The Sharpe ratio's market climate bias: theoretical and empirical evidence from US equity mutual funds

Sebastian Krimm, Hendrik Scholz, Marco Wilkens

Angaben zur Veröffentlichung / Publication details:

Krimm, Sebastian, Hendrik Scholz, and Marco Wilkens. 2012. "The Sharpe ratio's market climate bias: theoretical and empirical evidence from US equity mutual funds." *Journal of Asset Management* 13 (4): 227–42. https://doi.org/10.1057/jam.2012.11.

The Sharpe ratio's market climate bias: Theoretical and empirical evidence from US equity mutual funds

Sebastian Krimm

is a Research Assistant and a PhD student at the chair of finance and banking of the University of Augsburg, Germany. He studied business administration at the Catholic University of Eichstätt-Ingolstadt. His research focuses on performance analysis and empirical finance.

Hendrik Scholz

is a Professor of finance and banking at the Friedrich-Alexander-University Erlangen-Nürnberg, Germany. He received his PhD in finance from the Georg-August-University Göttingen and his habilitation from the Catholic University of Eichstätt-Ingolstadt. His research focuses on performance analysis, asset management, structured financial products and bank management. Hendrik Scholz has authored numerous articles in finance journals. He is a member of the American Finance Association, the Southern Finance Association, the German Finance Association and the German Academic Association for Business Research.

Marco Wilkens

is a Professor of finance and banking at the University of Augsburg, Germany. He received his PhD from the University of Hamburg and his habilitation from the Georg-August-University of Göttingen. His research focuses on performance analysis, interest rate risk, risk management in banking, valuation of financial instruments and financial engineering. Marco Wilkens has authored numerous articles in finance journals. He is a board member and treasurer of the German Finance Association and a member of the board of directors of the European Financial Management Association.

Correspondence: Hendrik Scholz, Finance and Banking, Friedrich-Alexander-University Erlangen-Nürnberg, Lange Gasse 20, Nürnberg 90403, Germany

ABSTRACT This article analyses the impact of market climates on the Sharpe ratios (SRs) of funds. On the basis of a common factor model, we derive analytically how market climates impact the SR – taking into account the abilities of fund managers. This applies especially to the mean of the market returns during the evaluation period: The performance of funds with relatively high unsystematic risk is biased upwards in outstandingly negative market climates, and vice versa. Our empirical study of US equity mutual funds supports these theoretical insights. We show that the SR of poorly diversified funds is biased upwards in bear markets, and vice versa. Subsequently, we confirm that actual fund SRs depend on especially the mean excess returns of the market. Thus, the SR does not provide a meaningful assessment of fund performance, especially in extraordinary times. We therefore suggest using the 'normalised' Sharpe ratio in future empirical research, in order to avoid the bias of SRs and rankings due to market climate.

Keywords: performance evaluation; equity mutual funds; Sharpe ratio; bear market; market conditions

INTRODUCTION

Since the seminal paper by Sharpe (1966), the Sharpe ratio (SR) has been widely used to assess the performance of funds in the finance literature and in practice. Private investors compare and choose funds using the SR, which is available through financial publications and information services on the Internet. The relevance of this performance measure is obvious even if measures following the approach of Jensen (1968) are increasingly popular in the literature. The SR is still referred to as the 'most common measure of risk-adjusted return' (Modigliani and Modigliani, 1997, p. 46) or as '(o)ne of the most commonly cited statistics in financial analysis' (Lo, 2002, p. 36).

Despite its common use, the SR has come under question, especially in the recent past. It is often stated that during periods of declining share prices, this measure leads to intuitively incomprehensible, if not even erroneous conclusions (see, for example, Tinic and West, 1979; Jobson and Korkie, 1981; Vinod and Morey, 2000; Ferruz and Sarto, 2004; Israelsen, 2005, 2009). Sharpe (1975, 1998) himself disputes issues of the original SR during bear markets. According to him, the SR is an appropriate performance measure, even for periods of decreasing share prices. The fund exhibiting the highest SR will also attain the highest average return when combined with a risk-free asset for any level of risk. This holds true in both bull and bear markets (see also Lobosco, 1999). McLeod and van Vuuren (2004) present another argument for the SR during declining markets. They show that the fund with the maximum SR is the fund with the highest probability of outperforming a risk-free investment.

In the light of the criticism of the SR during bear markets, several authors have proposed modifications of the SR. Israelsen (2003, 2005, 2009), Ferruz and Sarto (2004) and Chen *et al* (2008) have each developed measures that are supposed to yield more reliable inferences of fund performance during bear markets. Referring to an initial version of this article, Scholz (2007) finds plausible fund rankings based on the (single-factor) normalised SR in the only comparative study of refinements of the SR.

