment strategies, or at least appear to, and why people keep investing in these funds.

My paper is closely related to three strands of literature, though no other paper aims at reconciling theoretically the negative unconditional performance of actively managed funds with the good performance these funds realize in bad states of the economy. Empirical papers, such as Moskowitz (2000), Kosowski (2006), and Staal (2006), show that actively managed U.S. equity mutual funds perform significantly better during bad times than good times. First, Moskowitz (2000) estimates that over the 1975–1994 period, the average return associated to stock selection by mutual fund managers was 1% higher, on an annualized basis, in recessions than in non-recessions. Second, Kosowski (2006) estimates that over the 1962–2005 period, the average annualized four-factor alpha for equity mutual funds was 4.08% in recessions and -1.33% in non-recessions. Finally, Staal (2006) finds that over the 1962–2002 period, the average fund's risk-adjusted performance was negatively correlated with the Chicago Fed National Activity Index.[3] These

---

[2] Wermers (2000) finds similar results using data on mutual fund holdings.

[3] See also Avramov and Wermers (2006), Lynch and Wachter (2007), and Mamaysky, Spiegel, and Zhang (2007) for more evidence of predictability in mutual fund performance. Lynch and Wachter (2007)

papers postulate, explicitly or not, that unconditional performance measures understate the value actively managed funds create because these funds provide good realized performance when investors' marginal utility of consumption is thought to be high. However, these papers do not study the theoretical asset pricing mechanism underlying this postulate, or the origins of the observed state dependence in performance. My paper studies both elements through a theoretical model assuming a skilled fund manager facing rational investors. It highlights the conditions required for the above postulate to be valid and argues that these conditions should hold in practice. The paper also provides new empirical evidence consistent with the model's cross-sectional predictions.

Theoretical papers, such as Admati and Ross (1985), Dybvig and Ross (1985), Grinblatt and Titman (1989), Kothari and Warner (2001), Goetzmann, Ingersoll, Spiegel, and Welch (2007), and Mamaysky, Spiegel, and Zhang (2007), analyze the effects of active portfolio management on performance measurement. These papers either do not consider the delegation of portfolio management decisions at all, or they do not consider the related idea that a portfolio manager might choose his managerial activity, and the active return that should result, based on how much investors are anticipated to value this return (which, I argue, depends on the state of the economy). Other theoretical papers, including Brennan (1993), Basak, Pavlova, and Shapiro (2007), Cuoco and Kaniel (2007), Garcia and Vanden (2009), and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2010), analyze the delegation of portfolio management decisions but do not consider the effect on the measurement of risk-adjusted performance. My paper studies simultaneously the delegation of active portfolio management decisions and its effects on performance measurement using insights from Berk and Green (2004). Unlike them, I endogenize the production of active returns by a fund manager over different states of the economy. This feature partially explains why my model, but not theirs, rationalizes mutual fund investing even though risk-adjusted performance is expected to be negative unconditionally.

It is important to note that the mechanisms at work in my model are separate from the market-timing behaviors shown in Treynor and Mazuy (1966), Henriksson and Merton (1981), Ferson and Schadt (1996), or Savov (2009). Market timing consists of changing a portfolio's risk loadings over time with the intent of profiting from changes in predicted aggregate returns. The empirical tests of state-dependent mutual fund performance by Kosowski (2006) and Staal (2006), however, control for variations in risk exposure (i.e., time-varying betas).

Hence, the state dependence in fund performance that is central in this paper should not be the consequence of what the literature usually refers to as market-timing strategies.

The paper is organized as follows. Section 2 presents a simple model of the relationship between a fund manager and his investors. Section 3 derives the active management policy the fund manager will find optimal to implement. Section 4 derives how state variations in active returns will affect the unconditional risk-adjusted performance an econometrician measures. Section 5 presents quantitative implications of a parameterized and calibrated version of the model. Section 6 presents empirical evidence about the model's cross-sectional implications and Section 7 concludes.

2. Model

I begin by describing a model of the relationship between a fund manager and his rational investors. I study a one-period economy with a finite set of states of the world $s \in S$. This period can represent any horizon over which asset management is delegated by investors and contractual terms, such as the level of fees, do not change. For brevity, I add the subscript s to a random variable only when referring to a state-specific realization of this random variable.

2.1. Mutual fund manager

The model focuses on the optimal active management policy by a fund manager, taking investors' behavior as exogenous. The novelty here is in the way I model managerial ability. I assume that the fund manager can implement, at the beginning of the period, an investment strategy that will generate state-dependent excess returns over a passive portfolio. In order to generate a positive active return, denoted a_s, during the period if state s is realized, the fund manager needs to identify an investment strategy or portfolio that will perform sufficiently well if state s is realized without performing poorly in other states. Finding such an investment strategy or portfolio imposes a non-monetary cost or disutility on the fund manager at the beginning of the period.

The model itself is agnostic about the origins of active returns. The active management technology is a reduced-form specification that captures the superior skills or investment opportunities available to the fund manager, such as an ability to identify mispriced securities that will pay abnormal returns in specific states of the world, and the idea that the fund manager might consider optimal to focus his work toward outperforming a passive portfolio in some states of the world more than others. In this paper, I investigate how the fund manager's ability to generate state-dependent active returns influences the fee he will charge and the performance an econometrician will attribute to him.

Other agents do not possess the active management technology, which I assume to be non-tradable. The fund manager owns no capital and capital requirements

prevent him from investing independently in the market. He can, however, manage the wealth of other agents and charge them a fee f that is constant across all states of the world and that represents a fraction of assets under management at the beginning of the period.[4]

Since he owns the bargaining power in his relationship with investors, the fund manager collects the value he creates through f. Berk and Green (2004) convincingly advocate this assumption. The ability to generate positive active returns is the resource in scarce supply, thus, a fund manager who possesses this ability should set f such that he collects the rewards from his active management skills and investors are indifferent about owning mutual fund shares in their personal portfolios. I assume that, when indifferent, investors behave according to the fund manager's preference. Without this assumption, the fund manager would have to set his fee marginally below the level that makes investors indifferent about owning mutual fund shares to ensure their participation, thus sharing with them an infinitely small fraction of the value he creates.

I do not consider the time-dependent mechanisms of learning and fund flows as in Berk and Green (2004) or Pástor and Stambaugh (2010), but I instead focus on the state dependence of active management policies. I normalize the value of assets under management to be one dollar at the beginning of the period. This simpler setting aims at keeping the model tractable and intuitive.

I assume that the fund's realized return contains an idiosyncratic component v, which has mean zero and is independently distributed across states of the economy. The fund manager controls the active return a and the fee f he charges but not the idiosyncratic component v.

2.2. Equilibrium condition

Here, I describe how financial markets reach an equilibrium in terms of mutual fund investing. A financial market equilibrium implies no arbitrage, which itself implies the existence of at least one positive pricing kernel that prices all tradeable assets (see, e.g., Harrison and Kreps, 1979; Hansen and Richard, 1987; Cochrane, 2001). In order to reach an equilibrium, the excess return r^i between any two assets must satisfy the following condition:

$$E[mr^i] = 0, \tag{1}$$

where m is a pricing kernel ($m > 0$). Hansen and Richard (1987) show the existence of a unique portfolio yielding a payoff x^* that can serve as a pricing kernel. If a risk-free asset also exists, the excess return on the unique portfolio will be perfectly correlated with that of any risky portfolio belonging to the mean–variance frontier. The return r^i on a risky asset or portfolio will belong to the mean–variance

frontier if and only if a pair (γ_0, γ_1) exists such that $x_s^* = \gamma_0 + \gamma_1 r_s^i$ holds in all states of the world. If r^i does not belong to the mean–variance frontier, projecting any pricing kernel m on r^i and a constant will yield non-zero error terms (see Roll, 1977).

Next, I apply the equilibrium condition in Eq. (1) to mutual fund returns rather than to stock or bond returns as is standard in the literature. The equilibrium condition still relies on the fundamental idea in asset pricing theory that investors should be "satisfied" in equilibrium with the returns an asset provides, except this time the asset is a managed portfolio.

