NBER WORKING PAPER SERIES

PERFORMANCE EVALUATION WITH STOCHASTIC DISCOUNT FACTORS

Heber Farnsworth
Wayne E. Ferson
David Jackson
Steven Todd

Working Paper 8791
http://www.nber.org/papers/w8791

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2002

We are grateful to Mark Carhart, Martin Gruber, Campbell Harvey, Don Keim, Jennifer Koski, Robert Korajczyk and Rudi Schadt for help with data, and to Mark Grinblatt, Phil Dybvig, Andrew Karolyi, Paul Irvine, Peter Knez, Jeff Pontiff and Stephen D. Smith for comments. Bernard Yomtov worked closely with us on earlier drafts of the paper. This study has also benefited from research workshops at the following universities and conferences: Alberta, Arizona State, Berkeley, a Copenhagen Business School/Lund University joint workshop, Miami, North Carolina, the Norwegian School of Management, Ohio State, Southern California, Texas at Austin, Vienna, Washington, the Atlanta Finance Forum at Georgia State, the 1996 American Finance Association, the 1999 Asia Pacific Finance Association, the Fourth Conference on Pacific Basin Business, Economics and Finance, the 1995 National Bureau of Economic Research Summer Institute, the 1995 Northern Finance Association and the 2000 Utah Winter Finance Conference. Part of this research was conducted while Ferson was a visiting scholar at Arizona State University and a National Bureau of Economics Research Associate. Ferson also acknowledges financial support from the Pigott-PACCAR professorship at the.University of Washington. Previous drafts of the paper were entitled, "Conditional Performance Evaluation." The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Performance Evaluation with Stochastic Discount Factors
Heber Farnsworth, Wayne E. Ferson, David Jackson and Steven Todd
NBER Working Paper No. 8791
February 2002
JEL No. G12, G14, G23, C15

ABSTRACT

We study the use of stochastic discount factor (SDF) models in evaluating the investment performance of portfolio managers. By constructing artificial mutual funds with known levels of investment ability, we evaluate a large set of SDF models. We find that the measures of performance are not highly sensitive to the SDF model, and that most of the models have a mild negative bias when performance is neutral. We use the models to evaluate a sample of U.S. equity mutual funds. Adjusting for the observed bias, we find that the average mutual fund has enough ability to cover its transactions costs. Extreme funds are more likely to have good rather than poor risk adjusted performance. Our analysis also reveals a number of implementation issues relevant to other applications of SDF models.

Heber Farnsworth
John M. Olin School of Business
Washington University in St. Louis
Campus Box 1133
One Brookings Drive
St. Louis, MO 63130-4899
Tel: 314-935-4221
Fax: 314-935-6359
heberf@calvin.wustl.edu

Wayne E. Ferson
Department of Finance and Business Economics
School of Business Administration
University of Washington
Box 353200
Seattle, WA 98195-3200
Tel: 206-543-1843
Fax: 206-685-9392
and NBER

David Jackson
Carleton University School of Business
1125 Colonel By Drive
Ottawa K1S SB6
Ontario, Canada
Tel: 613-520-2600 x2383
Fax: 613-520-4427
djackson@business.carleton.ca

Steven Todd
Department of Finance
Loyola University
25 East Pearson, 12th Floor
Chicago, IL 60611
Tel: 312-915-7218
stodd@wpo.it.luc.edu

I. Introduction

The question of whether mutual fund managers can deliver expected returns in excess of naive benchmarks has long been controversial. If fund managers can "beat the market," it has implications for the efficiency of financial markets. If they underperform, it has implications for the structure of the fund management industry. From an investor's perspective, the problem is to choose from a large universe of investment alternatives. For these reasons, measuring the investment performance of fund managers remains an important research problem.

In this paper we study the use of stochastic discount factor (SDF) models in evaluating the investment performance of portfolio managers. With this approach abnormal performance is measured by the expected product of a fund's returns and a stochastic discount factor. Specifying the SDF corresponds to specifying an asset-pricing model.

A variety of models for SDFs have been developed in previous studies. Our goal is to provide empirical evidence on the performance of a wide set of models, using a common experimental design. This is important because inferences about abnormal performance will generally depend on the SDF chosen. Some models may attribute abnormal performance to particular types of naive trading strategies, and the power to detect truly superior performance will differ across models.

We evaluate the models on a set of artificial mutual funds, where we control the extent of market timing or security selection ability. We find the performance measures are not highly sensitive to the choice of the SDF, excluding the few models that perform poorly on our test assets. Many of the models are biased, producing small negative alphas when the true performance is neutral. The average bias is about 0.15% per month when the artificial fund selects stocks randomly. Previous evidence based on SDFs that U.S. equity mutual funds have negative abnormal returns, reflects a biased measure (Chen and Knez, 1996). Most of the models have sufficient power to detect truly superior ability, and no single model vastly out performs the others.

We use the models to evaluate performance in a monthly sample of 188 equity mutual funds. We find that the average mutual fund alpha, measured net of expenses and trading costs, is no worse than a hypothetical stock-picking fund with neutral performance. Since the artificial funds pay no transactions costs or management fees, our results suggest that the average mutual fund has enough ability to cover these costs. Thus, using a wide set of SDF models and more recent data, our findings are broadly consistent with Jensen (1968), who used the CAPM. We also find that extreme funds are more likely to have good, rather than poor, risk adjusted performance.

Our analysis produces a number of useful discoveries relevant to general applications of SDF models. First, the performance measure for a given fund is invariant to the number of funds in the system. Second, SDF models perform better when a risk-free asset is included as a test asset, as this helps to identify the conditional mean of the SDF. Third, when the SDF is based on traded factors, it is important to require the model to "price" the traded factors. Finally, compared to their unconditional counterparts, conditional models (i.e, those that used lagged instruments) deliver smaller average pricing errors for the returns of "dynamic strategies" based on the instruments. However, the cost of these smaller average pricing errors is larger variances for the pricing errors on the original test assets.

The remainder of this paper is organized as follows. Section II reviews performance evaluation with stochastic discount factors. Section III describes the models and empirical methods. Section IV describes how we construct the artificial mutual funds. Section V describes the data. Section VI presents results on the estimation of the stochastic discount factor models and Section VII presents our evaluation of the models in the context of fund performance evaluation. Section VIII uses the models to evaluate performance in a sample of mutual funds, and section IX offers concluding remarks.

II. Performance Evaluation with Stochastic Discount Factors

A central goal of performance evaluation is to identify those managers who possess investment information or skills superior to that of the general investing public, and who use these advantages to achieve superior portfolio returns. In order to identify superior returns, some model of "normal" investment returns is required; that is, an asset pricing model is needed. Modern asset pricing theory posits the existence of a *stochastic discount factor*, m_{t+1}, which is a scalar random variable, such that the following equation holds:

$$E\left(m_{t+1}R_{t+1} - \underline{1} \middle| \Omega_t\right) = 0, \qquad (1)$$

where R_{t+1} is the vector of primitive asset gross returns (payoff divided by price) and $\underline{1}$ is an N-vector of ones. Ω_t denotes an information set available at time t and $E\left(\cdot \middle| \Omega_t\right)$ denotes the conditional expectation. Virtually all asset pricing models may be viewed as specifying a particular stochastic discount factor, m_{t+1}. The elements of the vector $m_{t+1} R_{t+1}$ may be viewed as "risk adjusted" gross returns. The returns are risk adjusted by "discounting" them, or multiplying by m_{t+1}, so that the expected "present value" per dollar invested is equal to one dollar. Thus, m_{t+1} is called a stochastic discount factor (SDF). We say that a SDF "prices" the assets if equation (1) is satisfied.

A. Conditioning Information

Empirical work on conditional asset pricing uses predetermined information variables, Z_t, which are elements of the public information set Ω_t. By the law of iterated expectations, equation (1) holds when we replace Ω_t with Z_t, and we are interested in equation (2):

$$E(m_{t+1} R_{t+1} - \underline{1} \mid Z_t)=0. \qquad (2)$$

According to equation (2), it should not be possible to detect mispricings of the primitive assets using only the information in Z_t. A conditional approach to performance evaluation allows a researcher to set the standard for what is "superior" information by choosing the public information Z_t. When Z_t is restricted to a constant we have an unconditional measure. With an unconditional measure, *any* information about future returns is assumed to be superior information that may generate abnormal performance.

B. Measuring Performance

For a given SDF we define a fund's conditional alpha similar to Chen and Knez (1996) as:[1]

$$\alpha_{pt} \equiv E\left(m_{t+1}R_{p,t+1}|Z_t\right) - 1, \qquad (3)$$

where one dollar invested with the fund at time t returns $R_{p,t+1}$ dollars at time t+1. In the case of an open-end, no-load mutual fund, we may think of $R_{p,t+1}$ as the net asset value return. More generally, if the fund generates a payoff $V_{p,t+1}$ for a cost $c_{pt} > 0$, the gross return is $R_{p,t+1} \equiv V_{p,t+1}/c_{pt}$.

If the SDF prices the primitive assets, α_{pt} will be zero when the fund (costlessly) forms a portfolio of the primitive assets, provided the portfolio strategy uses only the public information at time t. In that case $R_{p,t+1} = x(Z_t)'R_{t+1}$, where $x(Z_t)$ is the portfolio weight vector. Then equation (2) implies that

$$\alpha_{pt} = \left[E\left(m_{t+1}x(Z_t)' R_{t+1}|Z_t\right)\right] = x(Z_t)'\left[E\left(m_{t+1}R_{t+1}|Z_t\right)\right] = x(Z_t)'\underline{1} - 1 = 0.$$

In many models m_{t+1} is the intertemporal marginal rate of substitution for a representative investor, and equation (2) is the Euler equation which must be satisfied in equilibrium. If the consumer has access to a fund for which the conditional alpha is not zero, he will wish to adjust his portfolio, purchasing more of the fund if alpha is positive and less if alpha is negative. This generalizes the interpretation of the traditional Jensen's alpha as a guide for marginal mean-variance improving portfolio choices.

---

[1] Chen and Knez (1996) study an excess return conditional-alpha, $E(m_{t+1}r_{p,t+1}|Z_t)$, where $r_{pt}=R_{pt}-R_{ft}$ and R_{ft} is a Treasury bill return. Our definition of α_p is based on raw returns. The excess return alpha for a fund is simply the difference between the raw return alpha of the fund and the raw return alpha of the Treasury bill. When the alpha of the bill is zero the two are equal.

The SDF alpha can correctly indicate abnormal performance, but not without further restrictive assumptions. In particular, α_{pt} depends on the SDF chosen, and the SDF is not unique unless markets are complete. Thus, different SDFs can produce different measured performance. This mirrors a problem in classical approaches to performance evaluation, wherein performance is sensitive to the benchmark. Roll (1978), Dybvig and Ross (1985), Brown and Brown (1987), Chen, Copeland and Mayers (1987), Lehman and Modest (1987) and Grinblatt and Titman (1989) address this issue. Given these ambiguities, it is important to assess the sensitivity of performance measures to the specification of the SDF. Our goal in this study is to compare a number of models in a unified setting.

III. The Models and Methods

A. The General Approach

We estimate SDFs using the Generalized Method of Moments (GMM, Hansen, 1982) on the following moment conditions, which follow from equation (2):

$$E\left[\{m_{t+1}R_{t+1} - 1\} \quad Z_t\right] = 0. \tag{4}$$

The stochastic discount factors that we evaluate are listed below.

