European Finance Review **3:** 205–232, 1999.

Asset Pricing Specification Errors and Performance Evaluation

JIA HE[1], LILIAN NG[2] and CHU ZHANG[3]*
[1] *Chinese University of Hong Kong, Hong Kong*
[2] *University of Wisconsin-Milwaukee, School of Business Administration, Wisconsin WI 53202-742, U.S.A.*
[3] *Hong Kong University of Science & Technology*

Abstract. Many evaluation techniques typically measure performance as deviations of average returns on actively managed funds from those predicted by some asset pricing model. Empirical evidence, however, has so far suggested that all asset pricing models lack empirical support, implying that the models contain mis-specification errors to various degrees. Evaluating mutual fund performance relative to any of these models thus becomes problematic. In this paper, we propose an approach to performance measurement that emphasizes minimizing explicitly the pricing error associated with an asset pricing function which is employed to compute performance measures. This approach is henceforth called the minimum specification-error (MSE) method. We also discuss the statistical properties for implementing MSE performance measures. To demonstrate the significance of the pricing error confounded in evaluation measurement, we contrast our methodology with the Grinblatt and Titman (1989) period weighting approach and with the empirical implementation of Chen and Knez (1996). We find that the greater the pricing error of passive assets, the larger the performance measures. Given the average pricing error generated from a collection of 163 diverse passive portfolios used in this analysis, the performance values assigned to a large number of the funds become statistically and economically insignificant.

* Jia He is from Chinese University of Hong Kong and Shanghai JiaoTong University, Lilian Ng is from University of Wisconsin, Milwaukee, and Chu Zhang is from Hong Kong University of Science and Technology and the University of Alberta. We thank Stephen Brown, Andy Chen, Zhiwu Chen, Wayne Ferson, Bruce Lehmann, Bob Litzenberger, Ravi Jagannathan, Sheridan Titman, Guofu Zhou, and participants at the Chinese University of Hong Kong, the HKUST, the National Taiwan University Conference on Finance in Taiwan, the 1996 FMA Meeting in New Orleans, the Seventh Annual Conference on Financial Economics and Accounting, Michigan State University, University of Southern California, and the 1998 Western Finance Association Meeting in Monterey for helpful comments and suggestions. We are also grateful to the encouragement from Financial Management Association Competitive Paper Award Committee. Jia He acknowledges assistance from Chinese University of Hong Kong's direct grant No. 2050150 and the H.K. Competitive Earmarked Research Grant No. CUHK 4041/99H. Lilian acknowledges assistance from the H.K. Competitive Earmarked Research Grant No. 9040239. Chu Zhang acknowledges partial financial support from the J. D. Muir fellowship at the University of Alberta.

Portfolio management has become an increasingly important industry not only in the U.S., but also in many other countries across the world. The performance of professional portfolio managers inevitably becomes of significant interest both to practitioners and academics and to many individual investors investing their wealth and pensions in mutual funds. The question is whether portfolio managers, using dynamic trading strategies for fund management, can demonstrate superior performance. Many traditional evaluation techniques generally measure performance as deviations of average returns on actively managed funds from those predicted by some asset pricing model. Mutual funds are assigned positive, zero, or negative performance measures, depending on whether their returns are greater, equal to, or less than, the risk-adjusted return determined by a specific asset pricing model used. Thus, the accuracy of performance measurement relies on the empirical plausibility of the asset pricing model.

The most popular asset pricing model is the Capital Asset Pricing Model (CAPM) and the associated performance measure is known as Jensen's measure. In the CAPM, expected asset returns are determined by the covariance of the asset with the return on the true market portfolio. Despite its popularity, numerous studies have provided mounting evidence that contradicts the CAPM. For example, the model cannot explain size and book-to-market effects (Fama and French (1992)). Further, there is concern that the true market portfolio is not observable. Roll (1977, 1978) argues that a proxy for the market portfolio may be inefficient. Thus, any deviation of the expected return on a fund from that determined by the CAPM may simply indicate model mis-specification due to the inefficient proxy for the market portfolio and not due to abnormal performance.

The lack of empirical support of the CAPM spurs studies to look to multifactor models. Empirically, however, these models require identification of all relevant factors. Theories of asset pricing do not explicitly specify the factors, or the number of factors, necessary to price risky assets. The choice of factors in empirical studies has been intuitively motivated (see, Chen, Roll, and Ross (CRR, 1986)). Unfortunately, evidence suggests that empirical multifactor models still cannot fully explain the cross-sectional difference in average asset returns.[1] The mis-specification of multifactor models inevitably causes problem for performance measures. Lehmann and Modest (1987) show that mutual fund performance measures based on multifactor models are very sensitive to the choice of the benchmark variables, and they find little similarity among performance rankings based on alternative models.

In this paper we present a methodology that stems from the stochastic discount-factor model.[2] The main feature of our approach is that we emphasize the importance of minimizing a metric of pricing errors explicitly before an estim-

[1] For example, He and Ng (1994) find that firm size and book-to-market equity have incremental explanatory power for average stock returns, in the presence of the five factors used by CRR.

[2] Such models have received a lot of attention in the asset pricing literature. See, for example, Hansen and Jagannathan (1991, 1997) and Hansen et al. (1995), and Cochrane (1996). Other related

ated asset pricing function is used to generate performance measures. We call this method the minimum specification-error (MSE) approach. In contrast to prior works on portfolio evaluation, our proposed performance measure is less affected by the inefficient benchmark problem. We provide evidence indicating that the effort in minimizing asset-pricing errors is worthwhile in order to derive a more reliable inference on mutual fund performance.

We also present a conditional version of the methodology. Previous empirical evidence has shown that asset returns can be reliably predicted using *ex ante* observable variables such as dividend yields, interest-rate spreads, and default spreads.[3] Such evidence implies that passive trading strategies should reflect publicly-available information. Any managed portfolio based on dynamic trading strategies that utilize public information should yield only a normal return and be assigned a zero performance value.[4] In this study, we demonstrate the implementation of both unconditional and conditional minimum specification-error measures and discuss their empirical results accordingly.

Finally, we develop consistent estimators of the unconditional and conditional MSE measures and provide a means to conduct statistical inference related to fund performance. In particular, we establish the consistency and asymptotic results for estimators of the average pricing error of mutual funds based on the discount factor proxy that yields the minimal pricing error of passive assets. The formal performance evaluation is then carried out by constructing test statistics under the null hypothesis that there is no difference in performance measures between mutual funds and passive strategies.

The paper is organized as follows. In the next section, we present the methodology. In Section 2, we derive the necessary statistical distribution for the estimators of our proposed unconditional and conditional performance measures. Based on these statistics, we can implement the approach to test the economic and statistical significance of the estimated performance values. Section 3 describes the data on mutual funds, benchmarks, and passive portfolios which we use as test assets for examining the magnitude of the mis-specification error. In Section 4, we compare the results of our measures with those of the Grinblatt and Titman (1989) period-weighting measures and the Chen and Knez (1996) non-parametric approach. The last section offers some concluding remarks.

works in the performance evaluation literature includes Grinblatt and Titman (1989, 1994), Glosten and Jagannathan (1994), and Chen and Knez (1996) which are cast in this framework.

[3] See, for example, Keim and Stambaugh (1986), Campbell (1987), Fama and French (1988, 1989), and Ferson (1989).

[4] Ferson and Schadt (1996) apply a conditional Jensen's measure and obtain results which are more plausible than those which are generated using the unconditional performance measures.

1. A Minimal Specification-Error Approach

1.1. BACKGROUND

For convenience, portfolios are assumed to be purchased at time t, and payoffs to the portfolios are received at time $t + 1$. Consider a universe of N financial assets, and let R_p denotes the N-vector of gross returns on these N financial assets at time $t + 1$ that are purchased at time t. Throughout this paper, returns refer to gross returns, unless otherwise stated. Here p signifies passive strategies. A stochastic discount factor (SDF) is a random variable m realizes at time $t + 1$ satisfying

$$E(R_p m) = 1_N, \tag{1}$$

where 1_N is the N-vector of ones. Hansen and Jagannathan (1991) develop the framework in which asset pricing models can be indexed by their corresponding SDFs. Thus, a multifactor asset pricing model can be indexed by an m which is a linear function of factors plus a constant (See Cochrane (1996)).

Suppose the set of all the information variables is denoted by Ω_t, and a piece of information is any random variable realized at time t. For now, assume that none of the information variables is publicly known. Thus, for simplicity, we drop the subscript t. An uninformed investor, by definition, can only form a portfolio whose weights are constants. From (1), the unconditional performance measure $\lambda_u(\cdot)$ for the passive portfolio whose return, $r_p = w'R_p$, where w is a constant N-vector satisfying $w'1_N = 1$, can be represented by

$$\lambda_u(r_p, m) \equiv E(r_p m) - 1 = 0. \tag{2}$$

This equation says that uninformed investors hold passive portfolios with the risk-adjusted return equal one.

