Why not use SDF rather than beta models in performance measurement?

Jonas Gusset · Heinz Zimmermann

Abstract This paper analyzes performance measurement based on stochastic discount factors, compared to beta models traditionally used in computing funds' (Jensen) alphas. From a theoretical point of view, standard alphas suffer from several limitations. Our paper addresses this issue from an empirical point of view using a sample of Swiss mutual funds from 2000 to 2011. Our results suggest that the key for a "fair" comparison between stochastic discount function (SDF) and beta models is the specification of the set of primitive assets used to calibrate the SDF function. Once this is established, the size of (absolute) performance differences considerably decreases between the two model families. However, there are sizeable performance deviations in the cross-section of funds if conditioning information is incorporated in the tests, up to some 20 basis points per month, or about 2.3 % per year. In almost all cases, the SDF-alphas are lower than the standard (Jensen) alphas. In absolute terms, the average SDF-based underperformance of the funds is way larger than the average total expense ratio (TER) of the funds, both in a conditional and unconditional setting.

Keywords Performance measurement · Conditioning information · Mutual funds

JEL Classification G12 · G23

1 Introduction

Performance measurement of actively managed portfolios has remained a topic of overwhelming interest in both research and practice over the past decades. While

J. Gusset · H. Zimmermann (✉)
WWZ Department of Finance, Wirtschaftswissenschaftliches Zentrum WWZ,
Universität Basel, Peter Merian Weg 6, 4003 Basel, Switzerland
e-mail: heinz.zimmermann@unibas.ch

in the asset management industry, regression-based "beta-pricing models" and their associated "alphas" (Jensen alphas) are still predominantly used to measure performance, much theoretical progress driven by the advances in asset pricing theory was achieved over the past decade, which widened the gap between the models used in academic research and those implemented in practice. A striking observation is that performance measurement based on stochastic discount functions (SDFs)[1] is almost inexistent in practical applications, although its merits vis-à-vis the beta-pricing models are well documented in the literature;[2] see Ferson (2012) for an extensive summary and discussion.

In the beta-pricing framework, the performance of an active manager is measured in terms of excess portfolio returns with respect to excess returns measured on one or several "benchmark" portfolios which are often used to control for the investment "style" of an investor. Typical beta-pricing models are the CAPM or multifactor-pricing models with excess returns on style-sorted portfolios. Given the popularity of these models, it is not surprising that the associated performance measures (Jensen alphas) enjoy widespread application as well. However, Ferson (2012) and Ferson and Lin (2012) discuss several methodological limitations of this approach compared to the SDF models. First, the sign of alphas has in general an ambiguous relationship to the marginal expected utility from adding or eliminating the assets from a portfolio.[3] The appropriate specification of an SDF eliminates this problem. Second, alphas lose their intuitive interpretation in the presence of conditioning information. And third, typical benchmark portfolios used in practice exhibit statistical properties which provide biased results; in the terminology of Aragon and Ferson (2006) they do not qualify as "appropriate benchmarks". A benchmark is appropriate if it has the same covariance (betas) with the SDF as the evaluated fund(s). As the authors show, in a mean-variance setting, this only holds under very restrictive assumptions, e.g., quadratic utility.

Therefore, from a theoretical point of view, traditional alphas are unreliable indicators of effective superior performance and do not unambiguously reveal whether a mispriced asset should be bought or sold. Given these limitations, it is surprising that the SDF methodology is not more widely used in empirical studies and practical applications. This may be due to the estimation framework which differs slightly from the beta models, and an alleged lack of intuition. However, the methodology has the

---

[1] Stochastic discount factor models have been used in performance measurement by Chen and Knez (1996), Dahlquist and Söderlind (1999), and Ferson et al. (2006).

[2] In order to avoid confusion, excess performance based on beta models is called "alpha" in this paper, while excess performance based on SDF-models is called "SDF-alpha" or SDF pricing error.

[3] The relationship is unambiguous only at the margin, i.e., for infinitesimal portfolio changes (Dybvig and Ross 1985b). However, as shown by Gibbons et al. (1989), in Table 6, an asset with a positive (Jensen) alpha can exhibit a negative weight in the optimal orthogonal portfolio (OOP). [The OOP is a portfolio allocated to a set of mispriced assets, which portfolio weights proportional to the (Jensen) alphas computed with respect to a (possibly inefficient) benchmark portfolio. The OOP can be regarded as the "optimal" orthogonal complement of the benchmark portfolio which, with appropriate weights, can be combined to an efficient portfolio; see Jobson and Korkie (1982) or MacKinlay (1995) for applications to active management and asset pricing.] Ferson (2010) concludes from this observation: "So, even if a positive alpha is attractive at the margin to a mean variance investor, it might not imply buying a positive alpha fund given a realistic discrete response." In contrast, the sign of the SDF-alpha is unambiguous with respect to the marginal expected utility from discrete changes in assets' weights.

same practical flair as the beta-pricing models: while the alphas derived from beta models are solely based on the idea that the risk exposures of a manager's portfolio can be (statically or dynamically) replicated by a benchmark portfolio and excess returns being compared, the alphas of SDF models are derived from the notion that a set of given assets (called "primitive assets") is assumed to be correctly[4] priced on a risk-adjusted basis. From a practitioner's perspective, this set of assets would eventually correspond to the "core" holdings in a core-satellite portfolio approach, or simply to the benchmark or passive assets. Unfortunately, the asset pricing restrictions imposed by the SDF approach differs from the standard beta-models, in particular in a conditional setting and depending on the specification of the primitive assets. The SDF approach goes at no substantial costs; while the traditional alpha is invalidated, other measures such as the widely used Sharpe ratio preserves its applicability (Ferson and Lin 2012).

Aside from these methodological considerations, the question arises to what extent the aforementioned limitations are empirically relevant in practical applications. This is the major focus of this paper. We address this question using a sample of Swiss equity mutual funds which has not yet been analyzed in the literature. There are relatively few empirical studies which allow a direct comparison of SDF- and beta-based performance measures in a conditional and unconditional setting. Bessler et al. (2009) perform a comparative study with German mutual funds and find substantial differences between SDF pricing errors and alphas. The paper does not, however, address the nature of these differences. We show in this paper that a meaningful comparison of performance tests derived from beta and SDF models requires a *consistent* specification of the asset universe underlying the SDF and beta models. Specifically, our results reveal that the key for a "fair" comparison between SDF and beta models is the consistent specification of the set of primitive assets used to calibrate the SDF function. Once this is established, the size of (absolute) performance differences considerably decreases between the two model families.

The paper does not contribute to the literature that addresses estimation issues related to the SDF and beta models, which are of course directly relevant for performance measurement, but which are not yet directly related to this topic. The early studies find that the generality of the SDF approach does not come at the cost of efficiency in parameter estimates or power in specification tests,[5] but more recent tests investigating the small sample properties of the test statistics (Garrett et al. 2011) suggest that the SDF approach is more likely to be less efficient but more robust than beta models. Overall, the findings suggest that the SDF approach where the stochastic deflator is specified by a single or multifactor model does not perform worse in small samples than standard beta models.

The rest of the paper is structured as follows: in Sect. 2, the empirical methodology of the two models, beta and SDF, is shortly presented. Section 3 describes the data used in this study. Section 4 contains the empirical results of the paper, and Sect. 5 summarizes our findings.

---

[4] Respectively, with minimum pricing error.

[5] See: Jagannathan and Wang (2002), and Kan and Zhou (2002).

2 Empirical models: beta vs. SDF

The empirical models used in this paper are described in the Appendix. The estimated beta models have the general structure given by (3b), where α_P are the conditional alphas being estimated. The alpha represents the average deviation between the manager's excess return and the excess return of a dynamic benchmark strategy which replicates the time-varying risk exposures of the manager's portfolio. In the empirical part, we implement three types of beta models:

- A CAPM-type single-beta model with excess returns on a broad market index: CAPM(SPI);
- The Fama–French multi-beta (three-factor) model with excess returns on the market index, a size-related spread portfolio and a value-growth spread portfolio: FF;
- A multi-beta model with excess returns on nine sector portfolios covering the same asset universe as the broad market index: 9SPI.

The SDF model has the general structure given by (4b), and the implied SDF-alphas are displayed in (3b). The alphas are zero if the fund can be replicated by a portfolio (at no cost) of a set of N primitive assets using public information as reflected in Z_t, and the SDF prices the primitive assets correctly, respectively, with minimum error. Performance measurement with SDF involves three steps:

(i) Model specification: The estimated SDF-alphas depend on the specification of the SDF which is not unique unless the market is complete. Thus, different SDF models are likely to produce different SDF-alphas. In the empirical part, we use three SDF model specifications: linear factor model SDFs, primitive-efficient SDFs, and the Bakshi–Chen SDFs. The models are described in "Model specification" of the Appendix. In the empirical part of the paper, we estimate two linear factor specifications of the SDF:
  - A single-factor model using gross market returns (SDF_CAPM);
  - A three-factor Fama–French type model with gross returns on the market, a size-related spread portfolio and a value-growth spread portfolio as factors (SDF_FF).

These specifications make it possible to directly compare the results to the beta models [CAPM(SPI), FF, 9SPI], where the excess returns on the benchmark portfolios are defined analogously. Notice that the risk-free rate is always included in the set of primitive assets.

(ii) Model calibration, i.e., estimating the model parameters. This requires the specification of a set of primitive (or benchmark) assets which are assumed to be correctly priced. In the empirical part, a set of nine sector indices (plus the risk-free security) is used. To make the performance results of the CAPM directly comparable with the corresponding estimates in the beta framework, we moreover add a CAPM-calibration where only the SPI market index (plus the risk-free security) is used as primitive asset; this model is denoted by SDF_CAPM(SPI) in contrast to SDF_CAPM(9SPI) which relies on the standard specification of the primitive assets (SPI sectors).

(iii) Performance estimation, i.e., fitting the SDF-alphas based on (i) and (ii). Given our model specifications and calibration, we get a total of five SDF models in our performance tests:

- SDF_CAPM(9SPI);
- SDF_FF;
- SDF_primitive-efficient;
- SDF_Bakshi–Chen;
- SDF_CAPM(SPI).

3 Data

The empirical part of the paper is based on 46 Swiss mutual funds over the period from January 2000 to August 2011 (140 months). The sample contains 37 equity funds and 9 mixed funds. To keep the sample homogenous, we exclude mixed funds that invest less than one-third in equities. In order to avoid issues related to currency risk, we furthermore restrict our sample to funds that are traded in CHF and invest primarily in the Swiss stock market. Moreover, the funds can be divided into general mutual funds and small- and mid-cap funds. According to the fund classification of Morningstar, our sample contains 37 general mutual funds and 9 small- and mid-cap funds.