Let R_o denote the gross risk-free rate and r^p denote the excess return on a passive portfolio or investment strategy such as buying and holding the Standard & Poor's (S&P) 500 or a mix of passive long-short portfolios like those in Carhart (1997). The passive portfolio return is $R_o + r^p$ and the fund's return is $R_o + r^p + a - f + v$. The fund's return over the passive portfolio return is $a - f + v$. In equilibrium, this excess return needs to satisfy the following condition:

$$E[m(a - f + v)] = 0. \tag{2}$$

If, instead, the left-hand side of Eq. (2) was higher than zero, then the demand for mutual fund services would be infinite and the fund manager would be able to improve his profits by increasing f marginally. Alternatively, if the left-hand side of Eq. (2) was lower than zero, then no one would invest in the mutual fund and the fund manager would collect no revenues. Hence, Eq. (2) has to hold in equilibrium.[5]

The random variable v has mean zero and is uncorrelated with the pricing kernel. From Eq. (2), the fee in equilibrium is $f = R_o E[ma]$, which represents the certainty equivalent of the value active management adds to a portfolio. This result differs from $f = E[a]$, derived by Berk and Green (2004) who do not allow active returns to vary systematically with the state of the economy. Instead, they assume that, for a given level of assets under management, volatility in realized active returns is purely idiosyncratic. As will become evident later, this difference explains why my model, but not theirs, rationalizes mutual fund investing even though expected risk-adjusted performance is negative from the point of view of an econometrician.

Note that, although I do not model fund flows explicitly here, the equilibrium condition in Eq. (2) could also be attained at the beginning of the period by keeping the fee fixed and allowing fund flows from investors to reach their optimal level as in Berk and Green (2004). In such a model, larger diseconomies of scale due to positive fund flows would bring fund returns down across all states of the world. The resulting returns would still have to be priced by investors and satisfy Eq. (2) and the intuition developed in the current model would follow.

---

[4] Golec (2003) argues that regulations from the U.S. Securities and Exchange Commission make alternative fee structures either illegal or unattractive to mutual fund companies. See also Christoffersen (2001), Golec and Starks (2004), and Kuhnen (2004) for more evidence about the popularity in the mutual fund industry of fees as fractions of assets under management.

---

[5] The equilibrium condition does not preclude the non-tradable active management technology to have a positive net present value.

2.3. Timeline and interpretation

The timeline of the model is summarized as follows. At the beginning of the period, the fund manager offers an active management policy $(f, \{a_s\}_{s \in S})$ to potential investors. Before knowing the state that will be realized, investors decide whether they commit to pay, at the end of the period, a constant fee f in exchange for the active return a_s the mutual fund will generate if state s is realized. Once an agreement has been reached, the fund manager implements, at a disutility, the investment strategy that will generate the state-specific active returns he promised to investors. The state of the economy is then realized, the mutual fund generates the state-specific active return (up to an idiosyncratic error term) and investors pay their fund manager the agreed-on fee.

In the current interpretation of the model, the fund manager picks at the beginning of the period an investment strategy that ensures that the state-dependent active return he promises to investors will be generated during the period. Although the model itself is agnostic about the origin of active returns, I now briefly suggest a possible way the fund manager could generate these returns. The fund manager could be able to acquire, at a disutility, superior information allowing him to know whether individual securities will perform abnormally well if some state of the world is realized. The fund manager would then identify cross-sections of securities likely to do abnormally well in each state of the economy and form a portfolio that would generate the state-specific active returns he promised for each state. In such a scenario, the fund manager would not have the ability to predict the state of the economy (i.e., to time the market), but he would have the ability to identify a group of mispriced securities and understand how their returns behave across different states of the economy. Naturally, as the fund manager exerts more effort to identify securities likely to do well in a given state, the expected active return the fund would produce if such state were to be realized would increase. The idiosyncratic component υ in the fund's return could then be interpreted as a mistake the fund manager makes when predicting the active return his portfolio can generate in each state.[6]

There also exists a more dynamic interpretation of the same model that could go as follows. Investors, facing frictions, delegate the management of their assets to a fund manager for a relatively long period of time. The fund manager commits to exert a state-dependent level of effort during the period, which will then generate an active return that depends on the level of effort. Throughout the period, information about the state of the economy becomes available and the fund manager adjusts how hard he tries to outperform a passive portfolio in the remainder of the period based on the anticipated state of the economy. He works harder when he anticipates some states of the economy to be realized rather than others and this state-dependent active management policy allows him to generate a state-dependent active return. This interpretation requires the fund manager to adjust his behavior based on signals disclosed throughout the period about the forthcoming state of the economy but assumes that, because of frictions or incomplete information, mutual fund investors do not use these signals as well as the fund manager does to time their investments in the fund. Although the timing is different in the two interpretations, these interpretations produce the same qualitative predictions and rely on a very similar economic intuition to do so. As will be evident later, what really matters in the model is how the fund's active return covaries with the true pricing kernel and, more specifically, how an econometrician in charge of measuring the fund's risk-adjusted performance accounts for this covariance.

3. Optimal active management

The fund manager acts in his own interests and maximizes his utility subject to an equilibrium condition, which is also the investors' participation constraint. The fund manager derives utility from consuming the fee he receives at the end of the period. However, as in Kihlstrom (1988), the fund manager also experiences disutility at the beginning of the period when exerting the effort required to identify and implement an investment strategy that will outperform the passive portfolio and deliver $\{a_s\}_{s \in S}$.

The disutility from generating a state-specific a_s does not only increase in the level of a_s but also in p_s, the probability that state s occurs. The more likely a state is to occur, the harder it should be for a fund manager to identify an investment strategy that will sustain a positive active return in that state of the world. For simplicity, I assume that the disutility function is separable and linear in probability for each state of the economy, i.e., the disutility from generating $\{a_s\}_{s \in S}$ is $\sum_{s \in S} p_s D(a_s)$, where $D(\cdot)$ is a state-independent function. For example, a fund manager willing to generate an active return of $a_s (> 0)$ in state s and zero otherwise would experience a disutility of $p_s D(a_s)$, whereas a fund manager willing to generate a constant active return of $\overline{a} (> 0)$ in every state of the world would experience a total disutility of $\sum_{s \in S} p_s D(\overline{a}) = D(\overline{a})$. The functional form assumed here will make the solution to the model simple and intuitive because the disutility from implementing an active management policy looks like an expectation ex ante.

To emphasize the timing of effort expenditure and fee collection, the fund manager's utility from consuming the fee, denoted $U(f)$, is discounted by an impatience parameter $\delta \in (0,1]$ in his objective function. The fund manager therefore offers investors a policy $(f^*, \{a_s^*\}_{s \in S})$ that maximizes the following objective function:

$$\delta U(f) - \sum_{s \in S} p_s D(a_s) \tag{3}$$

subject to the equilibrium condition: $f = R_o \sum_{s \in S} p_s m_s a_s$. The fund manager is maximizing the utility from consuming the highest fee investors accept to pay in exchange for the state-dependent active return, minus the disutility required to generate the return.

---

[6] I thank the referee for suggesting this interpretation of the active management technology.

Before deriving and analyzing the model's implications, I impose standard regularity conditions on $U(\cdot)$ and $D(\cdot)$. I assume that the utility function $U : \mathbb{R}^+ \to \mathbb{R}$ is twice-differentiable and concave and that the disutility function $D : \mathbb{R}^+ \to \mathbb{R}^+$ is twice-differentiable, strictly convex, and satisfies: $D(0)=0$, $D'(0)=0$, and $\lim_{a \to +\infty} D'(a) = +\infty$.

The following proposition derives the optimal policy $(f^*, \{a_s^*\}_{s \in S})$.

Proposition 1. *The optimal mutual fund policy $(f^*, \{a_s^*\}_{s \in S})$ satisfies*:

$$\delta U'(f^*)R_o m_s = D'(a_s^*) \tag{4}$$

in each state $s \in S$. Therefore, the optimal active return a^* is positively correlated with the pricing kernel m.

Proof. When inserting the equilibrium fee f into the fund manager's objective function (3), the fund manager's optimization problem becomes unconstrained and can be written as

$$\max_{\{a_s\}_{s \in S}} \delta U\left(R_o \sum_{s \in S} p_s m_s a_s\right) - \sum_{s \in S} p_s D(a_s). \tag{5}$$

For each state $s \in S$, the first-order condition with respect to a_s is $\delta U'(f^*)R_o p_s m_s = p_s D'(a_s)$ and is necessary and sufficient for an optimum given the assumptions made on $D(\cdot)$ and $U(\cdot)$. Canceling p_s on each side of the first-order condition yields Eq. (4).