However, none of these studies disclose how market climates in general – and not only during declining markets – affect the SR. As this question is unanswered as yet, the contradictions in the literature with respect to the interpretation of SRs in bear market periods remain. We fill this gap and assess the fundamental informational value of this prominent measure in finance. The burst of the new economy bubble at the beginning of the millennium and the financial crisis at the end of its first decade have resulted in commonly negative SRs since then. Therefore, the criticism cited above is especially relevant for the beginning of the present century.

The main purpose of this article is to examine to what degree the SR of funds depends on realizations of market returns during the evaluation period. We add new insights to the literature, as we do not only focus on bear markets but market climates in general. In the process, we answer the question of how far fund performance can be evaluated based on the SR. Furthermore, we extend the normalised SR, which overcomes the market climate bias in the context of a multi-factor world. Finally, we conduct the first broad empirical study of the market climate bias of the SR and methods to overcome this issue.

The remainder of this article is organized as follows: In our theoretical analysis in the section 'The market climates bias of the SR', we demonstrate that commonly specified *ex-post* SRs do not allow for a meaningful assessment of fund performance. We not only show that SRs are subject to random market climates, but reveal in particular the contribution of fund-specific risk, which is either positive or negative depending on the market climate. The section

'Data description' presents empirical results on the relevant impact of market climate on measured SRs based on a sample of US equity mutual funds. First, we highlight that, on average, the SR of funds exhibiting relatively high proportions of fund-specific risk is biased upwards in bear markets, and vice versa. Second, we ascertain that SRs of funds depend on especially the mean excess returns of the market. In the section 'Empirical relevance of the market climate bias', we recommend using the normalised SR for *ex-post* assessments of funds in order to overcome the impact of market climates. Furthermore, we employ the normalised SR based on a single-factor as well as on a multi-factor model to determine the performance of our funds sample and identify striking changes in rank compared with corresponding fund rankings based on the original measure. The final section concludes.

THE MARKET CLIMATE BIAS OF THE SHARPE RATIO

Impact of market climates on the Sharpe ratio

The SR of a fund i is usually calculated based on the fund's returns in excess of the risk-free rate. *Ex-post*, SR_i is the mean fund excess return μ_i divided by the standard deviation of the fund's excess returns σ_i, with both statistics estimated for a specific evaluation period:

$$SR_i = \frac{\mu_i}{\sigma_i} \qquad (1)$$

The mean fund excess return μ_i obviously depends on the abilities of the fund management. However, the SR is also affected by the general market conditions. Following the general approach of Pástor and Stambaugh (2002) and its application to this subject in Scholz (2007), we first assume a linear single-factor model as the return generating process of the excess returns of fund i:

$$er_{i,t} = \alpha_i + \beta_i er_{m,t} + \varepsilon_{i,t} \qquad (2)$$

The variable $er_{i,\,t}$ is the excess return of fund i in period t and $er_{m,\,t}$ is the excess return of the market in period t.[1] The coefficient β_i reflects the fund's market exposure. Positive (negative) selection ability is determined by a positive (negative) α_i. The fund's unsystematic risk is given by the residual term $\varepsilon_{i,\,t}$, which we assume to be normally, independently and identically distributed with mean zero. We denote the variance of the residual as $\sigma^2_{\varepsilon,\,i}$ and α_i, β_i and $\sigma^2_{\varepsilon,\,i}$ together as the fund-specific characteristics.

On the basis of Equation (2), market excess returns obviously influence excess returns of funds. To show how market climates impact the SR of funds, we assume the fund-specific characteristics in each evaluation period as given and coinciding with the true values. For a single-factor model, the SR then results as:

$$SR_i = \frac{\alpha_i + \beta_i \mu_m}{\sqrt{\beta_i^2 \sigma_m^2 + \sigma_{\varepsilon,i}^2}} \qquad (3)$$

μ_m and σ^2_m are the mean and the variance of the market excess returns. The SR is a function of these and therefore a random variable. To further break down the impact of the market climate, we decompose the SR into two components:

$$SR_i = \frac{\beta_i \mu_m}{\sqrt{\beta_i^2 \sigma_m^2}} + DSR_i \qquad (4)$$

The first component is the SR of the fund's asset allocation. It corresponds to the SR of a passive investment in the market with the constant fund-specific market exposure β_i. The second component indicates the skills of the fund manager, relative to her asset allocation. We denote it as differential

