January 2007

Conditional Performance Evaluation for German Mutual Equity Funds

WWZ Working Paper 22/07

Wolfgang Bessler, Wolfgang Drobetz, Heinz Zimmermann

The Author(s):

Prof. Dr. Wolfgang Bessler
Chair of Banking and Finance, Justus-Liebig-Universitey Giessen
Licher Strasse 74, D- 35394 Giessen
wolfgang.bessler@wirtschaft.uni-giessen.de

Prof. Dr. Wolfang Drobetz
Institute of Finance, University of Hamburg
Von-Melle-Park 5, D-20146 Hamburg
wolfgang.drobetz@wiso.uni-hamburg.de

Prof. Dr. Heinz Zimmermann
Center of Business and Economics (WWZ), University of Basel
Department of Finance
Holbeinstrasse 12, CH-4051 Basel
heinz.zimmermann@unibas.ch

A publication oft the Center of Business and Economics (WWZ), University of Basel.

Contact:
WWZ Forum | Petersgraben 51 | CH-4003 Basel | forum-wwz@unibas.ch | www.wwz.unibas.ch

Conditional Performance Evaluation for German Mutual Equity Funds

Wolfgang Bessler[a], Wolfgang Drobetz[b], and Heinz Zimmermann

First Version: January 2007

Abstract

We investigate the performance of a sample of German mutual equity funds over the period from 1994 to 2003. Our general finding is that mutual funds, on average, hardly produce excess returns relative to their benchmark that are large enough to cover their expenses. This conclusion is drawn from a variety of model specifications and is robust to many different benchmarks. Compared to unconditional measures, fund performance substantially deteriorates when we measure conditional alphas both in single-index and multi-factor models. We also measure fund performance in the Euler-equation framework and test several specifications of the stochastic discount factor using GMM. The result that funds underperform even before costs is even more pronounced. Overall, given the fact that stock returns are to some extent predictable by using publicly available information, conditional analysis raises the hurdle for active managers seeking abnormal positive performance, because it gives them no credit for exploiting readily available information.

Keywords: Mutual funds, stock return predictability, conditional performance measurement, stochastic discount factor.

EFM classification codes: 380, 370, 350, 310.

[a] Wolfgang Bessler, Chair of Banking and Finance, Justus-Liebig-University Giessen, Licher Strasse 74, Giessen, Germany, Phone: +49-641-9922460, Mail: wolfgang.bessler@wirtschaft.uni-giessen.de.

[b] Wolfgang Drobetz, Chair of Corporate Finance and Ship Finance, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany, Phone: +49-40-42838-5506, Mail: wolfgang.drobetz@wiso.uni-hamburg.de.

[c] Heinz Zimmermann, Department of Finance, University of Basel, Petersgraben 51, 4003 Basel, Switzerland, Phone: +41-61-2673316, Mail: heinz.zimmermann@unibas.ch.

1. Introduction

The public interest in the German mutual equity fund industry has grown rapidly in the past years. For example, in 1990 the total mutual fund holdings were worth € 128.9 billion. Ten years later, the holdings in mutual funds has increased sharply to € 923.1 billion. German funds experienced net capital inflows even during the bear periods on the stock markets and mutual fund holdings amounted to € 1'003.0 billion by the end of 2003 (BVI, 2004).[1] Equity funds (including balanced funds) account for 54% of total mutual fund holdings (Maurer, 2004). This strong interest in equity investments might be explained by the commonly held belief that they should play a major role in the development of the German pension system which is shifting away from pure defined benefit plans towards a system which puts more emphasis on private pension savings. Nevertheless, as measured by US standards, the German mutual funds industry is still fairly small. By the end the year 2000, the capital invested in mutual funds in the US amounted to € 27'570 per capita, while the corresponding figure in Germany was only € 5'154. Krahnen et al. (1997) and Theissen (2004) argue that the minor importance of mutual funds in Germany is a side-effect of its bank-based system its comparatively small stock market. A well-known feature of the German bank-based financial system is that a few big universal banks dominate the capital markets (e.g., Hackethal, 2004), and this is also inevitably the case in the German mutual fund industry. The largest mutual fund companies ("Kapitalanlagegesellschaften", KAGs) are bank subsidiaries and most of their retail business is done over the bank-counter. Banks are inclined to sell their own funds, and advise their customers in this direction. Therefore, the mutual fund industry is characterized by a restrictive distribution network, where the savings banks ("Sparkassen"), the credit-cooperatives ("Genossenschaften"), and the four large universal banks (Commerzbank, Deutsche Bank, Dresdner Bank, HypoVereinsbank) account for about 80% of all managed assets (Maurer, 2004).

Given the increased importance of the German mutual fund industry, performance measurement has also become a crucial issue at least from an investor's perspective. In most previous studies unconditional measures of performance compare the average return on an asset with an appropriate benchmark designed to control for the asset's average risk. The returns and betas are measured as averages over the evaluation period. These averages are taken uncondi-

---

1 These number include both retail funds ("Publikumfonds") and institutional funds ("Spezialfonds"). Institutional funds account for more than 50% of these figures.

tionally, i.e., no other information than past returns are used in the estimation. In contrast, the focus of this paper is on conditional performance measurement. Conditional performance measures relate to changes in the state of the economy and explicitly account for time variation in expected returns. The key ingredient of the models is that they include information on the correlation structure between a fund's beta and expected market returns, assuming semi-efficient markets and that both variables are dependent on publicly observable variables.

If expected returns and risks vary over time, an unconditional approach to performance measurement is likely to be unreliable. Already Jensen (1972) notes that the time-variation in risks and risk premiums will be confused with average performance. However, previous empirical studies interpret the variation in mutual fund risk and risk premiums as reflecting superior information or market timing (e.g., Merton, 1981). In a new strand of the literature, Ferson and Schadt (1996) were the first to empirically test the hypothesis that any model should not ascribe superior performance to a managed portfolio strategy that can be replicated using publicly available information.[2] Their analysis builds on the results of a set of empirical studies showing that the returns and risks of stocks and bonds are at least partly predictable over time, e.g., using dividend yields and interest rates (e.g., Cochrane, 1999; Schwert, 2002; Campbell and Thompson, 2005). If predictability reflects changing return expectations, valid performance measures also need to reflect this time variation. The empirical asset pricing literature provides convincing evidence that predictability based on ex ante observable instrument variables can (partly) be explained on the basis of time variation in expected returns (e.g., Ferson and Harvey, 1991, 1993; Ferson and Korajzyk, 1995). In addition, conditional versions of asset pricing models are better able to explain the cross-sectional variation in expected returns (e.g., Cochrane, 1996; Jagannathan and Wang, 1996).

Our results indicate that – compared to unconditional models – fund performance substantially deteriorates when we measure conditional alphas both in single-index and multi-factor models. For example, based on the full set of information variables, the mean conditional alpha for our sample of general funds is estimated to be -0.130% per month, or about 1.5% per year. In comparison, taking a simple average for all funds over the 1997 to 2003 period, the mean total expense ratio is only 1.06%. Accordingly, if we add back management fees, German mutual equity funds underperform already on a before fee basis. Given that stock returns

---

2 See already Breen, Glosten, and Jagannathen (1989).

are partly predictable using publicly available information, that part of fund performance that is attributable to time-variation in expected returns should be deducted from fund performance. In performance evaluation tests based on stochastic discount factor models, underperformance is even more pronounced and can even be as low as 4% per year. We conclude that conditional analysis raises the hurdle on active managers who seek abnormal positive performance because it gives them no credit for using readily available information, and this makes it more likely for funds to show no abnormal performance. The pronounced level shift in fund performance when switching from the beta-pricing framework to the stochastic discount factor framework (evening unconditional models) can be attributed to the fact that more "complicated" stochastic discount factor models and a larger number of primitive assets lead to stronger pricing conditions.

The remainder of this paper is as follows. In section 2 we introduce conditional performance evaluation techniques both in a beta pricing framework and in a stochastic discount factor (SDF) framework. Section 3 describes the data, and section 4 presents our empirical results.

2. Conditional performance measurement

2.1. Conditional performance evaluation in a beta pricing framework

Ferson and Schadt (1996) propose a measure of conditional performance evaluation that is consistent with Fama's (1970) semi-strong form of market efficiency. Their approach modifies the traditional Jensen's alpha to incorporate conditional information, as represented by a set of information variables. The model incorporates time-varying betas and exploits the (unconditional) correlation between these conditional betas and the public information variables. Time-variation in betas of managed portfolios may come from three different sources. First, the betas of the portfolio assets may change over time. Second, the weights of a passive investment strategy (e.g., buy-and-hold) vary as relative values change. And third, a manager can actively manipulate the portfolio weights by deviating from a buy-and-hold strategy. To capture the combined effect of these factors on the risk exposures, start with the following specification:[3]

(1a) $r_{P,t+1} = \beta_P(Z_t)\, r_{B,t+1} + \varepsilon_{P,t+1}$,

3 See Ferson and Schadt (1996), p. 429f.

(1b) $E\left[\varepsilon_{P,t+1} \middle| Z_t\right] = 0$,

(1c) $E\left[\varepsilon_{P,T+1} r_{B,t+1} \middle| Z_t\right] = 0$,

where Z_t in equation (1a) is a vector of information variables as a proxy for the full set of information available at time t, and $\beta_P(Z_t)$ denotes the time t conditional market betas of the excess return on portfolio P. Equation (1b) assumes semi-strong market efficiency, and equation (1c) implies that $\beta_P(Z_t)$ are conditional regression coefficients. This model implies that any unbiased forecast of the difference between the portfolio return and the product of the portfolio beta and the excess return on the market portfolio which differs from zero must be based on a set of information that is more informative than Z_t. If an intercept (alpha) is included into the model, it should be zero, and the error term is assumed to be unrelated to public information variables, i.e., their informational content is fully exploited in the regression. In empirical tests a specific function form of $\beta_P(Z_t)$ is required to estimate the model. Following Shanken (1990), Ferson and Schadt (1996) assume a linear form for the changing conditional beta of a managed portfolio:

(2) $\beta_{P,t+1} = \beta_P(Z_t) = b_{0P} + B'_P z_t$,

where $z_t = Z_t - E(Z)$ is the normalized vector of deviations of the information variables from their unconditional means, and B_P is a coefficient vector with the same dimension as Z_t. The coefficient b_{0P} can be interpreted as the “average beta”, i.e., the beta when all instrument variables are equal to their means. The elements of B_P measure the sensitivity of the conditional beta to deviations of the information variables from their means. The specification in equation (2) is consistent with the findings in Ferson, Sarkissian, and Simin (2003a,b), who suggest to de-mean autocorrelated information variables to reduce the risk of uncovering spurious predictive relations.

Combining equations (1) and (2) implies the following generating process for the returns of a managed portfolio:

(3a) $r_{P,t+1} = b_{0P} r_{B,t+1} + B'_P \left[z_t r_{B,t+1}\right] + \varepsilon_{P,t+1}$,

(3b) $E\left[\varepsilon_{P,t+1} \middle| Z_t\right] = 0$, (3c) $E\left[\varepsilon_{P,t+1} r_{B,t+1} \middle| Z_t\right] = 0$.

Empirical tests of the model in equation (3) require running an ordinary least square regression of a managed portfolio excess returns upon the market factor and the product of the market factor with lagged information:

(4) $r_{P,t+1} = \alpha_P + \delta_{1P} r_{B,t+1} + \delta'_{2P}\left(z_t r_{B,t+1}\right) + \varepsilon_{P,t+1}$.

Taking expected values of equation (4) reveals that the model implies $\alpha_P = 0$, $\delta_{1P} = b_{0P}$, and $\delta_{2P} = B_P$.[4] The products of the future benchmark return and the predetermined information variables capture the covariance between the conditional beta and the conditional expected market return (given z_t). The interaction terms serve as a control for common movements in a fund's conditional beta and the conditional expected benchmark return. The conditional alpha, α_P, is measured net of the effects of these risk dynamics. When $B_P = 0$, the managed portfolio beta is not a function of public information, hence, the conditional and unconditional betas are the same. Otherwise, an unconditional alpha contains a bias caused by the common variation in betas and expected market returns, and the covariance between beta and future market returns could falsely be interpreted as the result of a portfolio manager's superior performance (e.g., Grinblatt and Titman, 1989). In contrast, the model in equation (4) assumes that managers do not possess superior information. The null hypothesis of no abnormal performance explicitly allows for a covariance between beta and the future market returns because both variables depend on public information that is observable for the econometrician.[5]

In another strand of the asset pricing literature, Jagannathan and Wang (1996) demonstrate that the conditional capital asset pricing model can be interpreted as an unconditional model for average expected returns with more than one beta. This interpretation is also suitable for the model in equation (4), where (δ_{1P}, δ_{2P}) is a vector of regression coefficients or betas on the multiple factors, which are defined as ($r_{B,t+1}$, $z_t\, r_{B,t+1}$). The benchmark return is the first factor

4 Bansal and Harvey (1996) employ a similar approach, using a nonparametric benchmark. While the alpha in equation (4) reflects the average pricing error based on a conditional version of the conditional capital asset pricing model (parametric benchmark), their alpha represents the pricing error from a benchmark return that is constructed to be exactly efficient for a broad class of payoff which includes dynamic trading strategies.

5 Already Admati and Ross (1985) suggest that if the actions (portfolio choices) of the manager are observable ex post, then these actions should be regressed on the public information available when these choices were made, on the ex post observable returns, and on cross-products of the two.

and the product of the benchmark return and the lagged information variables are the additional factors. The latter can be interpreted as the returns to dynamic trading strategies, which hold z_t units of the market index, financed by borrowing or selling z_t units in treasury bills. Hansen and Jagannathan (1991) and Cochrane (1996) argue that scaled returns can be interpreted as the payoffs of actively managed portfolios.

2.2. Conditional performance evaluation in a stochastic discount factor framework

2.2.1. Admissible performance measures

Chen and Knez (1996) were the first to introduce performance evaluation in a stochastic discount factor (SDF) framework. They define a fund's conditional alpha for a given stochastic discount factor, denoted as m_{t+1}, as follows:

$$(5) \qquad \alpha_{Pt} \equiv E\left(m_{t+1} R_{P,t+1} \middle| Z_t\right) - 1,$$

where we denote α_{Pt} as the SDF alpha, and $R_{P,t+1}$ is the gross return of the fund at time t+1. The performance of the fund measures the difference between the expected risk adjusted gross return of the fund (conditional in Z_t) and its price, which is 1. If the performance is positive (negative), this implies that the fund offers a higher (lower) risk-adjusted excess return than expected, indicating superior (inferior) performance.[6] However, the theoretical results in Chen and Knez (1996) indicate several measurement problems with SDF alphas. Most important, unless the manager's fund lies in the payoff space generated by the primitive assets, any performance measure is possible.[7] Moreover, any relative ranking is possible for a whole set of managed funds lying outside uninformed investors' investment opportunity set. As a minimum requirement for performance measurement tests to deliver meaningful results, Farnsworth et al. (2002) require that if a given stochastic discount factor prices the set of N primitive assets, then α_{Pt} in equation (5) must be zero, provided that the fund forms a portfolio of these primitive assets (at no costs) and that the portfolio strategy uses only public information available at time t. In this case, we have $R_{P,t+1} = x(Z_t)' R_{t+1}$, where $x(Z_t)$ is a vector of portfolio

6 For the link to performance measures in a beta-pricing framework see Söderlind (1999).

7 This result ultimately rests upon Hansen and Jagannathan's (1991) projection argument for the (infinite) set of admissible stochastic discount factors.

weights depending on a set of information variables, and the corresponding α_{Pt} is an admissible performance measure, assigning zero performance to an uninformed manager's portfolio:[8]

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

This is, however, only a minimum requirement and cannot solve the more general measurement problems described in Chen and Knez (1996). In fact, equation (6) contains two strong assumptions: (i) m_{t+1} is already a valid stochastic discount factor, and (ii) the fund's return spans the benchmark return and, hence, is achievable by the uninformed investor. For this special case Chen and Knez (1996) prove that all admissible performance measures will assign zero performance to the fund.[9] However, in the more interesting case where a fund's return lies outside the uninformed investor's investment opportunity set, the arbitrariness of performance measures described above still remains an unsolved problem.