Now define the function $H(\cdot) \equiv D'^{-1}(\cdot)$ as the inverse of the marginal disutility of generating active returns. Due to the strict convexity of $D(\cdot)$, the function $H(\cdot)$ exists and is strictly increasing over $\mathbb{R}^+$. Since $cov(m, a^*) = cov(m, H(\delta U'(f^*)R_o m))$, Lemma 1 in Appendix A implies that $cov(m, a^*) > 0$ whenever $var(m) > 0$. □

Since $\delta U'(f^*)R_o$ is positive and constant across all states the world, the optimal active return a^* is always positive and positively correlated with m. Consequently, $a^* - f^* + v$, the fund's excess return over the passive portfolio, is also positively correlated with the pricing kernel.[7] The fund manager knows that investors value more the returns realized in bad states of the economy than in good states and it is optimal for him to focus on generating the active returns that investors are willing to pay more for. This prediction is consistent with Moskowitz's (2000) finding that the average return associated with successful stock selection by U.S. fund managers is 1% higher, on an annualized basis, in recessions than in non-recessions. Note that in the dynamic interpretation/timeline I mentioned

earlier, the prediction above would also be consistent with empirical findings by Glode, Hollifield, Kacperczyk, and Kogan (2010) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2010) that fund managers appear to be more active in bad states of the economy than in good states.

Using Eq. (2), I decompose the fee in the following way:

$$f^* = E[a^*] + R_o cov(m, a^*). \tag{6}$$

The fund manager is not only compensated for the level of active returns he produces but also for their covariance with the pricing kernel. Hence, my model suggests a novel source of cross-sectional differences in mutual fund fees (see Chordia, 1996; Christoffersen and Musto, 2002).

Similarly, the fund's expected excess return over the passive portfolio can be written as

$$E[a^* - f^* + v] = -R_o cov(m, a^*) \tag{7}$$

and is negative. Partially insuring investors against variations in the pricing kernel allows the fund manager to request a compensation that is higher than the active return he is expected to generate.

Note that these results do not rely on a specific parameterization of the pricing kernel or equivalently of the investors' utility function. The only assumption imposed on m is that the realized pricing kernel m_s is higher in bad states of the economy than in good states, similar to what most consumption-based models with risk aversion would predict.

4. Measuring the fund's unconditional performance

A fund's expected excess return over a passive portfolio return, as derived in Eq. (7), is not a valid measure of abnormal performance because it does not adjust for the fund's risk. Ideally, a fund's risk-adjusted performance would be measured by the fund's realized excess return, net of fees, minus a risk premium for the covariance between the fund's return and a pricing kernel. An econometrician, however, is unlikely to observe a true pricing kernel m and use it to measure fund performance. Instead, he proxies for it using $\hat{m} \equiv E[m|I]$, where I is the information available to him when trying to measure fund performance. This information set I is based on a coarser partition of the state space than the information sets of the mutual fund manager and his investors. By construction, the specification error ε ($\equiv m - \hat{m}$) satisfies $E[\varepsilon | \hat{m}] = 0$ for all values of $\hat{m}$. In other words, ε is mean independent of the pricing kernel proxy $\hat{m}$ and the pricing kernel proxy $\hat{m}$ is unbiased given the econometrician's information set. In Appendix B, I relax the assumption of mean independence of ε to $\hat{m}$ and replace it by the weaker assumption that ε is uncorrelated with $\hat{m}$. An example for the specification error ε would be an orthogonal labor income component that is omitted in the pricing kernel proxy. If the performance measure the econometrician uses relies on a pricing kernel proxy with non-zero error terms (i.e., $var(\varepsilon) > 0$), I characterize this performance measure as being *misspecified*. I assume (unsurprisingly) that the performance measure based on $\hat{m}$ assigns no abnormal

---

[7] Intuitively straightforward predictions would arise in the less frequent, yet possible, scenario of a fund manager being compensated through a performance-based fee of $f_0 + f_1 \cdot a$. The first-order condition for the fund manager's objective function would then be $(1 - f_1)$ times a first-order condition similar to that in Eq. (4) and f_1 times a first-order condition similar to that from a scenario where the fund manager simply consumes a fraction of portfolio returns. The fund manager would still have the incentive to implement an investment strategy that insures investors against bad states of the economy, but this incentive would be partly mitigated, if the fund manager were risk averse, by a willingness to reduce volatility in the revenues he consumes.

performance to passive investment strategies, including the passive strategy producing r^p. This condition is satisfied, for example, when the econometrician proxies for the true pricing kernel using a linear combination of passive returns, including r^p, as in Jensen (1968), Carhart (1997), Fama and French (2010), and Barras, Scaillet, and Wermers (2010), among others.

The following proposition derives the unconditional risk-adjusted performance $E[\alpha]$ that an econometrician using the pricing kernel proxy $\hat{m}$ is expected to attribute to a fund manager who generates any set of active returns $\{a_s\}_{s \in S}$ and who charges $f = R_o E[ma]$. The fund's expected measured performance $E[\alpha]$ is given by the fund's expected excess return over the risk-free rate minus the risk premium required given how the fund's return covaries with the pricing kernel proxy $\hat{m}$.

Proposition 2. *The expected risk-adjusted performance of the fund, as measured by the econometrician, is:*

$$E[\alpha] = -R_o cov(\varepsilon, a), \tag{8}$$

where ε denotes the specification error associated with $\hat{m}$, i.e., $m = \hat{m} + \varepsilon$. $E[\alpha]$ is negative whenever $var(\varepsilon) > 0$.

Proof. To measure $E[\alpha]$, I subtract from the fund's expected excess return over the risk-free rate the risk premium required given how the fund's return covaries with the pricing kernel proxy:

$$
\begin{aligned}
E[\alpha] &= E[r^p + a - f + v] + R_o cov(\hat{m}, r^p + a - f + v) \\
&= E[r^p] - R_o cov(m, a) + R_o cov(\hat{m}, r^p) + R_o cov(\hat{m}, a) \\
&= E[r^p] + R_o cov(\hat{m}, r^p) - R_o cov(\varepsilon, a). \tag{9}
\end{aligned}
$$

By assumption, the passive portfolio producing r^p is priced correctly by $\hat{m}$. Thus, $E[r^p] = -R_o cov(\hat{m}, r^p)$, canceling the first two terms in Eq. (9) and yielding the solution for $E[\alpha]$.

The covariance between ε and a can then be decomposed into:

$$cov(\varepsilon, a) = E[cov(\varepsilon, a | \hat{m})] + cov(E[\varepsilon | \hat{m}], E[a | \hat{m}]). \tag{10}$$

The results in Proposition 1 and Lemma 1 imply that $cov(\varepsilon, a^* | \hat{m})$ is positive whenever $var(\varepsilon) > 0$. Therefore, the first term on the right-hand side of the decomposition above is positive. The second term equals zero because the econometrician proxies for m using $\hat{m} \equiv E[m | I]$, which implies that $E[\varepsilon | \hat{m}] = 0$ for any possible level of $\hat{m}$. Hence, if $var(\varepsilon) > 0$, the covariance between ε and a^* is positive and $E[\alpha]$ is negative. □

The assumption that the return r^p is perfectly priced by the pricing kernel $\hat{m}$ ensures that the passive portfolio, which is exogenously given in my model, does not affect the measurement of the fund's abnormal performance. The expected measured performance is then equal to the risk premium that would be required if the specification error ε were a pricing kernel. The proposition also shows that, unless the performance measure is perfectly specified, the covariance between the specification error ε and the active return a^* is positive and the fund's expected risk-adjusted

performance $E[\alpha]$ is negative in the model, consistent with empirical findings by Jensen (1968), Malkiel (1995) and Fama and French (2010), among many others.