Informed investors, as many fund managers claim they are, possess some information in Ω and form portfolios whose weights are functions of the information they have. Let $r_a = w(a)'R_p$ be the gross return on a portfolio of an informed investor. Here the subscript a in r_a represents active portfolios. Since informed investors face a richer investment opportunity set, they can effectively exploit the information they possess and design active trading strategies. When discounted by the same discount factor used by uninformed investors, funds that are actively managed by informed investors are expected to generate a greater return than those of uninformed investors. Accordingly, the unconditional performance measure for actively-managed funds can be written as:[5]

$$\lambda_u(r_a, m) \equiv E(r_a m) - 1 > 0. \tag{3}$$

[5] Note that the performance measures can also be written in terms of excess returns such that

$$\lambda_u(r_j, m) = E(\tilde{r}_j m),$$

where $\tilde{r}_j$ represents the rate of return on portfolio j $(j = p,\ a)$ in excess of a riskfree rate.

We now present two examples of the SDF-based performance measures in the literature and explain the potential problems in their implementation. The first example is Grinblatt and Titman's (1989) positive period-weighting measure (PWM), whose probabilistic counterpart is defined as

$$\alpha_j^* = E(\tilde{r}_j m), \tag{4}$$

where $\tilde{r}_j$ is the excess return on portfolio j, m is a function of $\tilde{R}_b$, satisfying $E(\tilde{R}_b m) = 0$, and $\tilde{R}_b$ is the vector of excess returns on k efficient benchmark portfolios.[6] In a weaker sense than (1), a period weight function m is a stochastic discount factor satisfying $E(\tilde{R}_p m) = 0$. In implementing the PWM, Cumby and Glen (1990) assume the positive weighting measure to be the marginal utility of the return on an efficient portfolio, which in turn is a linear combination of k benchmark portfolios, R_b, including a riskfree asset. Let the marginal utility function be denoted by $f(\delta' R_b)$. For the benchmark portfolios to be efficient to uninformed investors, the weights $\tilde{\delta}$ must satisfy the first order condition of maximizing the expected utility,

$$E[R_b f(\tilde{\delta}' R_b)] = 1_k. \tag{5}$$

Under regularity conditions, a unique $\tilde{\delta}$ can be solved from the k-equation, and $f(\tilde{\delta}' R_b)$ is the positive weighting function. The problem in practice is that benchmark portfolios in empirical studies are rarely efficient such that the positive weighting function $y_{GT} = f(\tilde{\delta}' R_b)$ constructed from (5) cannot guarantee (1). As a result, (2) does not hold, and performance between passive and active strategies is not distinguishable.

The second example is Chen and Knez's (1996) nonparametric performance measure. Following Hansen and Jagannathan (1991), an SDF can be constructed as

$$m^* = R_p' [E(R_p R_p')]^{-1} 1_N. \tag{6}$$

Such an SDF involves only calculation, as opposed to parameter estimation, and it always satisfies (1) perfectly. As it appears, the problem of mutual fund performance evaluation is resolved. However, the computation of the SDF is nontrivial. Since R_p is the vector of gross rate of returns, for typical data frequency used in empirical studies, the magnitude of gross returns is generally close to one. Consequently, $E(R_p R_p')$ tends to the singular matrix $1_N 1_N'$. Even for a moderately large N, the calculated inverse becomes inaccurate. In the Chen and Knez study, the methodology is tried on a small set of portfolios, say R_b of dimension k, and the calculated "SDF" $y_{CK} = R_b' [E(R_b R_b')]^{-1} 1_k$ can only price (portfolios of) R_b

[6] Under Ross's (1978) k-fund separation condition, $\tilde{R}_p = B\tilde{R}_b + \varepsilon$ with $E(\varepsilon | \tilde{R}_b) = 0$. It can be verified that $E(\tilde{R}_p m) = 0$.

perfectly. When applied to the universe R_p, specification error still persists and the induced problem for performance measures remains unsolved.

1.2. UNCONDITIONAL MSE PERFORMANCE MEASURES

Given that the specification error cannot be completely eliminated from a discount factor proxy, it is imperative that one at least minimizes the error associated with the proxy. Suppose $y = f(R_b, \delta)$ is the proxy for a discount factor, where R_b is a given vector of benchmark portfolios (including a riskfree asset), δ is a vector of parameters, and f is a given functional form. We can consider optimizing over the parameter space to minimize the specification error.[7]

We first attempt to use a minimax rule. For a given δ, we calculate the maximal pricing error among all the linear combinations of returns, and given this maximal pricing error, we minimize over δ. Hansen and Jagannathan (1997) show that, for a given SDF proxy y, the maximal pricing error can be expressed in terms of the norm $\| m^* - y^* \|$, where $m^* = R_p[E(R_pR_p')]^{-1}R_p$ is the projection of any SDF m on the return space, and y^* the linear projection of y on the N-return space. The minimax rule can be, in principle, built on this framework, but unfortunately, this procedure cannot be easily implemented because it still requires the calculation of $[E(R_pR_p')]^{-1}$. We illustrate this by a special case in which the SDF proxy is assumed to be linear in R_b: $y = \delta' R_b$. In this case, $y^* = y = \delta' R_b$, and

$$\| m^* - y^* \|^2 = (ER_pR_b'\delta - 1_N)'[E(R_pR_p')]^{-1}(ER_pR_b'\delta - 1_N),$$

which is a quadratic function of δ. The minimax rule determines the optimal δ to be

$$\delta = [(ER_pR_b')'[E(R_pR_p')]^{-1}(ER_pR_b')]^{-1}(ER_pR_b')'[E(R_pR_p')]^{-1}1_N,$$

which involves $[E(R_pR_p')]^{-1}$.

We then consider another minimax approach in a principal-agent framework, formulated as follows. Consider a mutual fund manager (the agent) who fully understands how his performance is being evaluated, but does not possess any information. For a given SDF proxy $y = f(R_b, \delta)$, his portfolio w is evaluated by

$$\lambda_u(w'R_p, f(R_b, \delta)) = E[(w'R_p)f(R_b, \delta)] - 1.$$

Knowing that the SDF used is only a proxy and hence his performance will be nonzero, he has the incentive to manipulate the portfolio weights to make his performance looks good. He can take extreme long positions in assets that have

[7] It is also natural to consider optimization over the choice of R_b and the functional form f. However, this is beyond the scope of the current paper.

positive pricing errors and extreme short positions in assets that have low or negative pricing errors. For example, using the CAPM proxy typically yields a positive pricing error for long positions in small stocks and for short positions in large stocks. But he may eschew taking extreme positions, because by doing so, his actual return will become riskier as well. We assume here that he balances these two objectives by discounting the performance measure by a measure of concentration of his portfolio: $\| w \| \equiv (\sum_{i=1}^{N} w_i^2)^{1/2}$.[8] Therefore, for the given SDF proxy, the manager's problem is to solve

$$\max_{w:w'1_N=1} \lambda_u(w'R_p, f(R_b, \delta))/ \| w \| . \tag{7}$$

The task of a designer of mutual fund performance (the principal) is to preclude this manipulation by minimizing such gains to the fund manager, over the available choices of the SDF proxy. Thus, in an MSE approach, the SDF proxy is chosen to solve the following problem,[9]

$$\min_{\delta} \left[\max_{w:w'1_N=1} \lambda_u(w'R_p, f(R_b, \delta))/ \| w \| \right]. \tag{8}$$

The following proposition shows why such an approach is called the MSE.

PROPOSITION 1. The solution to problem (8) is the same as the following least square problem,

$$\min_{\delta}[E(R_p f(R_b, \delta)) - 1_N]'[E(R_p f(R_b, \delta)) - 1_N]. \tag{9}$$

Proof: See the Appendix.

If $f(R_b, \delta) = R_b'\delta$, a linear function of the benchmarks, then the MSE proxy of the SDF is $R_b'\delta^*$, where

$$\begin{aligned} \delta^* &= \mathrm{argmin}_{\delta}(ER_pR_b'\delta 1_N)'(ER_pR_b'\delta - 1_N) \\ &= [(ER_pR_b')'(ER_pR_b')]^{-1}(ER_pR_b')'1_N. \end{aligned} \tag{10}$$

The procedure seems similar to the approaches used for estimating and testing an asset pricing model. Our motivation, however, is different. In estimating and testing an asset pricing model, one hypothesizes that $y = f(R_b, \delta)$ is a true discount factor

[8] In the theory of industrial organization, $H = \sum_{i=1}^{N} w_i^2$ is known as the Herfindahl concentration index, satisfying $1/N \leq H \leq 1$ for $w_i \geq 0$, $\sum_{i=1}^{N} w_i = 1$. When there is no nonnegativity constraint on w_i, H has no upper bound.

[9] As we mentioned earlier, the minimization can be taken over the choices of R_b and the functional form f too, as a generalization.

and estimates the true parameter δ. Here we recognize that the set of benchmarks R_b may be allocationally inefficient and that y is only a discount factor proxy. So we select a parameter δ (and probably the functional form as well) to reduce the pricing error. It is important to point out that, in contrast to the approach of Chen and Knez (1996), (10) does not involve the calculation of $[E(R_p R_p')]^{-1}$. The matrix in (10) being inverted is of dimension k.

There is a key distinction between our approach and the PWM and nonparametric approaches. While the latter construct the SDF proxy to price a small number of benchmark portfolios R_b perfectly, without concerning about the pricing errors for the universe of returns R_p, we make an effort to minimize the pricing error of R_p directly. By construction, the pricing error resulted from the MSE approach is smaller than, or equal to, that generated from the PWM or the nonparametric approach. The difference depends on the efficiency of the benchmark portfolios. If the benchmarks are efficient, then the three methods yield the same pricing error. Otherwise, the PWM and nonparametric methods always give rise to a larger pricing error than does the MSE approach. The performance of the two measures as well as their pricing errors shall be empirically demonstrated later.