We collect monthly fund returns on a total return basis. Fund returns are net of management fees and expenses. Because the sample only contains surviving funds, survivorship bias is a potential problem. Adapting the procedure proposed by Malkiel (1995), we estimate the survivorship bias based on a sample of 112 (surviving and nonsurviving) Swiss mutual funds which existed during the 2000–2011 sample period. The sample contains 70 funds still active at the end of the observation interval, whereas the remaining 42 did not survive. The equally weighted average annual return of the 70 surviving funds is 1.078 %, in contrast to 0.447 % for the entire sample of 112 funds, which implies a survivorship bias of 63 basis points per year. This is slightly more than the bias of 45 and 40 basis points, respectively, reported by Bessler et al. (2009) and Griese and Kempf (2003) for comparable German mutual funds.

Benchmark returns We use the Swiss Performance Total Return Index, SPI TR, as benchmark index. It contains all 221 listed Swiss stocks and therefore captures the total investment opportunity set available to fund managers. For comparison, we also use the MSCI Switzerland TR Index for several tests. To estimate the Fama and French (1993) three-factor model, we use spread returns on size and value-growth style portfolios, denoted by SMB and HML, as additional factors. The construction of spread returns is based on the formulas published on Kenneth French's website, using raw returns on four MSCI Switzerland TR Style Indices (Small Cap, Large Cap, Value, Growth).

Primitive assets In an SDF framework, a set of primitive assets must be specified to calibrate the "benchmark" model; the parameters are estimated from a system of equations such that the sum of squared, weighted SDF pricing errors of the primitive assets is minimized. Theoretically, the set of primitive assets should reflect the major building blocks of the investment universe available to a fund manager. As stated above, the SPI family optimally represents our investment universe, and we therefore use its disaggregated nine (of its ten) sector indices as primitive assets in addition to

the risk-free security. The tenth sector is not included because its data are not available before 2006. The data are collected from Thomson Reuters Datastream, and summary statistics are shown in Table 1. For direct comparison of the empirical results between the SDF and beta models, we use the same sector return data as in multi-beta model (9SPI). To make the CAPM-based performance results more directly comparable with those of the corresponding beta model [CAPM(SPI)], we use the broad SPI market index (plus the risk-free security) as primitive asset in this specific test.

Information variables (instruments) Conditional performance evaluation requires the specification of a subset of measurable proxies for the public information on which investment decisions are based. We include three instruments which have performed well in predicting stock returns and risk in earlier studies: (i) the lagged log dividend yield from the Datastream Swiss TR Index, calculated as the log of the total value of dividends paid of all constituent parts of the index at time t minus the log of the total market value of the index at time t;[6] (ii) the lagged 1-month interest rate (Swiss interbank rate); and (iii) the lagged term spread, computed as the difference between the 10-year interest rate on the Datastream Swiss Benchmark Bond Index and the Swiss 3-month LIBOR.

The predictive power of the instruments is tested by regressing market (SPI) excess returns on the lagged instruments, individually and simultaneously. The results (which are not displayed) reveal that only the short-term interest rate and the term spread have predictive power with significant t-statistics of -2.42 and 2.46, respectively. In contrast, but similar to the findings of Bessler et al. (2009) for German mutual funds, the dividend yield is insignificant with a t-statistic of -0.2. For the simultaneous test of all three instruments, we apply Hansen (1982) GMM using a heteroskedasticity and autocorrelation consistent covariance weighting matrix. The estimation results reveal that the single-predictor results are not robust in the multivariate case. The t-statistics are insignificant for all three information variables. Nevertheless, a Wald test shows that the null hypothesis, where all three information variables are jointly equal to zero, can be rejected at the 5 % significance level. We therefore use the three instruments in our subsequent tests.[7]

4 Empirical results

In this section, the empirical results for the beta- and SDF-based performance tests are presented in two separate sub-sections, followed by a detailed comparison and discussion of the results in Sect. 4.3.

4.1 Performance estimates using beta models

4.1.1 Regression results

Table 2 shows the unconditional and conditional estimation results of the performance tests using a single-beta CAPM(SPI) representation and two multi-beta representations

---

[6] The logarithmic yield is computed because of superior time series properties (Lewellen 2004).

[7] The results are not different if the MSCI Switzerland TR market returns are used.

Table 1 Summary statistics of SPI and primitive assets

	SPI	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6	Sector 7	Sector 8	Sector 9
Summary statistics										
Mean	0.0018	0.0045	0.0046	0.0091	0.0011	0.0048	0.0024	0.0104	−0.0003	0.0004
Median	0.0096	0.0130	0.0147	0.0126	0.0007	0.0117	0.0033	0.0100	0.0058	−0.0034
Maximum	0.1182	0.1192	0.2674	0.2975	0.1348	0.2278	0.1989	0.2995	0.2236	0.5058
Minimum	−0.1393	−0.1365	−0.2648	−0.1713	−0.1686	−0.2542	−0.1422	−0.1748	−0.2039	−0.4486
Std. dev.	0.0443	0.0498	0.0847	0.0659	0.0405	0.0786	0.0500	0.0555	0.0725	0.1369
Skewness	−0.6233	−0.5750	−0.4469	0.1598	−0.2596	−0.3541	0.3244	0.7538	0.0219	0.4192
Kurtosis	3.6429	3.0818	4.3125	5.3093	4.8517	4.0098	4.9645	8.2420	3.7225	4.9996
JB test statistic	11.4764	7.7525	14.7088	31.7054	21.5748	8.8741	24.9678	173.5494	3.0561	27.4243
p value	0.0032	0.0207	0.0006	0.0000	0.0000	0.0118	0.0000	0.0000	0.2170	0.0000
Predictive regressions										
Log DY	−0.0200	−0.0285	−0.0498	−0.0238	−0.0018	−0.0210	0.0021	−0.0674	−0.0290	−0.0742
Short-term	−0.9086	−1.4681	−2.9357	−0.6836	0.1408	−2.0632	−0.3029	−0.9906	−1.3621	−3.3954
Term spread	0.0431	−0.7853	−1.0920	0.5736	0.9353	−0.2650	0.0490	−0.6441	−0.1035	−0.3095
R^2	0.0527	0.0397	0.0703	0.0287	0.0202	0.0600	0.0059	0.0672	0.0320	0.0550
adj. R^2	0.0318	0.0185	0.0498	0.0073	−0.0014	0.0393	−0.0160	0.0467	0.0107	0.0342

This table reports summary statistics of the SPI and the primitive assets that are used to estimate the parameters of the different SDF models. In addition to the risk-free security, we use the nine sector indices (on a total return basis) according to the Datastream classification as our set of primitive assets. The table also shows the coefficients of predictive regressions of the SPI and the sector returns on lagged values of the information variables (log dividend yield, short-term interest rate, and term spread). The sample period is from January 2000 to August 2011 (140 monthly observations) of predictive regressions. The R^2 denotes the explanatory power of these regressions.

Table 2 Beta models

	Unconditional models			Conditional models			
	α_P (%)	$t(\alpha_P)$	R^2	α_P (%)	$t(\alpha_P)$	R^2	p-val(F)
Panel A: CAPM(SPI)							
Mean values of individual fund regressions							
All funds	−0.082	−1.105	0.759	−0.026	−0.861	0.765	0.211 (21)
Small- and mid-cap	0.291	0.906	0.633	0.430	1.349	0.647	0.088 (5)
Large-cap	−0.173	−1.595	0.790	−0.137	−1.399	0.794	0.241 (16)
Median values of individual fund regressions							
All funds	−0.134	−1.489	0.748	−0.130	−1.137	0.760	0.108
Small- and mid-cap	0.358	1.136	0.622	0.534	1.621	0.630	0.103
Large-cap	−0.157	−1.625	0.833	−0.138	−1.376	0.837	0.125
Panel B: FF							
Mean values of individual fund regressions							
All funds	−0.203	−1.609	0.818	−0.139	−1.195	0.823	0.243 (19)
Small- and mid-cap	−0.094	−0.504	0.869	−0.038	−0.225	0.877	0.096 (5)
Large-cap	−0.230	−1.878	0.805	−0.163	−1.431	0.810	0.279 (14)
Median values of individual fund regressions							
All funds	−0.134	−1.772	0.850	−0.120	−1.217	0.860	0.193
Small- and mid-cap	−0.072	−0.403	0.891	−0.024	−0.145	0.893	0.017
Large-cap	−0.171	−2.065	0.837	−0.127	−1.403	0.850	0.262
Panel C: 9SPI							
Mean values of individual fund regressions							
All funds	−0.150	−1.173	0.810	−0.160	−1.144	0.842	0.017 (43)
Small- and mid-cap	0.084	0.193	0.815	0.011	−0.142	0.858	0.002 (9)
Large-cap	−0.207	−1.505	0.808	−0.201	−1.388	0.838	0.021 (34)
Median values of individual fund regressions							
All funds	−0.139	−1.348	0.818	−0.150	−1.294	0.857	0.000
Small- and mid-cap	0.172	0.731	0.816	0.054	0.226	0.860	0.000
Large-cap	−0.171	−1.766	0.819	−0.170	−1.597	0.855	0.001

The unconditional regression equation for the CAPM(SPI) is specified as follows: $r_{P,t+1} = \alpha_P + \beta_P r_{B,t+1} + \varepsilon_{P,t+1}$, where $r_{P,t+1}$ is the excess return of a fund and $r_{B,t+1}$ is the excess return on the benchmark portfolio. The benchmark index for all funds in the sample is the SPI on a total return basis. In the FF three-factor model we use the style portfolios SMB (Small Minus Big) and HML (High Minus Low) as our two additional factors. In the 9SPI model the SPI benchmark portfolio is replaced with nine SPI sectors as model factors. The conditional regression equation, where all predetermined information variables are used simultaneously, for the CAPM(SPI) is specified as follows: $r_{P,t+1} = \alpha_P + \beta_{1P} r_{B,t+1} + \beta'_{2P}(z_t r_{B,t+1}) + \varepsilon_{P,t+1}$, where the vector z_t is added to the model as a set of predetermined (lagged) information variables. These include the lagged log dividend yield, the lagged short-term interest rate, and the lagged term spread. As in the unconditional case, the model can easily be extended to multifactor models like the FF three-factor model and the 9SPI model. R^2 are the adjusted R-squares of the regression, and p-val(F) denotes the probability value of the Wald F-test for the null hypothesis that all coefficients on the interaction terms including the information variables are simultaneously equal to zero. The figures in brackets denote the number of funds where the null hypothesis is rejected to the 10 % significance level. All t-statistics are corrected for the effects of heteroskedasticity using the White (1980) methodology. Data run from January 2000 to August 2011 with monthly frequency (140 observations)

of returns: the Fama and French (FF) three-factor model, and a multi-beta model with nine SPI-sector returns as benchmarks (9SPI).

In the unconditional CAPM(SPI), the mean abnormal performance is -0.082% per month, or about -98 basis points per year. The small- and mid-cap funds clearly outperform the general funds. The three-factor FF model considerably deteriorates the mean abnormal performance to -0.203% per month, or about -241 basis points per year. The inclusion of the SMB style portfolio makes the difference between small-/mid-cap and large-cap funds less pronounced, which is also reflected in a relatively high R-squared value. The results of the multi-beta 9SPI-model shows that disaggregating the broad market index into nine sectors decreases the performance to -0.150% per month, or about -179 basis points per year.