Linear factor: $m^{LF}_{t+1} = a(Z_t) + b(Z_t)'F_{t+1},$ (5)

$= Z_t'B\,(1,F_{t+1}')'$.

Primitive-efficient: $m^{PE}_{t+1} = M(Z_t)'R_{t+1},$ (6)

$$= (M Z_t)'R_{t+1}.$$

Numeraire portfolio: $$m^N_{t+1} = [(AZ_t)'R_{2,t+1} + (1-(AZ_t)'\underline{1})R_{1t+1}]^{-1}, \qquad (7)$$

Bakshi-Chen: $$m^{BS}_{t+1} = \exp\{ Z_t'C \ln(R_{t+1})\}. \qquad (8)$$

In these equations, R_{t+1} is an N-element vector of the gross returns on a set of primitive assets and Z_t is an L vector of lagged instruments. A, B, C, and M denote the parameters of the various models. In equation (7) we partition $R_{t+1}=(R_{1,t+1},R_{2,t+1})$, where $R_{2,t+1}$ is an (N-1)-element vector. A brief description of each model appears below.

B. Linear Factor Models

Models in which m_{t+1} is linear in prespecifed factors are known as linear factor models. The Capital Asset Pricing Model is one such model in which m_{t+1} is a linear function of the market portfolio return (Dybvig and Ingersoll, 1982). Linear factor models can be unconditional or conditional. In the conditional versions of the models, we follow Dumas and Solnik (1995) and Cochrane (1996), and assume that the weights a(.) and b(.) in equation (5) are linear functions of Z_t. Thus, in equation (5) B is an L x (K+1) matrix of parameters, where K is the number of factors. To identify the parameters, we require $N \geq K + 1$.

We consider two sets of linear factor models: one based on nontraded factors (e.g. industrial production) and another based on traded factors (e.g. the S&P 500 index). For the traded factor models, we impose the restriction that the model price the traded factors.[2] For both the traded and non-traded factor models, we impose the restriction that the model price the risk-free asset.[3]

Typically, beta pricing models are estimated by least squares or maximum likelihood methods, while SDF models use the GMM. In SDF models with nontraded factors, Kan and Zhou (1999) show that GMM estimates of the risk premium for a factor are imprecise relative to OLS or MLE estimates, when the mean and variance of the factor are known. Jagannathan and Wang (2000) show that the efficiency of GMM estimates of factor premiums can be identical to OLS when additional moment conditions, identifying the mean and factor variance, are appended to the system. Our empirical evidence shows how the pricing accuracy of SDF models can be enhanced through ancillary moment conditions. We find, in experiments not reported in the tables, that the precision of the traded-factor models' parameters is lower and the pricing errors are much larger when the models are not forced to price the factors. The non-traded factor models, in particular, are much less accurate when they aren't forced to price the risk-free asset.

---

[2] For example, in the unconditional CAPM, $m_{t+1} = a + bR_{m,t+1}$, where $R_{m,t+1}$ is the gross market return. Requiring the model to price the market return and also a zero beta return we have:

$$E\{[a+bR_{m,t+1}]R_{m,t+1}\}=1 \text{ and } E\{[a+bR_{m,t+1}]R_{0t+1}\}=1.$$

These two conditions identify the parameters a(.) and b(.) in equation (5) as functions of the first and second moments of the market index and the zero beta return. Similar restrictions apply to multifactor and conditional models, as shown by Ferson and Jagannathan (1996).

[3] In other words, we impose the condition that $E(m_{t+1}RF_t - 1|Z_t)=0$, where RF_t is the gross risk-free return. Since RF_t is included in Z_t, this condition identifies the conditional mean of the SDF: $E(m_{t+1}|Z_t) = R_{ft}^{-1}$.

C. Primitive-Efficient Stochastic Discount Factors

Consider a conditional projection of an m_{t+1} that satisfies equation (2) onto the vector of primitive returns R_{t+1}. The solution is:

$$m^{PE}_{t+1} = \underline{1}' \, E(R_{t+1} R_{t+1}' \,|\, Z_t)^{-1} R_{t+1}. \qquad (9)$$

We call the stochastic discount factor m^{PE}_{t+1} a *primitive-efficient stochastic discount factor*. This term reflects the fact that m^{PE}_{t+1} is a linear function of a conditionally minimum-variance efficient portfolio.[4]

We follow Chen and Knez (1996) and Dahlquist and Soderlind (1999), who assume that the weights $M(Z_t) = \underline{1}' \, E(R_{t+1} R_{t+1}' \,|\, Z_t)^{-1}$ in equation (6) are linear functions of Z_t. Thus, M is an N x L matrix and L is the dimension of Z. The system is exactly identified, with NL parameters and NL orthogonality conditions.[5] With this assumption, the primitive-efficient model is comparable to linear factor models, in the sense that the conditional model is equivalent to a "scaled" unconditional model. By multiplying the primitive assets by the lagged instruments, we form "dynamic

[4] Specifically, the minimum variance portfolio has its target mean chosen to minimize the uncentered second moment of the portfolio return. Grinblatt and Titman (1989) propose an unconditionally mean-variance efficient portfolio as a benchmark for performance measurement. Chen and Knez (1996) develop primitive efficient SDFs for performance evaluation and Dahlquist and Soderlind (1999) study their sampling properties by simulation. He, Ng and Zhang (1998) specialize the approach to handle a large number of primitive assets.

[5] The solution for M may, in this case, be obtained in closed form by manipulating equation (4). The solution is given by:

$$Vec(M) = (HH')^{-1}(Z' \quad \underline{1}_N)\underline{1}_T,$$

where H is the NL x T matrix formed by putting the NL vectors $(Z_t \quad R_{t+1})$ in its columns, and Z is the T x K matrix of the Z_t's. We require T > NL to invert the (HH') matrix.

strategies." The conditional model will price these dynamic strategies in the sample, by construction.

D. Numeraire portfolios

Long (1990) proposes a model in which the SDF is the inverse of the gross rate of return on a "numeraire portfolio." He shows that if there are no arbitrage opportunities, then *some* numeraire portfolio exists. Kang (1995) uses a numeraire portfolio to evaluate fixed income mutual funds and Hentschell, Kang and Long (1998) use the approach for international bonds. We estimate numeraire portfolios using equation (7). R_{1t+1} is the first primitive asset, which we take to be the Treasury bill, and A is an (N-1)xL matrix of parameters. In the conditional version of the model, we assume that the weights are linear functions of the lagged instruments. Here, the SDF is a nonlinear function, and the conditional model is not equivalent to an unconditional model applied to the dynamic strategy returns, $(R_{t+1} \quad Z_t)$.[6]

E. The Bakshi-Chen Model

Bakshi and Chen (1998) propose a model in which the SDF is an exponential of a linear function of the log returns on the primitive assets. This formulation has the potential advantage that the SDF is constrained to be positive. The existence of *some*

[6] The moment conditions of (7) must be modified because the gross asset return matrix multiplied by the numeraire portfolio has a linear combination of columns that is nearly a vector of ones, and the GMM weighting matrix is singular. To resolve this problem we use N-1 primitive assets in the formation of the numeraire portfolio, and we ask the SDF to price N primitive assets. (In joint estimation with a fund, we use N-2 of the primitive assets in forming the numeraire portfolio.) We also follow Long (1990) and Kang (1995) by using nonlinear least squares, where the GMM weighting matrix is the identity matrix. Such estimates are consistent but not efficient. We experimented with a full GMM approach but found the estimates to be numerically unstable; this is one practical disadvantage of the numeraire portfolio approach.

strictly positive SDF is equivalent to a lack of arbitrage opportunities in a perfect market. (Lack of arbitrage does not require that *all* SDFs are strictly positive.) The Bakshi-Chen formulation, like the numeraire portfolio model, is nonlinear in the primitive asset returns.[7]

F. Measuring Mutual Fund Performance

Our approach for estimating alphas is to form a system of equations as follows:

$$u1_t = [m_{t+1}R_{t+1} - \underline{1}] \quad Z_t$$

$$u2_t = \alpha_p - m_{t+1}R_{p,t+1} + 1 \qquad (10)$$

The sample moment condition is $g = T^{-1}\sum_t (u1_t', u2_t')'$. We use the GMM to simultaneously estimate the parameters of the SDF model and the fund's alpha. The parameter α_p is the mean of the conditional alpha, defined by equation (3). Thus, we examine the average performance of a fund.[8]

A potential problem with this simultaneous approach is that the number of moment conditions grows substantially if many funds are to be evaluated. We therefore estimate the joint system separately for each fund.[9] Separate estimation is

[7] We experimented with imposing nonnegativity in the other models, using a differentiable approximation to the function Max(0,m) that was proposed by Bansal, Hsieh and Viswanathan (1993). This produced numerically unstable estimates and dramatically increased the pricing errors. We also examined exponential models, similar to Bakshi and Chen (1998), using our three traded and four nontraded factors. These models performed much worse than the models which use the primitive assets to form the SDF.

[8] For a discussion of time-varying conditional alphas, see Christopherson, Ferson and Glassman (1998).

[9] We evaluated a simpler two-step approach where the SDF is estimated in a first step, and the fund abnormal return is measured in the second step by simply multiplying the gross fund return by the fitted SDF and subtracting one. The

not restrictive, however. We show in the Appendix that the estimate of a fund's alpha and its standard error are invariant to the number of funds in the system. Thus, estimating the system for one fund at a time is equivalent to estimating a system with all of the funds simultaneously.

IV. Artificial Mutual Funds

We construct artificial mutual funds with varying amounts of known performance, using stock returns from the CRSP data files, for the sample period July 1963 through December, 1994. We characterize the funds as either stock pickers or market timers. A stock-picking fund gets noisy signals about the "non-market" component of the future returns of many stocks, and selects securities on the basis of these signals. An artificial market-timing fund receives a noisy signal about the future return of the Standard and Poors 500 index, and invests in the index or in Treasury bills.[10] When the signals are completely random, the managers have no ability. We use this case to evaluate the performance statistics under the null hypothesis of no abnormal performance. Then, we evaluate the various measures given known levels of ability.

---

average performance is the sample mean of the abnormal return. Such two-step estimators are consistent, but the standard errors do not account for the first stage estimation error in the parameters of the SDF. We found that the two-step approach is remarkably less efficient than the joint estimation procedure.

[10] This abstraction of market timing is consistent with Merton and Henriksson (1981) but it does not capture the range of trades that may actually be employed by market-timing funds. Our objective in drawing this characterization is partly to provide artificial returns that are likely to have different properties from those of our artificial stock pickers; hence, the extreme strategy of shifting between the market index and cash.

A. Artificial Stock Pickers

An artificial stock picker starts with a value weighted portfolio of 200 stocks randomly chosen from the largest 1,000 companies, ranked by market value as of January, 1963. We run the model from January through June and discard the first six months, using the returns for July, 1963-December, 1994 in the analysis. The portfolio is updated monthly, with buy and sell candidates chosen according to the following model:

$$\text{SIGNAL}_{i,t-1} = \gamma\ \varepsilon_{it} + (1-\gamma)\Phi\sigma(\varepsilon_{it}) \qquad (11)$$

Here, the signal received for stock *i* at time t-1 is a convex combination of an information term, ε_{it}, and a noise term, $\Phi\sigma(\varepsilon_{it})$. The quality of the signal is determined by the parameter, $\gamma\ (0 \le \gamma \le 1)$. The information term for stock *i* is the residual for the next month from a market model regression: $r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it}$ where r_{mt} is the CRSP value-weighted market return in excess of a one-month Treasury bill. The noise component of a stock picker's signal is the product of an independent N(0,1) random variable, Φ, and the standard error of the market model residual ε_{it} for firm i, $\sigma(\varepsilon_{it})$. A manager with perfect stock picking ability $(\gamma = 1)$ observes the market model residual one period ahead. A stock picker with no ability $(\gamma = 0)$ observes a random signal.