1.3. CONDITIONAL MSE PERFORMANCE MEASURES

We have so far formulated the specification-error performance measures under the presumption that only the fund managers have information in devising their dynamic trading strategies. Their performance is appraised relative to the performance of those non-professional investors, who have no available information so they hold fixed-composition passive portfolios. In practice, however, there exists public information which can be used in investment decision-making process. Further, it is also well documented that stock and bond returns are predictable over time. In the light of this, it is more appropriate to assign zero performance measure to portfolios based on public information.

For concreteness, we now use subscript t to indicate the time period. Recall Ω_t is the vector of all the relevant information variables at time t. Let a sub-vector of Ω_t, z_t, represents the publicly-available information variables at time t. We relax the assumption that the SDF is a constant parameter and let it be a function of z_t: $\delta_t = g(z_t)$. Therefore,

$$m_{t+1}^c = f(R_{b,t+1}, \delta_t) = f(R_{b,t+1}, g(z_t)).$$

For any portfolio based on z_t, $w(z_t)$, the SDF prices the portfolio according to

$$E_t[w(z_t)' R_{p,t+1} f(R_{b,t+1}, g(z_t))] = 1 \tag{11}$$

For constructing a (proxy of) SDF, it is impossible to experiment all such portfolios, but we can at least consider the following strategies. Let z_{jt} be the jth component of z_t with $1 \leq j \leq l$, and $R_{ip,t+1}$ be the return on ith asset with $1 \leq i \leq N$. The

(i, j)th strategy is to invest z_{jt} dollar in the ith asset. The pricing equation takes the form of

$$E_t[z_{jt} R_{ip,t+1} f(R_{b,t+1}, g(z_t))] = z_{jt}.$$

There are Nl such strategies, and they form a basis for strategies that are linear in z_t. In matrix notation, they can be written as

$$E_t[(z_t \otimes R_{p,t+1}) f(R_{b,t+1}, g(z_t))] = z_t \otimes 1_N. \tag{12}$$

Normalizing z_t by $Ez_t = 1$ and taking unconditional mean on the above equation gives

$$E[(z_t \otimes R_{p,t+1}) f(R_{b,t+1}, g(z_t))] = 1_{Nl}. \tag{13}$$

The conditional MSE approach to performance measures is to construct a conditional SDF proxy $f(R_b, g(z_t))$ that minimizes the least square of pricing errors over the space of parameters in the function $g(z_t)$. For example, if $g(z_t) = Dz_t$, then the conditional MSE proxy is obtained by

$$\min_D [E(z_t \otimes R_{p,t+1}) R'_{b,t+1} D z_t - 1_{Nl}]' [E(z_t \otimes R_{p,t+1}) R'_{b,t+1} D z_t - 1_{Nl}]. \tag{14}$$

Let the minimizer be denoted by D^*. The conditional MSE performance measure for the return, $r_{a,t+1}$, on an active portfolio will be

$$\lambda_c(r_{a,t+1}, R'_{b,t+1} D^* z_t) = E[(r_{a,t+1} R'_{b,t+1} D^* z_t)] - 1. \tag{15}$$

1.4. NON-NEGATIVITY CONSTRAINT ON STOCHASTIC DISCOUNT FACTORS

Based on Harrison and Kreps (1979), Hansen and Jagannathan (1991, 1997) develop the theory of stochastic discount factors with positivity constraint. In order to prevent arbitrage opportunities over the space of nonlinear functions of returns, such as derivative assets, the stochastic discount factor is required to be strictly positive with probability one. In the literature, there are two ways of constructing such SDFs. One way is to use a marginal utility function for the SDF, where the marginal utility function is everywhere positive. Cumby and Glen (1991) adopt such a method. The other approach is to truncate a linear function, say,

$$y^+ = \max\{R'_b \delta, 0\}.$$

This is initiated by Hansen and Jagannathan to calculate some functional of the discount factors and used by Chen and Knez (1996) in constructing SDFs for performance measures. The second method is not perfect, because the resultant discount factor (or the proxy) has a positive probability of being zero, as opposed to being strictly positive.[10] A nonnegative SDF, however, does meet the requirement of a weaker condition of no-arbitrage. (See Ingersoll (1987) for details.)

In the MSE approach, the nonnegativity constraint can also be imposed using either method described above.

2. Statistical Properties of the Estimators

In the empirical part of the paper, we will test the hypothesis that mutual fund managers do not have superior information and their performance is indistinguishable from passive strategies. In this section we develop consistency and asymptotic results for the estimators of the specification-error metric, as given by (10), that we can implement for making statistical inferences. Our starting point is to establish the distribution of the estimator for the pricing-error metric associated with passive assets. Next, we shall relate this to the specification-error measure used to evaluate mutual fund performance. We only consider two choices of the SDF proxy. One is linear in benchmark variables: $y = R_b'\delta$, and the other is truncated linear: $y^+ = \max\{R_b'\delta, 0\}$. Since the statement of the results are almost identical, we use the notation $f(R_b, \delta)$ to refer to either one of them. The proof of the former results is conventional, but that of the latter is not. Therefore, the proofs are given in terms of y^+ and are relegated to the Appendix.

Let $R_{a,t+1}$ be an M-vector of returns on actively managed mutual fund portfolios. To further simplify the notation, we will write $\lambda_u(R_p, \delta)$ and $\lambda_u(R_a, \delta)$ rather than $\lambda_u(R_p, f(R_b, \delta))$ and $\lambda_u(R_a, f(R_b, \delta))$. We begin by defining the following,

$$\lambda_u^2(R_p, \delta) = \lambda_u(R_p, \delta)'\lambda_u(R_p, \delta), \quad \text{where } \lambda_u(R_p, \delta) \equiv E[R_p f(R_b, \delta)] - 1_N, \tag{16}$$

$$\lambda_u^2(R_a, \delta) = \lambda_u(R_a, \delta)'\lambda_u(R_a, \delta), \quad \text{where } \lambda_u(R_a, \delta) \equiv E[R_a f(R_b, \delta)] - 1_M. \tag{17}$$

The sample analog estimators for (16)–(17) are defined by replacing the expectation operator E with time-series average $\frac{1}{T}\sum_{t=1}^{T}$, and are denoted by $\hat{\lambda}_{u,T}^2(R_p, \delta)$ and $\hat{\lambda}_{u,T}^2(R_a, \delta)$, respectively. It can be shown that

$$\Lambda_u^2(R_p) \equiv \min_{\delta} \lambda_u^2(R_p, \delta), \tag{18}$$

$$\hat{\Lambda}_{u,T}^2(R_p) \equiv \min_{\delta} \hat{\lambda}_{u,T}^2(R_p, \delta), \tag{19}$$

are well-defined with their solutions denoted by δ^* and $\hat{\delta}_T^*$, respectively. For further development, we establish the following assumptions.

[10] It is legitimate for Hansen and Jagannathan to use a truncated function, since they are interested in constructing functional of SDFs, rather than SDFs themselves.

ASSUMPTION 1. (i) $R_{t+1} = (R'_{p,t+1}, R'_{a,t+1})'$ is stationary and ergodic. (ii) $ER_{t+1}R'_{t+1} < \infty$. We can now state the consistency of the sample analog estimator sequences, $\hat{\Lambda}^2_{u,T}(R_p)$.

PROPOSITION 2. Under Assumption 1, $\hat{\Lambda}^2_{u,T}(R_p) \xrightarrow{\text{a.s.}} \Lambda^2_u(R_p)$, as $T \to \infty$. To examine the limiting distribution of the estimators, we make the following assumption.

ASSUMPTION 2. $\Pr[R_{p,t+1}f(R_{b,t+1}, \delta) = 0] = 0$, and

$$\sqrt{T}\begin{bmatrix} \frac{1}{T}\sum_1^T R_{p,t+1}f(R_{b,t+1}, \delta) - E[R_{p,t+1}f(R_{b,t+1}, \delta)] \\ \operatorname{vec}\left\{\frac{1}{T}\sum_1^T R_{p,t+1}R'_{b,t+1} \cdot I(R_{b,t+1}\delta > 0)\right\} - \operatorname{vec}\{E[R_{p,t+1}R'_{b,t+1} \cdot I(R_{b,t+1}\delta > 0)]\} \end{bmatrix},$$

converge in distribution to normally distributed random vectors with mean zero and its corresponding covariances, $\hat{V}$, where $\hat{V}$ can be degenerate, and $I(\cdot)$ is an index function.

PROPOSITION 3. Suppose $\Lambda^2_u(R_p) \neq 0$. Under Assumptions 1 and 2, $\sqrt{T}(\hat{\Lambda}^2_{u,T}(R_p) - \Lambda^2_u(R_p))$ converge to normally distributed random variables.

Assumption 2 is true if the conditions for central limit theorem are satisfied. With the assumption of $\Lambda^2_u(R_p) \neq 0$, the asymptotic distribution in Proposition 3 is a normal rather than a chi-square.[11] We have thus far established the distribution of the estimators for the pricing-error metric associated with the passive assets. Next, we turn to the distribution of the estimators for the corresponding metric associated with mutual fund performance. We will show below that $\sqrt{T}[\hat{\lambda}^2_{u,T}(R_a, \hat{\delta}^*_T) - \lambda^2_u(R_a, \delta^*)]$ converges to a normal random variable. To simplify the proof, we need some additional assumptions.