The SDF alpha in equation (5) will only produce valid inferences if the candidate stochastic discount factor satisfies the Euler equation for a set of N primitive assets. Therefore, we simultaneously estimate the parameters of the candidate stochastic discount factors and the SDF alpha of the fund using Hansen's (1982) Generalized Method of Moments (GMM). Following Farnsworth et al. (2002), we form the following system of equations, which is subject to our empirical tests:

(7) $$\begin{aligned} u1_t &= \left(m_{t+1}R_{t+1} - 1\right) \otimes Z_t \\ u2_t &= \alpha_P - m_{t+1}R_{P,t+1} + 1, \end{aligned}$$

where $u1_t$ denotes the vector of pricing errors relating to the primitive assets (or, reference assets), whose gross returns are collected in the vector R_{t+1}, and $u2_t$ is the pricing error of the fund with gross return $R_{P,t+1}$. The parameter α_P is the mean of the conditional SDF alpha, defined in equation (5). In the simplest case, where the vector of information variables, Z_t, only contains a constant, α_P measures an unconditional SDF alpha. Under the null hypothesis of no abnormal fund performance, α_P should be equal to zero. Farnsworth et al. (2002) use this sys-

8 See Farnsworth et al. (2002), p. 476.

9 See Chen and Knez, (1996), theorem 4, p. 525.

tem of equations to estimate the performance of U.S. mutual funds, and Fletcher and Forbes (2004) apply the same framework to assess the performance of U.K. unit trusts.[10]

In our empirical analysis, we exploit an additional feature of the stochastic discount factor framework. The Euler-equation implies that the mean of the stochastic discount factor should be equal to the inverse of the gross return on a risk-free security or, more generally, a zero-beta asset, hence, $E(m_{t+1}) = 1/E(R_{f,t+1})$. This condition can be imposed by including a proxy for the risk-free security or a zero-beta asset into the set of primitive assets. Dahlquist and Söderlind (1999) and Farnsworth et al. (2002) demonstrate that failure to impose this restriction on the stochastic discount factor can result in estimation of a valid discount factor which implies a mean-variance tangency portfolio that is not on the efficient frontier.[11] Therefore, in addition to 10 German sector portfolios as our set of primitive assets, we also include a moment condition for the risk-free security in our estimations to impose the mean restriction on the stochastic discount factor.

A problem in the GMM estimation procedure is that the number of moment conditions grows exponentially if many funds are evaluated at the same time. We therefore estimate the system of equations in (7) separately for each fund in our sample. Farnsworth et al. (2002) prove that this procedure is not restrictive. Estimating the system of equations for one fund at a time is equivalent to estimating a system with many funds simultaneously. The estimates of α_P and the standard errors for any subset of funds are invariant to the presence of another subset of funds in the system.[12]

2.2.2. Modeling the stochastic discount factor

We use the SDF alpha in the system of equations in (7) to estimate mutual fund performance. However, the SDF alpha depends on the specification of the stochastic discount factor, and this is not unique unless financial markets are complete. This indeterminacy implies that different models of the stochastic discount factor can measure fund performance differently and, hence, produce different SDF alphas. We employ the following three broad model classes: (i) linear factor model SDFs, (ii) primitive-efficient SDFs, and (iii) Bakshi-Chen SDFs. Linear

10 Chen and Knez (1996) and Dahlquist and Söderlind (1999) use a different methodology to estimate fund performance in a stochastic discount factor framework that is directly based on the minimum value of the quadratic form and the related *J*-statistic.

11 See also Hansen and Jagannthan (1997), Jagannthan and Wang (2002), and Dittmar (2003).

12 See Farnsworth et al. (2002), appendix, p. 499.

factor models can capture the traditional CAPM specification and the three-factor model of Fama and French (1993).

Linear factor model SDFs: Linear factor models are models where m_{t+1} is linear in prespecified factors. They can potentially capture already many different asset pricing models, including the capital asset pricing model and the three-factor model of Fama and French (1993). They can also represent capital asset pricing models with higher moments (Dittmar, 2003). In the most general form, we have that:

(8) $$m_{t+1} = a(Z_t) + b(Z_t)F_{t+1},$$

where F_{t+1} denotes the vector of (gross) returns at time t+1 on traded or non-traded factors. As before, Z_t is a vector of predetermined information variables with dimension L×1, where L denotes the number of information variables plus a constant. In the unconditional case, Z_t reduces to the constant and, hence, the weight parameters a and b are time-constant. A particularly interesting case to analyze is the capital asset pricing model, where m_{t+1} is a linear function of the market portfolio (Dybvig and Ingersoll, 1982). This notion can be generalized to multiple-beta pricing model (e.g., Chen, Roll, and Ross, 1986). In fact, Ferson and Jagannathan (1996) prove that any multiple-beta model can be expressed in the Euler-equation form with a particular specification of the stochastic discount factor.

We test both the capital asset pricing models and the Fama-French (1993) three factor specification of the stochastic discount factor. The stochastic discount factor in the unconditional version of the capital asset pricing model is given as:

(9) $$m_{t+1} = a + bR_{M,t+1},$$

where $R_{M,t+1}$ is the gross market return. The stochastic discount factor for the Fama-French (1993) three-factor model is specified as follows:

(10) $$m_{t+1} = a + b_1R_{M,t+1} + b_2HML_{t+1} + b_3SMB_{t+1},$$

where HML_{t+1} and SMB_{t+1} are the gross returns on the high minus low book-to-market and small minus big style portfolios, respectively. Plugging equations (9) and (10) into the Euler-equation and estimating the system of equations in (7) for the set of primitive assets plus the risk-free security identifies the a and b parameters.

Hansen and Richard (1987) argue that conditional factor models are not directly testable because the econometrician cannot observe an agent's information set. However, given a set of predetermined information variables, Cochrane (1996, 2001) suggests that conditional factor pricing model can be tested by scaling factors. Conditional factor pricing models allow that the parameters are time-varying, and again the simplest solution is to assume linearity. With a single factor and a single information variable (plus a constant), denoted by z_t, we have:[13]

(11) $$m_{t+1} = a(z_t) + b(z_t)F_{t+1} = a_0 + a_1 z_t + (b_0 + b_1 z_t)F_{t+1} = a_0 + a_1 z_t + b_0 F_{t+1} + b_1(z_t F_{t+1}).$$

Therefore, instead of a one-factor model with time-varying coefficients, we now have a three factor model (z_t, F_{t+1}, $z_t F_{t+1}$) with fixed coefficients (plus constant). Linearity is convenient, but not restrictive.[14] Cochrane (1996) suggests that one can simply add scaled factors and estimate the *unconditional* moments of the model as if conditioning information did not exist.

If there are K factors and L information variables (including a constant), there are $L \times (K+1)$ parameters to estimate in equation (11). To identify the parameters, the number of primitive assets (N) must suffice that $N \geq K+1$. Otherwise, the number of parameters is larger than the number of orthogonality conditions in GMM estimation.

Primitive-efficient SDFs: Hansen and Jagannathan (1991) that the solution of the conditional projection of a stochastic discount factor m_{t+1} on the vector of primitive asset returns R_{t+1} is:

(12) $$m_{t+1}^{PE} = 1' E\left(R_{t+1} R'_{t+1} \middle| Z_t\right)^{-1} R_{t+1}.$$

The stochastic discount factor in equation (12) is admissible and prices a given set of primitive assets by construction. m_{t+1}^{PE} is usually called a primitive-efficient stochastic discount factor, because it is a linear function of the returns on the set of primitive assets, where $1' E(R_{t+1} R'_{t+1})^{-1}$ provides the portfolio weights. Instead of computing the weights analytically, we follow Chen and Knez (1996), Dahlquist and Söderlind (1999), and Farnsworth, et al. (2002) and assume that the weight vector is a linear function of the information variables in the Z_t vector. Multiplying the primitive asset returns by lagged information variables expands the payoff space and results in "dynamic strategies" (Cochrane, 2001). Similar to the condi-

13 In this case, small capitalization letters do not indicate demeaned variables.
14 See Ferson and Siegel (2001) and Beakert and Liu (2004) for a discussion about optimal scaling.

tional linear factor models, this procedure results in scaled factor models that can be estimated using GMM as if the model was unconditional. With N primitive assets and L instrument variables (including a constant), there are $N \times L$ weight parameters to estimate and $N \times L$ orthogonality conditions, implying that the system of equations is exactly identified.

Bakshi-Chen SDFs: The existence of a strictly positive stochastic discount factor is equivalent to a no-arbitrage condition. Chen and Knez (1996) and Dahlquist and Söderlind (1999) impose positivity by cutting off the specification of the stochastic discount factor at zero. An alternative (ad hoc) approach that has been used extensively in the empirical asset pricing literature to make sure risk premiums are positive is to work with exponential functions.[15] Bakshi and Chen (1998) propose a model in which the stochastic discount factor is an exponential of a linear function of the logarithmic returns on the primitive assets. Farnsworth et al. (2002) also use this specification to evaluate the performance of U.S. mutual funds. Specifically, the Bakshi-Chen SDF is as follows:

$$(13) \qquad m_{t+1}^{BC} = \exp[C(Z_t) ln(R_{t+1})],$$

where C denotes an $N \times L$ matrix of coefficients. In this case, the stochastic discount factor is a nonlinear function of the primitive asset returns. The main advantage of this specification for the stochastic discount factor is that m_{t+1}^{BC} is guaranteed to take on only positive values. However, this comes at the cost that GMM estimation involving strong nonlinearities becomes more complicated. Again, with $N \times L$ coefficients and $N \times L$ orthogonality conditions, the system is exactly identified.

3. Data

Fund data: In this study we analyze the database of the Bundesverband Investment und Asset Management (BVI). Membership in the BVI is mandatory for all German mutual fund companies, institutional fund companies and asset management companies. Accordingly, the database contains all funds of its 74 members starting as early as 1950. We exclude funds that invest in foreign stock markets because they have different risk exposures that would require additional benchmarks to span the investment opportunity set. Most important, restricting the

---

15 For example, see Boudoukh, Richardson, and Smith (1993), Ostdiek (1998), De Santis and Gerard (1997, 1998), and Harvey (2001).

sample to domestic equity funds avoids the problems of dealing with currency risk (e.g., Paape, 2003). The funds are divided into regular equity mutual funds, small- and mid-cap funds, and growth funds. The sample contains monthly data on 98 general funds, 12 small- or mid-cap funds, and 6 growth funds, resulting in a total of 116 funds over the period from January 1980 to December 2003. The fund classification is taken from Hoppenstedt, and the investment objectives were double-checked by examining the fund fact sheets available from the websites of the fund management companies. We also exclude pure index funds, which passively track a specified index. The lengths of the available time series of these funds are very different (ranging from 288 months to 26 months), and in untabulated results we find that this characteristic of the sample can lead to biased results in our conditional performance measurement tests. Therefore, we finally use a reduced sample of 50 selected funds for which we have a full return history over the period from January 1994 to December 2003 (120 months). This smaller sample consists of 47 general funds and 3 small- and mid-cap funds. Although the number of funds may appear small in comparison with most U.S. studies, our sample is exhaustive for the German equity fund market. For comparison, Griese and Kempf (2003) also look at German equity mutual funds with an investment objective in German stocks. Their data are from Micropal and contain a slightly larger number of 123 funds. Bams and Otten (2002) analyze European equity mutual funds and include only 57 German funds (including index funds) in their sample.

We collect the mutual fund net asset values measured on a monthly basis form the BVI database and compute simple fund returns, assuming that total distributions (i.e., dividends and capital gain distributions, if any) were reinvested in the fund at the beginning of the following month. This implies that the returns are net of management fees and expenses but disregard load charges and exit fees (if any).

Survivorship bias: Previous studies have shown that survivorship issues can severely influence the results of performance measurement studies (e.g., Brown et al., 1992; Malkiel, 1995; Elton et al., 1996). The BVI database contains all funds for which return data exist in a given month. By including funds that do not survive until the end of the end of the sample period, we can control for a potential sample selection bias related to fund survival, i.e., our sample should not be afflicted from survivorship bias. Of the 98 general funds in our sample, only 75 have data available at December 31, 2003 (end of sample period). In the small- and mid-cap segment and the growth fund category there are 2 and 1 non-surviving funds, respectively.

The surprisingly high number of non-surviving general funds already indicates that neglecting the survivorship-bias might lead to inaccurate estimates of true fund performance.

We follow the approach in Malkiel (1995) to provide an estimate of the survivorship bias in our sample and use raw fund returns. Because the number of funds is only very small before 1994, we analyze our total sample of 116 (surviving and nonsurviving) funds that existed over the 1994-2003 sample period. The results are shown in table 1. In panel A, we report average annual returns of all funds in the different fund categories that existed continuously from the base year (1994, 1997, or 2000) through 2003. For example, the (equally-weighted) average annual return for the period from 1994 to 2003 was 4.97% for all continuously existing funds. This is the number one would obtain from normal databases if one asked what was the average annual return for all funds in existence at December 31, 2003 with a record of at least 10 years (120 months). All other percentages (for different base years and fund categories) can be interpreted accordingly. To assess the amount of survivorship bias, we also compute the average returns for all funds in existence in every year, starting from the base year until 2003, irrespective of whether the fund survived until December 31, 2003.[16] In our example given above, the equally-weighted average returns for *all* funds in our sample, including those that were liquidated during the 10 year period from 1994 to 2003, was only 4.52%, implying a survivorship bias of 45 basis points per year. This number is very close to the 40 basis points reported by Griese and Kempf (2003) for their sample of German mutual equity funds.

[Insert table 1 here]

In panel B of table 1 we take a slightly different perspective. In a first step, we calculate the mean return for all funds in existence each year from 1994 to 2002. In a second step, we compare the mean annual return from the funds that survived until 2003 with those that did not survive. We also report the mortality rate (attrition rate); mortality for a 1994 fund means that the fund was closed at some point in time between 1995 and 2003. The differences in the equally-weighted mean returns are substantial in absolute magnitude and in most instances also statistically significant, as indicated by a Welch *t*-test for equality of the mean returns of surviving and non-surviving funds. As expected, the mean return of surviving funds is always

16 Following Malkiel (1995), we dropped funds when they existed only for a partial year, but we included every fund that was in existence for the entire year even if it was closed later on.

higher than the mean return of non-surviving funds, i.e., non-surviving funs perform less well than surviving funds. In addition, we observe that the mortality rates are quite high. However, it decrease by construction as one moves down the respective column (i.e., as the length of the time period shrinks). The numbers imply an annual attrition rate of 2-3%, which is again consistent with the results by Griese and Kempf (2003).

Benchmark returns: To alleviate the benchmark problem (e.g., Roll, 1978; Roll and Ross, 1994, Grinblatt and Titman, 1994) we use a variety of benchmark indexes. Many funds in our sample explicitly use the DAX blue-chip index as their benchmark, as indicated in their fund prospectus. However, the DAX only consists of 30 blue-chip stocks and, hence, it may not capture the total investment opportunity set available to fund managers. Therefore, we also use the MSCI Germany Total Return Index Datastream Germany Total Return Index (with dividends reinvested) as benchmark portfolios and report the rank correlations of the estimated alphas. Many U.S. studies have applied much broader value-weighted stock market proxies. For example, Carhart (1997) used a value-weighted index of all NYSE, Amex, and Nasdaq stocks from the CRSP database. For Germany, the Karlsruher Kapitalmarktdatenbank (KKMDB) provides the DAFOX index, which represents a broad value-weighted market proxy of all stocks traded in the premier market segment ("Amtlicher Handel") of the Frankfurt Stock Exchange. The coefficients of correlation between the returns on the DAFOX index and the other indexes are very high; the correlation of the DAFOX index with the DAX 30 index is 0.97, and it is even 0.98 with both the MSCI and the Datastream indexes.