Taken together, results in Sections 3 and 4 imply that the *skilled* fund manager in my model will appear to perform poorly unconditionally. This prediction is completely opposite to what several financial economists have assumed in the past. It has been argued that fund managers, if skilled, should provide positive risk-adjusted performance to investors. Berk and Green (2004) instead argue that fund managers own the bargaining power in their relationship with investors and should collect the rewards from their active management skills, resulting in all funds being expected to generate zero risk-adjusted performance, net of fees. My paper extends the analysis by allowing for the costly production of state-specific active returns. The result is a simple rational explanation for the negative performance of mutual funds. In my model, a fund manager who can generate, at a disutility, an active return that is specific to one state of the economy will receive from investors a compensation that increases with the pricing kernel realization in that state. Therefore, the fund manager will find it optimal to generate active returns that covary positively with the pricing kernel, providing investors with (partial) insurance against bad states of the economy. An econometrician trying to evaluate the fund's risk-adjusted performance is, however, likely to use a performance measure that allows for a specification error. The misspecified performance measure will account for the covariance between the fund's return and the pricing kernel proxy but not for the covariance between the fund's return and the specification error. Consequently, the performance measure will be negatively biased and the fund manager in my model will appear to destroy value, even though the returns he generates are priced correctly in equilibrium by investors and the pricing kernel proxy is unbiased from the econometrician's point of view.

Negative expected performance arises in equilibrium because active returns covary positively with a component of the pricing kernel that the econometrician omits when measuring performance. If, instead, a fund manager were to promise active returns that did not covary with the true pricing kernel, my model would predict the same measured unconditional performance as in Berk and Green (2004), i.e., $E[\alpha] = 0$. Alternatively, if the econometrician were to perfectly account for the covariance between the active return and the true pricing kernel, the value the fund manager creates would be entirely captured by the performance measure and my model would, again, predict $E[\alpha] = 0$.

Note that Eq. (8) does not depend on derivations from Section 3. Hence, one could propose an alternative explanation for the funds' higher active returns realized in bad states of the economy than in good states and use Proposition 2 to rationalize the funds' negative unconditional performance.[8] Moreover, one could study the

---

[8] For example, Kacperczyk, Van Nieuwerburgh, and Veldkamp (2010) argue theoretically and find empirical support for the idea that portfolio managers should hold more distinct portfolios in high aggregate volatility states.

returns of a different asset class and use Proposition 2 to derive how state dependence in these returns might affect the unconditional performance an econometrician measures. When returns on an asset insure investors against bad states of the economy and are priced correctly in equilibrium by these investors, whatever the reason or asset, my paper shows that a misspecification in the pricing kernel proxy will lead to the measurement of negative unconditional performance.

5. Parameterization

In this section, I parameterize the model and derive explicit expressions for the fund's optimal active returns and the performance an econometrician will measure. Then, I calibrate the model to the U.S. economy and test whether its predictions are quantitatively sensible.

I assume that the disutility of generating an active return takes a quadratic form: $D(a) = (\theta/2)a^2$, for $a \geq 0$, where $\theta > 0$. Therefore, $D'(a) = \theta a$ and θ represents the slope of the marginal disutility function. The disutility of producing an active return increases with θ. The first-order condition in Eq. (4) becomes $a_s^* = (\delta U'(f^*)/\theta)R_o m_s$. The only way the manager's utility function $U(\cdot)$ appears in the model's predictions is through $\delta U'(f^*)/\theta$, which is positive and constant across all states of the world. Parameterizing and calibrating $U'(\cdot)$ will have no quantitative implication without a choice of δ/θ. The positive constant $\delta U'(f^*)/\theta$ can therefore be calibrated without deeper consideration for the values of δ, $U'(f^*)$, and θ, per se.

I assume the passive portfolio return r^p is equal to $\sum_{n=1}^{N} \omega_n r^n$, where each ω_n represents the exposure of the passive portfolio to a passive excess return (or factor) r^n and the econometrician uses the linear projection of m on the set of $\{r^n\}_{n=1}^{N}$ and a constant as his pricing kernel proxy ($\hat{m} = \gamma_0 + \sum_{n=1}^{N} \gamma_n r^n$). This last assumption yields a performance measure that is linear in the passive returns $\{r^n\}_{n=1}^{N}$ as in Carhart (1997). This specification ensures that the econometrician correctly prices the passive returns $\{r^n\}_{n=1}^{N}$ and r^p and that variations in these returns do not affect the measurement of the fund's abnormal performance.

5.1. Parameterized implications

The following proposition shows the fee a fund manager with a quadratic disutility function will charge and the performance an econometrician using a linear function of r^n as his pricing kernel proxy will measure.

Proposition 3. *In the parameterization, the fund's equilibrium fee satisfies:*

$$f^* = \frac{\delta U'(f^*)}{\theta}[1 + R_o^2 var(m)] \tag{11}$$

and the fund's expected risk-adjusted performance is

$$E[\alpha] = -\frac{\delta U'(f^*)}{\theta} R_o^2 var(\varepsilon). \tag{12}$$

$E[\alpha]$ is negative whenever $var(\varepsilon) > 0$.

Proof. Inserting $a_s^* = (\delta U'(f^*)/\theta)R_o m_s$ in Eqs. (6) and (8) yields the results. $\quad\square$

The model predicts that expected alpha is negative unless $var(\varepsilon) = 0$; any error in the pricing kernel proxy is expected to lead to the measurement of negative risk-adjusted performance. The unobservable $var(\varepsilon)$ represents the degree of misspecification in the econometrician's pricing kernel proxy (see Hansen and Jagannathan, 1997; Hodrick and Zhang, 2001). To have $var(\varepsilon) = 0$, the pricing kernel proxy must be perfectly specified. For example, when $N = 1$, realizations of ε can equal zero in all states of the world only if the passive excess return $r^p = \omega_1 r^1$ belongs to the mean–variance frontier. Otherwise, $var(\varepsilon)$ is positive, even if markets are complete. But as Roll (1977), Berk (1995), and Fama (1998) argue, we should not expect this situation to occur in empirical practice. Specifically, empirical findings by Bekaert and Hodrick (1992) suggest that $var(\varepsilon)$ is positive when an econometrician only uses domestic portfolios to proxy for the pricing kernel. They compute the maximal Sharpe ratio (SR) attainable with conditional trading strategies in international markets and find that international investing increases the Hansen and Jagannathan (1991) lower bound on $\sqrt{var(m)}$. Hence, $\sqrt{var(m)}$ has to be higher than the lower bound that domestic investing implies. When investors have access to financial instruments that allow for a better diversification than what the portfolio used in the pricing kernel proxy provides, then $var(\varepsilon)$ is positive.

Applying the implicit function theorem to Eq. (11) shows that, if the fund manager is risk-averse or risk-neutral around f^*, an increase in pricing kernel volatility will lead to active management generating more valuable insurance, and consequently, the fund manager charging a higher fee. Similarly, the model, when applied to a cross-section of funds, rationalizes why funds with poor unconditional performance charge high fees compared to other funds. All else being equal, fund managers with lower disutility parameters will implement investment strategies that insure investors better against pricing kernel variations. These fund managers will be able to collect higher fees from investors, but they will also appear to perform worse unconditionally when the performance measure the econometrician uses is misspecified. Hence, my model predicts that f^* and $E[\alpha]$ will move in opposite directions in a cross-section of funds, consistent with findings by Malkiel (1995), Gruber (1996), and Carhart (1997), among others. The paper offers a new channel to rationalize these cross-sectional variations through the prediction that realized risk-adjusted performance will move with the state of the economy, consistent with Kosowski (2006) and Staal (2006).

Proposition 4. *In the parameterization, the fund's realized risk-adjusted return in a given state s is*

$$\alpha_s = E[\alpha] + \frac{\delta U'(f^*)}{\theta} R_o \varepsilon_s + \upsilon_s. \tag{13}$$

The covariance between this return and the pricing kernel is positive whenever $var(\varepsilon) > 0$.