For later interpretation of the empirical results, it is useful to relate the parameter, γ, to a common measure of investment information. Active portfolio managers use the *information coefficient*, IC, defined as the correlation between the signal and subsequent returns. (See, for example, Grinold and Kahn, 1995.) Here, the correlation between the signal and the security specific return, ε_{it}, is $\gamma \Big/ \sqrt{\gamma^2 + (1-\gamma)^2}$. Thus, if $\gamma = 0$ the correlation is zero and if $\gamma = 0.8$ the correlation is 0.97. We find below that the best performing actual mutual funds (those in the upper 5%) have performance similar to artificial mutual funds with γ values of 0.20 to 0.25, or IC values of 0.24 to 0.32.

Each month, the artificial stock picker receives a signal for each stock in the available universe, which includes all NYSE, AMEX and NASDAQ stocks on the CRSP files with stock prices larger than $2.00 per share.[11] Each month, the portfolio manager turns over 6% of the value of the portfolio.[12] All stocks in the universe are first ranked on the basis of their signals. Stocks within the portfolio with the smallest SIGNAL values are sold, until 6% of the portfolio has been sold. Stocks outside the portfolio with the largest SIGNAL values are purchased. The total number of stocks in the portfolio is fixed at 200, so for each stock sold a new stock is added. When more than one stock is purchased in a given month, the weights on the purchased stocks are equal. The weights of stocks remaining in the portfolio evolve through time based on the returns earned, with dividends reinvested.

B. Artificial Market Timers

It is well known that classical measures of alpha are biased and otherwise difficult to interpret in the presence of market timing behavior [e.g. Grant (1977), Grinblatt and

[11] We exclude low price stocks to keep the universe representative of equity mutual fund holdings, and to avoid the extreme return patterns observed for such stocks. For missing data we assign a return of zero, with the following exceptions. For delisted stocks with missing return data, (CRSP codes of 500 and 520-584), we follow Shumway (1997) and assign a return of -30%. For liquidated stocks with missing return data (CRSP code 400), we assign a return of -100%.

[12] We arrived at this figure based on the 73% mean annual turnover of growth funds for the period January, 1976 through December, 1992 from the January, 1993 Morningstar, Inc. compact disk database.

Titman (1989)]. Specific market-timing models, such as Treynor and Mazuy (1966) and Merton and Henriksson (1981), as well as their conditional counterparts developed by Ferson and Schadt (1996) and Becker, et al (1999), rely on highly stylized assumptions. No previous study has examined the performance of a collection of SDF models in the context of known market timing ability.

Our artificial market timers invest in either the market index (Standard and Poors 500) or a one month Treasury bill. The portfolio return for month t is therefore equal to $RF_t + x_{t-1}\, r_{mt}$, where r_{mt} is the excess return of the S&P 500 over that of the one-month Treasury bill, RF_t. The portfolio weight x_{t-1} is a binary variable given by:

$$x_{t-1} = I\{0.5\ \ (1+\gamma\,) < \eta\}\, I\{r_{mt} \le 0\} + I\{0.5\ \ (1+\gamma\,) > \eta\}\, I\{r_{mt} > 0\}\,, \qquad (12)$$

where I{.} is the indicator function and η is an independent random variable, uniformly distributed on the interval (0,1).

The parameter γ, $(0 \le \gamma \le 1)$, measures the signal quality. When the information is perfect $(\gamma = 1)$, the manager invests the entire portfolio in the market index if the return on the market portfolio in the next period is greater than the return of the Treasury bill. Otherwise, the manager invests the entire portfolio in Treasury bills. When the manager has no market timing ability $(\gamma = 0)$, the portfolio is either 100% in the market or bills, based on a coin flip. γ may be interpreted by the correlation it implies between the portfolio weight, x_{t-1}, and the indicator for a positive market excess return, $I\{r_{mt}>0\}$. If $\gamma = 0$ the correlation is zero, and if $\gamma = 1$ the correlation is 1.0. In between these values the correlation depends on the value of $p=Prob\{r_m>0\}$, but the value of γ is a close approximation to the correlation.[13] We find below that the best performing actual mutual funds (those in the top 2.5% to 5%) have performance measures similar to an artificial market-timing fund with values of γ equal to 0.55 to 0.6.

V. The Data

We use four different data sets in our study: primitive assets, economic factors, instruments for public information, and mutual fund returns. A brief description of each data set appears below. Details are provided in the appendix.

A. Primitive Assets

Primitive assets should reflect the returns available to investors and fund managers. Of course, it is not practical to measure the entire universe of investment opportunities. For this study, we consider nine primitive assets: a short-term risk free security, two long-term bond returns, and stock portfolios that mimic large-cap, small-cap, value, growth, momentum and contrarian investment strategies. Value and growth portfolios are motivated by Fama and French (1993, 1996); momentum and contrarian portfolios are motivated by Grinblatt, Titman and Wermers (1995) and Ferson and Khang (2000).

---

[13] Let $q \equiv \text{Prob}\{x_{t-1} = 1\} = p\ (1+\gamma)/2 + (1-p)\ [1-(1+\gamma)/2]$. Choosing p=0.57 (the frequency of positive excess returns in our sample), the correlation between x_{t-1} and $I\{r_{mt}>0\}$ is given by $\gamma\sqrt{p(1-p)/q(1-q)} \approx \gamma$.

B. Economic Factors

We group the factors into "traded" and "nontraded" factors. For the nontraded factor models we consider four factors: the rate of inflation and measures of growth for money, industrial output and consumer spending. These factors are motivated by Lucas (1978), Breeden (1979), Ferson and Harvey (1991) and Chan, Foresi and Lang (1996).

For our traded factor models, we consider three factors: monthly excess returns on the S&P 500 index, a long-term government bond and a low-grade corporate bond portfolio. These factors are similar to those used by Chen, Roll and Ross (1986) and Ferson and Harvey (1991).

We also consider a three-factor model based on Fama and French (1993, 1996). Here the factors are a market index and the return differentials between small and large-cap stocks, and between high and low book-to-market stocks.

Finally, we consider an exact, three-factor version of the Arbitrage Pricing Model (Ross, 1976). Here the factors are the excess returns on the first three asymptotic principal components in a large sample of monthly stock returns. These factors are similar to those used by Connor and Korajczyk (1988) and Ferson and Korajczyk (1995).[14]

C. Information Variables

Previous studies have identified a number of variables that are useful in predicting security returns over time. We consider two instruments only: a short term interest

[14] The principal components data are courtesy of Robert A. Korajczyk. We do not expand the set of primitive assets to include the three principal components, in order to avoid the extreme colinearity that would result. Instead, when we estimate the APT models we replace the SP500, low-grade bond return and contrarian portfolio with the three principal components. We experimented with using the original nine primitive assets to estimate the APT models, and found that the models performed more poorly when the principal components were not included among the primitive assets.

rate and a stock market dividend yield. These two variables have figured most prominently in studies of mutual fund performance (see Ferson and Schadt (1996), Ferson and Warther (1996) and Becker, et al., 1999).[15]

Table 1 presents summary statistics of our data. The correlations (not shown) among the non traded factors range from -0.58 to 0.33; the correlation between the lagged instruments is -0.05; the correlations between the nontraded factors and the primitive asset returns range from 0.48 to -0.18. Most are in the range -.15 to +.20.

VI. Estimating the Stochastic Discount Factor models

In Tables 2 and 3 we evaluate the fit of the stochastic discount factor models in the sample of primitive assets. The models are estimated using monthly data for the period July, 1963 through December, 1994. The first row of the A panels reports results for a constant discount factor model, in which the SDF is assumed to be fixed over time, and equal to the inverse of the sample mean of the gross return of the one-month Treasury bill. A constant-SDF model can be motivated by risk neutrality, where the marginal rate of substitution of a risk-neutral investor (with time-additive, state-independent utility) is constant over time. For our purposes, this provides a simple point of comparison for the performance of the models.

[15] In an earlier draft of this paper we considered an expanded list of instruments, including a lagged measure of the slope of the U.S. Government term structure, a lagged yield spread in the corporate bond market (BAA versus AAA) and a dummy variable indicating that the next month is January.

A. Summary statistics of the Average Pricing Errors

Table 2 presents summary statistics for the time-series of the fitted SDF's. The means of most of the SDFs are close to the inverse of the mean of the gross Treasury bill return, as can be seen by comparison to the constant SDF model. Thus, including the one-month bill as a primitive asset is generally effective in controlling the mean of the SDF. The four nontraded-factor model is the exception, where the mean of the SDF is slightly above 1.0. As the complexity of the models increases (more factors are used, or we move from an unconditional to a conditional model), the standard deviation of the fitted SDF generally increases. Recall that for the linear factor and primitive efficient models, a conditional model is equivalent to an unconditional model fit to the primitive asset returns, and also the "dynamic strategy" returns obtained by multiplying the primitive returns by the lagged instruments. Hansen-Jagannathan (1991) show that the minimum variance of an SDF increases when the number of assets increases, because the mean variance frontier can only expand as more assets are included. Thus, it makes sense that the conditional SFD models could have larger standard deviations.

The SDFs have more negative values when more factors are used. The four nontraded factor model, for example, produces a large number. The conditional models also have more negative values. More frequent negative values are expected, other things equal, as the SDF becomes more volatile. However, negative values mean that the SDF assigns positive prices to negative payoffs at some points in time. The Bakshi-Chen model avoids negative SDF values, but at the cost of more variability. While a larger variance of an SDF is useful, according to the equity premium puzzle of

Mehra and Prescott (1985), a more volatile SDF implies lower power to detect abnormal performance.

The Hansen-Jagannathan (1997) distance measure is a summary of the mean pricing errors across a group of assets. The measure may be interpreted, analogous to Hotelling's T^2 statistic, as the maximum "t-ratio" of pricing errors for portfolios of the primitive assets. Its advantage in our setting is that the standard error of the "t-ratio" in question is not affected by estimation error in the SDF, as it depends only on the test asset returns. Thus, there is no penalty or advantage to a volatile SDF. The HJ measure may also be interpreted as the distance between the candidate SDF and one that would correctly price the primitive assets. When the lagged instruments, Z, are used to form dynamic strategy returns, and these are included, we have the conditional Hansen-Jagannathan distance measure, denoted by "HJcon" in the table. When Z is restricted to a constant and the dynamic strategies are not included, we have the unconditional measure denoted by "HJun."[16]

Using the unconditional HJ distance all of the unconditional models, except for the APT, have smaller pricing errors than would be obtained by the constant SDF model, discounting the returns at a fixed risk-free rate. The primitive-efficient and Bakshi-Chen models produce the smallest distances, essentially zero by construction. The numeraire portfolio also produces a small unconditional measure. This does not mean that these models will perform well for performance measurement. An out-of-sample evaluation is needed. The unconditional Fama-French model has a smaller

[16] For a given set of test asset gross returns R, instruments Z and a candidate stochastic discount factor m, let $r = R \quad \left(Z./E(Z)\right)$, where ./ denotes element-by-element division. The population value of the Hansen-Jagannathan distance measure takes the quadratic form: E(mr-1)'{E(rr')$^{-1}$} E(mr-1). We report the sample moment counterpart.

unconditional distance measure than the CAPM or three-factor model. The unconditional four-factor model also has a small distance measure.