To test the Fama and French (1993) three-factor model, we use the style portfolios *SMB* and *HML* as additional style portfolios to capture the size effect (e.g., Banz, 1981) and the value effect (e.g., Fama and French, 1992; Lakonishok et al., 1994), respectively. The *SMB* factor is proxied by the monthly return difference between the return on the SMAX index and the DAX 30 index.[17] The *HML* factor is constructed using the data published on the website of Kenneth French.[18] We use the market-to-book ratio as the defining criterion and compute monthly return differences between the value portfolio (low market-to-book ratio) and the growth portfolio (high market-to-book ratio).

---

17 The SMAX index used to be the small- and medium-cap index provided by Deutsche Börse AG. Deutsche Börse AG changed its index classification scheme as of 2004, and the SMAX index has been continued as a small-cap index of 50 small stocks of the prime segment, denoted as SDAX.

18 See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library.

Primitive assets: In a stochastic discount factor framework, a set of reference assets is needed to determine the coefficients of the stochastic discount factor. The parameters are estimated using the Generalized Method of Moments such that the stochastic discount factor model produces a "small" pricing error in a system of Euler-equations. In theory, all assets available should be included in the estimation. However, similar to the problem that only a proxy for the true market portfolio can be used in securities market analysis, the econometrician can only employ a subset of available assets. The decision of which assets to include is guided by the type of assets in which the fund invests, i.e., the choice of primitive assets should reflect a fund manager's investment universe. In addition to the risk free rate, our set of primitive assets is therefore represented by the 10 German sector indexes according to the Datastream classification.[19] Table 2 provides summary statistics of these sector indexes over the period from January 1994 to December 2003. The R^2 denotes the explanatory power of predictive regressions of the sector index returns upon lagged values of the three information variables described below.

[Insert table 2 here]

Information variables: We apply a set of public information variables that previous studies have shown are useful for predicting stock returns and risk over time. The instrument variables are (i) the lagged dividend yield on the Datastream German Total Return Index, (ii) the lagged level of the 1-month interest rate for Euro deposits (German Mark deposits before January 1, 2002) on the Eurocurrency market, and (iii) the lagged slope of the term structure. The time t dividend yield is computed as the average value of dividends paid over the last 12 months on the (value-weighted) Datastream index, divided by the index value at time t. Following Lewellen (2004), we use the natural log of the dividend yield rather than the raw series, because it has better time series properties. The term spread is calculated as the difference between the yield on long term government bonds (with maturity of at least 10 years) and the 3-month interest rate for Euro (German Mark) deposits on the Eurocurrency market.

In order for conditional performance measurement models to produce meaningful results, a necessary condition is that the information variables have predictive power. The econometric

---

19 Chen and Knez (1996) also use a set of sector portfolios for all stocks listed on the NYSE as primitive assets. Farnsworth et al. (2002) use a broad market index and a set of style portfolios, and Fletcher and Forbes (2004) use size portfolios as primitive assets.

method used in most prediction studies is an ordinary least squares (OLS) regression of stock returns on the lag of the predictor variables. Based on conventional critical values for the *t*-test, these studies have concluded that there is evidence for predictability. However, recent studies argue that the apparent stock return predictability based on standard *t*-tests might be spurious (e.g., Stambaugh, 1999; Torous et al., 2004; Lewellen, 2004; Amihud and Hurvich, 2004; Campbell and Yogo, 2005). A problem arises if a predictor variable is highly persistent and its innovations are strongly correlated with returns.[20]

In panel A of table 3 the excess returns on our four benchmark indexes are regressed on the lagged predictor variables separately one at a time using ordinary least squares. We report conventional *t*-statistics as well as biased-adjusted *t*-statistics following the approach by Amihud and Hurvich (2004). The results reveal that the dividend yield does not have predictive power for our German sample. This result is in sharp contrast to U.S. most of the evidence, but it might be explained by the fact that our sample contains the 1990's, a decade during which several studies claimed that the dividend yield lost its predictive power (e.g., Goyal and Welch, 2003; Rey, 2004). In contrast to the dividend yield, both the short-term interest rate and the term spread exhibit strong predictive power. A conventional *t*-test indicates statistical significance of the coefficients on the lagged predictor variables at the 5% level. Given the small sample bias with highly persistent regressors, these results might overstate the true predictive power of our information variables. We therefore report the bias-adjusted *t*-statistic based on the method proposed by Amihud and Hurvich (2004), and they indicate that the prediction power of the short rate and the term spread is in fact robust. The coefficients are generally estimated at the 5% significance level; only in two cases significance reduces to the 10% level.

[Insert table 3 here]

Because we also test specifications where all three information variables are used simultaneously to model expected returns in our conditional performance measurement tests, we show

---

20 All sample autocorrelations at the first lag are above 0.9, indicating strong persistence and possibly non-stationarity (unit root). The null hypothesis of a unit root cannot be rejected for all three series using both Augmented Dickey-Fuller and Phillips-Perron tests. However, the results of the KPSS-test (Kwiatkowski et al., 1992) are inconclusive and, hence, we conclude that all three series are not sufficiently informative to be sure whether they are stationary or integrated.

the results of multi-predictor regressions in panel B of table 3. We apply Hansen's (1982) Generalized Method of Moments (GMM) to estimate the multi-predictor regressions, using a heteroskedasticity and autocorrelation consistent covariance matrix as weighting matrix and a constant plus all three lagged information variables as instruments to specify the orthogonality conditions. Our estimation results reveal that the single-predictor results in panel A are robust in the multivariate case. Specifically, while the coefficient on the log dividend yield is again estimated insignificantly, we still find evidence for predictability with both the short rate and the term spread. A Wald test clearly rejects the null hypothesis that all three information variables are jointly equal to zero. Overall, therefore, we interpret these regression results as evidence that our predictor variables are appropriate to model the time variation in expected returns in conditional performance measurement tests.

4. Empirical results

4.1. Conditional performance measurement in beta-pricing framework

4.1.1. Measuring conditional alphas and betas

In what follows we use the reduced sample of 50 German mutual equity funds, consisting of 47 general funds and 3 small- and mid-cap funds, for which we have a full return history over the period from January 1994 to December 2003 (120 months). We start with the results for an unconditional model. First, we employ the Datastream Germany Total Return Index as the market proxy in the single-index model. We choose this benchmark because the time series for the dividend yield we use in the conditional tests is based on this index. For the small- and medium-cap funds in the single-index model we use the DAFOX as the market proxy. Second, we generally employ the DAXOF as the market proxy in the Fama and French (1993) three-factor models. Given the construction principles of the DAFOX as the most comprehensive index, we think that this approach is the most appropriate. In the analysis that follows, we are less interested in the absolute level of underperformance (which is lowest based on the DAFOX), but instead focus more on the resulting shift in the distribution of alphas when conditioning information is incorporated into the model.

The results of the unconditional performance evaluation are shown in table 4. Most important, the mean abnormal return in the single-index model for the subsample of general funds is -0.045% per month, or about 55 basis points per year. The negative alpha in the three-factor model is even -0.217% per month, or -2.6% per year. In contrast, small- and mid-cap funds slightly outperform the multi-factor benchmark. Both the mean and median abnormal returns are positive. However, the absolute values are quite low, and given that there are only three small- and mid-cap funds with a full 120 months return history, these results should again not be overemphasized.

[Insert table 4 here]

Tables 5 and 6 present the results from conditional performance evaluation models in the standard beta-pricing setup, as specified in equation (4). We start by testing conditional versions of the single-index model, and use each information variable separately one at a time in table 5 and all three information variables simultaneously in table 6. The information variables are (i) the lagged level of the log dividend yield, (ii) the lagged level of the 1-month interest rate, and (iii) the lagged slope of the term structure. Again, we employ the Datastream Germany Total Return Index as the market proxy for the subsample of general funds and the DAFOX for the small- and mid-cap funds. In table 5 in the column denoted "#sign", we also report the number of fund for which the coefficient on the interaction term of the benchmark return with the (lagged) single instrument variable is estimated significantly. In table 6 in the column labeled "pval(F)", we report the average probability values of Wald *F*-tests for the null hypothesis that the coefficients on all three interaction terms are jointly equal to zero. The figures in brackets denote the number of funds for which the *F*-test rejects the null hypothesis.

[Insert table 5 and 6 here]

In general, we observe that the distribution of conditional alphas shifts to the left, i.e., the performance of funds becomes worse when we control for public information. For example, the mean conditional alpha in table 5 for the subsample of general funds varies between -0.052% and -0.111% per month, depending on the information variable, and the average estimated alpha in table 6 is -0.130% per month, or about 1.5% per year. The null hypothesis that all interaction terms corresponding to the three instrument variables in table 6 are jointly equal to zero can be rejected for 36 funds, i.e., almost three quarters of our sample.

Our analysis is based on mutual fund returns net of costs, i.e., management fees were already deducted from the funds' returns. Taking the 1.5% annual underperformance as the benchmark and adding back the average total expense ratio over the 1997 to 2003 period of 1.06%, we find that funds underperform by roughly 45 basis points before fees. This clearly indicates that, on average, active fund management cannot add value to the investor. However, our results are less pronounced than those in Bams and Otten (2002), who document that the average fund in their smaller sample of German funds underperforms by even -1.32% per year on a conditional basis.

Overall, compared to an unconditional assessment of mutual fund performance on the basis of Jensen's alpha, our results seem to suggest that the performance of our sample of general funds appears even more unfavorable in a conditional framework. These findings are consistent with those in Dahlquist et al. (2000) for Swedish funds, in contrast with the original evidence in Ferson and Schadt (1996) for U.S. mutual funds and Silva et al. (2003) for European bond funds. The latter two studies reveal a tendency for better performance when predetermined information variables are incorporated into the analysis. In two follow-up studies with much larger samples, Christopherson et al. (1998) and Ferson and Qian (2004) no longer report any significant effect on the distribution of alphas, i.e., conditional alphas do not make the performance of the funds look better than unconditional alphas. Chen and Knez (1996) argue that only managers who use more than public information have the potential to be assigned a positive performance by a conditional performance measure. Conditional analysis raises the hurdle on managers seeking abnormal positive performance because it gives them no credit for using readily available information, and this makes it more likely for funds to show no abnormal performance.

To double-check our results, in a first step we run all regressions using all four alternative benchmark indexes as market proxies. Table 7 shows the rank correlation of the conditional versions of Jensen's alpha for the specification that uses all four instrument variables. Our results are robust to the choice of the benchmark indexes, and with one exception the rank correlations are all above 0.99. Again, we observe that the DAFOX index produces lower alphas than the other indexes, but the results are qualitatively the same.

[Insert table 7 here]

Our findings have interesting interpretations. In the sample of general funds, we find that the term $B_p' Cov[z_t, r_{B,t+1}]$ is a positive 0.000859 for the average fund. This finding implies that the correlation of fund betas with the expected market return that is attributable to the predetermined information variables tends to be positive, indicating that fund managers tend to increase their market betas when expected returns are high conditional on public information, and/or reduce their market betas when expected returns are low.[21] Traditional performance measurement studies in an unconditional setting and timing models, in particular, interpret the covariance between beta and future market returns solely as a result of a portfolio manager's superior (private) performance. In contrast, a conditional setup assumes that markets are semi-strong efficient, i.e., fund managers do *not* have superior information but rather exploit public information. To come up with a ”fair” performance measure, that part of the ex post abnormal return that is attributable to time variation in expected returns based on public information and changing betas in the "correct" direction must be excluded. As suggested by Chen and Knez (1996), using public information-based dynamic portfolios as performance references leads to "tougher" performance yardsticks.

An alternative (and somewhat agnostic) explanation for Ferson and Schadt's (1996) result is provided by Chen and Knez (1996). They argue that due to the nonuniqueness of admissible measures, switching from unconditional to conditional performance measures does not necessarily mean lowering the performance ranking of every fund. Due to the infinity of admissible conditional and unconditional measures, it is not the case that any conditional measure will automatically assign a performance value lower than what an unconditional measure does.[22]

4.1.2. Cross-sectional distribution of mutual fund alphas

To explore the shift in alphas when switching from an unconditional to a conditional framework more in detail, we now look at the entire distribution of abnormal returns instead of average returns only. Specifically, table 8 shows the cross-sectional distribution of *t*-ratios for individual fund alphas. Following Ferson and Quian (2004), we summarize the results by presenting the fractions of the individual *t*-ratios that lie between standard critical values for a normal distribution, which is the asymptotic distribution for the *t*-rations. Column (1) shows the fractions that would be expected under the null hypothesis of no abnormal performance, if the normal distribution provides a good approximation for the *t*-ratios. The table also shows

---

21 In contrast, for our sample of small and mid-cap funds, we find opposite results. They indicate that fund managers tend to reduce their market betas when public information implies relatively high expected market returns, and vice versa. Consequently, fund performance tends to look worse in a conditional framework. This finding is in line with those in the original Ferson and Schadt (1996) study. Ferson and Warther (1996) also report that the performance of some funds switches from negative to positive. To explain this finding, they suggest a new money-flow hypothesis based on the findings in Warther (1995).

22 See Chen and Knez (1996), p. 531, for more details.

the minimum and the maximum *t*-statistics for each model. Using these values, we test the hypothesis that all alphas are jointly equal to zero using the Bonferroni *p*-value. This is a one-tailed test of the null hypothesis that all alphas are zero against the alternative that at least one alpha is positive (Bonferroni *p*-value (+)) or negative (Bonferroni *p*-value (−)). It is computed as the smallest of the *p*-values for the individual tests (the most positive *t*-statistic and the most negative *t*-statistic) multiplied by the number of funds.

Similar to Ferson and Schadt (1996) and Ferson and Qian (2004), we find in column (2) of table 8 that the distribution of the unconditional alphas is centered slightly to the left of the distribution under the null hypothesis. There are 28 funds with insignificantly negative alphas and *t*-values between 0 and -1.645, whereas 3 funds exhibit negative alphas that are even statistically significant (with *t*-values below -1.645). There is 1 fund with a significantly positive alpha (with a *t*-value above 2.326). Nevertheless, none of the extreme *t*-ratios are significant based on the Bonferroni test, although unconditional performance measures suggest a negative performance, on average. Consistent with the regression results for the Fama and French (1993) three-factor model in table 4, column (7) of table 16 indicates that the cross-sectional distribution of *t*-ratios shifts even further to the left, and there are more extreme negative alphas for individual funds. For example, there are six funds with *t*-values below -2.326, and the Bonferroni *p*-value rejects the null hypothesis that all alphas are zero against the alternative that at least one is negative at the 10% level.[23]

[Insert table 8 here]

Overall, for the unconditional models we observe that funds have more negative than positive alphas. Of the significant alphas (with absolute *t*-statistics larger than 1.645), all are negative, with only one exception for the single-index model. Accordingly, a simple binomial test rejects the null hypothesis that 50% of the alphas are positive. The corresponding *t*-statistics are -1.697 for the one-factor model and -4.808 for the three-factor model. Although this result clearly indicates poor performance from an investor's perspective, a caveat is that it is diffi-

---

23 The nonnormality, per se, in the distribution of the t-ratios could have two reasons. First, our sample size could be too small for the asymptotic distribution to be accurate. Second, Koswoski et al. (2005) argue that the cross-section of mutual fund alphas has a complex, non-normal distribution which can result from (i) heterogeneous risk-taking and (2) non-normally distributed individual fund alphas. Using a bootstrap approach, they document that the superior alphas of star mutual fund managers, net of costs, are not attributable to luck and indicate genuine stock-picking skills.

cult to know where the distribution of the alphas should be centered under the hypothesis of no abnormal performance. Sharpe (1991) and Malkiel (2003) argue that transaction costs are deducted from the funds' returns but not from the benchmark returns and, hence, the alphas of fund returns should be centered to the left of zero. On the other hand, our reduced sample of funds selected for the conditional analysis suffers from survivorship bias, which shifts the distribution of alphas to the right.