Augmenting the tests for conditioning information has mixed effects: Performance is positively affected for the CAPM(SPI) and the FF three-factor model, where the conditional alphas are -0.026 and -0.139% per month, or -31 and -166 basis points per year, respectively. This finding is consistent with the results reported in Ferson and Schadt (1996) for US mutual funds and Silva et al. (2003) for European bond funds, but they are in contrast to the results in Dahlquist et al. (2000) for Swedish funds and Bessler et al. (2009) for German mutual funds. A possible interpretation of this finding is provided by Ferson and Schadt (1996), based on their observation that the term $B'_P \text{Cov}\{z_t, r_{B,t+1}\}$ is negative for the average fund. We compute a value of -0.0005572 for the CAPM(SPI), implying that fund managers decrease their market beta when expected market returns are high conditional on public information, and vice versa. Ferson and Schadt's explanation relies on the cyclicality of funds' inflows and outflows. If favorable business conditions drive expected market returns up, more new money flows into the funds. If fund managers need time to allocate this money, the funds temporarily increase their cash holdings, implying lower fund betas.

In contrast, the 9SPI sector model exhibits a moderate decrease of average performance after conditioning for public information: the conditional alpha is -0.160% per month, or -190 basis points per year. Thus, conditioning information does not unambiguously make fund performance look better, as suggested by the body of empirical literature. It moreover illustrates the well-known fact that performance measures are not immune with respect to the repackaging of the assets in the benchmark universe (here, SPI stocks), which is a particularly important point to recognize in a conditional setting.

The last column of Table 2 reports the average p values of the Wald F-tests for the null hypothesis that all interaction terms, i.e., the coefficients of the products of the factors and the predetermined information variables, are jointly equal to zero. The figures in brackets show the number of funds for which the null hypothesis is rejected at the 10 % significance level. For example, for the CAPM(SPI) the null hypothesis is rejected for 21 funds, i.e., almost half of the sample.[8]

---

[8] As a robustness test, we substitute the SPI by the MSCI Switzerland TR Index as benchmark portfolio in the CAPM and FF model. The results are very similar. The Spearman rank correlation coefficients between the alphas are 0.9842 and 0.9981 for the unconditional and conditional models, respectively.

The returns of our fund sample are net of costs, i.e., all management fees are deducted from the returns. In order to evaluate the manager's performance *before* fees, we have to add back the total expense ratio to our alphas; the average across all funds in our sample is 1.492 % per year. Taking the conditional annualized alpha of the CAPM(SPI) model of -0.312 %, the negative performance turns into an overperformance of 1.180 % before fees. In sharp contrast to our results, Bessler et al. (2009) find an underperformance before fees of roughly 32 basis points per year for their sample of German mutual funds. However, based on the three-factor FF model, the performance remains negative before fees.

4.1.2 Cross-sectional distribution of mutual fund alphas

Table 3 shows the cross-sectional distribution of t-statistics of the alphas estimated from the beta models discussed before. Following Ferson and Qian (2004), we display the distribution across intervals separated by quantiles of the standard normal distribution. Accordingly, column (1) shows the distribution of t-statistics that would be expected under the null hypothesis. Table 3 also presents the maximum and minimum t-statistics that are, in addition, used to calculate the Bonferroni p values to test the null hypothesis that all alphas are jointly equal to zero against the alternative of at least one positive [Bonferroni p-value (+)] or negative [Bonferroni p-value (−)] alpha. To compute the Bonferroni p-value, one simply has to multiply the maximum or minimum one-tailed p-value from the distribution of t-statistics across all funds by the number of funds.

Table 3 shows that the distribution of t-statistics is skewed to the left compared to the distribution under the null hypothesis. This finding is more pronounced compared to previous studies. For each model, we find at least five highly significant negative alphas with t-statistics below -2.326. The classical beta models [CAPM(SPI), FF] generate more extreme negative values compared to the 9SPI sector model. In contrast, for positive alphas, the Bonferroni p-value never rejects the null hypothesis that all alphas are zero against the alternative that at least one alpha is positive. The impact of conditioning information is equivalent to the results reported in the section before. The distribution of the alphas from the classical beta models [CAPM(SPI), FF] makes the performance of the funds look better, the number of negative alphas which are significant has decreased, and the conditional CAPM(SPI) even produces four positive alphas with significant t-statistics. Again, for the 9SPI sector model, conditioning information does not seem to shift the distribution of alphas in the same way.

4.2 Performance estimates using SDF models

4.2.1 Calibration and performance of SDF models

In a first step, we calibrate the SDF models based on the discount factor specifications discussed in the Appendix: a single-factor CAPM, a three-factor FF model, plus the primitive-efficient and Bakshi–Chen specifications. The returns of nine SPI sectors (plus the risk-free rate) are used as primitive asset returns to calibrate the models. To

Table 3 Cross-sectional distribution of t-statistics of the alphas

	CAPM(SPI)			FF		9SPI	
	(1) Null	(2) Uncond.	(3) Cond.	(4) Uncond.	(5) Cond.	(6) Uncond.	(7) Cond.
Minimum t-statistic		−4.108	−4.078	−4.078	−3.872	−3.257	−3.056
Bonferroni p-value (−)		0.000	0.000	0.000	0.005	0.051	0.106
$t < -2.326$	0.23	7	5	10	5	10	8
$-2.326 < t < -1.960$	0.92	6	5	9	3	4	4
$-1.960 < t < -1.645$	1.15	5	5	5	4	6	6
$-1.645 < t < 0.000$	20.70	17	20	18	28	15	19
$0.000 < t < 1.645$	20.70	11	7	4	6	11	9
$1.645 < t < 1.960$	1.15	0	3	0	0	0	0
$1.960 < t < 2.326$	0.92	0	1	0	0	0	0
$2.326 < t$	0.23	0	0	0	0	0	0
Bonferroni p-value (+)		4.692	1.610	25.967	18.055	7.420	6.205
Maximum t-statistic		1.636	2.110	0.576	0.855	1.401	1.495

This table shows the cross-sectional distribution of t-statistics of the estimated alphas using different model specifications. The individual t-statistics are divided into different fractions according to the standard critical values of the normal distribution. Column (1) shows the distribution that would be expected under the null hypothesis. For the CAPM(SPI) in column (2), the unconditional alphas are the intercepts in regressions of fund excess returns on the excess returns of the SPI Total Return Index. The conditional alphas in column (3) are estimated in regressions of fund excess returns on the excess returns of the benchmark index and the product of the benchmark index and the vector of information variables. In case of the FF three-factor model in columns (4) and (5), we implement the style portfolios SMB and HML as two additional factors in the regressions. In columns (6) and (7) we estimate the alphas using nine sectors of the SPI index as our model-factors instead of the SPI itself. The Bonferroni p-value is the maximum or minimum one-tailed p-value from the distribution of t-statistics across all funds multiplied by the number of funds. The entries in the middle-block of the table denote the number of funds for which the t-statistics for alpha falls within the range of critical values of a standard normal distribution. In the conditional models, we use the log dividend yield, the short-term interest rate, and the term spread as predetermined (lagged) information variables. Data run from January 2000 to August 2011 (140 monthly observations)

make the performance results of the CAPM directly comparable with the corresponding estimates in the beta framework, we moreover add a CAPM-calibration where only the SPI market index (plus the risk-free security) is used as primitive asset; this model is denoted by SDF_CAPM(SPI) in contrast to SDF_CAPM(9SPI) which relies on the standard specification of the primitive assets (SPI sectors).

Table 4 shows the results of the specification tests. The first row displays the results using a constant discount factor which can be motivated by risk-neutrality. Furthermore, the discount factor is equal to the inverse of the sample mean of the Swiss 3-month LIBOR which is 0.1157 % per month over the sample period. Accordingly, the inverse is 0.9988 [=1/(1 + 0.001157)], which is very close to the value of 0.9971 reported in the first line of Table 4. Panels A and B present the results of the unconditional and conditional models, respectively. In addition to the monthly mean of the SDF, the table also displays the monthly standard deviation, the first-order autocorrelation and the corresponding p-value, the minimum and maximum value of the SDF,

Table 4 SDF models

	$E(m)$	$SD(m)$	$\rho_1(m)$	(p-val)	$\min(m)$	$\max(m)$	$\#(m<0)$	HJun	HJcon	KRun	KRcon
Constant discount factor	0.9971	–	–	–	0.9971	0.9971	0	0.2354	0.4025	–	–
Panel A: unconditional SDF models											
SDF_CAPM(9SPI)	0.9981	0.0080	0.204	(0.015)	0.9770	1.0237	0	0.2364	0.4072	0.2400	0.3884
SDF_FF	0.9977	0.1584	0.199	(0.017)	0.5285	1.4377	0	0.2216	0.3964	0.2159	0.3954
SDF_primitive-efficient	0.9989	0.2679	0.020	(0.811)	−0.2231	1.6701	1	0.0000	0.0000	0.0000	0.0000
SDF_Bakshi–Chen	0.9988	0.2813	0.131	(0.118)	0.2585	1.8697	0	0.0000	0.0000	–	–
SDF_CAPM(SPI)	0.9988	0.0135	0.204	(0.015)	0.9634	1.0418	0	0.0000	0.0000	–	–
Panel B: conditional SDF models											
SDF_CAPM(9SPI)	0.9901	0.2535	0.167	(0.046)	−0.1485	1.6394	1	0.3812	0.3796	0.3082	0.3197
SDF_FF	0.9914	0.3786	0.187	(0.025)	−0.3143	2.0284	1	0.3620	0.3621	0.2758	0.2844
SDF_primitive-efficient	0.9988	0.6791	0.263	(0.002)	−1.1811	2.9165	8	0.0000	0.0000	0.0000	0.0000
SDF_Bakshi–Chen	0.9988	0.7667	0.209	(0.013)	0.0071	3.8656	0	0.0000	0.0000	–	–
SDF_CAPM(SPI)	0.9988	0.2534	0.129	(0.123)	−0.0675	1.6932	1	0.0000	0.0000	–	–

This table shows the results from specification tests for different SDF models using the following system of equations: $u1_t = (m_{t+1}R_{t+1} - 1) \otimes Z_t$, where $u1_t$ is the vector of pricing errors relating to the gross returns, R_{t+1}, of the N primitive assets. m_{t+1} denotes the SDF, and Z_t is the vector of predetermined information variables. The parameters of the SDFs are estimated using Hansen (1982) GMM, minimizing a quadratic form of the pricing errors. We correct the standard errors for the effects of heteroskedasticity using the White (1980) methodology. $E(m)$ is the sample mean, and $SD(m)$ is the sample standard deviation. $\rho_1(m)$ and p-val(Q) show the first-order autocorrelation of the estimated SDF and the corresponding p-value of the Q-statistics, respectively. HJun and HJcon denote the classical Hansen and Jagannathan (1997) distances (HJD) as a measure of misspecification. KRun and KRcon denote the so-called modified HJD measures of misspecification following the procedure of Kan and Robotti (2008). HJun and KRun are the unconditional measures, and HJcon and KRcon are the conditional measures, which use the lagged information variables in addition to the returns. As primitive assets, we use the nine sectors of the SPI index according to the Datastream classification along with the risk-free security. In case of the SDF_CAPM(SPI), only the SPI index and the risk-free security are used as primitive assets. In linear SDF models [SDF_CAPM(9SPI), SDF_FF, and SDF_CAPM(SPI)] the SPI index on a Total Return Basis is applied as benchmark portfolio. The set of predetermined information variables contains the log dividend yield, the short-term interest rate, and the term spread. The sample period is from January 2000 to August 2011 (140 monthly observations)

the number of negative monthly SDFs, and the results from two different methods to measure the unconditional and conditional Hansen and Jagannathan (1997) distance (HJD).