The conditional models generally produce larger unconditional HJ distances than their unconditional model counterparts. In attempting to price the dynamic strategies implied by the lagged instruments, the conditional models sacrifice some accuracy on the primitive returns. We also examine the mean absolute pricing errors on the primitive assets (not reported in the tables) and find they are larger for the conditional models. This is consistent with Ghysels (1998).

The conditional models are estimated with the objective of pricing both the primitive assets and the dynamic strategy returns. Given this objective, they should perform better, in the sample, according to the HJcon measure. Table 2 shows that this measure is smaller for five of the eight conditional models.

The conditional primitive-efficient and Bakshi-Chen models produce the smallest conditional distances, essentially zero by construction. The largest conditional distances appear for the numeraire portfolio, APT and Fama-French models. The conditional version of the Fama-French model does not register an improvement over the conditional CAPM, nor over the unconditional Fama-French model. This indicates that the conditional Fama-French model does a poor job pricing dynamic strategies, consistent with Ferson and Harvey (1999).

B. The Dynamic Performance of SDF Models

The HJ distances summarize the relative fit of different models to the cross-section of test assets, but they provide no insight into the economic magnitudes of the pricing errors for particular assets. Table 3 examines the dynamic performance of the models, focussing on the individual primitive assets. While R_{t+1} may be predictable based on Z_t, an SDF model implies that $m_{t+1}R_{t+1}$ should not be predictable using Z_t. We run time-series regressions of the model pricing errors, $m_{t+1}R_{t+1}$ -1, on the instruments Z_t, and report the sample standard deviations of the fitted values of the regressions. If the standard deviations are small, the model "explains" the predictable variation in the returns. Unlike the HJ distance measures, the standard deviations reflect no penalty for a model that gets the average return wrong. Thus, the standard deviation is a pure measure of the ability of the model to explain predictable variation in the returns, analogous to the variance ratios in studies such as Ferson and Korajczyk (1995).[17]

The constant discount factor, which explains none of the predictability, is shown as a point of reference. For example, the standard deviation of the predictable returns for the S&P 500 is shown as 1.05% per month, while the standard deviation of the raw return is about 4% (Table 1). This means that the R^2 in a regression of the S&P 500 return on the lagged instruments is about $(.01 / .04)^2$ = 6%.

Panel A summarizes results for the unconditional models. None of the models can explain the predictable variation in the fixed income returns (govt, junk, tbill) better than a constant discount factor model.[18] It is hard to find factors that can explain the dynamics of both stock and bond returns. A few of the unconditional models explain a good fraction of the predictability in the equity portfolio returns. The

[17] Kirby (1998) presents alternative tests based on restrictions to the coefficients of predictive regressions for returns and dynamic strategies, and shows they are closely related to the HJ distance measure of the previous section. He also finds that the Fama French (1996) model does not explain dynamic strategies very well.

[18] Recall that for the APT models, the SP500, junk bond and contrarian portfolio are replaced by the first three principal components.

unconditional four-factor model, APT, numeraire portfolio and primitive-efficient models perform poorly. The product $m_{t+1}R_{t+1}$ in these models has larger regression coefficients on Z_t than R_{t+1} does.

Panel B of Table 3 presents results for the conditional SDF models. The conditional primitive-efficient and Bakshi-Chen models produce standard deviations close to zero. This is because the models are fit to make the expected product of the errors with the lagged instruments equal to zero in the sample. In general, the conditional models perform better than their unconditional counterparts. The exceptions are the CAPM and Fama-French models. The Fama-French model deteriorates substantially in its conditional version, again illustrating that it performs poorly in the presence of dynamic strategies.

The results in Tables 2 and 3 suggest a refinement of the results of Ghysels (1998). In the context of beta pricing models, Ghysels finds that conditional models have larger mean squared pricing errors on the primitive assets than unconditional models. Tables 2 and 3 are consistent with this finding. However, conditional models have smaller pricing errors on the dynamic strategy returns, $\left(R_{t+1} \quad Z_t\right)$, and do a better job of controlling the predictable components of the primitive asset returns. More research is needed to compare the out-of-sample performance of conditional and unconditional models along these dimensions (see Simin, 2000).

VI. Evidence on Artificial Mutual Funds

Table 4 presents summary statistics for the returns of the artificial mutual funds. Panel A covers the stock pickers. When $\gamma = 0$, the stock picker's average return is 0.92% per month, similar to that of the market index.[19] With higher levels of ability, the average return rises to more than 6.4% per month, at $\gamma = 0.8$. At the same time, the standard deviation of return rises from about 5% to about 9% per month. The regression betas on the S&P 500 rise from 1.05 to 1.33 as ability moves from $\gamma = 0$ to $\gamma = 0.8$. As ability increases, the portfolio favors stocks with increasingly positive market model residuals. Such stocks tend to have larger standard deviations of return and also, higher betas. The first order autocorrelations are about 0.20 for all ability levels, similar to the autocorrelation of an equally weighted portfolio of small stocks.

Panel B summarizes the artificial market timers, for ability levels between $\gamma = 0.5$ and $\gamma = 1$. The beta of the timers' returns on the S&P 500 is always close to 0.5, since the strategies hold the S&P 500 about half of the months and the Treasury bill the other half. The standard deviations of return vary from 2.5% to 3.4% per month, and the mean returns vary from 0.7% to 2.3% per month across the ability levels. The effect of ability on the portfolio returns is not as great for the market-timing funds as for the stock-picking funds. This makes sense, as a stock-picking fund is informed about many stocks, and its errors can be partly diversified, while the timer has only a single, market-wide signal. Due to sampling variation, neither the mean nor the standard deviation of the timers' portfolio return is strictly monotonic in γ. Successful market timers may be considered to generate "underpriced" options (Merton and Henriksson, 1981). As timing ability increases, the market timers generate increasingly right-skewed return distributions; thus the returns get more option-like. Such skewness creates

[19] For $\gamma = 0$, there is randomness in the portfolio returns due to the signal, diminishing in its effect as γ approaches 1.0. To ensure representative values we run both the uninformed picker and timer for 100 trials, rank these on their sample average returns, and report the figures for the median-return artificial funds.

problems for traditional approaches to performance evaluation, as discussed by Jagannathan and Korajczyk (1986) and Leland (1999).

A. Performance of the Models with Artificial Funds

Table 5 summarizes the performance of the artificial mutual funds using the SDF models. Panels A and B show the alphas and their t-ratios for the stock pickers. Perhaps the most striking impression is that, with some exceptions, the performance results are remarkably similar across the SDF models. This is interesting in view of the sensitivity of beta pricing models to the performance benchmark, as documented by previous studies.

Most of the models have a mild bias, producing negative alphas when $\gamma = 0$ and there is no abnormal performance. The average alpha for an uninformed stock picker is -0.19% per month for the unconditional models, and -0.12% for the conditional models.[20] The typical standard error is about 0.10%. The conditional models tend to have smaller biases, excepting the Fama-French factors, where the bias is slightly larger in the model's conditional form. Among the linear factor models, the APT has the largest negative bias.

Chen and Knez (1996), using primitive-efficient models, find that mutual funds have insignificant but negative abnormal returns. Using unconditional models, they find the average alpha for 68 funds, 1968-89, is -0.09% per month. Our results

[20] When we discuss averages across the models, we exclude the numeraire portfolio models from the calculations. We repeat this exercise, where the artificial fund returns are computed without the Shumway (1997) adjustments for delisted stocks. In this case the average alpha is -0.18% per month with the unconditional models, and -0.08% with the conditional models.

using the artificial funds are similar, suggesting that their findings reflect a biased performance measure.[21]

Dahlquist and Soderlind (1999) study primitive efficient models using weekly Swedish data, 1986-95. They find no significant biases in the average pricing errors, but they do find size distortions, where tests for the hypothesis that $\alpha_p = 0$ reject the null hypothesis too often.

When $\gamma > 0$ the results in Table 5 provide information on the power of the models to detect superior fund performance.[22] Most of the models are able to detect superior performance at the higher ability levels $(\gamma \geq 0.25)$. The t-ratios in Panel B indicate that the conditional models often have slightly higher power than the unconditional models. The Fama-French model is an exception, where power is lower in its conditional form. The four-factor model generates smaller t-statistics than the other factor models when the ability levels are high. Also, the numeriare portfolio estimates have huge standard errors, and thus low power, relative to the other models.

Panels C and D of Table 5 show the performance measures for the artificial market timing funds. Again, with a few exceptions the results are not highly sensitive to the benchmark SDF. When the artificial timer switches randomly between the

---

[21] Chen and Knez (1996) compute alphas for excess returns, R_p-R_f, whereas we use raw returns. Since the primitive efficient models are exactly identified, the average pricing error of the Treasury bill is identically zero and the two alphas are the same. In overidentified models the two alphas can differ. For example, in our linear factor models the average pricing error for the one-month bill is 0.05% per month. This could explain a fraction of the negative bias that we find for these models, but not the similar bias we find in the primitive efficient models. Furthermore, the 0.05% average reflects a great deal of dispersion across the models, consistent with the large standard errors attached to the average pricing errors.

[22] Of course, a complete analysis of statistical power should use size-adjusted empirical critical values for the test statistics. Our use of the term "power" here is therefore imprecise.

market and cash, the average unconditional alpha is 0.005% per month, and the conditional alpha is -0.099%. The typical standard error is, again, about 0.10% per month. Thus, the model biases, especially for the unconditional models, are somewhat smaller in the face of uninformed market timing.

All of the models, except for the numeraire portfolio, can detect high levels of ability. However, compared with the stock-picker results, we find smaller t-ratios for a given IC. This is consistent with the smaller effect of γ on the market timers' returns. Most of the models have less power to detect market timing rather than stock picking ability. There are also some differences in power across the models. The primitive-efficient SDFs have slightly smaller t-ratios than the average model at the higher ability levels, and the four-factor models have markedly lower power.[23]

We conclude that no model for the SDF clearly dominates the others. The worst performing models are the numeriare portfolio and the model with the four nontraded factors. The poor performance of the latter model likely reflects the low correlation of monthly stock returns with the economic variables. Given the numerical instability of the numeraire portfolio model, its poor performance and large standard errors may be expected. The relatively high computation costs of this model present another disadvantage.

[23] Dahlquist and Soderlind (1999) find that primitive efficient SDF models have low power, when the alternative hypothesis of abnormal performance simply adds a constant amount to the portfolio return.