Looking at the distribution of *t*-ratios for the conditional performance model when all three information variables are used in column (6), the distribution becomes even more skewed to the left. Most important, 40 funds exhibit negative alphas, and there are 5 funds with *t*-ratios below -2.0. We cannot reject the null hypothesis that all fund alphas are jointly equal to zero against the alternative of superior performance, but the minimum *t*-ratio is close to being significantly negative when considering a 10% confidence level (with *p*-value = 0.1104). A binomial test again rejects the null hypothesis that 50% of the alphas are positive; the corresponding *t*-statistic is -4.243.

Although somewhat less pronounced, the results for the conditional regression specifications using only one information variable at a time in columns (3)-(5) are qualitatively similar (especially for the term spread and the short-term interest rate). This finding is again in contrast to the results in Ferson and Schadt (1996), who cannot reject the null hypothesis that 50% of the funds exhibit positive alphas in their conditional setup. They conclude that conditioning on public information removes the inference of the traditional approach that mutual fund alphas tend to be negative. As discussed above, our results for German mutual equity funds seem to suggest exactly the opposite conclusion. On average, the performance of the funds in our sample looks worse when we correct for the part of the ex post abnormal fund return that is attributable to time variation in expected returns based on publicly available information.

4.2. Performance measurement in a stochastic discount factor framework

To estimate unconditional and conditional SDF alphas, we again use the reduced sample of 50 German mutual equity funds for which we have a full return history over the period from January 1994 to December 2003 (120 months). In all our tests we employ the 10 German sector portfolios according to the Datastream classification and the risk-free security as the set of primitive assets. Consequently, in linear models of the stochastic discount factor we also use the Datastream Germany Total Return Index as the market proxy. This choice is further justified because the dividend yield that we employ in the conditional models is also based on this index. In addition, we also use the term spread and the short-term interest rates as information variables.[24]

4.2.1. Estimating stochastic discount factor models

We start with specification tests for the stochastic discount factor models. To provide a benchmark, the first line in table 9 shows the results from a constant discount factor model, in which the stochastic discount factor is assumed to be fixed over time and equal to the inverse of the sample mean of the gross 1-month interest rate.[25] A constant stochastic discount factor 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. The average monthly interest rate is 0.309% and, hence, the inverse of the 1-month gross return is 0.997 [= 1/(1+0.00309)]. As shown in the first line of table 9, the estimated value of the constant stochastic discount factor is 0.996, which is very close to the inverse of the sample mean of the gross 1-month interest rate. Furthermore, panels A and B of table 9 show the results of the different stochastic discount factor models. The means of the fitted values of all stochastic discount factor models are slightly below 1.00 as well, indicating that these models are effective in fixing at least the mean of the "true" stochastic discount factor. Figures 1 and 2 depict graphical illustrations of the fitted values from each estimated model. Similar to Farnsworth et al. (2002), a first observation is that the standard deviation increases as the stochastic discount factor model becomes more "complicated", i.e., as we move from a one-factor model to multi-factor models and, even more pronounced, from unconditional models to conditional models. This result can be explained by referring to Hansen and Jagannathan's (1991) volatility bounds for admissible stochastic discount factors and the underlying projection argument. They document that the minimum variance of an admissible stochastic discount factor in-

---

24 Using all information variables simultaneously can lead to an explosive number of orthogonality conditions in a stochastic discount factor framework. Therefore, the results we present for the conditional specifications of both the SDF-Bakshi-Chen model and the SDF-Primitive-efficient model are based on the short-term interest as the only information variable in the Generalized Method of Moments estimation. In contrast, for the SDF-CAPM and the SDF-Fama-French models we use the full set of information variables to scale factors.

25 Specifically, we estimate the system of equations in (7) using all primitive assets, but excluding the fund specific equation.

creases as the number of included assets increases.[26] Conditional stochastic discount factor models can be tested by adding scaled factors and simply estimating the unconditional moments of this extended model (e.g., Cochrane, 2001). New “assets” (interpreted as dynamic trading strategies) are added to the model by scaling the factors and/or the primitive assets by the lagged information variables. Therefore, it is perfectly reasonable to expect that the standard deviation of the estimated stochastic discount factors is higher in conditional models compared to unconditional models.

[Insert table 9 and figures 1 and 2 here]

A second (reassuring) observation is that the number of negative values in the time series of estimated stochastic discount factors is small in most specifications, as indicated in the last column of table 9. The Euler-equation suggests that a negative value of the stochastic discount factor implies it assigns positive prices to negative payoffs at some point in time. Obviously, as the standard deviation of the estimated stochastic discount factor increases, the number of instances where the estimated stochastic discount factor takes on a negative value increases as well. The value of a highly volatile stochastic discount factor is ambiguous. On the one hand, more volatility is necessary to explain the historical equity premium (e.g., Mehra and Prescott, 1985), or equivalently, to suffice the Hansen and Jagannathan (1991) volatility bound. On the other hand, a more volatile discount factor does not necessarily exhibit better pricing properties, and more volatility may also imply lower power to detect abnormal performance.

Following Farnsworth et al. (2002), we also explore the dynamic performance of the different stochastic discount factor models. In frictionless markets some transformation of the equilibrium price process should follow a martingale with respect to the information that market participants use to form expectations. Specifically, the discounted gross return, $m_{t+1}R_{t+1}$, should be a martingale and, hence, it should not be predictable based on the lagged information variables, collected in Z_t (e.g., Zimmermann, 1998; Cochrane, 2001). To test this prediction, we compute the pricing errors, $m_{t+1}R_{t+1} - 1 \equiv u1$, for each equation in our system and regress them

26 See De Santis (1995) for an in-dept discussion. Based on this intuition, he develops spanning tests in a stochastic discount factor framework. This methodology was used by Beakert and Hodrick (1992), Bekaert and Urias (1996), Errunza et al. (1999), and Drobetz (2003) to test for the benefits of international diversification.

upon the vector of lagged information variables, Z_t.[27] The predicted pricing errors should not be significantly different from zero using any information available at time t and, hence, the model implies that the regression coefficients should be zero. Even if there is predictability in a return, R_{t+1}, using lagged information variables, Z_t, this predictability should be removed when R_{t+1} is multiplied by the "correct" m_{t+1}. This is the sense in which conditional asset pricing models are required to "explain" predictable variation in asset returns. A related requirement is that the standard deviation of the fitted pricing errors of each equation should be small if the model captures the predictable variation in expected stock returns. This interpretation is similar to the notion of the variance ratio tests in Ferson and Harvey (1993) and Ferson and Korajczyk (1995) and the restrictions on the covariance between information variables and subsequent asset returns that must hold in efficient markets, as derived by Kirby (1998).

Table 10 shows the standard deviation of the fitted pricing errors in the different models for each primitive asset. The constant discount factor model can serve as a benchmark model, because it cannot explain any of the predictability. As an example, take the returns of sector 1. The monthly standard deviation of the fitted values for this sector is 2.91%, as shown in the first line of table 10. With a constant stochastic discount factor, this approach is equivalent to a regression of the gross returns of sector 1 multiplied by the estimated constant m minus 1 upon the lagged information variables. In comparison, the standard deviation of the monthly raw return on sector 1 is 14.25%, as shown in table 2. An R-square measures the explanatory power of a regression and indicates the percentage of variance explained. Accordingly, the R-square in an ordinary least squares regression of the returns on sector 1 upon the lagged information variables is roughly 4.2% [= $(0.02908/0.1425)^2$]. In fact, this value is exactly equal to the (unadjusted) R-square shown at the bottom of table 2 for the predictive regression involving sector 1 and the three information variables.

[Insert table 10 here]

Panel A in table 10 presents the results for the unconditional stochastic discount factor models. This class of models is unconditional in the sense that they do not exploit the predictive power of the information variables for raw returns, R_{t+1}. However, the stochastic discount factor, m_{t+1}, is time-varying and, hence, the product $m_{t+1}R_{t+1}$ could potentially follow a martin-

---

27 This procedure is similar to the widely used Sargan-test (Sargan, 1958) of overidentifying restrictions. See also Davidson and MacKinnonn (2004).

gale. Similar to Farnsworth et al. (2002), none of the models can explain the predictable variation in sector returns better than the constant stochastic discount factor model. The SDF-CAPM model explains a reasonable fraction of the predictability, implying that the standard deviations of the fitted values are relatively small. However, the other three models perform poorly and cannot capture the time variation in expected returns. In many cases, the standard deviations of the fitted pricing errors are even larger than the standard deviations of the sectors' raw returns. This result implies that the product of $m_{t+1}R_{t+1} - 1$ has even larger regression coefficients upon Z_t than R_{t+1} has, indicating particularly poor pricing performance.

Panel B of table 10 presents the results for the conditional stochastic discount factor models. The conditional SDF-primitive-efficient and the conditional SDF-Bakshi-Chen models produce standard deviations of the fitted pricing errors that are virtually zero. This result is similar those by Farnsworth et al. (2002). In general, we observe that the conditional models perform better than the unconditional ones. This finding can be explained by the nature of the GMM estimation procedure. Conditional models exploit the lagged instrument variables to form a set of orthogonality conditions, and the parameters are estimated to make the expected product of the pricing errors and the lagged instrument variables as close to zero as possible in the sample.[28]

4.2.2. Using the SDF alphas to measure mutual fund performance

Table 11 presents the cross-sectional distribution of individual fund alphas for the different stochastic discount factor models. To characterize this distribution of SDF alphas, we report the mean and the median alphas, the alphas of the bottom 3 and bottom 10 funds, and the alphas of the top 3 and top 10 funds. We also report the SDF alpha when the system of equations in (7) is estimated using an equally-weighted portfolio of all funds. Moreover, the table shows the Bonferroni p-values for the null hypothesis that all alphas are jointly zero.[29]

[Insert table 11 here]

28 A notable exception is the conditional SFD-Fama-French model, which performs even worse than the constant stochastic discount factor model in explaining the predictable variation in expected sector returns. This finding confirms previous results by Kirby (1998), who also reports that the three-factor model of Fama and French (1993) is unable to explain the returns on dynamic trading strategies (as proxied by scaled returns), which suggests that predictability is inconsistent with rational asset pricing.

29 If the adjusted p-value exceeds 1, it could also be set equal to 1.

In the SDF-CAPM model, the average estimated underperformance is 0.149% per month, or roughly 1.77% per year. The same result is obtained for an equally-weighted portfolio of all funds. The average expense ratio of our subsample of general funds is 1.29%, indicating that already under the simplest stochastic discount factor model the average fund strongly underperforms, even after adding back total expense ratios. The performance of the top 3 and bottom 3 funds indicates strong negative skewness of the distribution of individual fund alphas. Furthermore, this shape of the distribution is not strongly dependent on the stochastic discount factor model.

Table 12 shows the rank correlations between the estimated alphas. In many cases, the correlations are quite high, indicating that the relative performance of individual funds is quite robust to alternative specifications of the stochastic discount factor. Rank correlations range between 0.95 and 0.72 for the unconditional models in panel A, and between 0.91 and 0.64 for the conditional models in panel B. The lowest rank correlations generally involve the SDF-Fama-French model, which also tends to be the poorest performing model (see tables 9 and 10). In results not presented here, we also compute the rank correlations between unconditional and conditional versions of each model. These rank correlations range between 0.94 (for the SDF-Fama-French models) and 0.72 (for the SDF-CAPM models). Overall, we therefore conclude that the relative performance measured for the individual funds is reasonably correlated across different models of the stochastic discount factor.

[Insert table 12 here]

Nevertheless, compared to the unconditional SDF-CAPM model, the mean of the distribution generally shifts to the left in all other models of the stochastic discount factor. This finding is in contrast to the results reported by Farnsworth et al. (2002) and Fletcher and Forbes (2004), who report similar performance measures across all stochastic discount factor models. For example, in panel A of table 11 the average alpha for the unconditional SDF-primitive-efficient model is -0.177% per month, or about 2.10% per year. In contrast, the mean SDF alpha for the SDF-Fama-French three-factor model sharply deteriorates to -0.304% per month, or -3.59% per year. This large negative performance is confirmed when an equally-weighted portfolio of funds is used instead of estimating the system of equations separately for each fund and reporting the average SDF alpha. However, given the poor time series properties of the estimated stochastic discount factor for this model (see tables 9 and 10), this result should not be overemphasized.

As expected, the results in panel B of table 11 suggest that compared to an unconditional assessment the conditional performance of our funds appears even more unfavorably (with an exception of the SDF-Bakshi-Chen model). Compared to the unconditional specifications in panel A, the average conditional SDF alphas sharply decrease to values between -0.234% and -0.340% per month. This implies large underperformance between -2.77% and -4.00% per annum, depending on the model for the stochastic discount factor. The null hypothesis of no abnormal performance explicitly allows for a covariance between factor weights and future returns on the primitive assets because they depend on public information that is observable for the econometrician. A ”fair” performance measure excludes that part of the ex post abnormal return that is attributable to time variation in expected returns based on public information and changing factor weights in the “correct” direction. Using public information-based dynamic portfolios as performance references leads to “tougher” performance yardsticks.

A second, more technical, explanation for our findings is that a more “complicated” stochastic discount factor leads to more restrictive pricing conditions. In a conditional setup, the stochastic discount factor becomes more complicated in the sense that scaled returns (interpreted as dynamic trading strategies) are added. As shown by Hansen and Jagannathan (1991), imposing additional assets implies that the volatility bound for stochastic discount factors shifts upward and entails stronger pricing restrictions, i.e., the minimum variance of admissible stochastic discount factors increases. The ray from the origin to the minimum point of the volatility bound has the interpretation of a Sharpe ratio (e.g., De Santis, 1995; Drobetz, 2003) and, hence, conditional models specify a tougher benchmark, explaining why the performance of our funds looks particularly poor when the tests incorporate conditioning information.

As already mentioned, the distribution of our estimated SDF alphas is negatively skewed, and this property is also reflected in the Bonferroni *p*-values. With only one exception (the conditional SDF-Fama-French model), we cannot reject the null hypothesis that all SDF alphas are jointly equal to zero against the alternative hypothesis that there is at least one significantly positive alpha (Bonferroni *p*-value (+)). In contrast, in most instances we can reject the null hypothesis that all SDF alphas are jointly equal to zero against the alternative hypothesis that there is at least one significantly negative alpha (Bonferroni *p*-value (–)).

4.2.3. Comparing the results from beta-pricing models and SDF models

In a final step, it seems worthwhile to compare the estimated SDF alphas with the traditional Jensen’s alphas. A general observation is that the performance of our German mutual equity funds looks much worse in the Euler-equation framework than in the traditional beta-pricing setup. For example, the average unconditional Jensen's alpha in table 4 is -0.042% per month, or roughly -55 basis points per annum, whereas the average unconditional alpha in the SDF-CAPM specification in table 11 is -0.140% per month, or roughly -1.77% per year. The corresponding conditional performance measures are -0.116% (table 6) and -0.340% (table 11) per month, or about 1.38% and -4.00% per year, respectively. Roughly speaking, in this case the negative fund performance differs by a factor of three, but other specifications of the stochastic discount factor lead to somewhat less pronounced performance differences.