Proof. Using Eq. (6), the fund's excess return over the risk-free rate can be written as

$$r^p + a^* - f^* + v = r^p + a^* - E[a^*] - R_o cov(m, a^*) + v. \quad (14)$$

Inserting the parameterized result $a_s^* = (\delta U'(f^*)/\theta) R_o m_s$ and the parameterized pricing kernel proxy $\hat{m} = \gamma_0 + \sum_{n=1}^{N} \gamma_n r^n$ in the last equation yields

$$r^p + \frac{\delta U'(f^*)}{\theta} R_o \sum_{n=1}^{N} \gamma_n [r^n - E[r^n] - R_o cov(m, r^n)]$$
$$+ \frac{\delta U'(f^*)}{\theta} R_o \varepsilon - \frac{\delta U'(f^*)}{\theta} R_o^2 var(\varepsilon) + v. \quad (15)$$

Then, using $r^p = \sum_{n=1}^{N} \omega_n r^n$ to expand the first term, $E[r^n] = -R_o cov(m, r^n)$ to cancel the second and third terms within the summation's brackets, and $E[\alpha] = -(\delta U'(f^*)/\theta) R_o^2 var(\varepsilon)$ to replace the second-to-last term, the fund's excess return becomes

$$\sum_{n=1}^{N} \left[\omega_n + \frac{\delta U'(f^*)}{\theta} R_o \gamma_n \right] r^n + \frac{\delta U'(f^*)}{\theta} R_o \varepsilon + E[\alpha] + v. \quad (16)$$

When estimating the fund's risk-adjusted returns, the econometrician subtracts variations in fund returns that are due to the fund's exposure to passive returns. Here, the exposure of the fund's return on each passive return r^n is given by $[\omega_n + (\delta U'(f^*)/\theta) R_o \gamma_n]$ and is constant across any subset of states of the world. The econometrician therefore subtracts $\sum_{n=1}^{N} [\omega_n + (\delta U'(f^*)/\theta) R_o \gamma_n] r^n$ from the fund's excess return when adjusting for risk and the risk-adjusted return he measures in state s becomes

$$\alpha_s = E[\alpha] + \frac{\delta U'(f^*)}{\theta} R_o \varepsilon_s + v_s. \quad (17)$$

The covariance between α and any r^n is equal to zero and the covariance between α and m becomes the covariance between α and ε, which is $[(\delta U'(f^*)/\theta) R_o] var(\varepsilon)$. Consequently, the fund's realized risk-adjusted return is positively correlated with the pricing kernel whenever $var(\varepsilon) > 0$. $\square$

Proposition 4 shows that the level of returns $\{r^n\}_{n=1}^{N}$ and the passive portfolio's loadings $\{\omega_n\}_{n=1}^{N}$ do not affect the realized risk-adjusted performance an econometrician measures in a given subset of states of the economy. The slope coefficient in the regression of the fund's return on all the passive returns $\{r^n\}_{n=1}^{N}$ and a constant perfectly accounts for the effect of $\{r^n\}_{n=1}^{N}$ on fund returns. Hence, if the performance measure is perfectly specified, only the idiosyncratic term v is considered abnormal. However, if the performance measure is misspecified, realized risk-adjusted performance also moves with the omitted pricing kernel component ε. The second term in Eq. (13) is the fund's exposure to variations in the pricing kernel the econometrician omits and the first term is the negative risk premium associated with that omission. Consequently, the fund's risk-adjusted performance is positively correlated with the true pricing kernel, due to the omission of ε. Since the omitted pricing kernel component ε is independent from the pricing kernel proxy $\hat{m}$, it is more likely to be high in bad states of the economy than in good states. When comparing measured fund performance in recessions and non-recessions, as in

Moskowitz (2000), Kosowski (2006), and Staal (2006), the econometrician captures how mutual fund returns move with pricing kernel variations omitted by unconditional performance measures but captured, at least partially, by the aggregation of variables used in characterizing business cycles (e.g., real output or employment rates).

Proposition 4 also shows that the sensitivity of realized risk-adjusted returns to ε, and consequently to the pricing kernel m, increases with the ratio $\delta U'(f^*)/\theta$. Therefore, in a cross-section of funds, the funds charging high fees and generating low $E[\alpha]$ should also be those that provide investors with the best insurance against bad states of the economy and that exhibit more sensitivity of realized risk-adjusted returns to the state of the economy. In Section 6, I investigate whether cross-sectional differences in the state dependence of funds' realized risk-adjusted returns help to rationalize the negative relationship between fees and alphas that is observed in the data.

5.2. Calibration

I now verify whether my model, when calibrated to the U.S. economy, can reproduce quantitatively the measured underperformance of U.S. funds. I assume a representative fund manager who behaves as my model predicts and I use data from the Center for Research in Security Prices (CRSP) mutual fund database on 3,147 actively managed U.S. equity funds over the 1980–2005 period to calibrate the model. Appendix C provides a detailed description of the sample and Table 1 reports summary statistics for the main fund attributes.

Expressions $\delta U'(f^*)/\theta$ and $var(\varepsilon)$ are not directly observable in reality. Before addressing these unobservables, I tie down the moments that can be inferred directly from economic data. To simplify the calibration exercise, I set $N = 1$ and use the excess return on the value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks in the CRSP database as the return r^1 the passive portfolio loads on. The parameterized performance measure becomes a one-factor alpha based on a proxy for the market portfolio, as in Jensen (1968). In the projection $m = \gamma_0 + \gamma_1 r^1 + \varepsilon$, the coefficients become $\gamma_1 = -SR/[R_o \sqrt{var(r^1)}]$ and $\gamma_0 = 1/[R_o([1 + SR^2])]$, where SR denotes the Sharpe ratio of the passive strategy

Table 1
Summary statistics.

This table presents summary statistics about the main fund attributes for the sample of 3,147 actively managed U.S. equity mutual funds over the 1980–2005 period.

	Mean	S.D.	Median	25%	75%
Asset size ($M)	915.77	3669.51	159.93	47.78	551.54
Age (years)	13.15	14.30	8.00	4.00	16.00
Turnover (%, per year)	91.92	118.19	67.40	36.00	115.00
Expense ratio (%, per year)	1.30	0.49	1.23	0.98	1.54
Front-load fee (%)	1.72	2.37	0.00	0.00	3.47
Back-load fee (%)	0.53	0.97	0.00	0.00	0.85
Raw return (%, per quarter)	2.58	10.49	3.11	−2.31	8.65

Table 2

Moment values for model calibration.

This table presents moment values used to calibrate the model on an annual basis to the U.S. economy over the 1980–2005 period. Mutual fund fee is computed using the average of: expense ratio+(1/7)∗ front-load fee. Moment values for the passive return r^1 are calibrated based on the CRSP value-weighted portfolio of all NYSE, Amex, and Nasdaq stocks.

	Symbol	Value (%)
Mutual fund fee	f^*	1.55
Net risk-free rate	$R_o - 1$	5.97
Mean passive excess return	$E[r^1]$	7.99
Std. dev. of r^1	$\sqrt{var(r^1)}$	16.17

($SR = E[r^1]/\sqrt{var(r^1)}$). The passive portfolio's loading on r^1, denoted ω_1, need not be calibrated here because it does not enter the expressions for optimal fee, optimal active return, or measured risk-adjusted performance. The only place where ω_1 matters in the model is when measuring the fund's exposure to the passive return r^1. This exposure is given by $\omega_1 + cov(a^*,r^1)/var(r^1)$, which is smaller than ω_1 because the fund manager is in fact providing an insurance against bad states of the economy, including periods of low returns, on top of the passive portfolio generating $r^p = \omega_1 r^1$.

I calibrate the first two moments of r^1 and the mean risk-free rate R_o using data for the 1980–2005 period from Kenneth French's Web site. The mean annual expense ratio for my sample of funds is 1.30%. The expense ratio is the fraction of total investment that shareholders pay annually for the fund's operating expenses, including 12b-1 fees. To account for amortized loads, which are not included in the reported expense ratio, I calibrate the fee f^* by adding 1/7 of the average front-load fee to the average expense ratio as in Sirri and Tufano (1998) and Barber, Odean, and Zheng (2005).[9] Table 2 summarizes the calibration of the model's observable moments.

With the observable moments calibrated, I investigate the values for $\delta U'(f^*)/\theta$ and $var(\varepsilon)$ that allow the model's predictions to match selected empirical estimates. I solve for the relationship between $\delta U'(f^*)/\theta$ and $var(\varepsilon)$ that allows to match the average mutual fund fee in my sample. Then I compare the range of $var(\varepsilon)$ that produces reasonable predictions for $E[\alpha]$ in the current calibration to several empirical estimates of $var(\varepsilon)$.