The means of the SDFs are below 1.00 in all models, as implied by a positive interest rate, indicating a minimum reliability of the SDF estimates. Nevertheless, minimum and maximum values reveal that the intervals are large across all models, except for the unconditional CAPM. To illustrate this more in detail, Fig. 1 shows the time-variation of the fitted values from each model, except for the SDF_CAPM(SPI) which is nearly identical to the SDF_CAPM(9SPI) and is therefore skipped. Similar to Farnsworth et al. (2002) and Bessler et al. (2009), we find that the standard deviation increases with the complexity of the models, i.e., by adding factors or by moving from unconditional to conditional models. This is consistent with Hansen and Jagannathan (1991) who show that the minimum variance of an SDF increases, i.e., the HJ bound becomes more tight if the number of underlying assets increases (in absolute terms, or due to scaled factors or returns). Note that the value of a highly volatile SDF is ambiguous. While a more volatile SDF is necessary to explain the historical equity premium (Mehra and Prescott 1985), a larger variance of an SDF also implies lower power to detect abnormal performance.

The last four columns of Table 4 present the results of two different methods to estimate the Hansen and Jagannathan (1997) distance. The HJD is a useful diagnostic to test model adequacy and can be interpreted as the average pricing error across a group of primitive assets for a given ("candidate") SDF specification. It measures the distance between the candidate SDF and the true SDF which would correctly price the primitive assets. We report both the unconditional and conditional measure for each method.

Model diagnostics based on HJD. Following Lewellen et al. (2010), we first derive the traditional HJD labeled as HJun and HJcon in Table 4. Again, the pricing errors are defined as $u1 = E(m \cdot R - 1)$,[9] where m is a candidate SDF and R denotes a $N \times 1$ vector of gross returns of the primitive assets including the risk-free security. Hansen and Jagannathan (1997) show that the distance between a candidate SDF, m, and the corresponding true SDF, m^*, can be written as $D = min_m E\{(m^* - m)^2\}$. Term D is equivalent to $u1'H^{-1}u1$, where $H = E(RR')$ denotes the second moment matrix of gross returns of the primitive assets. To estimate the conditional measure of the HJD, we follow Farnsworth et al. (2002) and define $r = R \otimes \{Z./E(Z)\}$, where "./" denotes element by element division. Finally, replacing R by r in $u1'H^{-1}u1$ yields the following population value of the traditional HJD in quadratic form: $E(mr - 1)' \{E(rr')^{-1}\} E(mr - 1)$. Table 4 reports the corresponding sample moments under HJun and HJcon. If the vector of information variables, Z, is reduced to a constant, we simply estimate the unconditional measure.

The reported HJun estimates reveal that all unconditional models, except for the SDF_CAPM(9SPI), have smaller pricing errors compared to the model with a constant discount factor. However, the pricing error in the SDF_CAPM(9SPI) model is also only slightly higher. The pricing errors of the primitive-efficient and the Bakshi–Chen

---

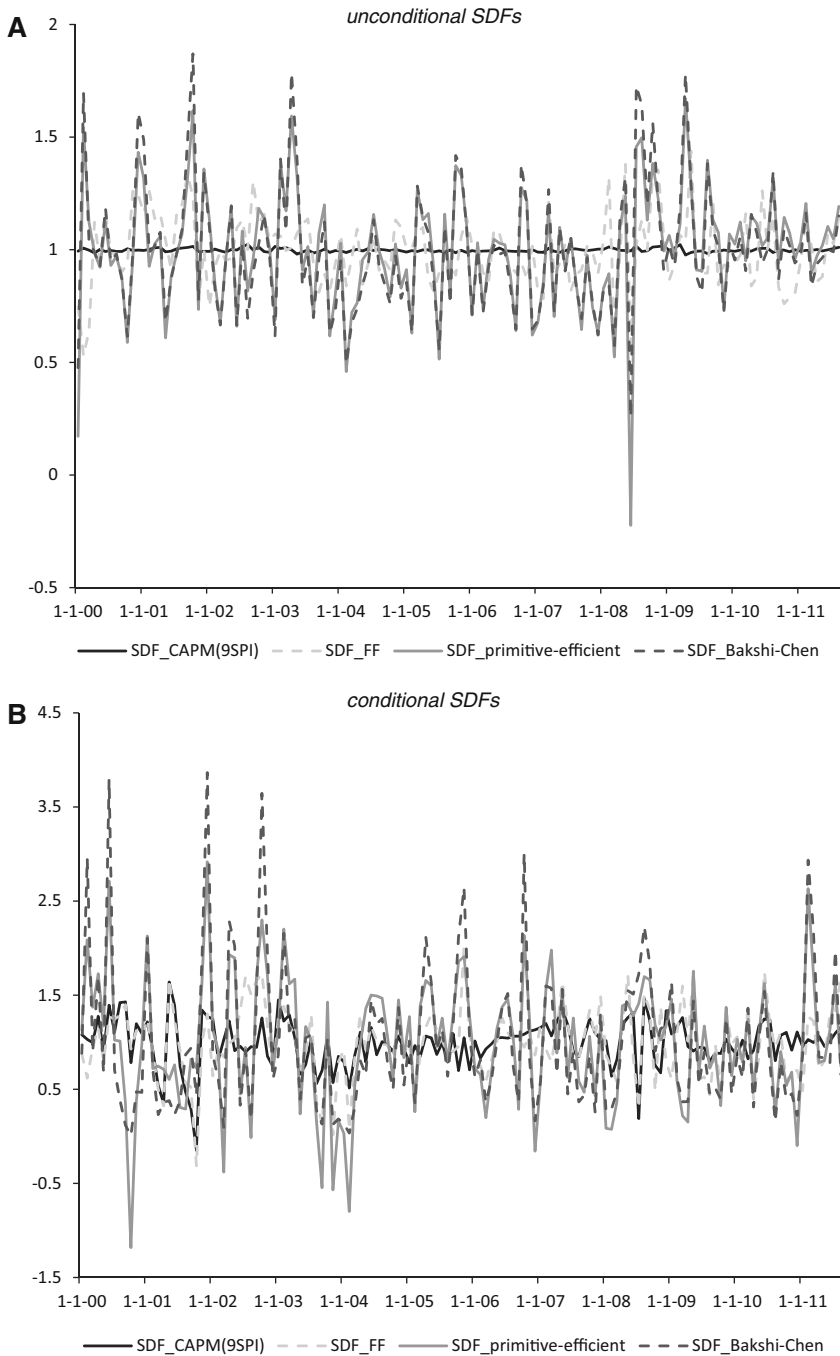[9] To simplify notation, we skip time indices in this section.

Fig. 1 Time-variation of unconditional and conditional SDFs. This figure compares the time-variation of the fitted values of each SDF using all models, except for the CAPM(SPI). The sample period is from January 2000 to August 2011 (140 monthly observations)

models are zero by construction. Thus, models with a time-varying SDF alleviate resulting pricing errors. Generally, conditional models exhibit larger unconditional HJDs compared to their unconditional counterparts. This difference arises due to the fact that conditional models price dynamic strategies implied by the lagged information variables, and thereby sacrifice some accuracy on the returns of primitive assets. However, because conditional models are estimated to price both the primitive assets and the dynamic strategies, they should produce smaller pricing errors according to the HJcon measure. This is the case for the SDF_CAPM(9SPI) single-factor and the SDF_FF three-factor model. Interestingly, we also find that, according to the HJcon measure, the conditional three-factor model performs better than the conditional SDF_CAPM(9SPI) and the unconditional three-factor model. Therefore, unlike in previous studies (e.g., Ferson and Harvey 1999; Farnsworth et al. 2002), the Fama and French three-factor model seems to do a good job pricing dynamic strategies.

Model diagnostics based on the modified HJD For comparative purposes, we also estimate a modified HJD, denoted by "KRun" and "KRcon" in Table 4, as proposed by Kan and Robotti (2008). This measure relies on excess returns instead of gross returns when the candidate SDF is specified. Kan and Robotti show that the traditional SDF representation (8) is problematic in this setting, because the pricing errors as well as the relative rankings between competing models are sensitive to affine transformations of the factors. The authors show that setting the constant equal to one and using *de-meaned* factors in the SDF specification

$$m_{t+1}^{\text{LF}} = 1 + \boldsymbol{b}'(\boldsymbol{F}_{t+1} - E\{\boldsymbol{F}_{t+1}\}) \tag{1}$$

results in robust misspecification measures that are invariant to affine transformations. They moreover show that this specification implies a modification in computing the HJD, namely, that the inverse variance–covariance matrix of (scaled) excess returns $\boldsymbol{V}^{-1}$ must be used as weighting matrix in $\boldsymbol{u}\boldsymbol{1}'\boldsymbol{V}^{-1}\boldsymbol{u}\boldsymbol{1}$ in place of the second moment matrix $\boldsymbol{H}$.

Equipped with excess returns, de-meaned factors, and the variance–covariance matrix as weighting matrix, we are able to estimate the modified HJD as an unconditional and conditional measure. Although there is a simple analytical relation to the original HJD via the riskless rate, it is interesting to see how close the measures are when empirically implemented in small samples. We estimate the modified HJD for the linear single-factor SDF_CAPM(9SPI), the FF three-factor model, and the primitive-efficient SDFs model augmented by a constant. Obviously, the alternative HJD cannot be applied to the Bakshi–Chen model where the SDF is a function of logarithmic factor returns which have now zero mean. The SDF_CAPM (SPI) is not estimated as well because it prices only the benchmark portfolio and the risk-free security, and the latter drops out in the modified setting.

Addressing the traditional HJD first, the empirical results in Table 4 indicate that unconditional models perform better according to the unconditional measure HJun (0.381 vs. 0.236, 0.362 vs. 0.222), and conditional models perform better according to the conditional measure HJcon (0.380 vs. 0.407, 0.362 vs. 0.396). However, the

absolute *size* of the pricing errors of the conditional SDF models is virtually identical under the conditional and unconditional measures (0.380 vs. 0.381, 0.362 vs. 0.362).[10]

The results derived from the modified HJD (denoted by KRuc and KRcon) are very similar and do not lead to different conclusions. Comparing KRun with HJun (Panel A) reveals that the unconditional measures are almost identical to the classical HJD (0.240 vs. 0.236, 0.216 vs. 0.222). However, when pricing dynamic strategies (Panel B) there is a clear discrepancy between the two models; the modified HJD produces smaller pricing errors for the conditional measure KRcon (0.320 vs. 0.380, 0.284 vs. 0.362) as well as for the unconditional measure KRuc.