A crucial question that arises is whether elegance and generality of the stochastic discount factor framework comes at the cost of estimation efficiency for risk premiums and testing power for model specifications. Kan and Zhou (1999) compare the stochastic discount factor methodology using the Generalized Method of Moments with maximum likelihood estimates of the static linear capital asset pricing model. Their results seem to suggest that the stochastic discount factor methodology performs much worse than the beta-pricing approach in specification tests. They explain this result by noting that the Euler-equation is merely a restriction on part of the first and second moments between the asset returns and the factors. However, without a fully specified model of asset returns, this implies ignoring many other first and second moments, thereby producing large estimation errors of the factor risk premiums. Intuitively, if the stochastic discount factor methodology was not very reliable in detecting even gross misspecifications of asset pricing models, this will clearly also have detrimental implications for performance evaluation tests.

However, this conclusion might be premature for two reasons. First, the specification tests in tables 9 and 10 indicate that, except for the SDF-Fama-French model, our stochastic discount factor models behave reasonably well. Second, Jagannathan and Wang (2002) document that Kan and Zhou's (1999) results are based on false assumptions. Kan and Zhou (1999) ignore the fact that the risk premium parameters in the two methods are not identical (albeit strictly related), and directly compare the asymptotic variances of the two estimators. Jagannathan and Wang (2002) present a more appropriate test setup that explicitly incorporates the transformation between the risk parameters in the two methods and conclude that the stochastic discount factor methodology is asymptotically as efficient as the beta method. Specification tests of asset pricing models based on the two methods are also equally powerful.

To gain some preliminary insight, we estimate the direct counterpart of a static Jensen's alpha in the stochastic discount factor framework and label this performance measure "SDF Jensen's alpha". Specifically, instead of using the full set of primitive assets in the system of equations in (7), we estimate the SDF-CAPM model by merely requiring that the stochastic discount factor prices the benchmark index and the risk-free security.[30] As in all estimations above, we employ the Datastream Germany Total Return Index as a proxy for the market portfolio. Following Dybvig and Ingersoll (1982), the stochastic discount factor is assumed to be a linear function of the market index, which is again represented by the Datastream index. In this simple setup the system to be estimated consists of only three equations. Since the stochastic discount factor is a linear combination of the Datastream index, the first equation formulates the pricing restriction on the index itself as a primitive asset, and the second equation contains the pricing restriction on the risk-free security. The third equation involves the pricing restriction on the specific fund to be evaluated and the corresponding SDF Jensen's alpha.

A graphical comparison of the results is depicted in panel A of figure 3. The figure displays alternative performance measures for all 50 funds in our sample: (i) the traditional Jensen's alpha, (ii) the SDF Jensen's alpha, and (iii) the SDF alpha from the SDF-CAPM model. Clearly, the means of all series match the values reported in various tables above. Specifically, the mean Jensen's alpha in our sample is -0.042% per month (see table 4), and the mean SDF-CAPM alpha is -0.149% per month (see table 11). Most important, the mean SDF Jensen's alpha is -0.042% per month, which is identical to its direct and more common counterpart in the beta-pricing formulation. In fact, as can be inferred from figure 3, the corresponding two lines with individual alphas exactly coincide. From this finding we conclude that the empirical results from the stochastic discount factor methodology and the beta-pricing method are in fact inherently related with each other, as suggested by asset pricing theory. In addition, the SDF-CAPM alphas are also highly correlated with the Jensen's alphas; the rank correlation is 0.91. Nevertheless, there is a notable level shift, which is responsible for our previous finding that funds tend to look worse, on average, in a stochastic discount factor framework.

In panel B of figure 3 we present the same results in a conditional setup, where all three information variables are used to account for time variation in expected returns. The average conditional SDF Jensen's alpha is -0.063% per month, or -76 basis points per annum, which

30 To be consistent with the beta-pricing framework we omit the information variables as individual "factors" to compute the SDF Jensen's alphas.

makes funds look better in this case compared to the Ferson and Schadt (1996) framework, for which we report a mean conditional alpha of -0.116% per month, or -1.380% per year (see table 6). Note, however, that the average conditional SDF Jensen's alpha is still slightly lower than its average unconditional analogue, which supports our general notion that funds look worse under conditional performance measures. Similar to the unconditional setup in panel A, the figure in panel B reveals that the conditional alpha of the SDF-CAPM model assigns the lowest performance to our funds, with a mean alpha of -0.340% per month (see table 11). As can be expected from a mere visual inspection, the rank correlations between the performance measures are again reasonably high; they range between 0.66 (between Jensen's alpha and SDF Jensen's alpha) and 0.81 (between Jensen's alpha and the SDF-CAPM alpha).

[Insert figure 3 here]

Overall, we find that the stochastic discount factor methodology and the beta-pricing method deliver closely related performance measures in our empirical tests, as suggested by asset pricing theory. We cannot conclude that either one approach leads to grossly misleading performance evaluation results. Nevertheless, our analysis seems to suggest that they depend on the model specifications and the estimation techniques. One immediate source of the differences in our results is the choice of the primitive assets. With a larger set of primitive assets, the number of equations in the system to be estimated increases, thereby imposing additional moment restrictions that affect the parameter estimates (Bekaert and Liu, 2004). This is mechanically true, although our 10 Datastream sector portfolios are merely subindexes of the Datastream aggregate market index. Intuitively, additional restrictions will always imply a tougher benchmark, leading to lower estimated fund performance.

Another noteworthy issue is that stochastic discount factors may be biased. Farnsworth et al. (2002) test the model using simulated trading strategies with different levels of ability and find that that the bias is about -0.19% per month for unconditional models and -0.12% per month for conditional models. While this seems economically large, they nevertheless report that this is less than two standard errors, as the typical standard error of their alpha measures is 0.10% per month. In contrast, studying primitive-efficient models in a slightly different setup using Swedish fund data, Dahlquist and Söderlind (1999) find no significant biases in the pricing errors but do find size distortions. Tests for the hypothesis of zero abnormal performance reject the null hypothesis too often.

4. Conclusion

Evaluating the performance of portfolio managers has received wide attention in the financial economics literature, presumably because a substantial part of the savings of investors is managed by professionals. The general idea behind performance evaluation is straightforward. In principle, the observer must assign the correct value to the cash flow (net of management fees) the manager generates from the amount entrusted to him by an investor. The difference between the assigned value and the amount entrusted to the manager is the value of the services provided by the manager. If this difference is positive, then the manager provides "valuable" service.

However, there are several theoretical difficulties in implementing this simple principle. Most important, financial economists still do not have a satisfactory valuation model that consistently prices arbitrary streams of cash flows sufficiently close to their market prices. Every asset pricing model that has been suggested in the literature has performed poorly at least with respect to one subset of the assets examined. Therefore, the econometrician who uses a particular valuation model has to be aware of the collection of assets for which the model performs satisfactorily to avoid false inferences on the performance of funds resulting from a joint-hypothesis problem.

In spite of these fundamental issues, the recent performance measurement literature has made substantial progress. We have addressed two important issues that have shaped the recent literature. First, we show how to extend the classical unconditional securities market line analysis to incorporate conditioning information and to compute conditional Jensen's alphas. This strand of the literature has been inspired by the influential studies by Ferson and Schadt (1996) and Christopherson et al. (1998). Second, given the deficiencies of securities market line analysis and following a general trend in the recent asset pricing literature, the focus switched from the classical beta-pricing framework to the more general stochastic discount factor framework. This development is based on the seminal work by Chen and Knez (1996), and more recently by Farnsworth et al. (2002).

We investigate the performance of a small, but survivorship bias controlled, sample of German mutual equity funds. Our general finding is that mutual funds, on average, hardly produce returns that are large enough to cover their expenses. This conclusion is drawn from a variety of model specifications that we have tested, and it is robust to many different benchmarks we have employed. Specifically, we start by measuring fund performance in the classical unconditional beta-pricing framework. Based on Jensen's alpha as a performance measure, mutual funds slightly underperform, but when expenses are added back their performance tends to be neutral. However, funds look much worse when we apply the three-factor model of Fama and French (1993) as the benchmark. Most important, fund performance substantially deteriorates when we measure conditional alphas both in single-index and multi-factor models. For example, based on the full set of information variables, the mean conditional alpha for our sample of general funds is estimated to be -0.130% per month, or about 1.5% per year. In comparison, taking a simple average for all funds over the 1997 to 2003 period, the mean total expense ratio is only 1.06%. Accordingly, if we add back management fees, German mutual equity funds underperform already on a before fee basis. Using the three-factor model of Fama and French (1993), fund performance tends to look even worse. This result should not come as a surprise: Given that stock returns are partly predictable using publicly available information, that part of fund performance that is attributable to time-variation in expected returns should be deducted from fund performance. We conclude that conditional analysis raises the hurdle on active managers seeking abnormal positive performance because it gives them no credit for exploiting readily available information, and this makes it more likely for funds to show no abnormal performance.

We then proceed to performance measurement in the Euler-equation framework and test several (unconditional and conditional) models of the stochastic discount factor. The result that funds underperform before costs is now even more pronounced, even in some of the unconditional stochastic discount factor models. Underperformance before fees can even be as low as 4% per year. To the best of our knowledge, we are the first to directly compare the results from performance measurement tests in the beta-pricing framework and the stochastic discount factor framework. We document that both methods are in fact inherently related, as suggested by asset pricing theory, and that the relative performance of our funds (i.e., their rankings within the sample) is highly correlated across models. However, there is a pronounced level shift when switching from the beta-pricing framework to the stochastic discount factor framework, which can be attributed to the fact that more "complicated" stochastic discount factor models and a larger number of primitive assets (which are required to estimate the parameters of the stochastic discount factor models) lead to stronger pricing conditions. This interpretation follows directly from Hansen and Jagannathan's (1991) seminar analysis on volatility bounds for admissible stochastic discount factors, an aspect which has been largely neglected in the previous literature.

Bibliography

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

Amihud, Y., and C. Hurvich, 2004, Predictive regressions: A reduced-bias estimation method, *Journal of Financial and Quantitative Analysis* 39, 813-841.

Bakshi, G., and Z. Chen, 1998, Asset pricing without consumption or market portfolio data, Working Paper, College Park, University of Maryland.

Bams, D., and R. Otten, 2002, European mutual fund performance, *European Financial Management* 8, 75-101.

Bansal, R., and C. Harvey, 1996, Performance evaluation in the presence of dynamic trading strategies, Working paper, Duke University.

Banz, R., 1981, The relationship between return and market value of common stock, *Journal of Financial Economics* 9, 3-18.

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

Bekaert, G., and J. Liu, 2004, Conditioning information and variance bounds on pricing kernels, *Review of Financial Studies* 17, 339-378.

Bekaert, G., and M. Urias, 1996, Diversification, integration, and emerging market closed end funds, *Journal of Finance* 51, 835-869.

Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts, 2004, On the importance of measuring payout yield: Implications for empirical asset pricing, NBER Working paper.

Breen, W., L. Glosten, and R. Jagannathan, 1989, Economic significance of predictable variations in stock returns, *Journal of Finance* 44, 1177-1190.

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

BVI (Bundesverband Investment und Asset Management), 2005, Aktuelle Fragen der Altersvorsorge mit Investmentfonds, available at: http://www.bvi.de.

Campbell, J., and S. Thompson, 2005, Predicting the equity premium out of sample: Can anything beat the historical average?, forthcoming in: *Journal of Financial Economics*.

Campbell, J., and M. Yogo, 2005, Efficient tests of stock return predictability, Working paper, Harvard University.

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

Chen, N., R. Roll, and S. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383-403.

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

Christopherson, J., W. Ferson, and D. Glassman, 1998, Conditioning manager alphas on economic information: Another look at the persistence of performance, *Review of Financial Studies* 11, 111-142.

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

Cochrane, J., 1999, New facts in finance, *Economic Perspectives*, 36-58.

Cochrane, J., 2001, Asset pricing, New Jersey: Princeton University Press.

Dahlquist, M., S. Engström, and P. Söderlind, 2000, Performance and characteristics of Swedish mutual fund, *Journal of Financial and Quantitative Analysis* 35, 409-423.

Dahlquist, M., and P. Söderlind, 1999, Evaluating portfolio performance with stochastic discount factors, *Journal of Business* 72, 347-384.

Davidson, R., and J. MacKinnon, 2004, Econometric theory and methods, New York: Oxford University Press.

De Santis, G., 1995, Volatility bounds for stochastic discount factors: Tests and implications from international financial markets, Working paper, University of Southern California.

De Santis, G., and B. Gerard, 1997, International asset pricing and portfolio diversification with time-varying risk, *Journal of Finance* 52, 1881-1912.

De Santis, G., and B. Gerard, 1998, How big is the premium for currency risk?, *Journal of Financial Economics* 49, 375-412.

Dittmar, R., 2003, Nonlinear pricing kernels, kurtosis preferences, and evidence from the cross-section of asset returns, *Journal of Finance* 62, 369-403.

Drobetz, W., 2003, Volatility bounds for stochastic discount factors on global stock markets, *Journal of Applied Social Science Study* (Schmollers Jahrbuch) 123, 1-30.

Dybvig, P., and J. Ingersoll, 1982, Mean-variance theory in complete markets, *Journal of Business* 55, 233-252.

Dybvig, P., and S. Ross, 1985a, The analytics of performance measurement using a security market line, *Journal of Finance* 40, 401-416.

Elton, E., M. Gruber, and C. Blake, 1996, Survivorship bias and mutual fund performance, *Review of Financial Studies* 9, 1097-1120.

Errunza, V., K. Hogan, M. Hung, 1999, Can the gains from international diversification be achieved without trading abroad?, *Journal of Finance* 54, 2075-2107.

Fama, E., 1970, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 383-417.

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

Fama, E., and K. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3-53.

Farnsworth, H., W. Ferson, D. Jackson, and S. Todd, 2002, Performance evaluation with stochastic discount factors, *Journal of Business* 75, 473-504.

Ferson, W., and C. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385-415.

Ferson, W., and C. Harvey, 1993, The risk und predictability of international equity returns, *Review of Financial Studies* 6, 527-566.

Ferson, W., and R. Jagannathan, 1996, Econometric evaluation of asset pricing models, in: Maddala, G., and C. Rao (eds.), Handbook of Statistics 14: Statistical Methods in Finance, Amsterdam: Elsevier, 1-30.

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

Ferson, W., and M. Qian, 2004, Conditional performance evaluation: Revisited, Working paper, Boston College.

Ferson, W., S. Sarkissian, and T. Simin, 2003a, Spurious regressions in financial economics?, *Journal of Finance* 58, 1393-1413.

Ferson, W., S. Sarkissian, and T. Simin, 2003b, Is stock return predictability spurious?, *Journal of Investment Management* 1, 1-10.

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

Ferson, W., and A.F. Siegel, 2001, The efficient use of conditioning information in portfolios, *Journal of Finance* 56, 967-982.

Ferson, W., and V. Warther, 1996, Evaluating fund performance in a dynamic mark*et, Financial Analysts Journal*, 20-28.

Fletcher, J., and D. Forbes, 2004, Performance evaluation of U.K. unit trusts within the stochastic discount factor framework, *Journal of Financial Research* 27, 289-306.

Goyal, A., and I. Welch, 2003, Predicting the equity premium with dividend ratios, *Management Science* 49, 639-654.

Griese, K., and A. Kempf, 2003, Lohnt aktives Fondsmanagement aus Anlegersicht? Ein Vergleich von Anlagestrategien in aktiv und passiv verwalteten Aktienfonds, *Zeitschrift für Betriebswirtschaft* 73, 201-224.

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

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

Hackethal, A., 2004, German banks and banking structure, in: Krahnen, J., and R. Schmidt (eds.), The German financial system, Oxford: Oxford University Press, 71-105.

Hansen, L., 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica* 50, 1029-1054.

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

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

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

Harvey, C., 2001, The specification if conditional expectations, *Journal of Empirical Finance* 8, 573-638.