VIII. Using the Models to Measure Mutual Fund Performance

We use the SDF models to measure performance in a sample of open-ended mutual funds that tries to control survivorship bias by including discontinued funds. The data come from Elton, Gruber and Blake (1996), and include all funds that were categorized as "common stock" funds in the 1977 edition of Wiesenberger's Investment Companies, and that had at least $15 million in total net assets under management at the beginning of 1977.[24] Variable annuity funds (which are usually tied to insurance products) and funds that place restrictions on the purchaser are excluded. The remaining 188 funds are followed through name changes and mergers for the period 1977 to 1993. (See Elton, Gruber and Blake, 1996.) We believe that few of these funds engage in active market timing. Therefore, we rely more on our results for the artificial stock-picking funds in evaluating the actual funds' performance. The monthly returns for the mutual funds reflect the reinvestment of dividends and capital gains, and are net of most expenses, except front-end load charges and exit fees. Summary statistics for the returns of the funds are presented in panel C of Table 1.

Table 6 summarizes the distribution of the mutual fund alphas under the various models. We report the values at selected fractiles of the distribution across the 188 funds. By comparing these with the alphas in Table 5, where the degree of ability is known, we interpret the results. Averaged across the models, the mean and median alphas both round to -0.1% per month and the 10% tail cutoffs are nearly symmetric. The 2.5% tails suggest mild left skewness. The Bonforoni p-values, based on the extreme t-ratios in the sample, suggest that the extreme performing fund is more likely to have a

[24] We thank Martin Gruber for graciously sharing these data with us, and Micropal for permission to use the data.

positive than a negative alpha.[25] Dahlquist and Soderlind (1999), in a weekly sample of Swedish funds for 1991-95, find small positive alphas and mild right skewness. Their sample suffers from survivorship bias, which is likely to affect both of these findings.

In Table 6 the distributions of the alphas for the actual mutual funds are not highly sensitive to the SDF model, similar to what we found for the artificial funds. Averaged across the models the average alpha for the median fund, is -0.06% per month using unconditional models, and -0.09% using conditional models. The average risk-adjusted performance of a typical mutual fund lies between that of an artificial stock-picking fund with neutral performance, and an uninformed market timer. With a typical standard error for alpha of about 0.1% per month, the differences are not statistically significant. Thus, the overall impression is that the average mutual fund performance is consistent with the null hypothesis of no ability.

These results are interesting in view of the fact that the artificial funds do not pay transactions costs, while actual funds do. Bogle (1994) suggests that turnover and expense ratios can be combined for a rough measure of the total expenses implicit in mutual fund returns. He argues that 1.2% is a conservative estimate of the costs of a round-trip trade. With turnover averaging 6% per month, he estimates typical trading costs to average about 0.9% per year. Based on Morningstar data, the average expense ratio for mutual funds over our sample period is about 1.08% per year. Adding these two figures gives 1.98% per year, or about 0.17% per month. Based on the average alphas, the typical mutual fund in our sample beats the uninformed stock picker, even after covering these transactions costs. If we add back the 0.17% per month for

[25] The Bonforoni p-value is the smallest tail area of the t-ratios for the 188 funds, multiplied by 188.

transactions costs, the typical fund's conditional alpha is significantly higher than either type of artificial fund. Thus, the evidence is consistent with the view that the average mutual fund manager has enough ability (or exerts enough effort) to cover transactions costs and the expense ratio, which includes the management fee.

While the distribution of the measured fund performance is not highly sensitive to the SDF model -- once the poorly performing models are excluded -- this does not imply that the relative performance for individual funds is robust. Individual fund alphas could shift position within the distribution, leaving the overall shape of the distribution intact. To investigate this issue we compute a correlation matrix of the individual alphas across the models. Each pair of models produces 188 pairs of "observations" on funds' alphas. We compute the sample correlations for these 188 observations. We find that the correlations are often quite high. For example, the correlation between the conditional and unconditional versions of the CAPM is 0.99. The correlation is 0.84 between the conditional CAPM and the conditional primitive efficient model. The lowest correlations are between the poorly performing models and the others. For example, the correlations between the four nontraded-factor models and the others are between 0.52 and 0.66. Most of the other correlations are in the 0.82-0.96 range. We conclude that the relative performance measured for the individual funds tends to be highly correlated across the SDF models.

IX. Concluding Remarks

This paper studies the stochastic discount factor (SDF) framework for evaluating the performance of mutual funds. We provide a comparison of a large number of asset pricing models in a uniform experimental design. We find that no model for the SDF clearly dominates the rest, but some models are clearly inferior. The worst performing models are the numeriare portfolio and a linear factor model with four nontraded economic factors. The choice between conditional and unconditional models presents a tradeoff. Conditional models can deliver smaller average pricing errors for dynamic strategies, and better control the predictability in pricing errors, but at the cost of larger variances of the pricing errors on the primitive assets of the model.

We evaluate the models using artificial "mutual funds," where we control the extent of market timing or security selection ability. The good news is that the measured performance is not highly sensitive to the specification of the SDF, excepting the few models we found to be clearly inferior. The bad news is that many of the SDF models are biased, producing negative alphas when stock-picking performance is neutral. The average bias is about -0.19% per month for unconditional models and -0.12% for conditional models. This is less than two standard errors, as a typical standard error is 0.1% per month. However, the magnitudes suggest that previous evidence (Chen and Knez, 1996) of negative abnormal mutual fund returns in SDF models reflects a bias in the measure.

We use the models to evaluate performance in a monthly sample of 188 equity mutual funds. These results update and generalize the evidence in the classic study of Jensen (1968), who used the CAPM to conclude that a typical fund has neutral performance, after accounting for expenses. We find that the average mutual fund alpha is no worse than the hypothetical stock-picking fund with neutral performance, although it is below the alpha of a hypothetical uninformed market timer by as much as

0.07% per month. These hypothetical funds pay no expenses. If we add back expenses of about 0.17% per month to the mutual fund alphas, the average fund's performance is higher than the hypothetical funds with no ability. Overall, we see no evidence that a typical fund has poor performance after we adjust for model biases and expenses.

Appendix

Invariance of Performance Measures to the Number of Funds

Referring to the system (10), partition g = (g_1', g_2')' where g_1 depends on only the parameters of the SDF and $g_2 = \alpha_p - h$, where $h = T^{-1}\sum\{m_{t+1}R_{p,t+1} - 1\}$, and the dimension of g_2 is the number of funds in the system. Conformably partition V, the asymptotic covariance matrix of g, where V_{11} is the upper left block, etc. The GMM weighting matrix is W = V^{-1} is also conformably partitioned. The GMM estimator for the system chooses the parameter vector θ to minimize g'Wg, which implies:

$$g' W \frac{\partial g}{\partial \theta} = 0 \qquad \text{(A.1)}$$

The structure of this problem implies that a partition of $\partial g/\partial\theta$ according to g_1 and g_2 (the rows) and the parameters of the SDF and α_p (the columns) is of the form:

$$\frac{\partial g}{\partial \theta} = \begin{bmatrix} gd_{11} & 0 \\ gd_{21} & 1 \end{bmatrix} \qquad \text{(A.2)}$$

where gd_{11} and gd_{12} are full matrixes. Solving for right hand partition of equation (A.1), corresponding to α_p we obtain:

$$\alpha_p = h - W_{22}^{-1} W_{21} g_1 \qquad \text{(A.3)}$$

Standard expressions for partitioned matrix inversion allow us to write $W_{22} = (V_{22} - V_{21}\ V_{11}^{-1}\ V_{12})^{-1}$, and $W_{12} = -\ V_{11}^{-1}\ V_{12}\ W_{22}$. Thus,

$$\alpha_p = h + V_{21} V_{11}^{-1} g_1 \qquad \text{(A.4)}$$

Equation (A.4) shows that α_p reduces to h, which is the two step estimator described in the text, when either of two conditions are satisfied. The conditions are (1) the covariance between the moment conditions g_1 and g_2, V_{12} is zero; or (2) the moment condition g_1=0, which is satisfied when the SDF model by itself is exactly identified. (This applies for the primitive efficient SDF model.) From (A.4) we can compute the asymptotic variance of the estimators:

$$\text{Avar}(\ _p) = \text{Avar}(h) + V_{21}\ V_{11}^{-1}\ V_{12} + 2\ V_{21}\ V_{11}^{-1}\ \text{Acov}(g_1,h). \qquad \text{(A.5)}$$

Equations (A.4) and (A.5) imply that the estimates of α_p and the standard errors for any subset of funds is invariant to the presence of another subset of funds in the system. To see this, partition $\alpha_p = \left(\alpha_p^1, \alpha_p^2\right)$, and examine the partition of equation (A.4) corresponding to α_p^1. The result is identical whether or not α_p^2 is present. Similarly, partition equation (A.5) to observe that the asymptotic variance of α_p^1 is exactly the same whether or not α_p^2 is present.

Data Description

Our short-term risk free security is the one-month Treasury bill, from Ibbotson Associates via the Center for Research in Security Prices (CRSP). To represent longer term fixed income assets we use the returns of an approximate 20 year U.S. Government bond, from Ibbotson Associates, and the return of a low-grade corporate bond from Blume, Keim and Patel (1991) series. We update this series (which goes from February of 1926 through January of 1990) with the Merrill Lynch High Yield Corporate Bond Index return, from the *Salomon Center Newsletter* (Spring and Summer, 1993), and from the *Wall Street Journal*.[26] Our large-cap strategy, and the market portfolio in the CAPM, is the return on the S&P 500 index.

We form five additional primitive assets from common stock portfolios constructed by Carhart, Krail, Stevens and Welch (1996).[27] For each month Carhart et al. group the common stocks on the CRSP tape into thirds according to each of three independent criteria, producing 27 portfolio return series. The grouping criteria are (1) the past return for months t-2 to t-12, (2) equity market capitalization, and (3) the ratio of book equity to market equity. We form a small-cap portfolio by equally weighting the nine portfolios with the lowest market capitalization. For the momentum (contrarian) strategy we use an equally weighted average of the nine portfolios with the highest (lowest) t-2 to t-12 returns, thereby controlling for book-to-market and firm size. For the value (growth) strategy we use an equally weighted average of the nine

---

[26] We are grateful to Don Keim for making the Blume, Keim and Patel (1991) data available to us.

[27] These data are courtesy of Mark Carhart.

portfolios with the highest (lowest) book-to-market ratio, thus controlling for firm size and past relative return effects.

Our nontraded risk factors include the rate of inflation and measures of growth for money, industrial production and consumer spending. The inflation rate is the monthly percentage change in the consumer price index, CPI-U, from Ibbotson Associates via CRSP. The growth of money is per-capita, inside money (Citibase FM2.monthly) less currency or M1 (FM1.monthly) divided by the total U.S. civilian noninstitutional population (P16.monthly), deflated by the consumer price index, and used in the form of the first difference of the logarithms. For industrial production we use the continuously-compounded growth rate of the seasonally adjusted index number, 1992=100 (Citibase IP.monthly). Our measure of consumer spending is the monthly real, per capita growth rate of aggregate personal consumption expenditures for consumer nondurable goods (Citibase GMCN) plus services (GMCS), divided by the population.

Our traded economic factors include monthly excess returns on the S&P 500 index, and the long term government and corporate bonds described above. The factors for the Fama and French model are formed from the small-cap, large-cap, value and growth portfolios described above.

We use two lagged instruments for public information. The one-month bill yield is from the CRSP riskfree files. The yield is calculated from the bid prices on the last trading day of the previous month. We subtract from the 1-month yield the average of its values for the previous 12 months, a simple form of stochastic detrending. The dividend yield is the price level at the end of the previous month on the CRSP value-weighted index of NYSE + AMEX firms, divided into the previous twelve months of dividend payments for the index.