ASSUMPTION 3.

$$\sqrt{T}\begin{Bmatrix} \hat{\delta}^*_T - \delta^* \\ \hat{\lambda}_{u,T}(R_a, \delta) - \lambda_u(R_a, \delta) \end{Bmatrix},$$

converge to normally distributed random vectors with mean zero. And

$$\frac{1}{T}\sum_1^T R_{a,t+1}R'_{b,t+1} \cdot I(A_{T_i}) \xrightarrow{\text{a.s.}} E[R_{a,t+1}R'_{b,t+1} \cdot I(A_i)], \quad \text{if } A_{T_i} \xrightarrow{\text{a.s.}} A_i,$$

for $i = 1, 2, 3$ where

$$A_{T_1} = (w : R'_{b,t+1}\hat{\delta}^*_T > 0, \quad \text{and } R'_{b,t+1}\delta^* \leq 0),$$

[11] See Heaton et al. (1995)

$$A_{T_2} = (w : R'_{b,t+1}\hat{\delta}^*_T < 0, \quad \text{and } R'_{b,t+1}\delta^* \geq 0),$$

$$A_{T_3} = (w : R'_{b,t+1}\hat{\delta}^*_T > 0, \quad \text{and } R'_{b,t+1}\delta^* > 0).$$

PROPOSITION 4. Suppose $\lambda^2_u(R_a, \delta^*) \neq 0$. Then, $\sqrt{T}[\hat{\lambda}^2_{u,T}(R_a, \hat{\delta}^*_T) - \lambda^2_u(R_a, \delta^*)]$ converge to normal random variables with mean zero, under Assumptions 1–3.
At this point we can test the null hypothesis that the average pricing error for passive assets is the same as the average pricing error for mutual funds. If we cannot reject the null hypothesis, this implies that, overall, mutual funds exhibit zero abnormal performance.

Suppose $\lambda^2_u(R_p, \delta^*) \neq 0$. Under the null hypotheses that

$$\frac{1}{M}\lambda^2_u(R_a, \delta^*) = \frac{1}{N}\lambda^2_u(R_p, \delta^*),$$

the corresponding $\sqrt{T}[\frac{1}{N}\hat{\lambda}^2_{u,T}(R_p, \hat{\delta}^*_T) - \frac{1}{M}\hat{\lambda}^2_{u,T}(R_a, \hat{\delta}^*_T)]$ converge to normally distributed random vectors with mean zero. The sample variance of $\sqrt{T}[\frac{1}{N}\hat{\lambda}^2_{u,T}(R_p, \hat{\delta}^*_T) - \frac{1}{M}\hat{\lambda}^2_{u,T}(R_a, \hat{\delta}^*_T)]$, as denoted by $\hat{V}$, is given by

$$\hat{V} = \frac{1}{T-1}\sum_{t=1}^{T}(\xi_{t+1} - \bar{\xi}_{t+1})^2,$$

where

$$\begin{aligned}
\xi_{t+1} = {} & \frac{1}{N}[R_{p,t+1}f(R_{b,t+1}, \hat{\delta}^*_T) - 1_N]' \\
& \times \left[\frac{1}{T}\sum_{1}^{T} R_{p,t+1}f(R_{b,t+1}, \hat{\delta}^*_T) - 1_N\right] \\
& - \frac{1}{M}[R_{a,t+1}f(R_{b,t+1}, \hat{\delta}^*_T) - 1_M]' \\
& \times \left[\frac{1}{T}\sum_{1}^{T} R_{a,t+1}f(R_{b,t+1}, \hat{\delta}^*_T) - 1_M\right].
\end{aligned}$$

Notice that $\sqrt{T}[\frac{1}{N}\hat{\lambda}^2_{u,T}(R_p, \hat{\delta}^*_T) - \frac{1}{M}\hat{\lambda}^2_{u,T}(R_a, \hat{\delta}^*_T)]$ is equal to $\sum_{t=1}^{T}\xi_{t+1}$. It can be shown that ξ_{t+1} is asymptotically independent across time, and therefore the sample variance of the estimator can be calculated as shown above.

The results so far have been developed in the unconditional framework. The derivation of the conditional counterparts to these results can be similarly carried out, except for some minor modification. Given the conditioning information, the payoff space would be naturally enlarged. Thus, the N dimensional space generated by the passive assets is expanded to $N \times l$. The modified statistics are then applied accordingly.

3. Data Description

3.1. MUTUAL FUND DATA

Mutual fund data are obtained from the Morningstar database. The data consist of monthly cash-distribution-adjusted returns on 1,067 mutual funds from January 1976 to December 1991, of which 259 funds have observations throughout the sample period. We concentrate on these 259 funds.[12] These funds can be classified into seven groups according to their investment objectives, as defined by Morningstar. They are small funds (SML), aggressive growth funds (AGG), growth funds (GRO), growth and income funds (G&I), equity income funds (EQI), balanced funds (BAL), and corporate bond mutual funds (CPB).

Table I presents the minimum, the lower quartile q_1, the median, the upper quartile q_3, and the maximum of the mean and standard deviation of monthly returns on the funds in each group. As indicated by the fund objectives, the small funds, the aggressive growth funds and the growth funds have relatively higher average returns and larger standard deviations. The corporate bond funds, however, have the lowest average return as well as the smallest standard deviation.

3.2. THE BENCHMARK

Following Sharpe (1992), we employ a broad passive six-index benchmark, including bonds and stocks. Four of them are returns on indexes of common stocks, compiled by Sharpe and BARRA. They are (1) large-capitalization value stock index, consisting of stocks in the S&P 500 stock index with high book-to-price ratios; (2) large-capitalization growth stock index, consisting of stocks in the S&P 500 stock index with low book-to-price ratios; (3) medium-capitalization stock index, consists of stocks in the top 80% of capitalization in the U.S. equity universe after the exclusion of stocks in the S&P 500 stock index; and (4) small-capitalization stock index, consisting of stocks in the bottom 20% of capitalization in the U.S. equity universe after the exclusion of stocks in the S&P 500 stock index. A detailed description on the construction of these four U.S. equity indexes is contained in Sharpe (1992). The other two indexes are bond indexes obtained from Ibbotson and Associates. They are (5) long-term government bond index with approximately 20 years of maturity; and (6) long-term corporate bond index.

As Sharpe (1992) explains, the benchmark is designed in such a way that the indexes are mutually exclusive, exhaustive, and have different market-capitalization-weighted returns. Particularly, the benchmark is formed not only on the basis of

[12] Do we want to discuss about survivorship bias?

Table I. Quartiles for means and standard deviations of annualized monthly returns (in %) on the mutual funds by fund groups. Sample period is 1976:1–1991:12. The table reports the means in the first row with standard deviations in the second row. The fund groups are: small funds, **SML**; aggressive growth funds, **AGG**; growth funds, **GRO**; growth and income funds, **G&I**; equity income funds, **EQI**; balanced funds, **BAL**; corporate bond funds, **CPB**.

Fund group	No.of funds	Quartiles min	q_1	Median	q_3	max
SML	14	12.200	14.506	17.731	20.006	23.887
		50.749	65.555	68.819	72.617	106.768
AGG	17	06.344	13.963	18.044	18.938	22.956
		54.298	72.747	77.338	85.935	116.309
GRO	111	11.175	14.550	16.150	18.494	27.063
		29.554	56.820	61.141	67.035	87.377
G&I	60	00.319	13.244	14.306	15.356	19.288
		40.858	48.188	53.026	56.497	68.170
EQI	11	11.338	13.056	14.556	15.338	17.850
		34.099	43.706	45.522	46.755	50.710
BAL	22	12.194	12.850	13.469	13.744	16.638
		31.354	36.243	39.614	43.011	57.546
CPB	24	09.087	09.688	10.181	10.412	11.019
		16.375	24.815	25.640	27.892	31.505

security characteristics, but also on their abilities to capture common economic and market-wide sources of security return variability. Notice that each index can be interpreted as a passive investment strategy that can be replicated at a low cost using an index fund. Inclusion of non-S&P 500 stocks is crucial in our analysis because they contribute significantly in assessing mutual fund performance.[13]

3.3. PASSIVE PORTFOLIOS

Ideally, a stochastic discount factor should be constructed using all the available individual risky assets. In practice, however, it is feasible to use only a limited number of assets, or portfolios of assets. These passive individual assets or asset portfolios are used to construct proxies of stochastic discount factors with minimum specification errors. It is therefore important that the passive assets used in the study capture a wide diversity of asset characteristics, so that mutual fund portfolios can be thought as not too different from being portfolios of these passive

[13] Other studies such as Elton et al. (1993) have shown that performance measures differ when non-S&P 500 stocks are included as a benchmark portfolio.

assets. We select 163 diverse passive assets that include the earlier six factors, 140 stock portfolios, and 17 bond portfolios. Along the lines of Grinblatt and Titman (1994), we form 120 stock portfolios based on the following six characteristics:

1. Market capitalization of a firm's equity.
2. Dividend yield.
3. Past stock returns based on the three prior years' observed returns.
4. Sensitivity of interest rates.
5. Co-skewness – the slope estimate of the squared excess return of the value-weighted market stock portfolio formed using NYSE, AMEX, and NASDAQ stocks.
6. Beta.