Jagannathan, R., and Z. Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-53.

Jagannathan, R., and Z. Wang, 2002, Empirical evaluation of asset pricing models, A comparison of SDF and beta methods, *Journal of Finance* 57, 2337-2367.

Jensen, M., 1972, Optimal utilization of market forecasts and the evaluation of investment performance, in: Szego, G., and K. Shell (eds.), Mathematical methods in investment and finance Elsevier: North-Holland.

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

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

Kosowski, R., A. Timmermann, R. Wermers, and H. White, 2005, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, Working paper, INSEAD.

Krahnen, J.P., F. Schmid, and E. Theissen, 1997, Performance and market share: Evidence from the German mutual fund industry, Working paper, Universität Frankfurt.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics* 54, 159-178.

Lakonishok, J., A. Shleifer, and R. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

Lewellen, J., 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74, 209-235.

Malkiel, B., 1995, Return from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 40, 549-572.

Malkiel, B., 2003, Passive investment strategies and efficient markets, *European Financial management* 9, 1-10.

Maurer, R., 2004, Institutional investors in Germany, in: Krahnen, J., und R. Schmidt (eds.), The German Financial System, Oxford: Oxford University Press, 106-138.

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

Merton, R., 1981, On market timing and investment performance I: An equilibrium theory of value for market forecasts, *Journal of Business* 54, 363-406.

Ostdiek, B., 1998, The world ex ante risk premium: An empirical investigation, *Journal of International Money and Finance* 17, 967-999.

Paape, C., 2003, Currency overlay and performance evaluation, *Financial Analysts Journal* 59, 55-68.

Rey, D., 2004, Stock market predictability and tactical asset allocation, Ph.D. thesis, University of St. Gallen.

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

Roll, R., and S. Ross, 1994, Cross-sectional relation between expected returns and betas, *Journal of Finance* 49, 101-121.

Sargan, J., 1958, The estimation of economic relationships using instrumental variables, *Econometrica* 26, 393-415.

Schwert, G.W., 2002, Anomalies and market efficiency, in: Constantinides, G., M. Harris, and R. Stulz (eds.), Handbook of the economics of finance, Elsevier: North Holland.

Shanken, J., 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Financial Economics* 45, 99-120.

Sharpe, W., 1991, The arithmetic of active management, *Financial Analysts Journal*, 7-9.

Silva, F., M. Ceu Cortez, and M. Armanda, 2003, Conditioning information and European bond fund performance, *European Financial Management* 9, 201-230.

Söderlind, P., 1999, An interpretation of SDF based performance measures, *European Finance Review* 3, 233-237.

Stambaugh, R., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375-421.

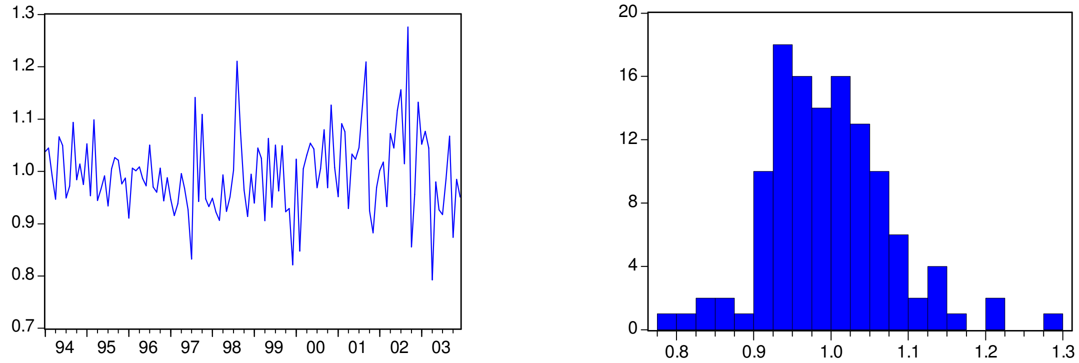
Theissen, E., 2004, Organized equity markets, in: Krahnen, J., und R. Schmidt (eds.), The German financial system, Oxford: Oxford University Press, 139-162.

Torous, W., R. Valkanov, and S. Yan, 2004, On predicting stock returns with nearly integrated explanatory variables, *Journal of Business* 77, 937-966.

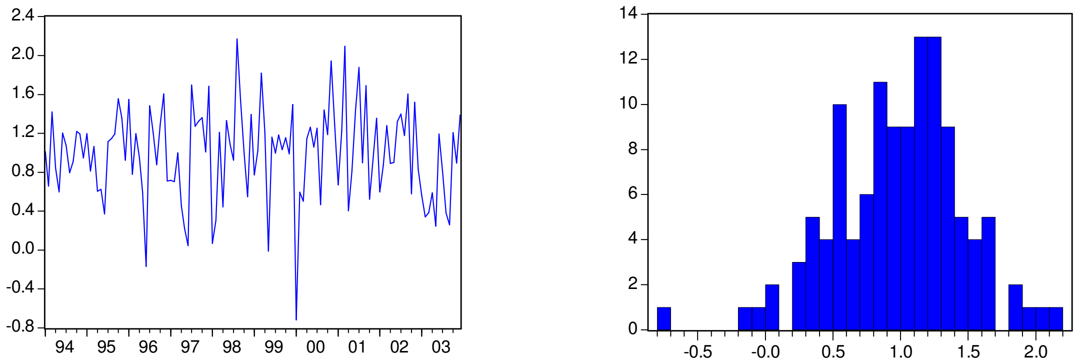
Zimmermann, H., 1998, State Preference Theorie und Asset Pricing: Eine Einführung, Heidelberg: Physica-Springer.

Figure 1: Unconditional stochastic discount factor (SDF) models

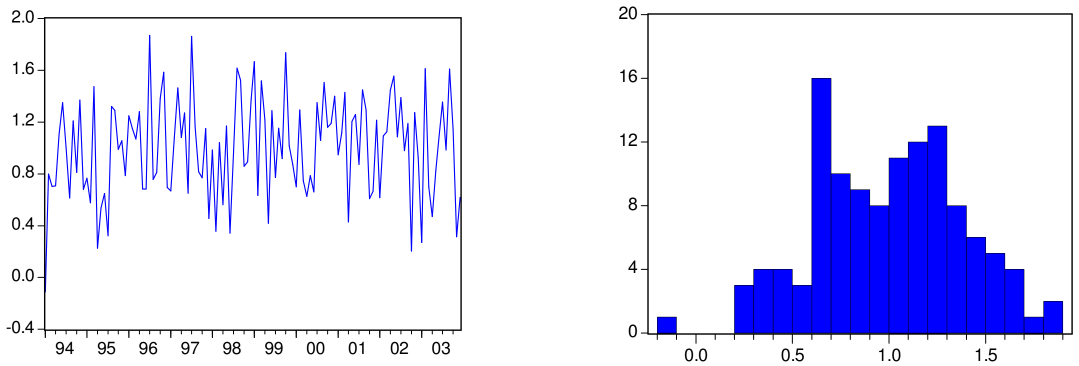
Panel A: SDF-CAPM



Panel B: SDF-Fama-French



Panel C: SDF-Primitive-efficient



Panel D: SDF-Bakshi-Chen

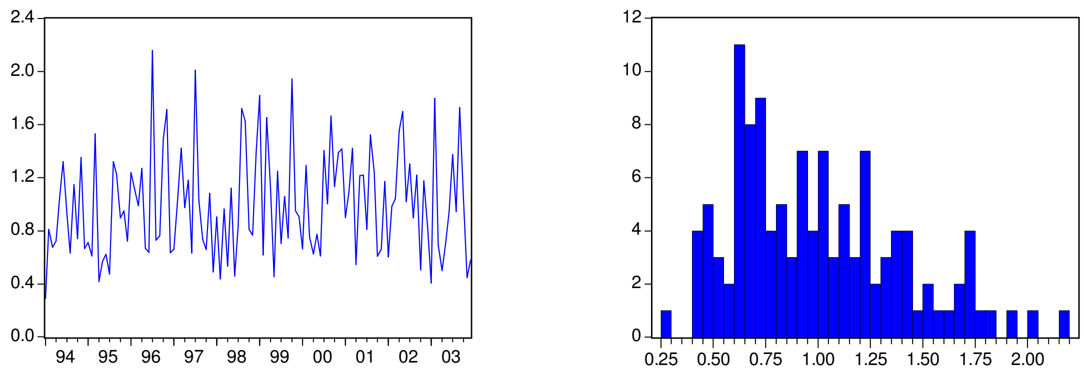
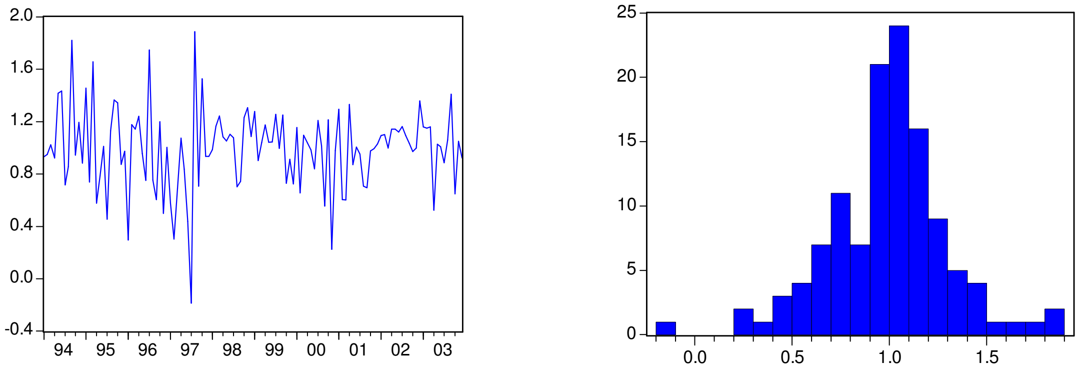
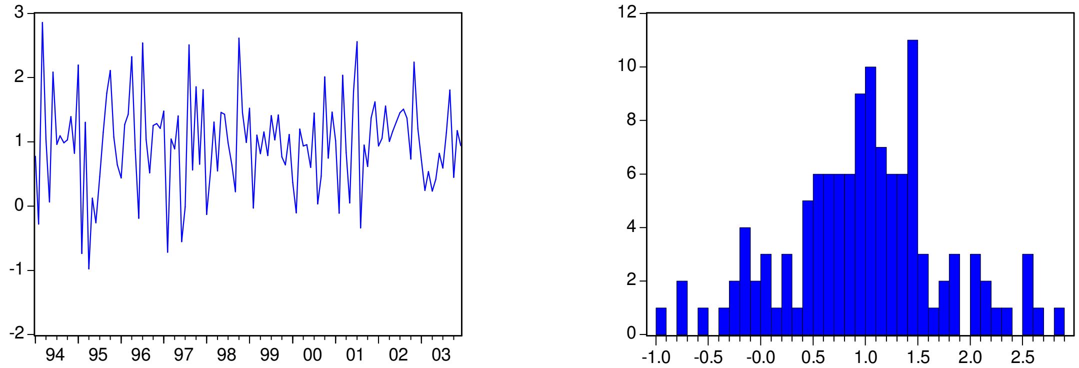


Figure 7: Conditional stochastic discount factor (SDF) models

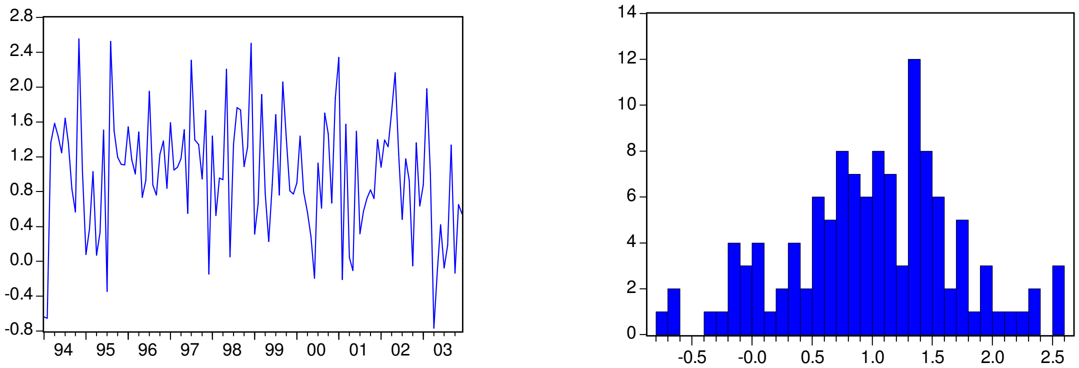
Panel A: SDF-CAPM



Panel B: SDF-Fama-French



Panel C: SDF-Primitive-efficient



Panel D: SDF-Bakshi-Chen

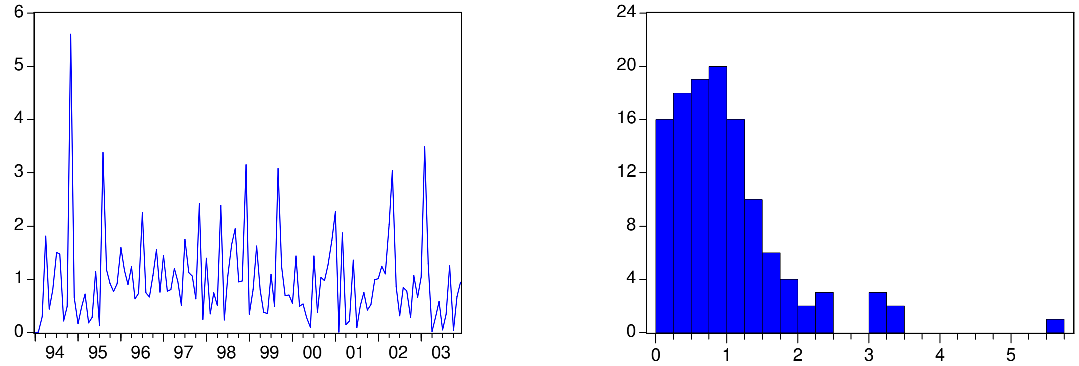
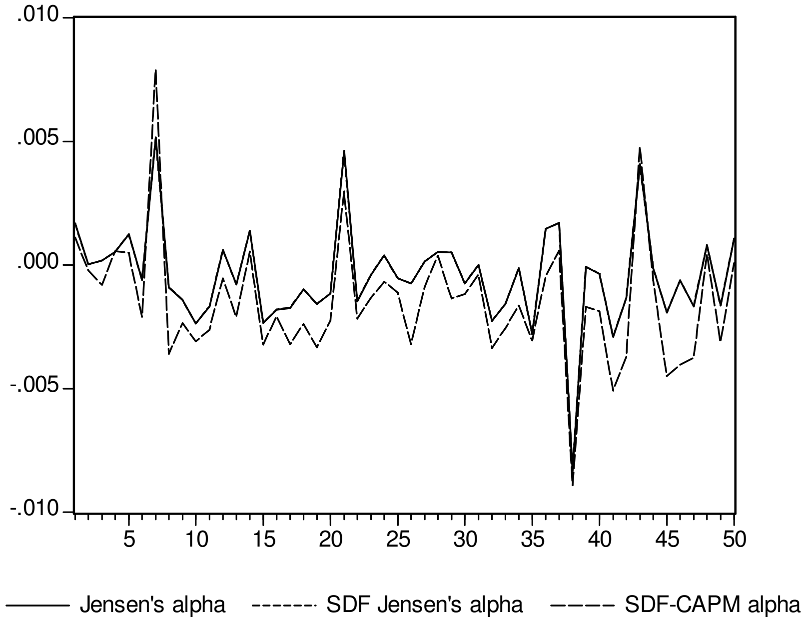
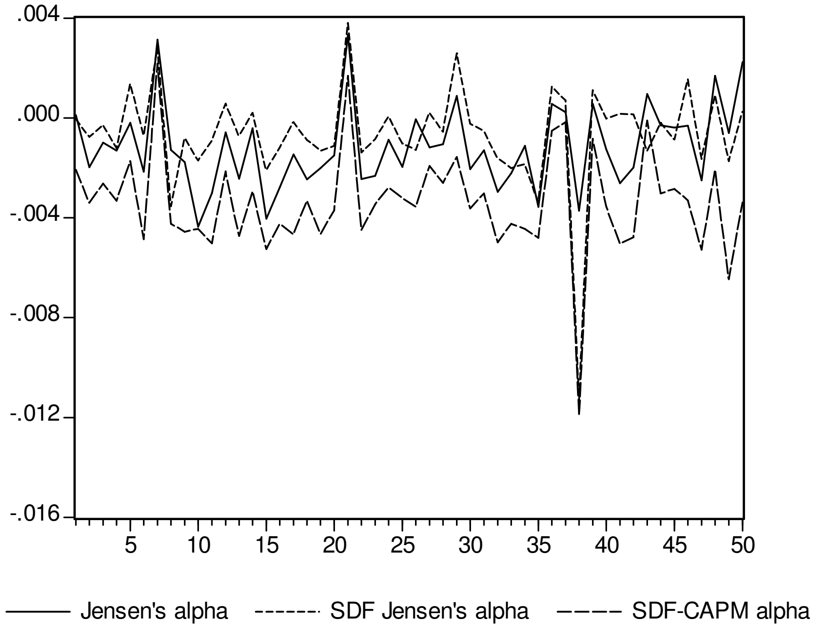


Figure 3: Beta-pricing models versus stochastic discount factor (SDF) models

Panel A: Unconditional performance measures



Panel B: Conditional performance measures



This figure displays alternative performance measures from tests in the beta-pricing framework and the stochastic discount factor (SDF) methodology for our reduced sample of 50 funds: (i) the traditional Jensen's alpha, (ii) the SDF Jensen's alpha, and (iii) the SDF alpha from the SDF-CAPM model. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds). The units are in percentages per month.