To satisfy Eq. (11), $\delta U'(f^*)/\theta$ needs to satisfy

$$\frac{\delta U'(f^*)}{\theta} = \frac{f^*}{1 + R_o^2 var(m)} = \frac{f^*}{1 + SR^2 + R_o^2 var(\varepsilon)}. \quad (18)$$

If it does, the expected risk-adjusted performance of mutual funds, for a given fee f^*, is

$$E[\alpha] = -\frac{R_o^2 f^*}{1 + SR^2 + R_o^2 var(\varepsilon)} var(\varepsilon). \quad (19)$$

Fig. 1 plots the relationship between $E[\alpha]$ and $var(\varepsilon)$ under the calibration of Table 2. The figure also identifies

---

[9] Barber, Odean, and Zheng (2005) provide two reasons for not including back-load fees into the computation of the total fee: back-load fees were not reported in the CRSP database prior to 1993 and are often waived if an investor holds a fund for a specific period of time.

the predicted $E[\alpha]$ for several levels of $var(\varepsilon)$ that empirical estimates of $var(m)$ would imply (see Bekaert and Hodrick, 1992; Chapman, 1997; Melino and Yang, 2003; Bansal and Yaron, 2004; Kan and Zhou, 2006). I use $var(m) = \gamma_1^2 var(r^1) + var(\varepsilon) = (SR^2/R_o^2) + var(\varepsilon)$, where SR denotes the Sharpe ratio of the passive return r^1, to derive the $var(\varepsilon)$ each empirical estimate of $var(m)$ implies.

The higher estimates of $var(m)$ from Chapman (1997) and Kan and Zhou (2006) allow my model to generate a one-factor alpha close to what Jensen (1968), Malkiel (1995), and Gruber (1996) report (i.e., around −1% per year). In a calibration where expected alpha is −1%, a two-standard-deviation shock in m results in a change of 1.40% in the (annual) active return a^* and a two-standard-deviation shock in ε results in a change of 1.33% in the (annual) risk-adjusted realized return α. If, instead, I use the lower estimates of pricing kernel volatility from Bekaert and Hodrick (1992), Melino and Yang (2003), and Bansal and Yaron (2004), the model still generates levels of underperformance that are economically significant. Hence, in light of this calibration, one should not be puzzled to observe significantly negative alphas for actively managed U.S. equity mutual funds given the insurance they provide against bad states of the economy.

6. Cross-sectional implications

This paper rationalizes mutual fund investing while linking together the negative unconditional risk-adjusted performance of actively managed U.S. equity funds and the funds' significantly better performance realized in bad states of the economy than in good states. These facts have been studied in aggregate for the mutual fund industry, but in this section I show new empirical evidence that can also help us understand the cross-section of mutual fund fees and returns. I use the same sample of funds used for the calibration (see Appendix C for the sample description).

My model predicts that the covariance between a fund's active return and the specification error in the econometrician's pricing kernel proxy should be negatively related with the fund's apparent unconditional performance. The parameterized version of the model predicts that, relative to other funds, funds with poor unconditional performance should charge high fees and exhibit more state dependence in realized risk-adjusted performance. How sensitive risk-adjusted returns are to the state of the economy (the term $(\delta U'(f^*)/\theta)R_o$ in the model) should help us understand the documented cross-sectional variations in mutual fund fees and unconditional performance. However, unlike the fund manager in the model, some "real life" managers might not be skilled enough and might waste investors' money by performing poorly in all states of the economy. Nonetheless, I investigate whether the documented state dependence in aggregate fund performance is mostly driven by funds with poor unconditional performance, as the comparative statics of my model would suggest, and whether insurance against bad states of the economy might partially explain the survival of some poorly performing funds.
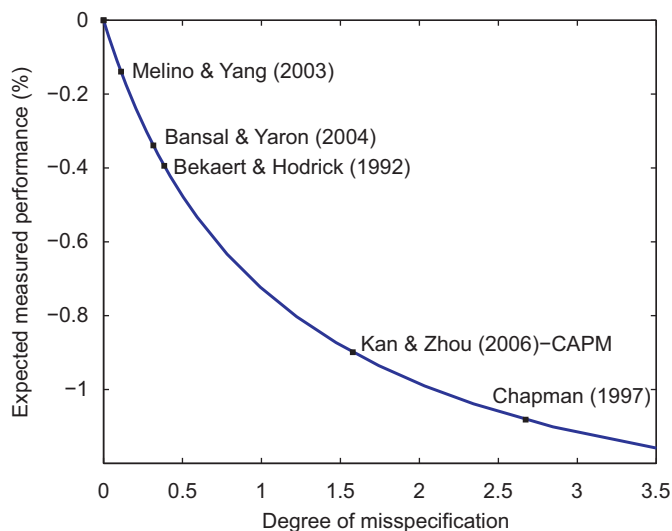
Fig. 1. Misspecification and measured fund performance. This figure plots the relationship between the expected measured performance $E[\alpha]$ and the degree of misspecification $var(\varepsilon)$ in the pricing kernel proxy when the model is calibrated on an annual basis to the U.S. economy over the 1980–2005 period (see Table 2). The figure also identifies the $E[\alpha]$ associated with each level of $var(\varepsilon)$ implied by the empirical estimates of $var(m)$ reported by Bekaert and Hodrick (1992), Chapman (1997), Melino and Yang (2003), Bansal and Yaron (2004), and Kan and Zhou (2006) (based on the CAPM).

As in Moskowitz (2000), Kosowski (2006), and Staal (2006), I use NBER recessions as a proxy of bad states of the economy (periods during which the pricing kernel is high). For each of the 2,075 funds in my sample that went through at least one NBER recession during the 1980–2005 period, I estimate the fund's average unconditional alpha, average expense ratio, and average total fee, that is, expense ratio+(1/7)*front-load fee as in Sirri and Tufano (1998) and Barber, Odean, and Zheng (2005), over the entire sample period using monthly data. I classify all funds into decile portfolios based on their average unconditional alpha during the sample period. Table 3 presents the mean alpha, expense ratio, and total fee for each decile portfolio. Then, for each fund, I run a time-series regression of the fund's excess returns on the NBER recession indicator, the risk factors, and the cross-products of the recession indicator and the factors. Even though my model abstracts from time variations in the fund exposure, I follow Kosowski (2006) and Staal (2006) and allow factor loadings to change with the state of the economy to ensure that empirical results are not driven by changes in the risk exposure of mutual funds, as a fund flow story along the lines of Ferson and Warther (1996) or a market-timing story along the lines of Treynor and Mazuy (1966), Henriksson and Merton (1981), or Ferson and Schadt (1996) could suggest. A fund's ordinary least-square coefficient associated to the recession indicator measures the difference between the fund's average risk-adjusted return realized in recessions and that realized in non-recessions. As long as business cycle variables capture some pricing kernel variations omitted by the performance measures used (e.g., a labor income component), movements in measured fund performance across the business cycle should proxy for the exposure of fund returns to ε. Hence, the difference between a fund's

average risk-adjusted returns realized in recessions and in non-recessions allows to proxy for the sensitivity of the fund's realized risk-adjusted return to the pricing kernel, which in the model is equal to its sensitivity to the specification error ε. The difference is predicted to be positively related in the cross-section of funds with the fund's fee and negatively related with the fund's unconditional risk-adjusted performance. Panels A, B, and C show results when unconditional and state-dependent performance are computed using the one-factor model of Jensen (1968), the three-factor model of Fama and French (1993), and the four-factor model of Carhart (1997), respectively.

Whether I adjust for front-load fees or not, the difference between the average fee of the first decile portfolio and that of the tenth decile portfolio is economically and statistically significant. Funds with poor unconditional performance charge high fees relative to other funds, consistent with empirical findings by Malkiel (1995), Gruber (1996), and Carhart (1997), among others. As Carhart (1997) shows, the empirical relationship between fees and alphas is, however, not strictly decreasing. In my sample, it is strictly decreasing for deciles with negative alphas but not for deciles with positive alphas. Table 3 shows that funds with poor unconditional alphas generate risk-adjusted returns that are, on average, more sensitive to the state of the economy than funds with good unconditional alphas. Differences between deciles 1 and 10 as well as between deciles 1–2 and 9–10 are, for the three performance measures, statistically significant at the 5% level, even though deciles 1–2 are likely to include some unskilled fund managers who simply perform poorly in all states of the economy. The empirical relationship studied here is again not strictly monotone. Yet, these empirical results suggest that, consistent with my model, the countercyclicality of performance is

Table 3
Unconditional fund performance and state dependence.