References

Bakshi, Gurdip and Zhi-wu Chen. 1998. Asset pricing without consumption or market portfolio data. Working paper: University of Maryland.

Bansal, Ravi, and S. Viswanathan. 1993, No arbitrage and arbitrage pricing: A new approach. *Journal of Finance* 48: 1231-1262.

Becker, C., W. Ferson, D. Myers and M. Schill. 1999. Conditional Market timing with Benchmark investors. *Journal of Financial Economics* 52: 119-148.

Blume, M., Donald B. Keim and S.A. Patel. 1991. Returns and volatility of low-grade bonds, 1977-1989. *Journal of Finance* 46: 49-74.

Bogle, John C. 1994. Bogle on Mutual funds: New perspectives for the intelligent investor. Irwin. New York.

Breeden, Douglas T. 1979. An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics* 7: 265-96.

Brown, Keith C. and Gregory D. Brown. 1987. Does the composition of the market portfolio really matter? *Journal of Portfolio Management* Winter: 26-32.

Carhart, Mark M., R.J. Krail, R.L. Stevens and K.D. Welch. 1996. Testing the conditional CAPM. working paper: University of Chicago.

Chan, Foresi and Lang. 1996. Does money explain asset returns? Theory and empirical evidence. *Journal of Finance* 51:345-362.

Chen, Nai-fu, T. Copeland and D. Mayers. 1987. A comparison of single and multifactor perforamance methodologies. *Journal of Financial and Quantitative Analysis* 22: 401-17.

Chen, Nai-fu, Roll, Richard R. and Ross, Stephen A. 1986. Economic Forces and the Stock Market. *Journal of Business* 59: 383-403.

Chen, Zhiwu, and Peter J. Knez. 1996. Portfolio Performance Measurement: Theory and Applications. *Review of Financial Studies* 9: 511-556..

Christopherson, Jon A., Wayne Ferson and Debra A. Glassman. 1998. Conditioning manager alpha on economic information: Another look at the persistence of performance. *Review of Financial Studies* 11: 111-142.

Cochrane, John. 1996. A cross-sectional test of a production-based asset pricing model. *Journal of Political Economy* 104(3): 572-621.

Connor and Korajczyk. 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics* 15: 373-394.

Connor, Gregory and Robert A. Korajczyk. 1988. Risk and return in an equilibrium APT: applications of a new test methodology. *Journal of Financial Economics* 21: 255-289.

Dahlquist, Magnus and Paul Soderlind. 1999. Evaluating portfolio performance with stochastic discount factors. *Journal of Business* 72: 347-384.

Dumas, Bernard, and Bruno Solnik. 1995. The world price of foreign exchange risk. *Journal of Finance* 50: 445-480.

Dybvig, Philip H., and J.E. Ingersoll, Jr. 1982. Mean-Variance Theory in Complete Markets. *Journal of Business* 55: 233-252.

Dybvig, Philip H., and Ross, Stephen A. 1985. Performance Measurement Using Differential Information and a Security Market Line. *Journal of Finance* 40: 383-99.

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake. 1996. Survivorship bias and mutual fund performance. *Review of Financial Studies* 9:1097-1120.

Fama, Eugene F. and French, Kenneth R. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3-56.

Fama, Eugene F. and French, Kenneth R. 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51: 55-87.

Ferson, W. and C. R. Harvey. 1991. The Variation of Economic Risk Premiums. *Journal of Political Economy* 99: 385-415.

Ferson, W. and C. R. Harvey. 1997. Fundamental Determinants of National Equity Market Returns: A perspective on Conditional asset pricing. *Journal of Banking and Finance* 21: 1625-1665.

Ferson, W. and C. R. Harvey. 1999. Conditioning Variables and Cross-section of Stock Returns, *Journal of Finance* 54: 1325-1360.

Ferson, W. and Ravi Jagannathan. 1996. Econometric Evaluation of Asset Pricing Models. in the *Handbook of Statistics* vol. 14: Statistical Methods in Finance, G.S. Maddala and C.R. Rao (editors), North Holland ISBBN: 0-444-81964-9.

Ferson, W. and Kenneth Khang. 2000. Conditional performance measurement using portfolio weights: Evidence for pension funds. Working paper: University of Washington.

Ferson, W. and Robert A. Korajczyk. 1995. Do arbitrage pricing models explain the predictability of stock returns? *Journal of Business* 68: 309-349.

Ferson, W. and Rudi Schadt. 1996. Measuring fund strategy and performance in changing economic conditions. *Journal of Finance* 51: 425-462.

Ferson, W. and Vincent A. Warther. 1996. Evaluating Fund Performance in a Dynamic Market. *Financial Analysts Journal* 52: 20-28.

Ghysels, Eric. 1998. On stable factors in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance* 53: 549-574.

Grant, Dwight. 1977. Portfolio performance and the "cost" of timing decisions. *Journal of Finance* 32: 837-846.

Grinblatt, Mark and Titman, Sheridan. 1989. Portfolio Performance evaluation: old issues and new insights. *Review of Financial Studies* 2: 393-422.

Grinblatt, Mark and Titman, Sheridan, and Russell Wermers. 1995. Momentum strategies, portfolio performance and herding: A study of mutual funds. *American Economic Review* 85: 1088-1105.

Grinold, R., Kahn, R. 1995. Active portfolio management. Irwin Press, Chicago.

Hansen, Lars P. 1982. Large sample properties of the generalized method of moments estimators. *Econometrica* 50: 1029-1054.

Hansen, Lars P. and Ravi Jagannathan. 1991. Implications of security market data for models of dynamic economies. *Journal of Political Economy* 99: 225-262.

Hansen, Lars P. and Ravi Jagannathan. 1997. Assessing specification errors in stochastic discount factor models. *Journal of Finance* 52: 557-90.

He, Jia, Lilian Ng and Chou Zhang. 1998. Asset pricing specification errors and performance evaluation. Working paper: Hong Kong University of Science and Technology.

Hentschel, Ludger, Jankoo Kang and John B. Long. 1998. Numeraire portfolio tests of international bond market integration. Working paper: University of Rochester.

Jagannathan, Ravi and Robert A. Korajczyk. 1986. Assessing the market timing performance of managed portfolios. *Journal of Business* 59: 217-236.

Jagannathan, Ravi and Z. Wang. 2000. Efficiency of the stochastic discount factor method for estimating risk premiums. Working paper: Northwestern University.

Jensen, M. 1968. The performance of mutual funds in the period 1945-1964. *Journal of Finance* 23: 389-46.

Kan, Raymond, and Guofu Zhou. 1999. A critique of the stochastic discount factor methodology. *Journal of Finance* 54: 1221-1248.

Kirby, Chris. 1998. The restrictions on predictability implied by asset pricing models. *Review of Financial Studies* 11: 343-382.

Kang, Jangkoo. 1995. Bond mutual fund performance evaluation: The numeraire portfolio appraoch. Working paper: University of Rochester.

Lehmann, Bruce and David M. Modest. 1987. Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance* 42: 233-65.

Leland, Hayne. 1999. Performance beyond mean-variance: Performance measurement in a nonsymmetric world. *Financial Analysts Journal* 55: 27-36.

Long, John B. 1990. The numeraire portfolio. *Journal of Financial Economics* 26: 29-70.

Lucas, Robert. 1978. Asset Prices in an Exchange Economy. *Econometrica* 46: 1429-1446.

Mehra, R. and E. Prescott. 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15: 145-162.

Merton, Robert C. and Roy D. Henriksson. 1981. On market timing and investment performance II: Statistical procedures for evaluating forecasting skills. *Journal of Business* 54: 513-534.

Roll, Richard. 1978. Ambiguity when performance is measured by the security market line. *Journal of Finance* 33: 1051-1069.

Ross, S. 1976. The arbitrage theory of capital market asset pricing. *Journal of Economic Theory* 13: 341-360.

Sharpe, William. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19: 425-42.

Shumway, T. 1997. The delisting bias in CRSP data. *Journal of Finance* 52: 327-340.

Simin, Timothy. 1999. The forecasting performance of asset pricing models. Working paper, University of Washington.

Treynor, Jack, and Kay Mazuy. 1966. Can mutual funds outguess the market? *Harvard Business Review* 44: 131-136.

Table 1
Summary Statistics

	MEAN	STD	MIN	MAX	ρ_1
panel A: The primitive asset rates of return					
sp500	0.0091133	0.0423373	-0.215200	0.165700	0.003994
govt. bond	0.0058926	0.0296860	-0.084100	0.152300	0.051763
low grade bond	0.0069866	0.0234560	-0.081000	0.130000	0.221632
one-month bill	0.0052717	0.0022346	0.002100	0.013500	0.937797
momentum	0.0147299	0.0532051	-0.288311	0.188122	0.114470
contrarian	0.0081780	0.0532749	-0.216222	0.324533	0.137965
value	0.0139884	0.0523659	-0.262444	0.304667	0.165299
growth	0.0085136	0.0529997	-0.265411	0.191722	0.130837
small	0.0119945	0.0574457	-0.294078	0.304522	0.205066
panel B: Economic variables					
non-traded factors:					
inflation rate	0.0042135	0.003228	-0.0046000	0.0181000	0.646433
industrial production	0.0024553	0.008288	-0.0424745	0.0334559	0.366964
personal consumption	0.0012924	0.004751	-0.0189459	0.0181018	-0.167308
real money supply	0.0007191	0.005490	-0.0163079	0.0318874	0.769596
lagged instruments:					
tbyld	0.0002039	0.013158	-0.0562000	0.0489558	0.799197
vwyld	0.0370277	0.008048	0.0251967	0.0612818	0.973259
panel C: Mutual fund returns: Equally-weighted portfolios sorted by sample mean returns (January, 1977 - December, 1993)					
Lower 2.5%	0.00336927	0.056832	-0.2556400	0.1675400	0.040739
Lower 2.5% - 5%	0.00814162	0.049622	-0.2543120	0.1451120	0.027890
Lower 5% - 10%	0.00929043	0.043067	-0.2118689	0.1257611	0.043084
Median fund	0.01191584	0.047639	-0.2159600	0.1653800	0.081600
Upper 10% - 5%	0.01505460	0.051957	-0.2324656	0.1461400	0.080769
Upper 5% - 2.5%	0.01585514	0.049884	-0.2325460	0.1408760	0.060769
Upper 2.5%	0.01694441	0.059543	-0.2683660	0.1681320	0.054545

Note.- The data are monthly from July of 1963 through December, 1994, a total of 378 observations (the lagged instruments are known at the end of the previous month). For the mutual funds the sample period is January, 1977 through December, 1993 and there are 204 observations. The units are decimal fraction per month. ρ_1 is the first order sample autocorrelation.