Sharpe ratio (DSR). We further simplify Equation (4) by assuming the fund's β_i to be positive, which is usually the case for equity mutual funds:[2]

$$SR_i = \frac{\mu_m}{\sigma_m} + DSR_i = SR_m + DSR_i \qquad (5)$$

According to Equation (5), the first component matches the SR of the market SR_m. Thus, ranking funds according to their SR or their DSR leads to identical results. However, varying market climates over time clearly affect the first component and therefore rankings of funds when their SRs are measured over different evaluation periods. Irrespective of the sign of β_i, we further split up the DSR in two components:

$$DSR_i = \underbrace{\frac{\alpha_i}{\sqrt{\beta_i^2\sigma_m^2 + \sigma_{\varepsilon,i}^2}}}_{\text{total-risk-adjusted performance}} + \underbrace{\left(\frac{1}{\sqrt{\beta_i^2\sigma_m^2 + \sigma_{\varepsilon,i}^2}} - \frac{1}{\sqrt{\beta_i^2\sigma_m^2}}\right)\beta_i\mu_m}_{\text{contribution of unsystematic risk}} \qquad (6)$$

The first component is the α_i adjusted by the fund's total risk (total-risk-adjusted performance, henceforth TRAP). It is a linear transformation of α_i, thus positive α's increase the SR and vice versa. The second component is the contribution of unsystematic risk (CUR). For a perfectly diversified portfolio, the term in brackets is zero and so is CUR. Otherwise, this term is negative and increases *ceteris paribus* with the fund's proportion of unsystematic risk (PUR) on total risk. However, the sign of CUR is determined by the market climate. Depending on the realization of μ_m during the evaluation period, CUR is either negative or positive for a given market risk exposure of a fund.

This dependence on market climates of the SRs of funds will thus be referred to as market climate bias. It leads to an overestimation of the performance of funds exhibiting relatively high proportions of unsystematic risk in extraordinarily negative market climates, but it also results in an underestimation of these funds' performance in outstandingly positive market climates. With this finding, our article contributes to the fund literature, which until now criticises the use of SRs for declining markets only.

Moreover, because of this market climate impact in the form of CUR, funds with successful selection activities do not necessarily outperform the market during positive market climates. Such an outperformance requires that the positive contribution of α to the SR (TRAP > 0) overcompensates for the penalty of the associated unsystematic risk (CUR < 0). During negative market climates, even funds with negative α can outperform the market: The positive CUR in bear markets (CUR > 0) may outweigh the negative performance (TRAP < 0). In short, the more positive the market climate, the more complicated it is for fund managers to achieve SRs superior to the market, and vice versa.

Such a simple statement cannot be made with respect to the effect of the unsystematic risk of funds. While higher unsystematic risk always increases the market climate bias, its impact on TRAP may either be positive or negative depending on the sign of the fund's α. The latter also applies to the effect of the variance of the market excess returns on CUR in connection with the sign of the mean market excess return. Thus, one cannot determine with certainty how σ_m^2 influences the SRs of funds. We finally stress that this Equation (6) strictly holds and always applies to *ex-post* SRs.

The normalised Sharpe ratio

To overcome the market climate bias of the SR, we recommend using a measure based on Equation (3) to evaluate the 'pure' fund performance while controlling for the impact of market climates. For this measure, fund characteristics and distribution parameters of the market excess return are estimated using different data samples. To estimate the fund-specific characteristics, one commonly uses the return of funds for a time series of a few years, as long-ranging data often does not exist, especially for new funds. With respect to the distribution parameters of the market, investors should use long-term time frames – which are regularly available, unlike for fund returns – and employ the resulting estimates of the mean and the variance of the market excess return to calculate the normalised Sharpe ratio (nSR) as follows:

$$nSR_i = \frac{\alpha_i + \beta_i \mu_{Lm}}{\sqrt{\beta_i^2 \sigma_{Lm}^2 + \sigma_{\varepsilon,i}^2}} \tag{7}$$

μ_{Lm} and σ_{Lm}^2 indicate long-term estimates of the mean and the variance of the market.[3] The nSR can be interpreted as the risk-adjusted performance measure of a fund that results from its fund-specific characteristics during a 'normal' market climate. The considerable advantage of the nSR is that it is not affected by random and exceptional market climates, which allows for an undistorted assessment of fund performance. Scholz (2007) applies the nSR in a comparative study and finds its results superior to those of other enhancements of the original SR.