The remaining 20 stocks are formed based on firms' book-to-market equity ratios (BM). This grouping is motivated by Fama and French's (1992) findings that firm size and BM are the two key variables for explaining a cross section of average stock returns. In forming these 140 portfolios, we rank all NYSE stocks based on each of the seven characteristics and then find twenty breakpoints based on each characteristic. All NYSE, AMEX, and NASDAQ stocks are then assigned to each portfolio using the predetermined NYSE breakpoints. Subsequently, the monthly equally-weighted returns on each portfolio are calculated for the year. This procedure is updated yearly.

In addition, we include three U.S. Government bond portfolios from the Center of Research in Security Prices at the University of Chicago and 14 corporate bond composites from Ibbotson's corporate bond module that have return data starting from January 1976 to December 1991. The three Government bonds consist of the short term, intermediate term, and long term bonds, while the 14 corporate bond composites contain the long term, intermediate term, and various industry bond composites.

3.4. INFORMATION VARIABLES

The information variables we use in this study are the same as those employed by CK. Drawn from existing empirical evidence, these variables are selected based on their abilities to forecast stock returns. They are the nominal one-month Treasury bill rate, the dividend yield of the CRSP value-weighted NYSE stock index, and the difference between yields on long-term corporate bonds of more than 15 years of maturity and bonds of 5 to 15 years of maturity. The first two variables are available from CRSP, while the last variable is obtained from the Ibbotson corporate bond module.

4. Empirical Results

We construct unconditional and conditional SDF proxies that are derived from the two sets of benchmark variables. Analogously, we compute PWMs and compare them with our proposed MSEs. Since the pricing error, $\sqrt{\lambda'\lambda/N}$, is estimated using a wide collection of 163 diverse passive assets, as described in Section 3.3, it allows us to gauge how good a performance measure is.

We use linear SDF proxies, as well as those with positivity constraints, as in Cumby and Glen (1991) and Chen and Knez (1996). But only the results for linear proxies are reported here, because the patterns we observe are very similar, with or without positivity constraints, and because there are no derivative assets explicitly involved here, the positivity constraints are less important.

4.1. UNCONDITIONAL PERFORMANCE EVALUATION

Estimating performance measures of Table II relies on the assumption of an unconditional CAPM, while generating those of Table III is based on an unconditional six-factor model (the six factors are described in Section 3.2). Each table has two panels. Panel A presents estimated MSEs, and Panel B reports estimated PWMs. Each panel contains the minimum, the lower quartile q_1, the median, the upper quartile q_3, and the maximum of the performance measures for each of the seven groups of mutual funds.

The two approaches produce distinctly disparate results. Mutual fund performance is found to vary with the type of measure we use. In almost all cases, the period weighting method assigns a greater performance value to each group of funds than does the pricing error approach. In Panel A of Table II, more than half of the MSEs are negative, indicating that the majority of the mutual funds exhibit poor performance. The negative performance is in accord with that reported by, for example, Jensen (1968) and Elton et al. (1993).[14] The latter show that returns on both S&P and non-S&P stocks and on bonds are crucial factors for performance measurement. After incorporating these factors, the four authors find no evidence that mutual funds earn abnormal returns.

In contrast to MSEs, their PWM counterparts shown in Panel B are mainly positive. Even if the PWMs are negative, they are typically smaller in absolute value than the MSEs. For example, the median MSEs in Panel A of Table II are positive only for SML, AGG, and GRO, while negative for those of the remaining funds. These contrast with the median PWMs indicated in Panel B of the same table, where the values are all positive. For instance, the MSEs computed for the GRO fund group range from −0.24% to 0.98%[15] (Panel A), while its estimated PWMs vary from 0.26% to 1.53% per month (Panel B). Evidently, the distribution of the PWMs tends to be more skewed to the right.

[14] These results differ from Ippolito's (1989) finding that the estimated abnormal risk-adjusted return is positive even after adjusting for transaction costs and expenses. He attributes this positive performance to the management's informed actions.

[15] The returns on these mutual funds are not adjusted for transaction costs such as different front-end loads and management fees charged by various investment fund companies. Such an adjustment of fees would only exacerbate the values.

Table II. Quartiles for fund performance measures (in % per month) by fund groups, using the MSE approach and the period weighting measure (PWM) approach. Both methods assume an unconditional CAPM model. Sample period is for 1976:1–1991:12. groups are: small funds, **SML**; aggressive growth funds, **AGG**; growth funds, **GRO**; growth and income funds, **G&I**; equity income funds, **EQI**; balanced funds, **BAL**; corporate bond funds, **CPB**. z-ratio tests the hypothesis that there is no difference between the average pricing error of passive assets and the average specification error-based measure for active mutual fund returns.

Fund	No. of	Quartiles				
group	funds	min	q_1	Median	q_3	max
A: The unconditional MSE approach, $y = R_b'\delta^*$						
Average pricing error, $\sqrt{\lambda_u^2(R_p, \delta^*)/N} = 0.248$.						
Average performance measure, $\sqrt{\lambda_u^2(R_a, \delta^*)/M} = 0.287$						
$H : \lambda_u^2(R_p, \delta^*)/N - \lambda_u^2(R_a, \delta^*)/M = 0$, z-test $= -0.378$.						
SML	14	−0.273	−0.127	0.102	0.390	0.723
AGG	17	−0.757	−0.089	0.180	0.317	0.592
GRO	111	−0.252	−0.043	0.100	0.299	0.979
G&I	60	−1.159	−0.107	−0.025	0.064	0.382
EQI	11	−0.284	−0.162	−0.091	0.032	0.250
BAL	22	−0.193	−0.144	−0.104	−0.051	0.148
CPB	24	−0.470	−0.418	−0.375	−0.356	−0.301
B: The unconditional PWM approach, $y = R_b'\tilde{\delta}$						
Average pricing error, $\sqrt{\lambda_u^2(R_p, \tilde{\delta})/N} = 0.716$.						
Average performance measure, $\sqrt{\lambda_u^2(R_a, \tilde{\delta})/M} = 0.633$						
SML	14	0.285	0.438	0.666	0.934	1.270
AGG	17	−0.202	0.452	0.752	0.861	1.163
GRO	111	0.261	0.493	0.633	0.824	1.531
G&I	60	−0.648	0.409	0.489	0.583	0.911
EQI	11	0.247	0.378	0.417	0.558	0.783
BAL	22	0.330	0.380	0.420	0.449	0.683
CPB	24	0.060	0.111	0.156	0.171	0.226

Generally, the performance values computed using the PWM approach tend to be larger than the MSEs calculated using the pricing-error technique. As shown in Table II, the average MSE is about 0.29% per month, whereas the average PWM is about 0.63% per month. At first glance, one would conclude that mutual funds, on average, exhibit superior performance. However, these values need to be evaluated relative to the pricing errors induced by discount factor proxies and by inefficient benchmarks employed. We find that both the average MSE and average PWM are

Table III. Quartiles for fund performance measures (in % per month) by fund groups, using the MSE approach and the period weighting measure (PWM) approach. Both methods assume an unconditional 6-factor model. Sample period is 1976:1–1991:12. The fund groups are: small funds, **SML**; aggressive growth funds, **AGG**; growth funds, **GRO**; growth and income funds, **G&I**; equity income funds, **EQI**; balanced funds, **BAL**; corporate bond funds, **CPB**. z-ratio tests the hypothesis that there is no difference between the average pricing error of passive assets and the average specification error-based measure for active mutual fund returns.

Fund group	No. of funds	Quartiles min	q_1	Median	q_3	max
A: The unconditional MSE approach, $y = R_b'\delta^*$						
Average pricing error, $\sqrt{\lambda_u^2(R_p, \delta^*)/N} = 0.154$						
Average performance measure, $\sqrt{\lambda_u^2(R_a, \delta^*)/M} = 0.234$						
$H : \lambda_u^2(R_p, \delta^*)/N - \lambda_u^2(R_a, \delta^*)/M = 0$, z-test = - 0.165						
SML	14	−0.646	−0.270	−0.191	0.012	0.198
AGG	17	−1.561	−0.430	−0.176	−0.013	0.180
GRO	111	−0.827	−0.244	−0.114	0.056	0.523
G&I	60	−1.456	−0.279	−0.211	−0.123	0.301
EQI	11	−0.394	−0.262	−0.192	−0.144	0.118
BAL	22	−0.276	−0.223	−0.157	−0.103	0.145
CPB	24	−0.275	−0.173	−0.115	−0.086	0.081
B: The unconditional PWM approach, $y = R_b'\tilde{\delta}$						
Average pricing error, $\sqrt{\lambda_u^2(R_p, \tilde{\delta})/N} = 0.427$						
Average performance measure, $\sqrt{\lambda_u^2(R_a, \tilde{\delta})/M} = 0.362$						
SML	14	−0.135	0.040	0.175	0.433	0.052
AGG	17	−0.813	0.014	0.347	0.571	0.963
GRO	111	−0.493	0.079	0.211	0.401	1.254
G&I	60	−1.208	−0.048	0.076	0.170	0.447
EQI	11	−0.326	−0.174	0.102	0.137	0.416
BAL	22	−0.887	−0.010	0.028	0.097	0.448
CPB	24	−0.742	−0.311	−0.249	−0.195	−0.071

no different from their respective estimated pricing errors,[16] suggesting evidence of no abnormal performance. The finding of no significant abnormal performance is consistent with previous results obtained by many studies, such as Lehman

[16] The Morningstar database contains only historical data on survived U.S. mutual funds. Brown et al. (1992) show that survivorship bias may lead to biases in the performance measurement results. Given the lack of evidence of abnormal performance in this study and Grinblatt and Titman's (1988) finding of a small magnitude of this bias, our results are unlikely to be affected by survivorship bias.

and Modest (1987), Connor and Korajczyk (1991), and CK. Unlike our pricing-error approach and CK's, both Lehman and Modest and Connor and Korajczyk adopt parametric models that include the CAPM and the multifactor model in their analyses. Although their results are similar to those summarized here, their performance measures may be muddled by errors due to inefficient benchmarks used in their models, while CK's measures may be distorted by errors induced by a discount-factor proxy, the inefficient benchmark assets they select, or both.