Model diagnostics based on dynamic performance The HJD from above summarizes the relative fit of a specific SDF model to the cross-section of primitive assets. However, the HJD provides no information about the economic magnitude of pricing errors in the underlying primitive assets. Therefore, following Farnsworth et al. (2002), we also investigate the dynamic performance of the different SDF models in Table 5, in their ability to price the underlying primitive assets. Recall that an SDF model implies that the term $m_{t+1}R_{t+1}$ should not be predictable using public information, Z_t, even though R_{t+1} might be predictable based on Z_t. To check this, we compute the pricing errors, $u1 = m_{t+1}R_{t+1} - 1$, of individual primitive assets and regress these errors on the vector of lagged instruments, Z_t. Using the "correct" m_{t+1} implies that these predicted pricing errors should be close to zero. Accordingly, the reported standard deviations of these fitted values of the regressions should be small if the model captures the predictable variation in expected returns (e.g., Ferson and Harvey 1993; Ferson and Korajczyk 1995). In contrast to the HJD, the standard deviations from Table 5 do not penalize a model that gets the average return wrong. Thus, the standard deviation can be interpreted as a pure measure of the ability of a model to explain the variation of expected returns.

The first row in Table 5 shows the standard deviations of the fitted pricing errors when assuming a constant discount factor. This serves as a benchmark case because of its inability to explain any of the predictability, which can be shown, according to Farnsworth et al. (2002), by the following example using the returns of sector 1. As reported in the first line of Table 5, the monthly standard deviation of the fitted pricing error for sector 1 is 0.990%. In contrast, Table 1 shows that the standard deviation of the monthly raw return on sector 1 is 4.984%. Because the R-square measures the percentage of variance explained, we obtain an R-square of roughly 0.0395 [$=(0.00990/0.04984)^2$] when we regress the returns on sector 1 upon the vector of lagged information variables. The value of 0.0395 is nearly identical to the unadjusted R-square reported in Table 1. The same can be shown for the remaining primitive assets.

Panel A in Table 5 shows the results of unconditional SDF models. In contrast to the results of Farnsworth et al. (2002) and Bessler et al. (2009), we find that the unconditional SDF_CAPM(9SPI) model performs slightly better, on average across all primitive assets, than the constant discount factor model. Also, the three-factor SDF_FF model exhibits relatively small standard deviations and, therefore, performs

---

[10] Remember that the conditional SDF refers to the scaling of *risk factors*, while the un/conditional HJD refers to the scaling of *primitive returns*.

Table 5 Dynamic performance of SDF models

	Sector 1	Sector 2	Sector 3	Sector 4	Sector 5	Sector 6	Sector 7	Sector 8	Sector 9	Risk-free	SPI
Constant discount factor (%)	0.990	2.240	1.112	0.574	1.919	0.384	1.436	1.295	3.201	0.093	–
Panel A: unconditional SDF models											
SDF_CAPM(9SPI) (%)	0.830	2.089	0.947	0.441	1.765	0.272	1.373	1.131	3.047	0.256	–
SDF_FF (%)	1.967	0.664	1.792	2.540	1.069	2.911	2.471	1.676	0.690	3.159	–
SDF_primitive-efficient (%)	5.837	6.608	6.100	5.849	6.596	6.170	4.367	6.278	6.503	5.768	–
SDF_Bakshi–Chen (%)	6.477	6.133	6.522	6.975	6.316	7.225	5.510	6.743	5.527	7.249	–
SDF_CAPM(SPI) (%)	–	–	–	–	–	–	–	–	–	0.370	0.662
Panel B: conditional SDF models											
SDF_CAPM(9SPI) (%)	2.521	2.418	2.481	2.353	2.440	2.505	2.641	2.353	1.870	2.585	–
SDF_FF (%)	2.275	2.256	2.632	2.301	2.384	2.417	2.261	2.405	2.411	2.483	–
SDF_primitive-efficient (%)	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	–
SDF_CAPM(SPI) (%)	–	–	–	–	–	–	–	–	–	0.000	0.000

This table shows the standard deviations of the fitted pricing errors from the following system of equations: $u1_t = (m_{t+1} R_{t+1} - 1) \bigotimes Z_t$, where $u1_t$ denotes the vector of pricing errors relating to the gross returns, R_{t+1}, of the N primitive assets. The fitted pricing errors are the fitted values of regressions from $u1_t$ on Z_t using different SDF models. The table shows the corresponding standard deviations for all primitive assets. Hansen's (1982) GMM is used to estimate the parameters of the model, minimizing a quadratic form of the pricing errors. The primitive assets are the nine SPI sectors according to the Datastream classification and the risk-free security. In case of the CAPM(SPI), only the SPI index and the risk-free security are used as primitive assets. In linear SDF models [SDF_CAPM(9SPI), SDF_FF, and SDF_CAPM(SPI)] the SPI index on a Total Return Basis is applied as benchmark portfolio. The set of predetermined information variables contains the log dividend yield, the short-term interest rate, and the term spread. The sample period is from January 2000 to August 2011 (140 monthly observations)

well in explaining the predictable variation in returns. The remaining two comparable models, i.e., the primitive-efficient model and the Bakshi–Chen model, do poorly in capturing the time-variation of expected returns. The standard deviations of the fitted pricing errors are large, in many cases even larger than the standard deviation of the sectors' raw returns, implying that the product $m_{t+1}R_{t+1} - 1$ produces larger regression coefficients on Z_t than R_{t+1} does.

Panel B in Table 5 presents the results of conditional models. We observe that they perform better on average compared to their unconditional counterparts. The primitive-efficient model and the Bakshi–Chen model now perform extremely well producing standard deviations of the fitted pricing errors close to zero. This is not surprising because conditional models are constructed to make the expected product of the pricing errors and the lagged information variables equal to zero. However, the SDF_CAPM(9SPI) and three-factor SDF_FF model perform worse in a conditional setting—with the SDF_FF model performing only slightly worse than in the unconditional setting, which confirms our results from the HJD measures. Therefore, the Fama–French SDF-specification does not deteriorate substantially in the conditional version, implying a good performance in pricing dynamic strategies; this finding is in contrast to previous studies, e.g., Ferson and Harvey (1999) and Farnsworth et al. (2002).

4.2.2 Mutual fund performance: SDF models

Based on the different specifications of the discount factor, Table 6 presents some cross-sectional characteristics of the estimated SDF-alphas: the mean and median, and the mean of the top and bottom 3, respectively 10, funds. It also shows the Bonferroni p values for the null hypothesis that all alphas are jointly equal to zero against the alternative of at least one positive [Bonferroni p-value (+)] or one negative [Bonferroni p-value (−)] alpha.

The unconditional results in Panel A reveal that the average performance based on the SDF_CAPM(9SPI) model is −0.190 % per month, or −226 basis points per year. The total expense ratio on average of all funds is 1.492 % from 2000 to 2011, indicating that the average fund clearly underperforms even before fees. This observation is even more pronounced when the SDF_FF three-factor model is estimated, whereas it is less pronounced using the primitive-efficient model and the Bakshi–Chen model. In contrast to the results from Farnsworth et al. (2002) and Bessler et al. (2009), we generally find a positively skewed distribution of individual fund alphas, except for the three-factor model.

The results of the conditional models are presented in Panel B. The average and median SDF-alphas estimated from the SDF_CAPM(9SPI) model are substantially lower than in the unconditional setting: the mean of −0.258 % per month implies an underperformance of 305 basis points per year. Interestingly, the deterioration of average performance is not reflected in the top and bottom performing funds. The SDF_FF three-factor model, the primitive-efficient model, and the Bakshi–Chen model imply that introducing information variables has almost no effect on the performance of the average fund. The corresponding distributions of SDF-alphas shift only slightly to the left. Interestingly, the single-factor SDF_CAPM(SPI) exhibits a positive effect

Table 6 Performance of funds using SDF models

	Bonferroni p-value (−)	Bottom 3 (%)	Bottom 10 (%)	Mean (%)	Median (%)	Top 10 (%)	Top 3 (%)	Bonferroni p-value (+)
Panel A: unconditional SDF models								
SDF_CAPM(9SPI)	0.005	−0.586	−0.497	−0.190	−0.214	0.163	0.298	11.022
SDF_FF	0.000	−0.648	−0.536	−0.234	−0.215	0.038	0.145	25.438
SDF_primitive-efficient	0.014	−0.558	−0.476	−0.150	−0.139	0.187	0.330	8.864
SDF_Bakshi–Chen	0.014	−0.568	−0.481	−0.159	−0.145	0.163	0.315	10.115
SDF_CAPM(SPI)	0.000	−0.433	−0.366	−0.082	−0.134	0.353	0.482	9.499
Panel B: conditional SDF models								
SDF_CAPM(9SPI)	0.000	−0.574	−0.513	−0.258	−0.329	0.180	0.335	0.423
SDF_FF	0.000	−0.543	−0.472	−0.237	−0.227	0.024	0.176	7.986
SDF_primitive-efficient	0.023	−0.481	−0.425	−0.155	−0.136	0.096	0.268	7.811
SDF_Bakshi–Chen	0.000	−0.589	−0.505	−0.166	−0.168	0.109	0.224	0.506
SDF_CAPM(SPI)	0.000	−0.356	−0.283	−0.013	−0.115	0.520	0.663	1.030

This table shows the distribution of mutual fund SDF-alphas for different SDF models using the following system of equations: $u1_t = (m_{t+1} R_{P,t+1} - 1) \otimes Z_t$, $u2_t = \alpha_P - m_{t+1} R_{P,t+1} + 1$, where $u1_t$ denotes the vector of pricing errors relating to the gross returns, R_{t+1}, of the N primitive assets. m_{t+1} denotes the SDF, and Z_t is the vector of predetermined information variables. α_P is the SDF-alpha and depends on the model specification. Hansen's (1982) GMM is used to estimate the parameters of the model, minimizing a quadratic form of the pricing errors. We correct the standard errors for the effects of heteroskedasticity using the White (1980) methodology. The primitive assets are the nine SPI sectors according to the Datastream classification and the risk-free security. In case of the CAPM(SPI), only the SPI index and the risk-free security are used as primitive assets. In linear SDF models [SDF_CAPM(9SPI), SDF_FF, and SDF_CAPM(SPI)] the SPI index on a Total Return Basis is applied as benchmark portfolio. The set of predetermined information variables contains the log dividend yield, the short-term interest rate, and the term spread. The Bonferroni p-value is the maximum or minimum one-tailed p-value from the distribution of t-statistics across all funds multiplied by the number of funds. The sample period is from January 2000 to August 2011 (140 monthly observations)

from conditioning information, confirming our findings from the CAPM(SPI) in the beta framework. The next sub-section addresses this comparison in more detail. Again, the distribution of fund alphas is in general positively skewed with the Bakshi–Chen model as the only exception. Finally, the Bonferroni p values clearly suggest that the extreme performing fund is more likely to have a negative than a positive SDF-alpha. Both in the unconditional and conditional setup and independent of the applied model, the null hypothesis that all alphas are jointly equal to zero can always be rejected against the alternative that there is at least one significantly negative alpha. On the other hand, we cannot reject the null hypothesis against the alternative that at least one significantly positive alpha exists.