This table presents the mean unconditional risk-adjusted performance (alpha), annual expense ratio, total annual fee, and a proxy for the dependence of performance to the state of the economy for ten decile portfolios of mutual funds sorted by unconditional alpha. State dependence in performance is proxied by a fund's average risk-adjusted return in NBER recessions minus the fund's average risk-adjusted return in non-recessions, as computed over the entire sample period. Panels A, B, and C show results when risk-adjusted performance is computed using Jensen's (1968) one-factor model, Fama and French's (1993) three-factor model, and Carhart's (1997) four-factor model, respectively. I use monthly data for the 2,075 actively managed U.S. equity mutual funds in my sample that went through at least one recession over the 1980–2005 period to compute alpha and the level of state dependence in performance of each fund over the entire sample period. Total annual fee is computed using: annual expense ratio+(1/7)∗front-load fee. Estimates are in % terms. The differences between the averages of deciles 1 and 10 are reported with their standard errors. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Decile (alpha)	Alpha (%, per month)	Expenses (%)	Total fee (%)	State dep. (%, per month)
Panel A: One-factor model				
1	− 0.89	1.72	1.96	0.53
2	− 0.38	1.43	1.74	0.18
3	− 0.23	1.35	1.62	− 0.02
4	− 0.13	1.34	1.61	− 0.02
5	− 0.05	1.23	1.47	0.04
6	0.03	1.22	1.46	− 0.03
7	0.13	1.27	1.54	− 0.01
8	0.26	1.31	1.54	0.13
9	0.46	1.35	1.51	0.04
10	1.07	1.49	1.67	− 0.11
1 minus 10	− 1.96	0.23	0.29	0.64
	[0.05]***	[0.05]***	[0.06]***	[0.28]**
Panel B: Three-factor model				
1	− 0.87	1.70	1.97	0.32
2	− 0.37	1.40	1.65	0.19
3	− 0.24	1.35	1.59	− 0.02
4	− 0.17	1.32	1.55	0.01
5	− 0.10	1.28	1.52	− 0.07
6	− 0.04	1.28	1.54	− 0.09
7	0.03	1.28	1.51	− 0.21
8	0.11	1.25	1.50	− 0.17
9	0.25	1.37	1.61	− 0.36
10	2.02	1.47	1.67	− 0.47
1 minus 10	− 2.89	0.23	0.30	0.79
	[1.26]**	[0.05]***	[0.06]***	[0.23]***
Panel C: Four-factor model				
1	− 1.19	1.70	1.93	0.81
2	− 0.37	1.44	1.73	0.34
3	− 0.24	1.37	1.63	− 0.01
4	− 0.17	1.37	1.61	− 0.03
5	− 0.11	1.27	1.49	0.02
6	− 0.05	1.26	1.51	− 0.06
7	0.01	1.29	1.54	− 0.27
8	0.10	1.23	1.45	− 0.32
9	0.22	1.29	1.54	− 0.26
10	0.67	1.49	1.70	− 0.07
1 minus 10	− 1.86	0.21	0.23	0.88
	[0.21]***	[0.05]***	[0.06]***	[0.43]**

stronger for funds with poor unconditional alphas and high fees.

In unreported tests, I investigate the possibility that the observed insurance that funds provide might be driven by time variations in the level of mutual fund fees rather than by a state dependence in the active return that fund managers generate. Consistent with Kuhnen (2004), I find very little time-series variations in funds' fees (e.g., less than a two basis points difference in average annual expense ratios, net of waivers and reimbursements, between recessions and non-recessions). This finding suggests that state variations in funds' realized returns are mostly driven by state variations in active returns rather than in fees. From a theoretical perspective, this finding also suggests that the mutual fund returns we observe in a given period are mostly driven by ex post realizations of the state of the economy rather than by variations in the ex ante empirical moments used at the time of the fee-setting decision, like the pricing kernel volatility anticipated for the (potentially long) period during which the fee is fixed.[10]

The findings I document for the cross-section of funds are, to the best of my knowledge, novel to the literature and give credit to the theory I develop in the paper. When an econometrician uses a misspecified performance measure, the unconditional risk-adjusted performance of an asset priced correctly in equilibrium should be negatively related to the level of insurance against pricing kernel variations this asset provides. Consequently, observing simultaneously negative unconditional performance and abnormally good performance in bad states of the economy, for any asset class, can be the result of a misspecification in the performance measure rather than the result of an irrational mispricing for the whole asset class. The sensitivity of measured performance to proxies of bad states of the economy could therefore be used to measure the degree of misspecification in a given asset pricing model, a role similar to that played by firm size in Berk (1995).

7. Conclusion

In this paper, I derive a parsimonious model that reproduces the negative unconditional risk-adjusted performance of actively managed U.S. equity mutual funds and the funds' significantly better realized performance in bad states of the economy than in good states. The model focuses on the optimal active management policy of a fund manager able to generate active returns that depend on the state of the economy. Facing rational investors, the fund manager will optimally focus his work toward realizing good performance in bad states of the economy, when investors' marginal utility of consumption is high. He will therefore generate active returns that covary positively with the pricing kernel. My model shows that a performance measure that does not allow for a perfect specification of the pricing kernel will underestimate the value active management creates when the active return that results is positively correlated

---

[10] In unreported regressions, I also investigate whether the cross-sectional differences in the insurance funds provide can be driven by a fund size effect. I find that, on average, funds larger than the median and funds smaller than the median, in a given quarter, produce similar insurance against bad states of the economy in the following quarter.

with the pricing kernel. Consequently, the skilled fund manager in my model will (wrongly) appear to under-perform passive investment strategies net of fees, yet mutual fund investing will be rational.

I calibrate the model to the U.S. economy and repro-duce quantitatively the measured underperformance of U.S. funds. I also use data on 3,147 actively managed U.S. equity mutual funds over the 1980–2005 period and find new empirical evidence consistent with the model's cross-sectional predictions. Relative to other funds, funds with poor unconditional performance tend to charge high fees and generate highly countercyclical risk-adjusted returns. This finding might explain the survival of several funds with poor unconditional performance and suggest the existence of a recession-related misspecification in popular performance measures.

Appendix A. Covariance between a random variable and its strictly increasing transformation

Here, I present a lemma needed to derive and analyze the model's predictions.

Lemma 1. *Let z be a random variable with $var(z) > 0$ and $G : \mathbb{R} \rightarrow \mathbb{R}$ be a strictly increasing function. Then $cov(z,G(z)) > 0$.*

Proof.

$$
\begin{aligned}
cov(z,G(z)) &= E[(z-E[z])(G(z)-E[G(z)])] \\
&= E[(z-E[z])(G(z)-G(E[z]))] \\
&\quad + E[(z-E[z])(G(E[z])-E[G(z)])] \\
&= E[(z-E[z])(G(z)-G(E[z]))].
\end{aligned} \tag{20}
$$

Since $G(\cdot)$ is strictly increasing, it follows that $cov(z,G(z)) > 0$ whenever $var(z) > 0$.   □

Appendix B. Relaxing the assumption of mean independence

Here, I relax the assumption that ε is mean independent from the level of $\hat{m}$ (i.e., $E[\varepsilon|\hat{m}] = 0$ for all levels of $\hat{m}$). I replace this assumption with the less restrictive assump-tion that ε is uncorrelated with the proxy $\hat{m}$ (i.e., $E[\varepsilon] = E[\hat{m}\varepsilon] = 0$). The following result ensues.

Proposition 5. *Using a second-order Taylor expansion for the disutility function $D(\cdot)$ around a given value $\bar{a}$, the measured fund performance can be approximated by*

$$
E[\alpha] \approx -\frac{\delta U'(f^*)}{D''(\bar{a})} R_o^2 var(\varepsilon), \tag{21}
$$

which is negative whenever $var(\varepsilon) > 0$.

Proof. Since $D(\cdot)$ is twice-differentiable, $D(a_s^*)$ can be approximated around a given $\bar{a}$ by

$$
D(a_s^*) \approx D(\bar{a}) + D'(\bar{a})(a_s^* - \bar{a}) + \frac{1}{2}D''(\bar{a})(a_s^* - \bar{a})^2. \tag{22}
$$

The first-order condition (4) becomes $\delta U'(f^*)R_o m_s \approx D'(\bar{a}) + D''(\bar{a})(a_s^* - \bar{a})$, or equivalently $a_s^* \approx \bar{a} - (D'(\bar{a})/D''(\bar{a})) + (\delta U'(f^*)/D''(\bar{a}))R_o m_s$.