Table 1: Estimates of the survivorship bias based on raw returns

Panel A: Estimates of survivorship bias

	Base year: 1994		*Base year: 1997*		*Base year: 2000*	
	All funds in existence each year	Surviving funds	All funds in existence each year	Surviving funds	All funds in existence each year	Surviving funds
General	4.52%	4.97%	3.94%	4.35%	-13.43%	-11.94%
Small & mid-cap	4.61%	6.91%	6.01%	6.91%	-13.32%	-10.67%
All funds	4.53%	5.08%	4.15%	4.35%	-13.78%	-11.94%

Panel B: Differences in rates of return of surviving and nonsurviving funds

	Total funds in existence in base year		*Total number of fund surviving until 2003*		*Funds that did not survive until 2003*			*T-test for difference in means**
	Mean return	Number	Mean return	Number	Mean return	Number	Mortality rate	
1994	-4.83%	63	-4.69%	49	-5.35%	14	22.22%	0.41
1995	3.43%	70	4.66%	53	-0.42%	17	24.29%	2.20**
1996	23.72%	71	25.31%	57	17.22%	14	19.72%	2.93***
1997	38.53%	79	40.29%	66	29.56%	13	16.46%	2.44**
1998	13.96%	86	15.07%	73	7.72%	13	15.12%	1.75*
1999	34.06%	90	35.39%	75	27.41%	15	16.67%	1.71*
2000	-4.98%	94	-3.24%	81	-15.86%	13	13.83%	2.79**
2001	-23.42%	99	-21.54%	85	-34.84%	14	14.14%	2.43**
2002	-41.87%	94	-41.79%	90	-43.70%	4	4.26%	0.24

* The *t*-test is based on a Welch test for equality of the mean returns of surviving and nonsurviving funds. It assumes that the variances of surviving and nonsurviving funds are unknowns and unequal.

Table 2: Summary statistics of primitive assets

	Sector 1	*Sector 2*	*Sector 3*	*Sector 4*	*Sector 5*	*Sector 5*	*Sector 7*	*Sector 8*	*Sector 9*	*Sector 10*
Mean	0.0324	0.0087	0.0079	0.0082	0.0092	0.0081	0.0069	0.0096	0.0103	0.0046
Maximum	0.5135	0.2349	0.1638	0.1949	0.1750	0.1194	0.1739	0.3832	0.1075	0.2682
Minimum	-0.3936	-0.1646	-0.1638	-0.1930	-0.2579	-0.1377	-0.2213	-0.3183	-0.0533	-0.2744
Standard deviation	0.1425	0.0743	0.0662	0.0673	0.0790	0.0458	0.0675	0.1086	0.0303	0.0741
JB test statistic	6.9247	3.9532	1.1102	2.1235	6.0573	8.7276	1.3090	16.5177	17.3392	51.0554
p-value	0.0314	0.1385	0.5740	0.3459	0.0484	0.0127	0.5197	0.0003	0.0002	0.0000
R^2	0.0420	0.0674	0.0293	0.0581	0.0184	0.0104	0.0314	0.0458	0.0074	0.0485
Adjusted R^2	0.0172	0.0433	0.0042	0.0338	-0.0070	-0.0152	0.0063	0.0211	-0.0183	0.0239

The table reports the summary statistics of the primitive assets that are used to estimate the parameters of the stochastic discount factor model in performance measurement tests within the Euler-equation framework. We use the 10 sector indexes (on a total return basis) according to the Datastream classification as the set of primitive assets (in addition to the risk-free security to fix the mean of the stochastic discount factor). The sample period runs from 1994.01 to 2003.12 (120 monthly return observations). The units are in percentages per month. The units are in percentages per month

Table 3: Regressions of stock market excess returns on lagged information variables

	MSCI Germany	*Datastream Germany*	*DAX 30*	*DAFOX*
Panel A: Single-predictor regressions				
Log dividend yield				
Coefficient	0.023	0.022	0.024	0.011
t-statistic	0.789	0.872	0.789	0.504
Amihud-Hurvich *t*-statistic	0.739	0.825	0.727	0.478
R^2	0.005	0.006	0.005	0.002
Short-term interest rate				
Coefficient	-0.174	-0.170	-0.177	-0.151
t-statistic	-2.074**	-2.300**	-2.033**	-2.169**
Amihud-Hurvich *t*-statistic	-2.000**	-2.202**	-1.917*	-2.087**
R^2	0.035	0.042	0.033	0.038
Term spread				
Coefficient	0.019	0.018	0.020	0.014
t-statistic	2.250**	2.444**	2.309**	2.071**
Amihud-Hurvich *t*-statistic	2.036**	2.212**	2.076**	1.855*
R^2	0.041	0.048	0.043	0.035
Panel B: Multi-predictor regressions				
Log dividend yield				
Coefficient	-0.006	-0.005	-0.006	-0.012
t-statistic	-0.180	-0.180	-0.187	-0.418
Short-term interest rate				
Coefficient	-0.130**	-0.126**	-0.129**	-0.122**
t-statistic	-2.347	-2.576	-2.239	-2.470
Term spread				
Coefficient	0.015**	0.014**	0.017**	0.012*
t-statistic	2.215	2.230	2.195	1.937
Adjusted R^2	0.034	0.046	0.035	0.033
Wald test (χ^2-statistic)	18.710 (0.000)	21.271 (0.000)	18.052 (0.000)	19.059 (0.000)

The table reports the estimation results of predictive regressions of excess benchmark returns on the information variables in single- and multi-predictor specifications. The information variables are (i) the lagged log dividend yield on the Datastream German Total Return Index, (ii) the lagged level of the 1-month interest rate for Euro deposits (German Mark deposits before January 1, 2000) on the Eurocurrency market, and (iii) the lagged slope of the term structure, computed as the difference between the yield on long term government bonds (with maturity of at least 10 years) and the 3-month interest rate for Euro (German Mark) deposits on the Eurocurrency market. The sample period is from January 1994 to December 2003 (120 months). For each regression specification the estimated coefficient and the standard *t*-statistic are reported. In addition, for the single-predictor specifications in panel A we provide the bias-adjusted *t*-statistic following the method proposed by Amihud and Hurvich (2004). The multi-predictor regressions in table B apply Hansen's (1982) Generalized Method of Moments (GMM), using a heteroskedasticity and autocorrelation consistent covariance matrix as weighting matrix and a constant and all three lagged information variables to specify the orthogonality conditions. The χ^2-statistic tests the null hypothesis that all three information variables are jointly zero.

Table 4: Measures of performance using unconditional models (50 selected funds)

	Unconditional CAPM					Fama-French 3-factor model								
	α_P	$t(\alpha_P)$	β_P	$t(\beta_P)$	R^2	α_P	$t(\alpha_P)$	β_{1P}	$t(\beta_{1P})$	β_{2P}	$t(\beta_{2P})$	β_{3P}	$t(\beta_{3P})$	R^2
Panel A: Mean values of individual fund regressions														
General	-0.045	-0.286	1.037	32.32	0.894	-0.217	-1.329	1.030	23.13	-0.008	0.283	-0.078	-2.080	0.889
Small- & mid-cap	0.008	0.005	0.851	14.82	0.694	0.293	0.999	1.062	17.12	-0.207	-3.170	0.329	4.141	0.769
All funds	-0.042	-0.269	1.026	31.27	0.882	-0.187	-1.189	1.032	22.77	-0.020	0.076	-0.053	-1.707	0.882
Panel B: Median values of individual fund regressions														
General	-0.061	-0.358	1.066	33.37	0.931	-0.261	-1.517	1.045	23.58	0.021	0.788	-0.111	-2.393	0.926
Small- & mid-cap	0.081	0.271	0.848	13.94	0.702	0.377	1.264	1.078	17.43	-0.186	-3.254	0.357	4.980	0.768
All funds	-0.058	-0.354	1.060	32.48	0.930	-0.243	-1.487	1.049	23.42	0.004	0.130	-0.101	-2.266	0.923
Panel C: Results for equally-weighted portfolios														
General	-0.045	-0.391	1.037	52.37	0.963	-0.217	-1.557	1.030	31.88	-0.008	-0.317	-0.077	-2.273	0.948
Small- & mid-cap	0.008	0.003	0.851	17.76	0.740	0.293	1.182	1.062	20.03	-0.207	-3.406	0.329	-3.406	0.816
All funds	-0.034	-0.303	1.025	53.42	0.963	-0.186	-1.329	1.032	31.92	-0.020	-0.743	0.053	-1.538	0.947

The unconditional CAPM tests the following regression model (market model):

$$r_{P,t+1} = \alpha_P + \beta_P r_{B,t+1} + \varepsilon_{P,t+1},$$

where $r_{P,t+1}$ is the excess return of a fund and $r_{B,t+1}$ is the excess return on the benchmark index. Benchmark indexes are the Datastream Germany Total Return Index for general funds, and the DAFOX for small- and mid-cap funds. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds). The Fama-French (1993) three-factor model is specified as follows:

$$r_{P,t+1} = \alpha_P + \beta_{1P} r_{B,t+1} + \beta_{2P} HML_{t+1} + \beta_{3P} SMB_{t+1} + \varepsilon_{P,t+1},$$

where the DAFOX is used as the market proxy for all funds, irrespective of the segment they belong to, *HML* is the return on a portfolio of high book-to-market stocks minus low book-to-market stocks, and *SMB* is the return difference between the DAX 30 and the SDAX indexes. The units are in percentages per month. All *t*-ratios are adjusted for heteroscedasticity using the White (1980) covariance matrix.

Table 5: Measures of performance using conditional models (1)

	α_P	$t(\alpha_P)$	δ_{1P}	$t(\delta_{1P})$	δ_{2P}	$t(\delta_{2P})$	*# sign.*	R^2
Panel A: Information variable = Term spread								
Mean values of individual fund regressions:								
General	-0.102	-0.786	1.047	34.059	0.052	1.278	25	0.896
Small- & mid-cap	0.090	0.232	0.838	14.117	-0.091	-1.019	0	0.697
All funds	-0.091	-0.725	1.035	32.862	0.044	1.140	25	0.884
Median values of individual fund regressions:								
General	-0.137	-0.882	1.080	34.730	0.077	1.604	25	0.935
Small- & mid-cap	0.126	0.375	0.831	12.317	-0.109	-1.189	0	0.703
All funds	-0.106	-0.777	1.071	34.003	0.066	1.497	25	0.932
Results for equally-weighted portfolios:								
General	-0.102	-0.947	1.047	57.013	0.052	1.843		0.864
Small- & mid-cap	0.090	0.281	0.838	17.740	-0.078	-1.169		0.740
All funds	-0.084	-0.751	1.033	57.080	0.046	1.618		0.964
Panel B: Information variable = Short-term interest rate								
Mean values of individual fund regressions:								
General	-0.111	-0.831	1.026	34.049	-0.738	-1.757	29	0.898
Small- & mid-cap	0.075	0.179	0.870	15.129	0.879	1.088	1	0.699
All funds	-0.100	-0.771	1.016	32.914	-0.641	-1.587	30	0.886
Median values of individual fund regressions:								
General	-0.115	-0.717	1.053	35.270	-0.913	-2.257	29	0.938
Small- & mid-cap	0.150	0.385	0.851	13.622	0.566	0.546	1	0.714
All funds	-0.111	-0.682	1.045	34.377	-0.793	-2.166	30	0.937
Results for equally-weighted portfolios:								
General	-0.111	-1.114	1.026	60.864	-0.739	-3.020		0.965
Small- & mid-cap	0.075	0.231	0.870	17.666	0.879	1.350		0.745
All funds	-0.095	-0.910	1.014	60.546	-0.671	-2.798		0.965
Panel C: Information variable = Log dividend yield								
Mean values of individual fund regressions:								
General	-0.052	-0.391	1.034	32.733	3.323	0.921	23	0.900
Small- & mid-cap	0.015	0.014	0.860	14.394	-7.785	-0.805	1	0.701
All funds	-0.048	-0.367	1.023	31.633	2.657	0.817	24	0.888
Median values of individual fund regressions:								
General	-0.067	-0.444	1.057	33.557	4.633	0.976	23	0.937
Small- & mid-cap	0.101	0.277	0.844	13.902	-2.548	-0.225	1	0.722
All funds	-0.059	-0.425	1.049	33.158	4.305	0.855	24	0.935
Results for equally-weighted portfolios:								
General	-0.052	-0.515	1.034	53.117	3.324	1.102		0.963
Small- & mid-cap	0.015	0.046	0.860	16.644	-7.785	-1.122		0.740
All funds	-0.044	-0.380	1.021	54.195	2.603	0.847		0.963

This table shows the results from the conditional CAPM tests, where each predetermined information variable is used separately one at a time in the following regression model:

$$r_{P,t+1} = \alpha_P + \delta_{1P} r_{B,t+1} + \delta_{2P}\left(z_t r_{B,t+1}\right) + \varepsilon_{P,t+1} ,$$

where $r_{P,t+1}$ is the excess return of a fund, $r_{B,t+1}$ is the excess return on the benchmark index, and z_t denotes a predetermined (lagged) information variable. Benchmark indexes are the Datastream Germany Total Return Index for general funds, and the DAFOX for small- and mid-cap funds. As predetermined information variables we use the term-spread (panel A), the short-term interest rate (panel B), and the log dividend yield (panel C) separately one at a time. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds). R^2 denotes the adjusted R-squares of the regressions, and *#sign* denotes the number of funds with significantly estimated coefficients (δ_{2P}) on the interaction term including the lagged information variable. The units are in percentages per month. All *t*-ratios are adjusted for heteroscedasticity using the White (1980) covariance matrix.