Overall, and in contrast to the beta framework, conditioning information seems to make the performance of funds look less favorably in the SDF framework. Chen and Knez (1996) suggest that this strongly negative effect of conditioning information arises through "tougher" performance yardsticks when dynamic strategies based on information variables are used. Scaling factors or returns is equivalent to augmenting the asset space by dynamic strategies (with conditional weights linearly related to the instruments). Based on Hansen and Jagannathan (1991), adding assets implies an upward shifting SDF minimum volatility bound and equivalently a maximum higher Sharpe ratio of the scaled benchmark universe, which entails stronger pricing restrictions lowering the funds' performance.

We also compute rank correlations between the SDF-alphas of all five models. The results are not shown here. We find that the relative performance of individual funds is quite robust against different specifications of the SDF. In the unconditional version, the rank correlations range between 0.999 and 0.944. However, the range is between 0.967 and 0.741 for the conditional models. The rank correlations between the unconditional and conditional version of each model vary between 0.971 for the SDF_CAPM(9SPI) and 0.835 for the Bakshi–Chen model. Compared to the findings of Bessler et al. (2009), we find even larger rank correlations in our sample of funds, indicating that the relative performance is reasonably correlated across different model specifications.

4.3 Comparison of CAPM-based performance results

In this section we compare the performance results from the beta framework with those from the SDF framework, and rely on the CAPM specifications in doing this. While, in the beta framework, the specification of the CAPM model is unambiguous once the relevant market index (benchmark portfolio) is defined, the SDF framework furthermore requires the specification of a set of primitive or benchmark assets. Interestingly, none of the previous studies addresses the impact of this choice on empirical performance measures. In our empirical tests displayed in Tables 4, 5, and 6, we have selected a common set of nine SPI sectors as primitive assets for all analyzed models. By comparing the SDF-alphas from the so-calibrated SDF_CAPM(9SPI) model with the alphas from the standard single-beta CAPM(SPI), we observe much worse performance results in the SDF setting than in the beta-pricing setup; this is particularly the case when conditioning information is introduced. The CAPM(SPI) alpha yields

an average monthly performance of −0.082 %, respectively −0.026 %, in the unconditional and conditional case. In contrast, the average monthly performance from the SDF_CAPM(9SPI) model is −0.190 %, respectively −0.258 %, in the unconditional and conditional setting. Thus, we find quite large performance discrepancies between the beta and SDF frameworks.[11] It is doubtful whether the discrepancy is of statistical nature, since Jagannathan and Wang (2002) show that the SDF methodology is asymptotically as efficient as the beta method. Also, specification tests of asset pricing models based on the two frameworks are equally powerful.

We argue that the just conducted way in comparing the performances from SDF and beta models is not appropriate; the amount of information used to calibrate the SDF model—the set of primitive assets—is superior to the information reflected in the benchmark (market) returns in the beta model. Or in statistical terms: the wider range of primitive assets which is often used to calibrate the SDF model imposes additional pricing restrictions which implies a "tougher benchmark" compared to the simple traditional CAPM relying on a single benchmark (market) portfolio. The implications for performance measurement are straightforward: a "fair" comparison requires that either

- the traditional alphas from the CAPM(SPI) model being compared to a single-factor SDF where the nine sectors are replaced by the SPI index as the only primitive asset (in addition to the risk-free security), i.e., SDF_CAPM(SPI). The results—which are not reported here—confirm that the relative performance of the individual fund is virtually identical; the rank correlations are 1.000 and 0.956 in the unconditional and conditional setting, respectively. Moreover, the performance differences are extremely small in both cases and, hence, the models perform similarly;
- or that the SDF_CAPM(9SPI) performance measures being compared to alphas estimated from a multi-beta model where the market (SPI) excess returns are replaced by those on the nine sectors, i.e., 9SPI.

Figure 2 displays the unconditional and conditional alphas for 3 models (two beta- and one SDF-model) across our cross-section of 46 funds:

 (i) the alpha from the 9SPI beta model;
 (ii) the SDF-alpha from SDF_CAPM(9SPI); and
(iii) the standard alpha from the CAPM(SPI) beta model.

Our main focus is on the "fair" comparison between (i) and (ii). The unconditional results in Panel A reveal that the individual funds perform similarly in the SDF and beta frameworks. The average fund in the 9SPI model (i) outperforms the average fund in the SDF_CAPM(9SPI) model (ii), with an average alpha of −0.150 % compared to −0.190 % per month, which is about half a percentage on an annual basis. As expected from a visual inspection of the results, the rank correlation is a high 0.985, implying that a small-level shift explains the performance discrepancies.

---

[11] However, these performance discrepancies strongly depend on the applied model, i.e., the performance differences are much less pronounced when we use, for example, the Fama and French three-factor model.
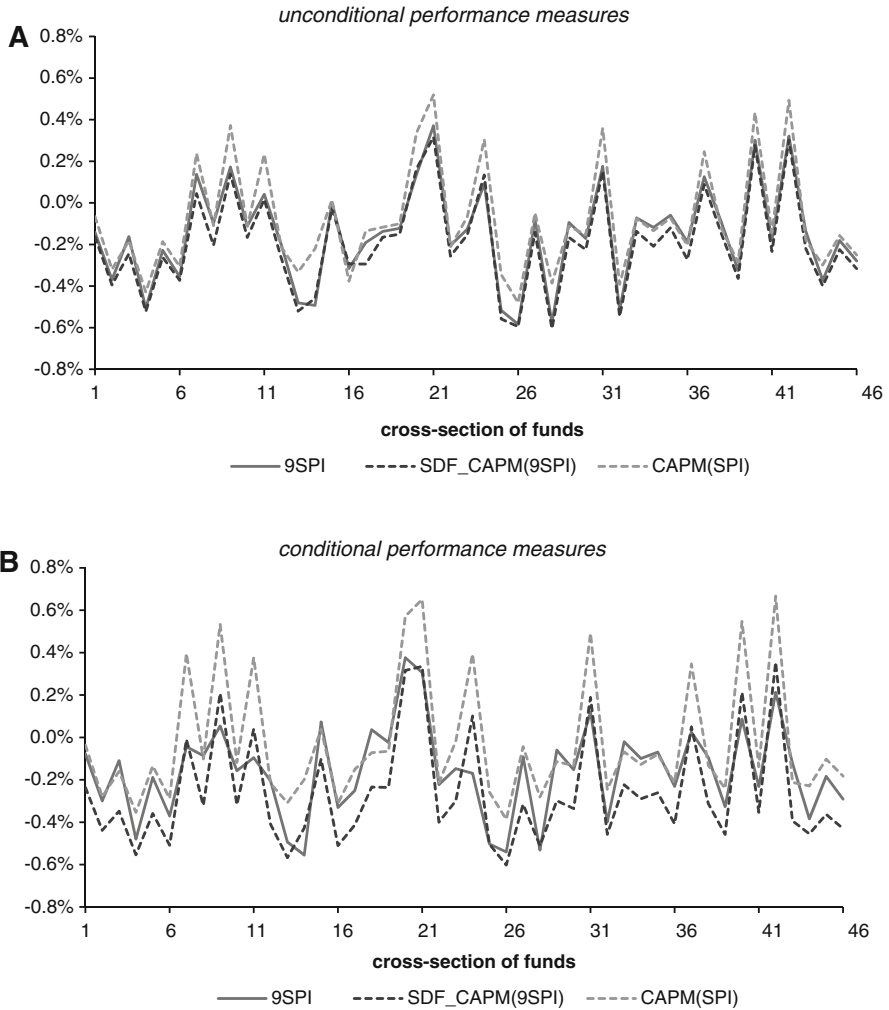
Fig. 2 Performance of Beta-pricing models (Jensen Alphas) versus SDF models (SDF-alphas). This figure shows the comparison of performance measures using two beta models [9SPI and CAPM(SPI)] and the SDF_CAPM(9SPI) from the SDF framework. We contrast each fund's alpha in the SDF_CAPM(9SPI) with its direct counterpart in the beta setup, i.e., the 9SPI model. For comparative purposes, we also show the classical Jensen's alpha estimated with the CAPM(SPI) in the beta framework. The sample period is from January 2000 to August 2011 (140 monthly observations). The units are in percentages per month

Panel B shows the comparison in the conditional setting. The performance differences are apparently quite large. The average fund in the beta setup, 9SPI, clearly outperforms the average fund in the SDF setup, SDF_CAPM(9SPI), with average alphas of -0.160 and -0.258% per month, respectively. The rank correlation is 0.937; hence, there is a notable level shift when we switch from the beta to the SDF framework. Interestingly, a closer look at the data (see Fig. 2) reveals that the level shift in favor of the beta setup tends to occur primarily in bad months, i.e., when the performance of

a fund is clearly negative. On the other hand, funds exhibiting a positive alpha seem to perform better in the SDF framework.

Second, the "inappropriate" comparison of the two dotted lines, i.e., the CAPM(SPI) and the SDF_CAPM(9SPI), reveals again that the performance strongly deteriorates when we switch from the beta to the SDF setup. Interestingly, the relative performance of the individual funds is stable, which is confirmed by the high rank correlations of 0.973 and 0.972 in the unconditional and conditional setup, respectively.

Overall, we find that the results from the SDF methodology and the beta-pricing setup are closely related if compared appropriately. Nevertheless, our findings suggest that the performance discrepancies depend on the model specification. Estimating the simplest CAPM version in the SDF framework, where the SDF is only required to price two primitive assets, SDF_CAPM(SPI), leads to almost identical results in both unconditional and conditional versions. In contrast, when a broader range of primitive assets is used, such as the nine SPI sectors (plus the risk-free security), we find relatively large performance discrepancies and, hence, an economically non-negligible shift in the level of alphas. As a consequence, performance differences arise through the selected set primitive assets in the SDF framework. A larger set of primitive assets increases the number of equations in the system, thereby imposing additional moment restrictions that affect the parameter estimates (Bekaert and Liu 2004). Therefore, our findings suggest that, despite that the nine sectors merely represent a finer segmentation of the broad SPI index, the additional restrictions caused by adding primitive assets in the SDF setup imply a tougher benchmark and thereby lower fund performance.

5 Conclusion

This paper analyzes performance measurement based on stochastic discount factors, compared to beta models traditionally used in computing funds' (Jensen) alphas. From a theoretical point of view, standard alphas suffer from several limitations which are particularly relevant in a conditional setting (e.g., Ferson 2012; Ferson and Lin 2012). These deficits can be avoided in an SDF setting, fortunately, without scarifying the intuition which might explain the popularity of the standard alphas. While the alphas derived from beta models are based on the idea that the risk exposures of a manager's portfolio can be (statically or dynamically) replicated by a benchmark portfolio and excess returns being compared, the alphas of SDF models are derived from the notion that a set of given assets—practitioners would eventually associate it with "core" portfolio holdings—is correctly priced. Unfortunately, the asset pricing restrictions imposed by the two approaches are different, in particular in a conditional setting. But how much do they differ?