A similar pattern also prevails in Table III, where an unconditional six-factor model is employed. We also note that all the median MSEs reported in Panel A of Table III are negative, while all the median PWMs in Panel B of the same table are positive. In comparison with the performance measures reported in Table II, those summarized in Table III are generally smaller. Average pricing errors of Table II are also greater than those of Table III. The average pricing errors indicated in Panel A are 0.25% (Table II) and 0.15% per month (Table III), and those in Panel B are 0.72% (Table II) and 0.43% per month (Table III). These results illuminate two observations. One, as compared to those contained in Table III, the larger pricing errors in Table II are consistent with the widely-documented empirical evidence that a multifactor model performs better than a single factor model such as the CAPM. Since the six-factor model better captures the cross-sectional variation in the stock returns on the passive portfolios, its resulting pricing error would be expected to be smaller than that of the CAPM. As a result, we observe that almost all of the performance measures based on the CAPM are greater than those based on the six-factor model. Two, the average pricing errors obtained using the PWM approach are at least twice the average pricing errors generated from the pricing-error method. The vast difference in the pricing error of the two approaches is merely attributable to the inefficiency of the commonly-selected benchmarks employed in this study as well as in many other existing studies. Given the bias generated from the pricing errors, it would be difficult to infer that the average fund performance value constitutes superior performance, or to conclude that the performance measures provide reliable evidence of superior or inferior information or management skills.

4.2. CONDITIONAL PERFORMANCE EVALUATION

Tables IV and V contain the conditional counterparts of Tables II and III. The pattern of the conditional measures is generally consistent with that of the unconditional measures.

Table IV shows that the majority of the MSEs are negative, with the lowest performance value of -1.49% (G&I) and the highest of 0.73% (GRO), whereas most of the PWMs are positive, with abnormal returns ranging from -0.55% (G&I) to 1.62% per month (GRO). It is obvious that the assignment of performance values by each method differs markedly. For instance, the MSEs computed for the GRO fund group range from -0.67% to 0.73%, whereas its estimated PWMs vary from

Table IV. Quartiles for fund performance measures (in % per month) by fund groups, using the MSE approach and period weighting measure (PWM) approach. Both methods assume a conditional CAPM model. Sample period is for 1976:1–1991:12. The fund groups are: small funds, **SML**; aggressive growth funds, **AGG**; growth funds, **GRO**; growth and income funds, **G&I**; equity income funds, **EQI**; balanced funds, **BAL**; corporate bond funds, **CPB**. z-ratio tests the hypothesis that there is no difference between the average pricing error of passive assets and the average specification error-based measure for active mutual fund returns.

Fund group	No. of funds	Quartiles min	q_1	median	q_3	max
A: The conditional MSE approach, $y = R_b' D^* z_t$						
Average pricing error, $\sqrt{\lambda_c^2(R_p, D^*)/Nl} = 0.227$.						
Average performance measure, $\sqrt{\lambda_c^2(R_a, D^*)/M} = 0.352$.						
$H : \lambda_c^2(R_p, D^*)/Nl - \lambda_c^2(R_a, D^*)/M = 0$, z-test = −0.034.						
SML	14	−0.475	−0.349	−0.106	−0.023	0.473
AGG	17	−1.200	−0.411	−0.241	0.037	0.325
GRO	111	−0.673	−0.310	−0.152	0.035	0.734
G&I	60	−1.493	−0.435	−0.294	−0.203	0.237
EQI	11	−0.415	−0.369	−0.320	−0.209	0.040
BAL	22	−0.424	−0.357	−0.285	−0.219	−0.045
CPB	24	−0.531	−0.399	−0.363	−0.292	−0.263
B: The conditional PWM approach, $y = R_b' \tilde{D} z_t$						
Average pricing error, $\sqrt{\lambda_c^2(R_p, \tilde{D})/Nl} = 0.793$.						
Average performance measure, $\sqrt{\lambda_c^2(R_a, \tilde{D})/M} = 0.708$.						
SML	14	0.354	0.580	0.748	1.046	1.398
AGG	17	−0.052	0.528	0.905	0.978	1.278
GRO	111	0.323	0.582	0.710	0.903	1.622
G&I	60	−0.552	0.464	0.583	0.655	1.002
EQI	11	0.292	0.414	0.495	0.606	0.865
BAL	22	0.366	0.438	0.472	0.522	0.771
CPB	24	0.035	0.102	0.129	0.170	0.218

0.32% to 1.62% per month. And the monthly MSEs for the G&I group lie between −1.49% and 0.24%, while their monthly PWMs are between −0.55% and 1.0%. A similar observation is made in Table V, where most of the MSEs assigned to the funds are negative as compared to the positive PWMs reported in Panel B.

The average MSEs presented in Panel A of both Tables IV and V are quite close in magnitude, which is approximately 0.35% per month. This contrasts with the order of magnitude for the average PWMs – Table IV lists an average monthly value

Table V. Quartiles for fund performance measures (in % per month) by fund groups, using the MSE approach and the period weighting measure (PWM) approach. Both methods assume a conditional 6-factor model. Sample period is 1976:1–1991:12. The fund groups are: small funds, **SML**; aggressive growth funds, **AGG**; growth funds, **GRO**; growth and income funds, **G&I**; equity income funds, **EQI**; balanced funds, **BAL**; corporate bond funds, **CPB**. z-ratio tests the hypothesis that there is no difference between the average pricing error of passive assets and the average specification error-based measure for active mutual fund returns.

Fund group	No. of funds	Quartiles min	q_1	Median	q_3	max
A: The conditional MSE approach, $y = R_b' D^* z_t$						
Average pricing error, $\sqrt{\lambda_c^2(R_p, D^*)/Nl} = 0.180$.						
Average performance measure, $\sqrt{\lambda_c^2(R_a, D^*)/M} = 0.340$.						
$H: \lambda_c^2(R_p, D^*)/Nl - \lambda_c^2(R_a, D^*)/M = 0$, z-test = - 0.309.						
SML	14	−0.782	−0.395	−0.125	0.181	0.563
AGG	17	−1.579	−0.459	−0.086	0.219	0.870
GRO	111	−1.047	−0.165	−0.026	0.243	0.715
G&I	60	−1.015	−0.299	−0.185	−0.061	0.831
EQI	11	−0.445	−0.397	−0.340	−0.307	0.396
BAL	22	−0.372	−0.328	−0.192	−0.079	0.255
CPB	24	−0.495	−0.258	−0.148	−0.067	0.118
B: The conditional PWM approach, $y = R_b' \tilde{D} z_t$						
Average pricing error, $\sqrt{\lambda_c^2(R_p, \tilde{D})/Nl} = 0.426$.						
Average performance measure, $\sqrt{\lambda_c^2(R_a, \tilde{D})/M} = 0.284$.						
SML	14	−261	−0.153	0.021	0.118	0.863
AGG	17	−1.939	−0.293	−0.046	0.023	0.418
GRO	111	−0.774	−0.084	0.012	0.120	0.709
G&I	60	−1.388	−0.094	−0.010	0.089	0.355
EQI	11	−0.082	−0.045	0.019	0.076	0.394
BAL	22	−0.101	0.015	0.133	0.171	0.565
CPB	24	−0.014	0.139	0.201	0.253	0.470

of 0.71%, while Table V reports 0.28% per month. Again, however, these values have to be evaluated relative to the average pricing error of the model employed.

?'From Tables IV and V, we find that the magnitude of average pricing errors in Panel A is about half the size of that in Panels B. The MSE method yields an average pricing error of 0.23% per month (Panel A of Table IV), while the period weighting approach produces an average monthly error of 0.79% (Panel B of Table IV). Similarly, their counterparts in Table V are 0.18% and 0.43%, respectively. We attribute the difference in pricing errors generated from the two approaches to the inefficient benchmarks employed. As mentioned earlier, the two methodologies would engender the same size of pricing errors, if the benchmarks used are efficient. Otherwise, the more inefficient the benchmark proxies, the larger the disparity in the pricing error. Thus, we cannot reject the null that the average MSE (or average PWM) is the same as the average pricing error. The test statistics suggest that the funds on average do not outperform the passive assets.

Interestingly, we also observe that the larger the pricing error for the passive assets, the greater the size of the performance measures. The pricing errors appear to be positively correlated with the performance measures.

Overall, there is lack of evidence that mutual funds perform better than the passive portfolios. Our evidence of no abnormal conditional performance is in line with that of CK's conditional measures, but contradicts a recent finding by Ferson and Schadt (1996), who use conditional parametric models to evaluate fund performance. Like any other parametric model, the validity of their conditional models also depends on a particular parametrization of the model as well as on the efficiency of the benchmarks they adopt.