Table 6: Measures of performance using the conditional CAPM (2)

	α_P	$t(\alpha_P)$	δ_{1P}	$t(\delta_{1P})$	$\delta_{TS,P}$	$t(\delta_{TS,P})$	$\delta_{SR,P}$	$t(\delta_{SR,P})$	$\delta_{DY,P}$	$t(\delta_{DY,P})$	R^2	*pval(F)*
Panel A: Mean values of individual fund regressions												
General	-0.130	-1.002	1.032	40.511	0.025	0.790	-0.636	-1.540	0.491	0.309	0.904	0.084 [35]
Small- & mid-cap	0.111	0.288	0.858	14.281	-0.066	-0.771	0.508	0.669	-5.012	-0.413	0.705	0.423 [1]
All funds	-0.116	-0.924	1.021	38.937	0.019	0.697	-0.567	-1.407	0.160	0.266	0.892	0.104 [36]
Panel B: Median values of individual fund regressions												
General	-0.131	-1.074	1.047	41.457	0.044	1.078	-0.635	-1.841	1.118	0.304	0.943	0.010
Small- & mid-cap	0.169	0.494	0.826	14.026	-0.104	-1.163	0.294	0.301	-0.186	-0.015	0.726	0.584
All funds	-0.129	-1.034	1.043	40.563	0.041	0.990	-0.619	-1.769	1.011	0.289	0.939	0.011
Panel C: Results for equally-weighted portfolios												
General	-0.130	-1.228	1.032	72.646	0.025	1.026	-0.637	-2.749	0.488	0.201	0.964	0.019
Small- & mid-cap	0.111	0.340	0.858	16.157	-0.066	-0.870	0.508	0.822	-5.012	-0.641	0.738	0.383
All funds	-0.111	-0.997	1.019	70.227	0.020	0.828	-0.603	-2.635	-0.047	-0.018	0.964	0.042

This table shows the results from the conditional CAPM tests, where all predetermined information variables are used simultaneously in the following regression model:

$$r_{P,t+1} = \alpha_P + \delta_{1P} r_{B,t+1} + \delta'_{2P}\left(z_t r_{B,t+1}\right) + \varepsilon_{P,t+1} ,$$

where $r_{P,t+1}$ is the excess return of a fund, $r_{B,t+1}$ is the excess return on the benchmark index, and z_t denotes the vector of predetermined (lagged) information variable. Benchmark indexes are the Datastream Germany Total Return Index for general funds, and the DAFOX for small- and mid-cap funds. As predetermined information variables we use the term-spread (*TS*), the short-term interest rate (*SR*), and the log dividend yield (*DY*) simultaneously. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds). R^2 are the adjusted R-squares of the regressions, and *pval(F)* denotes the probability value of the *F*-test for the null hypothesis that all coefficients on the interaction terms including the predetermined variables (collected in the vector δ_{2P}) are simultaneously equal to zero. The figures in brackets in the last column denote the number of funds where the null hypothesis is rejected. The units are in percentages per month. All *t*-ratios are adjusted for heteroscedasticity using the White (1980) covariance matrix.

Table 7: Rank correlation of conditional alphas using alternative benchmarks

	MSCI Germany	*Datastream Germany*	*DAX 30*	*DAFOX*
MSCI Germany	1			
Datastream Germany	0.9962	1		
DAX 30	0.9991	0.9952	1	
DAFOX	0.9903	0.9913	0.9875	1

The table contains the Spearman rank correlations of alphas in the conditional CAPM for individual funds using alternative benchmarks for the sample of 50 selected funds. The regression specification is identical to table 14, where the term-spread, the short-term interest rate, and the log dividend yield are used simultaneously as the set of predetermined information variables. The sample period is from January 1994 to December 2003.

Table 8: Cross-sectional distribution of *t*-statistics for the alphas

		Single-index model					*Three-factor model*	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
	Null	*Uncond. model*	*Term spread*	*Short-term rate*	*Dividend yield*	*All information variables*	*Uncond. model*	*All information variables*
Minimum *t*-statistic		-1.9087	-2.6894	-2.8722	-2.5358	-2.9042	-2.9799	-2.368171
Bonferroni *p*-value (–)		1.4676	0.2051	0.1210	0.3135	0.1104	0.0877	0.4909
$t < -2.326$	0.25	0	4	4	1	5	6	1
$-2.326 < t < -1.960$	1.00	0	0	1	2	0	6	2
$-1.960 < t < -1.645$	1.25	3	5	3	2	12	9	8
$-1.645 < t < 0$	22.50	28	27	32	30	23	22	29
$0 < t < 1.645$	22.50	18	13	9	14	9	7	10
$1.645 < t < 1.960$	1.25	0	0	0	0	0	0	0
$1.960 < t < 2.326$	1.00	0	1	1	0	1	0	0
$t > 2.326$	0.25	1	0	0	1	0	0	0
Maximum *t*-statistic		2.6711	2.2885	2.1881	2.8804	2.0984	1.2873	1.558308
Bonferroni *p*-value (+)		0.2154	0.5976	0.7663	0.1181	0.9514	5.0129	3.0521

This table shows the cross-sectional distribution of heteroscedasticity-consistent *t*-values for the estimated alphas using different model specifications. For the single-index model in column (2), the unconditional alphas are the intercepts in regressions of fund excess returns on the excess returns on the Datastream Germany Total return index and the DAFOX for the subsamples of general funds and small- and mid-cap funds, respectively. The conditional alphas in columns (3)-(6) are the intercepts in regressions of fund excess returns on the benchmark index and the product of the index with the vector of predetermined (lagged) information variables. The unconditional alphas in the Fama and French (1993) three-factor model in column (7) are the intercepts in the regressions of the excess returns of the funds on the DAFOX and the *HML* and *SMB* long-short portfolios. The conditional alphas in the three-factor model are the intercepts when fund excess returns are regressed on the factors and the products of the factors with the vector of predetermined (lagged) information variables. The entries in the middle block of the table indicate the number of funds for which the *t*-values for the alphas fall within the range of critical values of a standard normal distribution. The Bonferroni *p*-value is the maximum or minimum one-tailed *p*-value from the distribution of *t*-values, across all funds, multiplied by the number of funds. As predetermined information variables we use the term-spread, the short-term interest rate, and the log dividend yield simultaneously. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds).

Table 9: Stochastic discount factor models

	E(*m*)	*SD*(*m*)	$\rho_1(m)$	*Minimum*	*Maximum*	# (*m* < *0*)
Constant discount factor	0.9961	0.0000	0.0000	0.9963	0.9961	0
Panel A: Unconditional stochastic discount factor models						
SDF-CAPM	0.9965	0.0784	0.015	0.7922	1.2763	0
SDF-Bakshi-Chen	0.9970	0.4007	-0.071	0.2893	2.1600	0
SDF-Fama-French	0.9905	0.4745	0.055	-0.7183	2.1711	3
SDF-Primitive-efficient	0.9969	0.3883	-0.062	-0.1113	1.8680	1
Panel B: Conditional stochastic discount factor models						
SDF-CAPM	0.9910	0.3112	-0.146	-0.1860	1.8866	1
SDF-Bakshi-Chen	0.9970	0.8410	-0.057	0.0012	5.6034	0
SDF-Fama-French	0.9800	0.7361	-0.153	-0.9771	2.8601	13
SDF-Primitive-efficient	0.9969	0.6953	0.043	-0.7683	2.5521	10

This table shows the results from various models for stochastic discount factors, denoted as m_{t+1}, using the following system of equations:

$$u1_t = \left(m_{t+1}R_{t+1} - 1\right) \otimes Z_t$$

where $u1_t$ denotes the vector of pricing errors relating to the *N* primitive assets, whose gross returns are collected in the vector R_{t+1}, and Z_t is the vector of predetermined information variables. The parameters of the stochastic discount factors are estimated using Hansen's (1982) Generalized Method of Moments (GMM), minimizing a quadratic form of the pricing errors. The standard errors of the estimated coefficients are corrected for the effects of heteroscedasticity using the White (1980) methodology. *E*(*m*) is the sample mean, *SD*(*m*) is the sample standard deviation, and $\rho_1(m)$ is the first-order autocorrelation of the estimated stochastic discount factor (fitted values). The primitive assets used in estimating the models are the 10 sector portfolios according to the Datastream classification and the risk-free security. In linear models of the stochastic discount factors (SDF-CAPM and SDF-Fama-French) the Datastream Germany Total Return Index is used as the market proxy. The predetermined information variables are the term-spread, the short-term interest rate, and the log dividend yield. To avoid an explosive number of orthogonality conditions, the results for the conditional specifications of both the SDF-Bakshi-Chen model and the SDF-Primitive-efficient model are based on the short-term interest as the only information variable in the estimation. In contrast, for the SDF-CAPM and the SDF-Fama-French models we use the full set of information variables to scale factors. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds).

Table 10: Dynamic performance of stochastic discount factor models

	Sector 1	*Sector 2*	*Sector 3*	*Sector 4*	*Sector 5*	*Sector 5*	*Sector 7*	*Sector 8*	*Sector 9*	*Sector 10*	*Risk-free*
Constant discount factor	2.908%	1.920%	1.128%	1.616%	1.067%	0.466%	1.192%	2.314%	0.259%	1.626%	0.076%
Panel A: Unconditional stochastic discount factor models											
SDF-CAPM	2.647%	1.874%	0.949%	0.451%	1.074%	1.768%	0.917%	0.490%	2.031%	0.812%	2.060%
SDF-Fama-French	8.775%	10.083%	9.751%	9.239%	9.516%	11.163%	10.003%	8.317%	11.262%	9.587%	11.276%
SDF-Primitive-efficient	8.384%	7.275%	8.464%	8.833%	8.336%	7.711%	9.086%	9.723%	7.900%	9.107%	7.727%
SDF-Bakshi-Chen	6.171%	5.965%	7.087%	7.366%	6.894%	6.488%	7.681%	8.165%	6.808%	7.689%	6.632%
Panel B: Conditional stochastic discount models											
SDF-CAPM	1.806%	1.725%	1.191%	1.272%	1.451%	1.242%	1.435%	1.661%	1.367%	1.026%	1.120%
SDF-Fama-French	3.041%	3.397%	3.592%	3.773%	3.577%	3.542%	3.456%	3.668%	3.999%	4.254%	3.642%
SDF-Primitive-efficient	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
SDF-Bakshi-Chen	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

This table shows the standard deviations of the fitted pricing errors from the following system of equations:

$$u1_t = \left(m_{t+1}R_{t+1} - 1\right) \otimes Z_t$$

where $u1_t$ denotes the vector of pricing errors relating to the N primitive assets, whose gross returns are collected in the vector R_{t+1}, m_{t+1} is the stochastic discount factor, and Z_t is the vector of predetermined information variables. The fitted pricing errors are the fitted values of regressions of $u1_t$ on Z_t, using different models of the stochastic discount factor, and the entries in the table show the corresponding standard errors for all primitive assets. The parameters of the stochastic discount factors are estimated using Hansen's (1982) Generalized Method of Moments (GMM), minimizing a quadratic form of the pricing errors. The standard errors of the estimated coefficients are corrected for the effects of heteroscedasticity using the White (1980) methodology. The primitive assets used in estimating the models are the 10 sector portfolios according to the Datastream classification and the risk-free security. In linear models of the stochastic discount factors (SDF-CAPM and SDF-Fama-French) the Datastream Germany Total Return Index is used as the market proxy. The predetermined information variables are the term-spread, the short-term interest rate, and the log dividend yield. To avoid an explosive number of orthogonality conditions, the results we present for the conditional specifications of both the SDF-Bakshi-Chen model and the SDF-Primitive-efficient model are based on the short-term interest as the only information variable in the estimation. In contrast, for the SDF-CAPM and the SDF-Fama-French models we use the full set of information variables to scale factors. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds).

Table 11: Performance of funds using the stochastic discount factor framework

	Bonferroni p-value (–)	*Bottom 3*	*Bottom 10*	*Mean*	*EW-Portf.*	*Median*	*Top 10*	*Top 3*	*Bonferroni p-value* (+)
				Panel A: Unconditional SDF models					
SDF-CAPM	0.005	-0.448	-0.322	-0.149	-0.149	-0.179	0.007	0.112	0.942
SDF-Fama-French	0.094	-0.838	-0.476	-0.304	-0.304	-0.275	-0.059	0.089	0.935
SDF-Primitive-efficient	0.136	-0.414	-0.348	-0.177	-0.192	-0.205	-0.031	0.297	1.528
SDF-Bakshi-Chen	0.068	-0.456	-0.399	-0.238	-0.210	-0.256	-0.103	0.073	1.354
				Panel B: Conditional SDF models					
SDF-CAPM	0.000	-0.528	-0.478	-0.340	-0.360	-0.343	-0.207	-0.018	2.788
SDF-Fama-French	0.000	-0.924	-0.537	-0.330	-0.320	-0.280	-0.108	0.043	0.000
SDF-Primitive-efficient	0.007	-0.692	-0.347	-0.234	-0.272	-0.223	-0.042	0.023	6.965
SDF-Bakshi-Chen	0.001	-0.519	-0.407	-0.243	-0.290	-0.221	-0.076	0.006	5.703

This table shows the distribution of mutual fund SDF alphas for different models of the stochastic discount factor, m_{t+1}, from estimations of the following system of equations:

$$u1_t = \left(m_{t+1}R_{t+1} - 1\right) \otimes Z_t$$
$$u2_t = \alpha_P - m_{t+1}R_{P,t+1} + 1,$$

where $u1_t$ denotes the vector of pricing errors relating to the N primitive assets, whose gross returns are collected in the vector R_{t+1}, $u2_t$ is the pricing error of the fund with gross return $R_{P,t+1}$, m_{t+1} is the stochastic discount factor, and Z_t is the vector of predetermined information variables. α_P is the SDF alpha, depending on the model of the stochastic discount factor. All parameters are estimated using Hansen's (1982) Generalized Method of Moments (GMM), minimizing a quadratic form of the pricing errors. The standard errors of the estimated coefficients are corrected for the effects of heteroscedasticity using the White (1980) methodology. The primitive assets used in estimating the models are the 10 sector portfolios according to the Datastream classification and the risk-free security. In linear models of the stochastic discount factors (SDF-CAPM and SDF-Fama-French) the Datastream Germany Total Return Index is used as the market proxy. The predetermined information variables are the term-spread, the short-term interest rate, and the log dividend yield. To avoid an explosive number of orthogonality conditions, the results we present for the conditional specifications of both the SDF-Bakshi-Chen model and the SDF-Primitive-efficient model are based on the short-term interest as the only information variable in the estimation. In contrast, for the SDF-CAPM and the SDF-Fama-French models we use the full set of information variables to scale factors. The Bonferroni p-value is the maximum or minimum one-tailed p-value from the distribution of t-values, across all funds, multiplied by the number of funds. The sample period is from January 1994 to December 2003, and the full return history (120 months) is available for 47 general funds and 3 small- and mid-cap funds (50 surviving funds).

Table 12: Rank correlation of SDF alphas

	SDF-CAPM	*SDF-Fama-French*	*SDF-primitive efficient*	*SDF-Bakshi-Chen*
Panel A: Unconditional model specifications				
SDF-CAPM	1			
SDF-Fama-French	0.7279	1		
SDF-primitive efficient	0.9476	0.7940	1	
SDF-Bakshi-Chen	0.8553	0.7219	0.9405	1
Panel B: Conditional model specifications				
SDF-CAPM	1			
SDF-Fama-French	0.6415	1		
SDF-primitive efficient	0.6791	0.6623	1	
SDF-Bakshi-Chen	0.7203	0.7540	0.9102	1

The table contains the Spearman rank correlations of estimated SDF alphas for the sample of 50 selected funds using different specifications for the stochastic discount factor. The SDF alphas are estimated from the system of equations (18) using Hansen's (1982) Generalized Method of Moments (GMM). The sample period is from January 1994 to December 2003.