Inserting this approximation into Eq. (8) gives

$$
E[\alpha] \approx -\frac{\delta U'(f^*)}{D''(\bar{a})} R_o^2 var(\varepsilon). \tag{23}
$$

Since $D(\cdot)$ is strictly convex and twice-differentiable, $D''(\bar{a})$ is positive. Hence, the approximation of $E[\alpha]$ is negative if $var(\varepsilon) > 0$ and equals zero otherwise.   □

Under the assumption that ε is uncorrelated with $\hat{m}$, expected risk-adjusted performance can be approximated by a number that is no greater than zero. This approxima-tion is exact when $D(\cdot)$ is quadratic. Hence, derivations in Section 5, which rely on a quadratic disutility function, would still hold even if I were to replace the assumption that ε is mean independent from $\hat{m}$ with the less restrictive assumption that ε is uncorrelated with $\hat{m}$.

Appendix C. Sample description

I use the CRSP survivorship-bias-free mutual fund database to build my sample of funds. The database includes information on funds' returns, fees, investment objectives, and other characteristics, such as assets under management and turnover.

I use a sample similar to that of Kacperczyk, Sialm, and Zheng (2008) and Glode, Hollifield, Kacperczyk, and Kogan (2010) covering the 1980–2005 period and focusing on actively managed open-end domestic diversified equity mutual funds. I eliminate balanced, bond, money market, international, sector, and index funds. I identify index funds by filtering fund names for the words "index," "idx," "S&P," "Vanguard," "DFA," and "program" and then check manually the remaining funds for omissions. I exclude funds that hold less than ten stocks and those that have less than 80% of their assets invested in equity.[11] For funds with multiple share classes, I compute fund-level variables by aggregating across the different share classes and eliminating duplicate share classes. Evans (2006) reports a bias in the CRSP mutual fund database (see also Elton, Gruber, and Blake, 2001). Fund families occasionally incubate several private funds—the track records of the surviving funds are made public, but the track records of terminated funds are kept private. To address this bias, I try to exclude all observations of funds in their incubation period. I exclude observations for which the observation year precedes the reported fund starting year and observations with missing fund name. Since incubated funds tend to be small, I also exclude funds that had less than $5 million in assets under management at the beginning of the quarter.

The sample includes 3,147 distinct funds and 77,281 fund-quarter observations. Table 1 reports summary statistics for the main fund attributes.

References

Admati, A., Ross, S., 1985. Measuring investment performance in a rational expectations equilibrium model. Journal of Business 58, 1–26.

---

[11] I thank Marcin Kacperczyk and Amit Seru for giving me access to part of their data.

Avramov, D., Wermers, R., 2006. Investing in mutual funds when returns are predictable. Journal of Financial Economics 81, 339–377.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzle. Journal of Finance 59, 1481–1509.

Barber, B., Odean, T., Zheng, L., 2005. Out of sight, out of mind: the effects of expenses on mutual fund flows. Journal of Business 78, 2095–2120.

Barras, L., Scaillet, O., Wermers, R., 2010. False discoveries in mutual fund performance: measuring luck in estimated alphas. Journal of Finance 65, 179–216.

Basak, S., Pavlova, A., Shapiro, A., 2007. Optimal asset allocation and risk shifting in money management. Review of Financial Studies 20, 1–39.

Bekaert, G., Hodrick, R., 1992. Characterizing predictable components in excess returns on equity and foreign exchange markets. Journal of Finance 47, 467–509.

Berk, J., 1995. A critique of size-related anomalies. Review of Financial Studies 8, 275–286.

Berk, J., Green, R., 2004. Mutual fund flows and performance in rational markets. Journal of Political Economy 112, 1269–1295.

Brennan, M., 1993. Agency and asset pricing. Unpublished working paper. University of California, Los Angeles.

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

Chapman, D., 1997. Approximating the asset pricing kernel. Journal of Finance 52, 1383–1410.

Chordia, T., 1996. The structure of mutual fund charges. Journal of Financial Economics 41, 3–39.

Christoffersen, S., 2001. Why do money fund managers voluntarily waive their fees? Journal of Finance 56, 1117–1140.

Christoffersen, S., Musto, D., 2002. Demand curves and the pricing of money management. Review of Financial Studies 15, 1499–1524.

Cochrane, J., 2001. Asset Pricing. Princeton University Press, Princeton, NJ.

Cuoco, D., Kaniel, R., 2007. Equilibrium prices in the presence of delegated portfolio management. Unpublished working paper. University of Pennsylvania.

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

Elton, E., Gruber, M., Blake, C., 2001. A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund database. Journal of Finance 56, 2415–2430.

Evans, R., 2006. Does alpha really matter? Evidence from mutual fund incubation, termination and manager change. Unpublished working paper. University of Virginia.

Fama, E., 1998. Market efficiency, long-term returns and behavioral finance. Journal of Financial Economics 49, 283–306.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

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

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

Ferson, W., Warther, V., 1996. Evaluating fund performance in a dynamic market. Financial Analysts Journal 52, 20–28.

Garcia, D., Vanden, J., 2009. Information acquisition and mutual funds. Journal of Economic Theory 144, 1965–1995.

Glode, V., Hollifield, B., Kacperczyk, M., Kogan, S., 2010. Is investor rationality time varying? Evidence from the mutual fund industry. Unpublished working paper. Carnegie Mellon University.

Goetzmann, W., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures. Review of Financial Studies 20, 1503–1546.

Golec, J., 2003. Regulation and the rise in asset-based mutual fund management fees. Journal of Financial Research 26, 19–30.

Golec, J., Starks, L., 2004. Performance fee contract change and mutual fund risk. Journal of Financial Economics 73, 93–118.

Grinblatt, M., Titman, S., 1989. Portfolio performance evaluation: old issues and new insights. Review of Financial Studies 2, 393–421.

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

Hansen, L., Richard, S., 1987. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. Econometrica 55, 587–613.

Hansen, L., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. Journal of Political Economy 99, 225–262.

Hansen, L., Jagannathan, R., 1997. Assessing specification errors in stochastic discount factor models. Journal of Finance 52, 557–590.

Harrison, J., Kreps, D., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20, 381–408.

Henriksson, R., Merton, R., 1981. On market timing and investment performance II: statistical procedures for evaluating forecasting skills. Journal of Business 54, 513–533.

Hodrick, R., Zhang, X., 2001. Evaluating the specification errors of asset pricing models. Journal of Financial Economics 62, 327–376.

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

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

Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L., 2010. Attention allocation over the business cycle. Unpublished working paper. New York University.

Kan, R., Zhou, G., 2006. A new variance bound on the stochastic discount factor. Journal of Business 79, 941–961.

Kihlstrom, R., 1988. Optimal contracts for security analysts and portfolio managers. Studies in Banking and Finance 5, 291–325.

Kosowski, R., 2006. Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions. Unpublished working paper. INSEAD.

Kothari, S., Warner, J., 2001. Evaluating mutual fund performance. Journal of Finance 56, 1985–2010.

Kuhnen, C., 2004. Dynamic contracting in the mutual fund industry. Unpublished working paper. Northwestern University.

Lynch, A., Wachter, J., 2007. Does mutual fund performance vary over the business cycle? Unpublished working paper. University of Pennsylvania.

Malkiel, B., 1995. Returns from investing in equity mutual funds 1971 to 1991. Journal of Finance 50, 549–572.

Mamaysky, H., Spiegel, M., Zhang, H., 2007. Estimating the dynamics of mutual fund alphas and betas. Review of Financial Studies 21, 233–264.

Melino, A., Yang, A., 2003. State dependent preferences can explain the equity premium puzzle. Review of Economic Dynamics 6, 806–830.

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

Pástor, L., Stambaugh, R., 2010. On the size of the active management industry. Unpublished working paper. University of Pennsylvania.

Roll, R., 1977. A critique of the asset pricing theory's tests part I: on past and potential testability of the theory. Journal of Financial Economics 4, 129–176.

Savov, A., 2009. Free for a fee: the hidden cost of index fund investing. Unpublished working paper. University of Chicago.

Sirri, E., Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance 53, 1589–1622.

Staal, A., 2006. Essays in empirical finance. Doctoral dissertation. Northwestern University.

Treynor, J., Mazuy, K., 1966. Can mutual funds outguess the market? Harvard Business Review 44, 131–136.

Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transactions costs, and expenses. Journal of Finance 55, 1655–1703.