Table 2
Stochastic Discount Factor Models

	E(m)	sd(m)	ρ_1(m)	min(m)	max(m)	num(m<0)	HJun	HJcon
Constant discount factor	0.995	0.000	0.000	0.995	0.995	0	0.144	0.319
UNCONDITIONAL MODELS								
SDF-CAPM (ucapm)	0.994	0.071	0.004	0.733	1.37	0	0.136	0.311
Three traded factors (u3fac)	0.994	0.081	0.084	0.688	1.44	0	0.134	0.309
Fama-French (uff)	0.989	0.225	0.234	-0.204	1.68	2	0.078	0.235
APT (uapt)	0.995	0.145	0.091	0.024	1.59	0	0.329	0.329
Four nontraded factors (u4fac)	1.010	1.220	0.208	-2.760	6.95	71	0.099	0.296
Primitive-efficient (upem)	0.995	0.380	0.094	-0.695	2.65	3	1.2E-25	1.2E-25
Numeraire portfolio (unum)	0.995	0.521	-0.031	0.444	7.43	0	0.008	0.300
Bakshi-Chen (ubc)	0.995	0.423	-0.009	0.249	4.77	0	3.7E-10	0.052
CONDITIONAL MODELS								
SDF-CAPM (ccapm)	0.989	0.218	-0.048	-0.576	1.80	3	0.172	0.270
Three traded-factors (c3fac)	0.994	0.291	-0.039	-0.532	2.09	6	0.179	0.266
Fama-French (cff)	0.979	0.565	0.098	-5.450	2.43	11	0.164	0.284
APT (capt)	0.985	0.228	0.128	-0.214	1.78	3	0.346	0.474
Four non-traded factors (c4fac)	1.020	2.060	-0.016	-6.440	7.79	96	0.035	0.068
Primitive-efficient (cpem)	0.995	0.566	0.242	-1.890	2.47	16	1.7E-19	1.7E-19
Numeraire portfolio (cnum)	0.995	9.220	0.001	-117.0	75.2	53	6.9E-7	1.080
Bakshi-Chen (cbc)	0.995	0.677	0.166	0.007	5.67	0	7.2E-10	1.9E-08

Note. - Various models for stochastic discount factors (SDFs) are estimated using the equations in the text and monthly data for July of 1963 through December 1994 (378 observations). The units of the returns are monthly decimal fractions. E(m) is the sample mean, sd(m) is the sample standard deviation and ρ_1(m) is the first order autocorrelation of the estimated stochastic discount factor. The primitive assets used in estimating the models are the Standard and Poors 500, a long term government bond, a low-grade corporate bond, a one-month Treasury bill, and five portfolios grouped as described in the text, according to lagged returns (momentum, contrary), book-to-market ratios (value, growth) and market capitalization (small stocks). For the APT models, three asymptotic principal components replace the sp500, low-grade bond and contrarian portfolios. HJun and HJcon are the Hansen-Jagannathan measures of misspecification. HJcon is the conditional measure, which uses the returns and the lagged instruments, while HJun uses no lagged instruments. The lagged instruments are the one-month Treasury bill yield and the dividend yield of the CRSP value-weighted stock index.

Table3
Dynamic Performance of Stochastic Discount Factor Models

SDF model	Primitive Assets								
	sp500	**govt**	**junk**	**momentum**	**contrary**	**value**	**growth**	**small**	**tbill**
PANEL A: UNCONDITIONAL MODELS									
Constant discount factor	1.05	0.496	0.553	1.37	1.30	1.29	1.39	1.55	0.18
ucapm	0.79	1.46	1.27	0.46	0.55	0.54	0.46	0.30	1.75
u3fac	1.16	1.84	1.65	0.82	0.94	0.91	0.83	0.65	2.13
uff	0.27	1.03	0.78	0.25	0.20	0.16	0.25	0.24	1.29
uapt	1.83	1.18	1.28	1.32	2.07	2.02	2.09	2.24	0.95
u4fac	7.80	7.38	7.50	8.15	8.47	8.31	8.27	8.64	6.94
upem	4.46	5.45	5.00	3.70	4.40	3.98	4.05	3.75	5.83
unum	4.90	5.98	5.62	4.24	4.96	4.60	4.57	4.33	6.44
ubc	4.10	4.81	4.42	3.62	4.07	3.79	3.83	3.63	5.05
PANEL B: CONDITIONAL MODELS									
ccapm	1.22	1.15	1.10	0.91	1.09	0.92	1.06	0.78	1.19
c3fac	1.14	1.53	1.34	1.42	1.18	1.40	1.23	1.52	1.19
cff	7.20	6.48	6.75	7.47	8.24	8.14	7.56	8.13	6.66
capt	0.55	0.64	0.51	0.22	0.17	0.32	0.27	0.26	0.49
c4fac	2.15	1.99	2.07	2.24	2.30	2.24	2.35	2.33	2.04
cpem	7.8E-9	7.7E-9	7.7E-9	8.0E-9	7.9E-9	7.9E-9	7.9E-9	8.0E-9	7.5E-9
cnum	0.38	0.12	0.25	0.21	0.21	0.15	0.19	0.34	0.35
cbc	6.3E-8	6.4E-8	6.5E-8	6.3E-8	6.3E-8	6.4E-8	6.2E-8	6.3E-8	6.6E-8

Note.-Various models for stochastic discount factors (SDFs), denoted by m, are estimated using the equations in the text. The standard deviations of the fitted pricing errors are shown in the table. The fitted pricing errors are the fitted values of a regression of $m_{t+1}R_{t+1}$-1 on Z_t, where m_{t+1} is the SDF, R_{t+1} is the particular asset gross return and Z_t is the vector of lagged instruments. The primitive assets used in the SDF models are the Standard and Poors 500 index (sp500), the long term government bond (govt), a low-grade corporate bond (junk), a one-month Treasury bill (tbill), and five portfolios grouped as described in the text, according to lagged returns (momentum, contrary), book-to-market ratios (value, growth) and market capitalization (small). In the case of the APT models, the SP500, junk bond and contrarian portfolios are replaced by the three asymptotic principal components. The lagged instruments are the one-month Treasury bill yield and the dividend yield of the CRSP value-weighted stock index. The symbols denoting the various models are the same as shown in Table2.

Table 4
Summary Statistics of Artificial Mutual Funds

Ability()	Mean	Std	Min	Max	ρ_1	SP500beta
PANEL A: ARTIFICIAL STOCK PICKERS						
0.00	0.0094926	0.0527125	-0.205859	0.222765	0.144480	1.05901
0.10	0.0116396	0.0572638	-0.282882	0.169699	0.212161	1.11825
0.20	0.0152671	0.0665882	-0.325791	0.266646	0.219904	1.20855
0.25	0.0207156	0.0748688	-0.353136	0.316370	0.241416	1.27060
0.30	0.0253968	0.0813125	-0.357061	0.303963	0.224202	1.33252
0.35	0.0315279	0.0833675	-0.301477	0.300201	0.259412	1.30099
0.40	0.0386342	0.0854730	-0.322398	0.296783	0.251474	1.30930
0.50	0.0530149	0.0896819	-0.337816	0.460395	0.198480	1.28819
0.60	0.0574245	0.0892648	-0.336422	0.406289	0.214116	1.33832
0.70	0.0644359	0.0906787	-0.251885	0.353727	0.200964	1.28475
0.80	0.0642396	0.0904048	-0.302697	0.364462	0.190315	1.32745
PANEL B: ARTIFICIAL MARKET TIMERS						
0.50	0.0072781	0.0321722	-0.215200	0.165722	-0.023750	0.572041
0.55	0.0088169	0.0285112	-0.116979	0.165722	-0.087784	0.448242
0.60	0.0102663	0.0293961	-0.108228	0.165722	-0.023375	0.482613
0.65	0.0131346	0.0337522	-0.215200	0.165722	0.005960	0.650919
0.70	0.0126723	0.0297808	-0.108228	0.134300	0.001528	0.506642
0.75	0.0162635	0.0296212	-0.090300	0.165722	-0.085124	0.528737
0.80	0.0149506	0.0283264	-0.116979	0.165722	-0.138601	0.475052
0.85	0.0187121	0.0281571	-0.108228	0.165722	-0.038624	0.510130
0.90	0.0205448	0.0258198	-0.075913	0.165722	-0.030016	0.465304
0.95	0.0208255	0.0247428	-0.089058	0.134300	-0.038660	0.437822
1.0	0.0232471	0.0259912	0.002120	0.165722	-0.069138	0.513261
S&P500	0.0091133	0.0423373	-0.215200	0.165700	0.003994	1.00

Note. - Monthly artificial mutual fund returns are generated for 378 months from July, 1963 through December of 1994. Panel A summarizes the properties of the returns for artificial stocks pickers with varying degrees of ability, γ . When $\gamma = 0$, there is no ability and when $\gamma = 1$ there is perfect ability, as described in the text. Panel B presents summary statistics for the artificial market timing mutual funds. For comparison purposes the Standard and Poors 500 index return is shown on the last line. Std is the standard deviation of the monthly return, ρ_1 is the first order autocorrelation and SP00 beta is the regression coefficient of return on the Standard and Poors 500 index return.

Table 5
Estimates of Performance for Artificial Mutual Funds

Results of joint estimation of a stochastic discount factor model and the performance of an artificial mutual fund for the July, 1963-December, 1994 period (378 months). The ability level of the artificial fund is indicated by the parameter , as explained in the text. Umean and Cmean are the averages taken across the unconditional and conditional models, respectively, excluding the numeraire portfolio models.

PANEL A: STOCK PICKERS: Estimates of Alpha

γ =	0	0.10	0.20	0.25	0.30	0.35	0.40	0.50	0.60	0.70	0.80
Ucapm	-0.0017	-0.00079	0.00119	0.00517	0.00925	0.01529	0.02158	0.0354	0.0394	0.04649	0.04616
Ccapm	-0.0112	-0.00089	0.00098	0.00483	0.00867	0.01435	0.02104	0.0372	0.0390	0.04666	0.04579
U3fac	-0.0017	-0.00079	0.00118	0.00519	0.00925	0.01529	0.02163	0.0355	0.0394	0.04657	0.04622
C3fac	-0.0012	-0.00097	0.00133	0.00534	0.00873	0.01463	0.02145	0.0373	0.0392	0.04695	0.04561
UFF	-0.0019	-0.00058	0.00149	0.00532	0.00993	0.01575	0.02209	0.0360	0.0393	0.04644	0.04631
CFF	-0.0007	-0.00128	0.00129	0.00491	0.00901	0.01417	0.01957	0.0375	0.0397	0.04627	0.04605
UAPT	-0.0021	-0.00165	-0.00002	0.00428	0.0093	0.01572	0.02151	0.0356	0.0397	0.04636	0.04752
CAPT	-0.0025	-0.00205	-0.00048	0.00384	0.0086	0.01397	0.02095	0.0355	0.0397	0.04534	0.04791
U4fac	-0.0022	0.00028	0.00130	0.00303	0.00779	0.01482	0.01899	0.0321	0.0376	0.04056	0.04248
C4fac	0.0009	-0.00126	-0.00154	-0.00149	0.00526	0.00958	0.01549	0.0322	0.0345	0.04177	0.03845
Upem	-0.0018	-0.00086	0.00125	0.00529	0.00936	0.01519	0.02158	0.0357	0.0394	0.04637	0.04617
Cpem	-0.0013	-0.00105	0.00109	0.00473	0.00869	0.01438	0.02146	0.0378	0.0397	0.04610	0.04578
Unum	0.0000	0.00000	0.00000	0.01165	0.00764	0.02249	0.02157	0.0353	0.0400	0.04828	0.04835
Cnum	-0.2085	-0.26184	-0.02144	-0.03494	-0.19293	-0.22835	0.00510	0.0000	0.0000	0.00000	0.00000
UBC	-0.0021	-0.00091	0.00121	0.00506	0.00948	0.01551	0.02181	0.0357	0.0392	0.04634	0.04609
CBC	-0.0009	-0.00018	0.00177	0.00521	0.00967	0.01447	0.02212	0.0378	0.0403	0.04665	0.04707
Umean	-0.00192	-0.00075	0.00108	0.00476	0.00919	0.01536	0.02131	0.03514	0.03914	0.04559	0.04585
Cmean	-0.00124	-0.00109	0.00063	0.00391	0.00837	0.01365	0.02029	0.03647	0.03887	0.04568	0.04524