We decompose the nSR in the same manner as in Equations (5) and (6) which leads to:

$$\begin{aligned} nSR_i &= \frac{\mu_{Lm}}{\sigma_{Lm}} + nDSR_i \\ &= nSR_m + \underbrace{\frac{\alpha_i}{\sqrt{\beta_i^2 \sigma_{Lm}^2 + \sigma_{\varepsilon,i}^2}}}_{\text{normalised TRAP}} \\ &\quad + \underbrace{\left(\frac{1}{\sqrt{\beta_i^2 \sigma_{Lm}^2 + \sigma_{\varepsilon,i}^2}} - \frac{1}{\sqrt{\beta_i^2 \sigma_{Lm}^2}} \right) \beta_i \mu_{Lm}}_{\text{normalised CUR}} \end{aligned} \tag{8}$$

The components of the nSR follow the economic interpretation of their counterparts in the context of the original SR. As Equation (8) is based on the long-term market parameters μ_{Lm} and σ_{Lm}^2, normalised total-risk-adjusted performance (nTRAP) and normalised contribution of unsystematic risk (nCUR) represent the TRAP and the CUR the fund would have shown in a normal market period given its fund-specific characteristics.

The Sharpe ratio and its market climate bias in a multi-factor world

Instead of a single-factor model as in Equation (2), we now assume a linear multi-factor model with K risk factors as the return-generating process of the fund excess returns:

$$er_{i,t} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{f}_t + \varepsilon_{i,t} \tag{9}$$

Vector $\mathbf{f}_t$ contains the realizations of the K market factors $f_{1t}, f_{2t}, \ldots, f_{Kt}$ in period t. Vector $\boldsymbol{\beta}_i$ reflects the fund's exposure to these factors. α_i discloses the selection performance, the residual term $\varepsilon_{i,t}$ the unsystematic risk. In the context of the K-factor model, the SR of a fund i is:

$$SR_i = \frac{\alpha_i + \boldsymbol{\beta}_i' \boldsymbol{\mu}_{\mathbf{f}}}{\sqrt{\boldsymbol{\beta}_i' \mathbf{V}_{\mathbf{f}} \boldsymbol{\beta}_i + \sigma_{\varepsilon,i}^2}} \tag{10}$$

Vector $\boldsymbol{\mu}_{\mathbf{f}}$ stands for the mean realizations and matrix $\mathbf{V}_{\mathbf{f}}$ for the variance–covariance matrix of the market factors during the evaluation period. Similar to the decomposition for the single-factor case in Equation (4), we further decompose the SR of a fund i based on its asset allocation given by its factor loadings $\boldsymbol{\beta}_i$.

$$SR_i = \frac{\boldsymbol{\beta}_i' \boldsymbol{\mu}_{\mathbf{f}}}{\sqrt{\boldsymbol{\beta}_i' \mathbf{V}_{\mathbf{f}} \boldsymbol{\beta}_i}} + DSR_i \tag{11}$$

The first component is the SR of a passive investment in the market factors with constant fund-specific factor loadings $\boldsymbol{\beta}_i$. Accordingly, the DSR can be interpreted as fund performance in excess of the performance of this passive investment alternative.[4] We further split up the DSR as follows:

$$DSR_i = \underbrace{\frac{\alpha_i}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V_f}\boldsymbol{\beta}_i + \sigma_{\varepsilon,i}^2}}}_{\text{total}-\text{risk}-\text{adjusted performance}} + \underbrace{\left(\frac{1}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V_f}\boldsymbol{\beta}_i + \sigma_{\varepsilon,i}^2}} - \frac{1}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V_f}\boldsymbol{\beta}_i}}\right)\boldsymbol{\beta}_i'\boldsymbol{\mu}_\mathbf{f}}_{\text{contribution of unsystematic risk}} \quad (12)$$

The decomposition is structurally identical to the one in the context of the single-factor model. The first component TRAP is the α of the fund adjusted by its total risk. The second component CUR depicts again the influence of market climates. As in the single-factor case, CUR is zero for a perfectly diversified portfolio. In all other cases, the sign of CUR is determined by $\boldsymbol{\beta}_i'\boldsymbol{\mu}_\mathbf{f}$. Depending on the market climate, that is, the realization of $\boldsymbol{\mu}_\mathbf{f}$ during the evaluation period, this component is either negative or positive for a given risk exposure $\boldsymbol{\beta}_i'$ of a fund.