5. Concluding Remarks

We introduce a new performance measure based on the consideration of minimizing specification errors of asset pricing models and also derive the necessary statistical distribution for the estimator of this measure. Contrary to many existing performance evaluation techniques, this approach *a priori* assumes that the pricing error induced by a discount factor proxy, or by an inefficient benchmark proxy employed, is strictly positive. Our proposed minimum specification error-based measures, both unconditional and conditional, are constructed from a stochastic discount-factor proxy that has the minimal pricing error over a large number of passive portfolios. We apply these measures to evaluating mutual fund performance. To emphasize the importance of our contribution, we contrast our methodology with Grinblatt and Titman's (1989) and Chen and Knez's (1996) implementation methods, which rely on the efficiency of a small number of benchmark variables and may contain large specification errors for a large number of passive assets.

Results show that estimates of period-weighting measures are generally larger than those of minimum specification error-based measures. But the pricing error for passive portfolios are also higher for period-weighting measure than for the minimum specification error-based measure. This evidence suggests that large performance measures reflect a large specification error, rather than the true performance. Results also show that irrespective of the measures used, the performance values increase with the size of the errors derived from the passive assets. After taking into consideration the average pricing error resulted from the benchmark measure, the absolute performance values assigned to more than 75%

of the mutual funds in the sample become relatively small and economically insignificant. Even when we explicitly minimize the specification errors contained in asset pricing models, we cannot reject the null hypothesis that the pricing error of passive assets and the performance of mutual funds are the same. Given the observed large pricing errors, we argue that a plethora of performance measures used in the existing studies which do not take into account of specification errors, are not informative, nor are they meaningful.

Appendix: Proof of Propositions 1–4

Proof of Proposition 1.
?'From the Cauchy–Schwarz inequality,

$$\begin{aligned}\lambda_u(w'R_p, f(R_b,\delta)) &= w'E[R_p f(R_b,\delta)] - 1\\ &= w'[E(R_p f(R_b,\delta)) - 1_N]\\ &\leq \| w \| \| E(R_p f(R_b,\delta)) - 1_N \|\end{aligned}$$

with equality if and only if w is proportional to $E(R_p f(R_b,\delta)) - 1_N$. Therefore, for the given $f(R_b,\delta)$, the w that maximizes $\lambda_u(w'R_p, f(R_b,\delta))/ \| w \|$ is

$$w^* = k[E(R_p f(R_b,\delta)) - 1_N]$$

where $k = 1/[E(R_p f(R_b,\delta)) - 1_N]'1_N$ if the denominator is nonzero, or indeterminate if it is. In the former case, the portfolio is a usual one satisfying $w^{*'}1_N = 1$, and in the latter it is a zero cost portfolio. Now, substituting w^* back to λ_u, we have

$$\lambda_u(w^{*'}R_p, f(R_b,\delta))/ \| w^* \| = \| ER_p f(R_b,\delta) - 1_N \|$$

Therefore, minimizing $\lambda_u(w^{*'}R_p, f(R_b,\delta))/ \| w^* \|$ is the same as minimizing $\| ER_p f(R_b,\delta) - 1_N \|$, which is the same as minimizing $\| ER_p f(R_b,\delta) - 1_N \|^2$. Q.E.D.

Proof of Proposition 2.
For $y^+ = \max\{R_b'\delta, 0\}$, consistency of the estimator for $\Lambda_u^2(R_p)$ depends on the fact that $\lambda_u^2(R_p,\delta)$ and $\hat{\lambda}_{u,T}^2(R_p,\delta)$ are convex functions of δ and that sets of minimizers of $\lambda_u^2(R_p,\delta)$ and $\hat{\lambda}_{u,T}^2(R_p,\delta)$ are compact. The remaining proof of this proposition follows the proof of Proposition 2.1 of Hansen et al. (1995) Q.E.D.

Proof of Proposition 3.
Since $\sqrt{T}[\hat{\lambda}_{u,T}^2(R_p,\delta) - \lambda_u^2(R_p,\delta)]$ can be written as

$$\sqrt{T}\left[\frac{1}{T}\sum_1^T R_{p,t+1}y_{t+1}^+(\delta) + ER_{p,t+1}y_{t+1}^+(\delta) - 2\cdot 1_N\right]'$$

$$\times \left[\frac{1}{T} \sum_{1}^{T} R_{p,t+1} y_{t+1}^{+}(\delta) - E R_{p,t+1} y_{t+1}^{+}(\delta) \right],$$

from Assumptions 1–2,

$$\sqrt{T}[\hat{\lambda}_{u,T}^2(R_p, \delta) - \lambda_u^2(R_p, \delta)) \longrightarrow^D \mathcal{N}(0, \sigma^2), \tag{20}$$

where

$$\sigma^2 = 4[E R_{p,t+1} y_{t+1}^{+}(\delta) - 1_N]' V_1 [E R_{p,t+1} y_{t+1}^{+}(\delta) - 1_N],$$

with V_1 being the covariance of the limiting distribution of

$$\sqrt{T} \left[\frac{1}{T} \sum_{1}^{T} R_{p,t+1} y_{t+1}^{+}(\delta) - E[R_{p,t+1} y_{t+1}^{+}(\delta)] \right].$$

Since $\hat{\delta}_T^*$ and δ^* are solutions of $\min_\delta \hat{\lambda}_{u,T}^2(R_p, \delta)$ and $\min_\delta \lambda_u^2(R_p, \delta)$, we can make the following decomposition:

$$\begin{aligned}\sqrt{T}[\hat{\Lambda}_{u,T}^2(R_p) - \Lambda_u^2(R_p)] &= \sqrt{T}[\hat{\Lambda}_{u,T}^2(R_p) - \hat{\lambda}_{u,T}^2(R_p, \delta^*)] \\ &\quad + \sqrt{T}[\hat{\lambda}_{u,T}^2(R_p, \delta^*) - \Lambda_u^2(R_p)].\end{aligned}$$

The second term converges to a normal random variable, as we just proved in (20) for $\delta = \delta^*$. The rest of the proof shows that the first term converges to zero almost surely.

By definition, $\hat{\Lambda}_u^2(R_p) = \hat{\lambda}_{u,T}^2(R_p, \hat{\delta}_T^*) < \hat{\lambda}_{u,T}^2(R_p, \delta^*)$. Since $\hat{\lambda}_{u,T}^2(R_p, \delta)$ is a convex function of δ, we have

$$\hat{\lambda}_{u,T}^2(R_p, \hat{\delta}_T^*) - \hat{\lambda}_{u,T}^2(R_p, \delta^*) \geq q'(\hat{\delta}_T^* - \delta^*),$$

where q' is any subgradient of $\hat{\lambda}_{u,T}^2(R_p, \delta)$ at $\delta = \delta^*$. We choose a particular subgradient which can be represented by

$$2 \left[\frac{1}{T} \sum_{1}^{T} R_{p,t+1} R_{b,t+1}' \cdot I(y_{t+1}^{+}(\delta^*) > 0) \right]' \left[\frac{1}{T} \sum_{1}^{T} R_{p,t+1} y_{t+1}^{+}(\delta^*) - 1_N \right].$$

?'From Assumption 2, $A = (w\colon R_{b,t+1}'\delta = 0)$ is a set with zero measure for any given δ. Since $\lambda_u(R_p, \delta)^2$ is differentiable, the first order condition for δ^* to be minimal is

$$\frac{\partial}{\partial \delta} \lambda_u^2(R_p, \delta^*) = [E R_{p,t+1} R_{b,t+1}' I(y_{t+1}^{+}(\delta^*) > 0)]' [E R_{p,t+1} y_{t+1}^{+}(\delta^*) - 1_N] = 0,$$

Thus, by adding a zero,

$$\sqrt{T}q = 2\sqrt{T}\left[\left(\frac{1}{T}\sum_{1}^{T} R_{p,t+1}R'_{b,t+1}I(y^{+}_{t+1}(\delta^*) > 0)\right)'\right.$$

$$\left.\times\left(\frac{1}{T}\sum_{1}^{T} R_{p,t+1}y^{+}_{t+1}(\delta^*) - 1_N\right) - \frac{\partial}{\partial\delta}\lambda^2_u(R_p,\delta^*)\right]$$

$$= 2\sqrt{T}\left[\left(\frac{1}{T}\sum_{1}^{T} R_{p,t+1}R'_{b,t+1}I(y^{+}_{t+1}(\delta^*) > 0)\right.\right.$$

$$\left. - ER_{p,t+1}R'_{b,t+1}I(y^{+}_{t+1}(\delta^*) > 0)\right)'$$

$$\times\left(\frac{1}{T}\sum_{1}^{T} R_{p,t+1}y^{+}_{t+1}(\delta^*) - 1_N\right) + \left(ER_{p,t+1}R'_{b,t+1}I(y^{+}_{t+1}(\delta^*) > 0)\right)'$$

$$\left.\times\left(\frac{1}{T}\sum_{1}^{T} R_{p,t+1}y^{+}_{t+1}(\delta^*) - E[R_{p,t+1}y^{+}_{t+1}(\delta^*)\right)\right].$$

By Assumptions 1 and 2, $\sqrt{T}q$ converges to a normally distributed random vector. Hence,

$$\sqrt{T}q'(\hat{\delta}^*_T - \delta^*) \xrightarrow{\text{a.s.}} 0,$$

¿From $0 \geq \sqrt{T}[\hat{\lambda}^2_{u,T}(R_p,\hat{\delta}^*_T) - \hat{\lambda}^2_{u,T}(R_p,\delta^*)] \geq \sqrt{T}q'(\hat{\delta}^*_T - \delta^*)$, we have

$$\sqrt{T}[\hat{\Lambda}^2_{u,T}(R_p) - \hat{\lambda}^2_{u,T}(R_p,\delta^*)] \xrightarrow{\text{a.s.}} 0.$$

Q.E.D.