Our paper addresses this issue from an empirical, and rather pragmatic, point of view. For illustrative purposes, we use monthly returns of a representative sample of Swiss mutual funds from 2000 to 2011. Our results reveal that the key for a "fair" comparison between SDF and beta models is the specification of the set of primitive assets used to calibrate the SDF function. Once this is established, the size of (absolute) performance differences considerably decreases between the two model families. However, there are sizeable performance deviations in the cross-section of

funds if conditioning information is incorporated in the tests, up to some 20 basis points per month, or about 2.3 % per year. In almost all cases, the SDF-alphas are lower than the standard (Jensen) alphas.

In absolute terms, the annualized average (median) performance of the sample of funds[12] is $-2.3\,\%\,(-2.6\,\%)$ on an unconditional basis, dropping to $-3.1\,\%\,(-4.0\,\%)$ if conditioning information is used. The figures include management fees; however, the underperformance is much larger than the average total expense ratio (TER) reported by the funds. Therefore, evaluated as a group, the funds were unable to "beat" the benchmark in an SDF-based framework. The conclusion would be less pessimistic in the traditional setting using beta models.

Appendix: Empirical models

Performance evaluation with beta-pricing models

The unconditional alpha, α_P, of a managed portfolio or fund P is estimated by a simple linear regression equation,

$$r_{P,t+1} = \alpha_P + \beta_{1,P} \cdot r_{B,t+1} + \varepsilon_{P,t+1}, \tag{2}$$

where $r_{P,t+1}$ and $r_{B,t+1}$ are excess returns of the managed portfolio and the excess return of the benchmark portfolio, respectively. The Swiss 3-month LIBOR is used as the riskless asset. $\beta_{1,P}$ denotes the unconditional beta, and $\varepsilon_{P,t+1}$ is the regression error. The intercept in equation (2) was originally proposed by Jensen (1968) as measure of abnormal performance, using a proxy for the market portfolio as benchmark. However, the model can also easily be extended to a multiple-factor model, such as the Fama and French (1993) three-factor model, in which $\beta_{1,P}$ and $r_{B,t+1}$ can be reinterpreted as vectors.

Equation (2) is miss-specified if expected returns and mangers' betas are correlated and change over time. Accordingly, to measure conditional performance, Ferson and Schadt (1996) propose a modification of the traditional Jensen's alpha in Eq. (2), and assume that the information which is publicly available can be represented by a set of information variables (subsequently called "instruments") summarized by a vector Z_t with dimension $L \times 1$, where L corresponds to the number of instruments. They assume a linear structure for the time-varying conditional beta of managed portfolios,

$$\beta_{1,P}(Z_t) = \delta_{P0} + \delta'_P z_t, \tag{3a}$$

where $z_t = Z_t - E(Z)$ is the vector of innovations of the instruments. δ_P is a vector of sensitivities with dimension equal to the dimension of Z_t, whereas the coefficient δ_{P0} is an "average beta". Replacing $\beta_{1,P}$ in Eq. (2) by its conditional counterpart from (3a) leads to an augmented regression equation for the conditional alpha, α_P:

---

[12] The subsequent figures are based on the SDF_CAPM(9SPI) specification.

$$r_{P,t+1} = \alpha_P + \beta_{1,P} \cdot r_{B,t+1} + \boldsymbol{\beta}'_{2,P}(z_t \cdot r_{B,t+1}) + \varepsilon_{P,t+1} \tag{4a}$$

with $\beta_{1,P} = \delta_{P0}$ and $\boldsymbol{\beta}_{2,P} \equiv \boldsymbol{\delta}_P$. The product of the benchmark returns and lagged instruments, also known as "scaled" or "actively managed" returns, captures the covariance between the conditional beta and the conditional expected market return, given z_t. These interaction terms are used to control for common movements in a fund's conditional beta and the conditionally expected benchmark returns. As in the unconditional setup, the conditional model can be easily extended to a multiple-factor model by rewriting Eq. (3a) as

$$\beta_{1,P}(\mathbf{Z}_t) = \delta_{P0} + \boldsymbol{\Delta}'_P \mathbf{Z}_t, \tag{3b}$$

where The the vector $\boldsymbol{\delta}_P$ is replaced by an $L \times K$ matrix $\boldsymbol{\Delta}_P$ summarizing the sensitivities of the K factor betas to the L instruments. The augmented regression equation is then

$$r_{P,t+1} = \alpha_P + \boldsymbol{\beta}_{1,P} \cdot \boldsymbol{r}_{B,t+1} + \boldsymbol{\beta}'_{2,P}(z_t \otimes \boldsymbol{r}_{B,t+1}) + \varepsilon_{P,t+1}, \tag{4b}$$

where $\boldsymbol{\beta}_{1,P} = \boldsymbol{\delta}_{P0}$ and $\boldsymbol{\beta}_{2,P} \equiv \text{vec}(\boldsymbol{\Delta}_P)$ which is the vectorized matrix of sensitivities; $\otimes$ denotes the Kronecker product between the instruments and benchmark returns, the multifactor extension of scaled benchmark returns. Hence, with L instruments and K factors the number of parameters to be estimated equals $(L+1)K+1$ (including the constants).

Performance evaluation with SDF models

The SDF, m_{t+1}, is a random variable which relates the future value of securities to their current value or price. Standardized by the current price, the basic valuation equation in a conditional setting is

$$E(m_{t+1} \cdot R_{t+1}|Z_t) = 1, \tag{5a}$$

where $\boldsymbol{R}_{t+1}$ are gross returns (the end-of period value divided by the current price). It is assumed that the deflator prices a pre-specified set of N basis (or primitive) assets without error,

$$E(m_{t+1} \cdot \boldsymbol{R}_{t+1} - \mathbf{1}|Z_t) = 0, \tag{5b}$$

R_{t+1} is the vector of gross returns of the N primitive assets including the risk-free security;[13] $\mathbf{1}$ and $\mathbf{0}$ are the unit and zero vectors with dimension N. The cross-products $m_{t+1} \cdot \boldsymbol{R}_{t+1}$ may be viewed as risk-adjusted gross returns on the benchmark assets.

---

[13] Notice that $E(m_{t+1}) = 1/E(R_{f,t+1})$, where $R_{f,t+1}$ is the gross return on the risk-free security. Including the risk-free asset in the set of primitive assets is suggested by Dahlquist and Söderlind (1999) and Farnsworth et al. (2002). They show that the incorporation of the risk-free security in the set of primitive assets helps to correctly identify the conditional mean of the SDF.

Z_t denotes the vector of instruments including a constant. The unconditional setting is a special case where the vector of instruments collapses to a constant.

Performance measures

Following Chen and Knez (1996), who first introduced performance evaluation using SDF models, a fund's conditional SDF-alpha, $\alpha_{P,t}$, is defined by[14]

$$\alpha_{P,t} = E(m_{t+1} \cdot R_{P,t+1} | Z_t) - 1, \tag{6}$$

where $R_{P,t+1}$ is the gross return of a managed fund at time $t+1$. Thus, $\alpha_{P,t}$ measures the deviation between the expected risk-adjusted gross return of a fund and its failure, which is 1 by definition. As shown by Farnsworth et al. (2002), $\alpha_{P,t}$ is zero if the fund forms a portfolio (at no cost) of the primitive assets using public information as reflected in Z_t, and the SDF prices the set of N primitive assets correctly. Under these assumptions we have $R_{P,t+1} = x(Z_t)' R_{t+1}$, where $x(Z_t)$ is a vector of portfolio weights depending on the set of instruments. As a consequence, Eq. (6) implies that:

$$\begin{aligned}
\alpha_{P,t} &= \left\{ E(m_{t+1} \cdot x(Z_t)' R_{t+1} | Z_t) \right\} - 1 \\
&= x(Z_t)' \left\{ E(m_{t+1} \cdot R_{t+1} | Z_t) \right\} - 1 \\
&= x(Z_t)' \mathbf{1} - 1 = 0.
\end{aligned} \tag{7}$$

The portfolio of an uninformed manager therefore yields a performance of zero.

The inferences of an estimated SDF-alpha in Eq. (6) only hold if the pricing equation (5b) holds for a candidate SDF using a set of N primitive assets. Following Farnsworth et al. (2002), the following system of equations is specified to simultaneously estimate the parameters of a candidate SDF and a fund's SDF-alpha using Hansen's (1982) generalized method of moments (GMM):

$$\begin{aligned}
u1_t &= (m_{t+1} \cdot R_{t+1} - 1) \otimes Z_t \\
u2_t &= \alpha_P - m_{t+1} \cdot R_{P,t+1} + 1,
\end{aligned} \tag{8}$$

where $u1_t$ is the vector of pricing errors relating to the set of N primitive assets, and $u2_t$ denotes the fund specific pricing error. R_{t+1} and $R_{P,t+1}$ denote the gross returns of the primitive assets and the funds, respectively. Thus, the first N equations define a set of orthogonality conditions involving pricing errors and instruments, and the final equation estimates the parameter α_P which is the mean of the conditional SDF-alpha defined in Eq. (6) and should be equal to zero under the null hypothesis of no abnormal fund performance. The sample moment condition is defined as $g = T^{-1} \sum_t (u1_t', u2_t')'$ and allows to simultaneously estimate the system of equations (8) using Hansen's

---

[14] In general, the alphas from the beta model and the alphas from the SDF model in (6) differ in the unconditional as well as in the conditional setting. The issue is discussed in detail in Söderlind (1999). We denote both alphas by α_P to keep notation simple.

(1982) GMM.[15] The moment conditions are constructed such that the expectation at their true values is zero. Accordingly, GMM minimizes the sample moment condition g such that the difference between the estimated parameters and the true values, i.e., the implied pricing error, is minimized. This implies that the value of an estimated candidate SDF comes as close as possible to the true value of SDF.

Model specification

We compute performance based on three SDF model specifications:
Linear factor models: In linear factor models, e.g., the CAPM and Fama and French (1993) three-factor model, m_{t+1} is defined as

$$m_{t+1}^{\text{LF}} = a(\mathbf{Z}_t) + \mathbf{b}(\mathbf{Z}_t)' \mathbf{F}_{t+1}, \tag{9}$$

where $\mathbf{F}_{t+1}$ denotes the vector of factors gross returns[16] at time $t + 1$, and $\mathbf{b}(\mathbf{Z}_t)$ represents a vector of the respective sensitivities of the SDF. In the unconditional case, $\mathbf{Z}_t$ collapses to a constant and the parameters a and $\mathbf{b}$ are constant over time. For example, the SDF for the unconditional CAPM is

$$m_{t+1}^{\text{LF}} = a + b R_{B,t+1}, \tag{10}$$

where $R_{B,t+1}$ denotes the gross market return, or for the unconditional Fama and French three-factor model the SDF is

$$m_{t+1}^{\text{LF}} = a + b_1 R_{B,t+1} + b_2 \text{SMB}_{t+1} + b_3 \text{HML}_{t+1}, \tag{11}$$

where SMB_{t+1} and HML_{t+1} denote the gross returns on the small minus big and the high minus low book-to-market style portfolios, respectively.