Table V, page 2

PANEL B: STOCK PICKERS: T-ratios for Alpha

γ =	0	0.10	0.20	0.25	0.30	0.35	0.40	0.50	0.60	0.70	0.80
Ucapm	-1.782	-0.810	1.0306	3.392	5.3597	7.3395	10.474	13.652	16.870	17.662	17.841
Ccapm	-1.234	-0.969	0.8889	3.254	5.1544	7.1033	10.535	14.699	17.061	18.333	18.101
U3fac	-1.801	-0.818	1.0370	3.418	5.4096	7.3992	10.594	13.697	16.999	17.816	17.954
C3fac	-1.346	-1.059	1.2211	3.600	5.1798	7.2096	10.754	14.366	17.204	18.498	17.860
UFF	-1.892	-0.575	1.2941	3.496	5.6197	7.4028	10.756	14.004	16.754	17.465	18.046
CFF	-0.773	-1.239	1.1099	3.091	5.0309	6.4928	8.8363	14.114	16.446	16.976	17.193
UAPT	-2.059	-1.660	-0.022	2.897	5.6794	7.5496	10.514	13.640	17.550	17.376	18.676
CAPT	-2.776	-2.209	-0.447	2.703	5.4236	7.0883	11.141	14.421	18.043	17.888	19.472
U4fac	-1.504	0.085	0.6646	0.941	1.9786	4.0348	3.946	6.529	5.6697	3.416	7.895
C4fac	-0.429	-0.527	-0.5638	-0.336	1.1439	1.5609	2.613	5.663	5.4207	5.179	5.501
Upem	-1.625	-0.815	1.0358	3.452	5.2338	6.9908	9.807	14.035	16.443	16.795	17.226
Cpem	-1.199	-0.998	0.8826	2.849	4.6608	6.5091	9.137	13.681	15.667	15.948	16.099
Unum	0.000	0.000	0.0000	0.006	0.0035	0.0103	0.010	0.0144	0.0170	0.0221	0.0235
Cnum	-0.002	-0.002	-0.0001	-0.000	-0.0015	-0.0009	0.000	0.0000	0.0000	0.0000	0.0000
UBC	-1.852	-0.862	0.9977	3.375	5.2879	7.2364	9.879	13.763	16.401	17.033	17.343
CBC	-0.883	-0.184	1.4774	3.299	5.4932	7.0509	10.485	14.291	17.479	17.496	17.499
Umean	-1.788	-0.779	0.8625	2.996	4.9384	6.8504	9.4243	12.76	15.2409	15.366	16.426
Cmean	-1.236	-1.026	0.6527	2.637	4.5838	6.1449	9.0716	13.034	15.3315	15.759	15.961

PANEL C: MARKET TIMERS: Estimates of Alpha

γ =	0	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.0
Ucapm	0.00058	0.0018	0.0034	0.0046	0.0056	0.00866	0.0094	0.0117	0.0143	0.0139	0.0169
Ccapm	-0.00094	0.0021	0.0028	0.0057	0.0054	0.00649	0.0075	0.0097	0.0121	0.0135	0.0150
U3fac	0.00058	0.0019	0.0034	0.0046	0.0056	0.00872	0.0095	0.0117	0.0143	0.0139	0.0169
C3fac	-0.00104	0.0023	0.0025	0.0056	0.0052	0.00669	0.0076	0.0096	0.0122	0.0137	0.0150
UFF	0.00052	0.0022	0.0039	0.0046	0.0055	0.00872	0.0097	0.0117	0.0144	0.0136	0.0169
CFF	0.00018	0.0011	0.0027	0.0065	0.0057	0.00664	0.0072	0.0099	0.0124	0.0125	0.0147
UApt	0.00089	0.0015	0.0038	0.0039	0.0061	0.00801	0.0089	0.0118	0.0135	0.0132	0.0161
CApt	-0.00083	0.0011	0.0028	0.0042	0.0057	0.00667	0.0073	0.0098	0.0119	0.0129	0.0148
U4fac	-0.00097	0.0000	0.0038	0.0064	0.0067	0.00750	0.0071	0.0095	0.0120	0.0125	0.0148
C4fac	-0.00310	-0.0023	0.0076	0.0072	0.0097	0.00629	0.0080	0.0098	0.0116	0.0134	0.0149
Upem	0.00071	0.0019	0.0035	0.0050	0.0056	0.00873	0.0091	0.0117	0.0142	0.0138	0.0169
Cpem	-0.00084	0.0018	0.0035	0.0060	0.0056	0.00686	0.0077	0.0103	0.0123	0.0131	0.0152
Unum	0.00000	0.0039	0.0033	0.0044	0.0057	0.01017	0.0076	0.0118	0.0130	0.0135	0.0157
Cnum	-0.25733	-0.0291	-0.0039	-0.2497	0.0000	-0.19672	-0.0200	0.0000	0.0000	0.0000	0.0000
UBC	-0.00196	0.0017	0.0037	0.0045	0.0053	0.00934	0.0099	0.0123	0.0147	0.0141	0.0175
CBC	-0.00042	0.0026	0.0026	0.0063	0.0057	0.00716	0.0084	0.0095	0.0129	0.0138	0.0157
Umean	0.00005	0.0016	0.0036	0.0048	0.0058	0.00853	0.0091	0.0115	0.0139	0.0136	0.0166
Cmean	-0.00099	0.00124	0.0035	0.0059	0.0061	0.00669	0.0077	0.0098	0.0122	0.0133	0.01504

Table V, page 4

PANEL D: MARKET TIMERS: T-ratios for Alpha

γ =	0	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.0
Ucapm	0.5398	1.709	3.191	4.535	5.448	8.949	9.636	12.916	17.657	17.605	24.114
Ccapm	-1.6068	2.267	2.949	6.044	5.983	7.464	8.624	12.273	16.823	19.967	24.607
U3fac	0.5313	1.708	3.149	4.466	5.374	8.997	9.649	12.796	17.396	17.404	23.856
C3fac	-1.0687	2.386	2.616	5.752	5.484	7.590	8.417	11.658	16.377	20.129	24.186
UFF	0.4604	1.896	3.472	4.546	4.958	8.571	9.972	12.069	16.486	15.493	22.860
CFF	0.1707	1.005	2.364	6.119	5.214	6.866	7.694	10.506	15.076	14.803	21.684
UApt	0.7992	1.409	3.493	3.714	5.777	7.849	8.668	12.733	15.805	15.377	21.491
CApt	-0.8498	1.084	2.847	4.363	6.060	7.518	8.087	12.298	15.801	17.589	23.104
U4fac	-0.5175	0.017	2.174	2.931	3.385	4.692	4.747	4.4030	6.2784	9.6236	8.9699
C4fac	-1.3931	-0.765	3.011	2.421	3.536	3.099	3.600	4.5043	6.4141	6.3667	8.7328
Upem	0.5271	1.423	2.659	4.303	4.270	7.233	8.265	10.177	14.317	13.047	19.888
Cpem	-0.7176	1.512	2.963	5.418	4.835	6.579	7.620	10.170	14.689	14.857	22.035
Unum	0.0000	0.003	0.003	0.004	0.005	0.008	0.006	0.0098	0.0107	0.0113	0.0129
Cnum	-0.0035	-0.000	-0.000	-0.004	0.000	-0.006	-0.000	0.0000	0.0000	0.0000	0.0000
UBC	-2.0128	1.205	2.524	4.003	3.443	7.805	9.702	10.684	15.562	11.639	21.708
CBC	-0.4825	2.437	2.343	5.889	5.507	7.223	8.849	9.5886	16.574	19.231	24.016
Umean	0.0467	1.338	2.952	4.071	4.665	7.728	8.663	10.825	14.786	14.312	20.4124
Cmean	-0.8497	1.418	2.727	5.144	5.2313	6.619	7.5558	10.143	14.536	16.135	21.1949

Table 6

Tests with Mutual Fund Returns

Tests are based on 188 equity mutual funds and monthly data for the January, 1977 to December, 1993 period. The number of observations is 204. The distribution of funds' alphas for each model are summarized by the values of alpha at various fractiles of the distribution for the 188 funds. Umean and Cmean are the averages, taken across the unconditional and conditional models, respectively, but not including the numeraire portfolio models. The symbols denoting the various models are the same as in Table III. The symbols denoting each model are the same as in Table III. Bonfor. Max. (Min>) are the p-values based on the Bonferoni inequality using the t-ratios for the alphas and assuming a t distribution. These are the one-tailed areas, or p-values, associated with the maximum (minimum) of the 188 p-values, multiplied by 188.

	Bonfor. Min.	Left 2.5%	Left 5%	Left 10%	Mean	Median	Right 10%	Right 5%	Right 2.5%	Bonfor. Max.
Ucapm	0.089	-0.008	-0.005	-0.003	-0.001	-0.0007	0.001	0.002	0.002	0.000
Ccapm	0.519	-0.008	-0.005	-0.003	-0.001	-0.0008	0.001	0.002	0.002	0.000
U3fac	0.059	-0.008	-0.005	-0.003	-0.0007	-0.0006	0.001	0.002	0.002	0.000
C3fac	0.551	-0.008	-0.006	-0.004	-0.001	-0.001	0.0009	0.001	0.002	0.000
UFF	0.007	-0.009	-0.005	-0.003	-0.0006	-0.0004	0.002	0.002	0.003	0.000
CFF	0.154	-0.010	-0.006	-0.004	-0.001	-0.0008	0.002	0.002	0.002	0.000
UAPT	0.161	-0.008	-0.005	-0.003	-0.0008	-0.0006	0.0013	0.002	0.003	0.000
CAPT	0.032	-0.009	-0.006	-0.003	-0.0011	-0.0008	0.0013	0.002	0.003	0.000
U4fac	1.000	-0.006	-0.005	-0.004	-0.001	-0.0007	0.002	0.003	0.004	0.450
C4fac	1.000	-0.010	-0.008	-0.005	-0.002	-0.001	0.002	0.003	0.004	0.002
Upem	0.746	-0.00781	-0.005	-0.003	-0.0008	-0.00067	0.0015	0.002	0.0024	0.000
Cpem	0.580	-0.00950	-0.006	-0.004	-0.0012	-0.00094	0.0013	0.002	0.0023	0.000
Unum	1.000	-0.00208	-0.0005	0.000	0.0012	0.000	0.0051	0.006	0.0064	1.000
Cnum	1.000	-0.00250	0.0114	0.024	0.177	0.207	0.2780	0.286	0.290	1.000
UBC	0.183	-0.00852	-0.0051	-0.003	-0.0006	-0.00044	0.0018	0.002	0.0026	0.000
CBC	0.066	-0.00956	-0.0062	-0.004	-0.0012	-0.00095	0.0015	0.002	0.0021	0.000
Umean	-0.00790	-0.00501	-0.00314	-0.00075	-0.00058	0.001514	0.002	0.002714		
Cmean	-0.00915	-0.00617	-0.00385	-0.00121	-0.00089	0.001428	0.002	0.002488		