Proof of Proposition 4.
Given

$$\sqrt{T}[\hat{\lambda}^2_{u,T}(R_a,\hat{\delta}^*_T) - \hat{\lambda}^2_{u,T}(R_a,\delta^*)]$$

$$= \sqrt{T}\left[\frac{1}{T}\sum_{1}^{T} R_{a,t+1}(y^{+}_{t+1}(\hat{\delta}^*_T) + y^{+}_{t+1}(\delta^*)) - 2\cdot 1_M\right]'$$

$$\times \left[\frac{1}{T} \sum_{1}^{T} R_{a,t+1}(y_{t+1}^{+}(\delta_T^*) - y_{t+1}^{+}(\delta^*)) \right],$$

in order to show that $\sqrt{T}[\hat{\lambda}_{u,T}^2(R_a, \hat{\delta}_T^*) - \hat{\lambda}_{u,T}^2(R_a, \delta^*)]$ converges to a normal random variable with mean zero, we only need to show that $\sqrt{T}[\frac{1}{T}\sum_1^T R_{a,t+1}(y_{t+1}^{+}(\hat{\delta}_T^*) - y_{t+1}^{+}(\delta^*))]$ converges to a normal random vector. Note that

$$\frac{1}{T} \sum_{1}^{T} R_{a,t+1}(y_{t+1}^{+}(\hat{\delta}_T^*) - y_{t+1}^{+}(\delta^*)) = \frac{1}{T} \sum_{t \in T_1} R_{a,t+1} y_{t+1}^{+}(\hat{\delta}_T^*)$$

$$- \frac{1}{T} \sum_{t \in T_2} R_{a,t+1} y_{t+1}^{+}(\delta^*)$$

$$+ \frac{1}{T} \sum_{t \in T_3} R_{a,t+1} R'_{b,t+1}(\hat{\delta}_T^* - \delta^*),$$

where

$$T_1 = \{t : y_{t+1}^{+}(\hat{\delta}_T^*) > 0,\ y_{t+1}^{+}(\delta^*) = 0\},$$

$$T_2 = \{t : y_{t+1}^{+}(\hat{\delta}_T^*) = 0,\ y_{t+1}^{+}(\delta^*) > 0\},$$

$$T_3 = \{t : y_{t+1}^{+}(\hat{\delta}_T^*) > 0,\ y_{t+1}^{+}(\delta^*) > 0\}.$$

For $t \in T_1$, $0 < y_{t+1}^{+}(\hat{\delta}_T^*) = R'_{b,t+1}\hat{\delta}_T^*$, and $R'_{b,t+1}\delta^* \leq 0$. Hence,

$$0 < \frac{1}{T} \sum_{t \in T_1} R_{a,t+1} y_{t+1}^{+}(\hat{\delta}_T^*) \leq \frac{1}{T} \sum_{t \in T_1} R_{a,t+1} R'_{b,t+1}(\hat{\delta}_T^* - \delta^*).$$

Assumption 3 suggests that

$$\frac{1}{T} \sum_{t \in T_1} R_{a,t+1} R'_{b,t+1} \xrightarrow{\text{a.s.}} \lim_{T \to \infty} E[I(A_{T_1}) R_{a,t+1} R'_{b,t+1}].$$

Since $\hat{\delta}_T^* \longrightarrow^{\text{a.s.}} \delta^*$,

$$\lim_{T \to \infty} E[I(A_{T_1}) R_{a,t+1} R'_{b,t+1}] = 0.$$

As $\sqrt{T}(\hat{\delta}_T^* - \delta^*)$ converges to a normal random vector by Assumption 3,

$$0 < \sqrt{T} \left[\frac{1}{T} \sum_{t \in T_1} R_{a,t+1} y_{t+1}^{+}(\hat{\delta}_T^*) \right] \leq \left[\frac{1}{T} \sum_{t \in T_1} R_{a,t+1} R'_{b,t+1} \right]$$

$$\sqrt{T}(\hat{\delta}_T^* - \delta^*) \xrightarrow{\text{a.s.}} 0.$$

Similarly, we can show that

$$\frac{1}{T}\sum_{t\in T_2} R_{a,t+1} y_{t+1}^{+}(\delta^*) \xrightarrow{\text{a.s.}} 0.$$

Based on these results, $\sqrt{T}[\frac{1}{T}\sum_1^T R_{a,t+1}(y_{t+1}^{+}(\hat{\delta}_T^*) - y_{t+1}^{+}(\delta^*))]$ has the same limiting distribution as $(\frac{1}{T}\sum_{t\in T_3} R_{a,t+1}R'_{b,t+1})\sqrt{T}(\hat{\delta}_T^* - \delta^*)$, and based on Assumption 3,

$$\frac{1}{T}\sum_{t\in T_3} R_{a,t+1}R'_{b,t+1} \xrightarrow{\text{a.s.}} E[I(A_{T_3})R_{a,t+1}R'_{b,t+1}].$$

Since $\sqrt{T}(\hat{\delta}_T^* - \delta^*)$ converges to a normal random vector, $[\frac{1}{T}\sum_{t\in T_3} R_{a,t+1}R'_{b,t+1}]$ $\sqrt{T}(\hat{\delta}_T^* - \delta^*)$ converges to a normal random vector. Q.E.D.

References

Brown, S., Goetzmann, W., Ibbotson, R., and Ross, S. (1992) Survivorship bias in performance studies, *Review of Financial Studies* **5**, 553–580.

Campbell, J. Y. (1987) Stock returns and the term structure, *Journal of Financial Economics* **18**, 373–400.

Chen, Z. and Knez, P. J. (1996) Portfolio performance measurement: Theory and applications, *Review of Financial Studies* **9**, 511–555.

Cochrane, J. H. (1996) A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* **104**, 572–621.

Chen, N. F., Roll, R. R., and Ross, S. A. (1986) Economic forces and the stock market, *Journal of Business* **59**, 383–403.

Connor, G. and Korajczyk, R. (1991) The attributes, behavior, and performance of U.S. mutual funds, *Review of Quantitative Finance and Accounting* **1**, 5–26.

Cumby, R. E. and Glen, J. D. (1990) Evaluating the performance of international mutual funds, *Journal of Finance* **45**, 497–521.

Elton, E., Gruber, M., Das, S., and Hlavka, M. (1993) Efficiency with costly information: A reinterpretation of evidence from managed portfolios, *Review of Financial Studies* **6**, 1–22.

Fama, E. F. and French, K. R. (1988) Dividend yields and expected stock returns, *Journal of Financial Economics* **22**, 3–26.

Fama, E. F. and French, K. R. (1989) Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* **25**, 23–49.

Fama, E. F. and French, K. R. (1992) The cross-section of expected stock returns, *Journal of Finance* **47**, 427–465.

Ferson, W. (1989) Changes in expected security returns, risk and the level of interest rates, *Journal of Finance* **44**, 1191–1218.

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

Glosten, L. and Jagannathan, R. (1994) A contingent claim approach to performance evaluation, *Journal of Empirical Finance* **1**, 133–166.

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

Grinblatt, M. and Titman, S. (1994) A study of monthly mutual fund returns and performance evaluation techniques, *Journal of Financial and Quantitative Analysis* **29**, 419–444.

Hansen, L., Heaton, J., and Luttmer, E. (1995) Econometric evaluation of asset pricing models, *Review of Financial Studies* **8**, 237–274.

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

Hansen, L. and Jagannathan, R. (1997) Assessing specification errors in stochastic discount factor models, *Journal of Finance*, forthcoming.

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

He, J. and Ng, L. (1994) Economic forces, fundamental variables, and equity returns, *Journal of Business* **67**, 600–609.

Ingersoll, J. (1987) *Theory of Financial Decision Making*, Rowman and Littlefield.

Ippolito, R. (1989) Efficiency with costly information: A study of mutual fund performance, *Quarterly Journal of Economics* **104**, 1–23.

Jensen, M. C. (1968) The performance of mutual funds in the period 1945–1964, *Journal of Finance* **23**, 89–416.

Keim, D. B. and Stambaugh, R. F. (1986) Predicting returns in the stock and bond markets, *Journal of Financial Economics* **17**, 357–390.

Lehmann, B. and Modest, D. (1987) Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons, *Journal of Finance* **42**, 233–265.

Roll, R. (1977) A critique of the asset pricing theory's test; Part I: On past and potential testability of the theory, *Journal of Financial Economics* **4**, 129–176.

Roll, R. (1978) Ambiguity when performance is measured by the securities market line, *Journal of Finance* **33**, 1051–1069.

Ross, S. A. (1978) A simple approach to the valuation of risky streams, *Journal of Business* **51**, 453–475.

Sharpe, W. F. (1992) Asset allocation: Management style and performance measurement, *Journal of Portfolio Management* 7–19.