Conditional factor pricing models can be most easily implemented by using scaled factor returns (Cochrane 1996, 2001) by assuming linearity of $a(\mathbf{Z}_t)$ and $\mathbf{b}(\mathbf{Z}_t)$ in $\mathbf{Z}_t$. In the simplest setting of one factor and one instrument, the following specification results

$$\begin{aligned} m_{t+1}^{\text{LF}} &= a(Z_t) + b(Z_t)' F_{t+1} \\ &= a_0 + a_1 Z_t + (b_0 + b_1' Z_t) F_{t+1} \\ &= a_0 + a_1 Z_t + b_0 F_{t+1} + b_1 (Z_t F_{t+1}) \end{aligned} \tag{12a}$$

---

[15] A potential problem of the simultaneous GMM procedure is that the number of moment conditions grows substantially with the number of funds to be evaluated. Thus, we estimate the system (7) separately for each fund. Nevertheless, Farnsworth et al. (2002) prove that separate estimation is not restrictive as the estimated fund's alpha and its standard error are invariant to the number of funds in the system. Estimating the system for one fund at a time is therefore equivalent to estimating the system with all funds in the sample simultaneously.

[16] We only specify models where factors are returns or excess returns on traded assets (spread portfolios, mimicking portfolios).

which is a three-factor model (Z_t, F_{t+1}, and $Z_t F_{t+1}$) with fixed coefficients. In the general case of K factors and L instruments, with $a(\boldsymbol{Z}_t) = a_0 + \boldsymbol{a}_1' \boldsymbol{Z}_t$ and $\boldsymbol{b}(\boldsymbol{Z}_t) = \boldsymbol{b}_0 + \boldsymbol{B}_1' \boldsymbol{Z}_t$, the SDF specification becomes

$$m_{t+1}^{\mathrm{LF}} = a_0 + \boldsymbol{a}_1' Z_t + \boldsymbol{b}_0' \boldsymbol{F}_{t+1} + \boldsymbol{b}_1'(\boldsymbol{Z}_t \otimes \boldsymbol{F}_{t+1}) \tag{12b}$$

where $\boldsymbol{b}_1 \equiv \mathrm{vec}(\boldsymbol{B}_1)$ is the vectorized matrix of $\boldsymbol{B}_1$-coefficients. Equation (12b) implies that the generalized model requires the estimation of $L \times (K + 1)$ parameters (including a constant). To identify the parameters, we need a minimum number of $N \geq K + 1$ primitive assets.

Primitive-efficient SDFs: Hansen and Jagannathan (1991) show that the projection of the SDF, m_{t+1}, on the space of primitive gross asset returns, $\boldsymbol{R}_{t+1}$, is given by

$$m_{t+1}^{\mathrm{PE}} = 1' E(\boldsymbol{R}_{t+1} \boldsymbol{R}_{t+1}' | \boldsymbol{Z}_t)^{-1} \boldsymbol{R}_{t+1} \tag{13}$$

and exhibits the same pricing properties as the "true" SFD which prices the primitive assets correctly. In our context, the projection is called "primitive-efficient", m_{t+1}^{PE}, because the first two terms, $\boldsymbol{M}(\boldsymbol{Z}_t) \equiv \boldsymbol{1}' E(\boldsymbol{R}_{t+1} \boldsymbol{R}_{t+1}' | \boldsymbol{Z}_t)^{-1}$, define a vector of conditionally efficient portfolio weights. Following Chen and Knez (1996), Dahlquist and Söderlind (1999), and Farnsworth et al. (2002), we assume that the weights are a linear function of the instruments, $\boldsymbol{M}(\boldsymbol{Z}_t) = \boldsymbol{Z}_t' \boldsymbol{M}$, which implies:

$$m_{t+1}^{\mathrm{PE}} = \boldsymbol{Z}_t' \boldsymbol{M} R_{t+1} = \boldsymbol{m}'(\boldsymbol{Z}_t \otimes \boldsymbol{R}_{t+1}), \tag{14}$$

where $\boldsymbol{M}$ is an $L \times N$ matrix and $\boldsymbol{m} = \mathrm{vec}(\boldsymbol{M})$. Apparently, this specification also implies a scaled return model which can be estimated unconditionally, with an expanded asset universe, using GMM. With N primitive assets and L instruments (including a constant), there are $N \times L$ orthogonality conditions and parameters to be estimated, implying that the system of equations is exactly identified.

Bakshi–Chen SDFs: Bakshi and Chen (1998) propose an exponential model that defines a strictly positive SDF, as implied by no-arbitrage. The SDF is specified as follows:

$$m_{t+1}^{\mathrm{BC}} = \exp\left\{\boldsymbol{Z}_t' \boldsymbol{C} ln(\boldsymbol{R}_{t+1})\right\}, \tag{15}$$

where $\boldsymbol{C}$ (similar to $\boldsymbol{M}$ above) is an $L \times N$ matrix that denotes the parameters of the model. Again, using N primitive assets and L instruments (including a constant), the system of equations is exactly identified.

References

Aragon, G., Ferson, W.: Portfolio performance evaluation. Found. Trends Financ. **2**, 83–190 (2006)

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

Bekaert, G., Liu, J.: Conditioning information and variance bounds on pricing kernels. Rev. Financ. Stud. **17**, 339–378 (2004)

Bessler, W., Drobetz, W., Zimmermann, H.: Conditional performance evaluation for German mutual equity funds. Eur. J. Financ. **15**, 287–316 (2009)

Chen, Z., Knez, P.: Portfolio performance measurement: theory and applications. Rev. Financ. Stud. **9**, 511–555 (1996)

Cochrane, J.: A cross-sectional test of an investment-based asset pricing model. J. Polit. Econ. **104**, 572–621 (1996)

Cochrane, J.: Asset Pricing. Princeton University Press, New Jersey (2001)

Dahlquist, M., Söderlind, P.: Evaluating portfolio performance with stochastic discount factors. J. Bus. **72**, 347–384 (1999)

Dahlquist, M., Engström, S., Söderlind, P.: Performance and characteristics of Swedish mutual funds. J. Financ. Quant. Anal. **35**, 409–423 (2000)

Dybvig, P., Ross, S.: The analytics of performance measurement using and a security market line. J. Financ. **40**, 401–416 (1985b)

Fama, E., French, K.: Common risk factors in the returns on stocks and bonds. J. Financ. Econ. **33**, 3–56 (1993)

Farnsworth, H., Ferson, W., Jackson, D., Todd, S.: Performance evaluation with stochastic discount factors. J. Bus. **75**, 473–504 (2002)

Ferson, W., Lin, J.: Alpha and performance measurement: the effect of investor heterogeneity. Working Paper (2012)

Ferson, W., Qian, M.: Conditional performance evaluation: revisited. Working Paper, Boston College (2004)

Ferson, W.: Investment Performance: a review and synthesis. In: Constantinides, G., Harris, M., Stulz, R. (eds.) Handbook of Economics and Finance (2012)

Ferson, W., Harvey, C.: The risk and predictability of international equity returns. Rev. Financ. Stud. **6**, 527–566 (1993)

Ferson, W., Korajczyk, R.: Do arbitrage pricing models explain the predictability of stock returns? J. Bus. **68**, 309–349 (1995)

Ferson, W., Schadt, R.: Measuring fund strategy and performance in changing economic conditions. J. Financ. **51**, 425–461 (1996)

Ferson, W., Harvey, C.: Conditioning variables and cross-section of stock returns. J. Financ. **54**, 1325–1360 (1999)

Ferson, W., Henry, T., Kisgen, D.: Evaluating government bond fund performance with stochastic discount factors. Rev. Financ. Stud. **19**, 423–456 (2006)

Ferson, W.: Investment performance evaluation. Ann. Rev. Financ. Econ. **2**, 207–234 (2010)

Garrett, I., Hyde, S., Lozano, M.: Trade-offs between efficiency and robustness in the empirical evaluation of asset pricing models. Working Paper (2011)

Gibbons, M., Ross, S., Shanken, J.: A test for the efficiency of a given portfolio. Econometrica **57**, 1121–1152 (1989)

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

Hansen, L.: Large sample properties of generalized method of moments estimators. Econometrica **50**, 1029–1054 (1982)

Hansen, L., Jagannathan, R.: Implications of security market data for models of dynamic economies. J. Polit. Econ. **99**, 225–262 (1991)

Hansen, L., Jagannathan, R.: Assessing specification errors in stochastic discount factor models. J. Financ. **52**, 557–590 (1997)

Jagannathan, R., Wang, Z.: Empirical evaluation of asset pricing models: a comparison of the SDF and beta methods. J. Financ. **57**, 2337–2367 (2002)

Jensen, M.: The performance of mutual funds in the period 1945–1964. J. Financ. **23**, 389–416 (1968)

Jobson, J., Korkie, B.: Potential performance and tests of portfolio efficiency. J. Financ. Econ. **10**, 433–466 (1982)

Kan, R., Zhou, G.: Empirical asset pricing: the beta method versus the stochastic discount factor method. Working Paper. University of Toronto and Washington University, St. Louis (2002)

Kan, R., Robotti, C.: Specification tests of asset pricing models using excess returns. J. Empir. Financ. **15**, 816–838 (2008)

Lewellen, J.: Predicting returns with financial ratios. J. Financ. Econ. **74**, 209–235 (2004)

Lewellen, J., Nagel, S., Shanken, J.: A skeptical appraisal of asset pricing tests. J. Financ. Econ. **96**, 175–194 (2010)

MacKinlay, C.: Multifactor models do not explain deviations from the CAPM. J. Financ. Econ. **38**, 3–28 (1995)

Malkiel, B.: Returns from investing in equity mutual funds 1971 to 1991. J. Financ. **50**, 549–572 (1995)

Mehra, R., Prescott, E.: The equity premium: a puzzle. J. Monet. Econ. **15**, 145–162 (1985)

Silva, F., Cortez, M.C., Armada, M.: Conditioning information and European bond fund performance. Eur. Financ. Manag. **9**, 201–230 (2003)

Söderlind, P.: An interpretation of SDF based performance measures. European Finance Review **3**, 233–237 (1999)

White, H.: A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica **48**, 817–838 (1980)

Jonas Gusset is a doctoral candidate at the Department of Finance at the University of Basel (Switzerland). He holds an M.Sc. in Business and Economics with a major in Monetary Economics and Financial Markets. After graduating in 2012, he started working part-time in the Research and Investment Solutions unit at Basellandschaftliche Kantonalbank (BLKB) besides conducting his Ph.D. studies at the University of Basel. His research interests are empirical asset pricing and monetary economics.

Heinz Zimmermann is a professor of Finance at the Faculty of Economics and Business at the University of Basel (Switzerland). His research interest includes empirical asset pricing, portfolio management, derivatives and corporate finance. Before joining the University of Basel, he was a professor at the University of St. Gallen (Switzerland), from 1989 to 2001. He studied economics at the University of Bern (Switzerland) and held post-doc positions at various US universities. He is an Honorary Member of the Swiss Society for Finance Research. From 2004–2005, he served as Co-President (together with Wolfgang Drobetz) of the European Financial Management Association.