As for the single-factor model, investors can use nSRs based on long-term estimates of the means and the variance–covariance matrix of the market factors to calculate (multi-factor) nSRs that do not suffer from a market climate bias:

$$nSR_i = \frac{\alpha_i + \boldsymbol{\beta}_i'\boldsymbol{\mu}_{L\mathbf{f}}}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V}_{L\mathbf{f}}\boldsymbol{\beta}_i + \sigma_{\varepsilon,i}^2}} \quad (13)$$

Vector $\boldsymbol{\mu}_{L\mathbf{f}}$ contains long-term estimates of the means and matrix $\mathbf{V}_{L\mathbf{f}}$ a long-term estimate of the variance–covariance matrix of the market factors. We can decompose the nSR as before and calculate its components as:

$$nSR_i = \frac{\boldsymbol{\beta}_i'\boldsymbol{\mu}_{L\mathbf{f}}}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V}_{L\mathbf{f}}\boldsymbol{\beta}_i}} + \underbrace{\frac{\alpha_i}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V}_{L\mathbf{f}}\boldsymbol{\beta}_i + \sigma_{\varepsilon,i}^2}}}_{\text{normalised TRAP}} + \underbrace{\left(\frac{1}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V}_{L\mathbf{f}}\boldsymbol{\beta}_i + \sigma_{\varepsilon,i}^2}} - \frac{1}{\sqrt{\boldsymbol{\beta}_i'\mathbf{V}_{L\mathbf{f}}\boldsymbol{\beta}_i}}\right)\boldsymbol{\beta}_i'\boldsymbol{\mu}_{L\mathbf{f}}}_{\text{normalised CUR}} \quad (14)$$

Until now, the question of whether the market climate bias is empirically relevant remains unanswered. How much do rankings of funds based on SRs really vary as a result of differing market climates? How strong is the impact of changing market climates on SRs of funds? The following examines actual US equity mutual funds, and how their SRs and rankings based on the SR are impacted by market climates.

DATA DESCRIPTION

Fund data set

The daily fund returns in our sample are from the CRSP Survivor-Bias-Free US Mutual Fund Database. Our evaluation period is from 4 January 1999 to 31 December 2009. We select funds for our sample based on three criteria: First, a fund has to belong to one of six Lipper objective codes, which imply it is a US domestic equity fund.[5] We exclude funds changing between these codes or classified otherwise at any time. Second, we ignore funds whose return observations are not available for the full sample period without gaps. This restriction allows us to analyse all funds for the same evaluation period, but it causes a substantial survivorship bias[6] in our sample. As we do not aim to judge the economic value added by fund managers but focus on the impact of market climates on the SR, we consider this drawback to be acceptable. Third, the time series of the funds' returns may show no obviously implausible values.[7] Finally, we exclude all but one of the multiple share classes of a fund, if those exist, and only keep the oldest share class.

In the previous section, we assumed that funds only perform selection activities as a precondition for determining the fund-specific characteristics according to Equation (2). We use the squared-regression approach proposed by Treynor and Mazuy (1966) and the dummy variable regression approach developed by Henriksson and Merton (1981) to test for timing activities during the evaluation period.[8] We exclude all funds whose timing activities differ statistically from zero at the 5 per cent level.[9] Hence, potential timing activities of funds should not be a serious problem for our data sample.[10] Our final sample of fund returns consists of 605 funds which each have a full daily return history in our evaluation period from 4 January 1999 to 31 December 2009. Henceforth, the term 'full evaluation period' denotes that estimations are based on data using this period of time.

To demonstrate how the market climate bias of the SR varies over time and to allow for changing characteristics of funds, parts of our analysis are based on 101 rolling evaluation windows (henceforth 'rolling windows') defined as follows: Beginning 4 January 1999, these windows iteratively roll over by 25 days with each window covering 250 consecutive trading days. We perform all calculations separately for each fund and for each window whenever our analysis is based on the latter.

Market factors and market climates

We use the value-weighted index of all NYSE, AMEX and NASDAQ stocks from CRSP to estimate the fund-specific characteristics in the single-factor case. For the multi-factor model, we also consider factors for the size- and value-effect according to Fama and French (1993) and the momentum effect according to Carhart (1997).[11] The existence and the extent of market climates during our evaluation period is a crucial element of our empirical analysis, as these climates may have a serious impact on investment decisions based on the SR. To depict the market climates, Figure 1 presents charts of the mean and the standard deviation of the market factors over the rolling windows defined before.

Obviously, all market factors show distinct market climates. These are quite sizable for the market index, as our evaluation period covers the climax and burst of the new technology bubble at the beginning of the century as well as the housing boom and the financial crisis at the end of the decade. The Fama-French-Carhart factors exhibit market climates as well. To the extent these factors impact fund returns, their climates also potentially affect investor decisions based on original SRs.

With respect to the nSR, in Table 1 we compare descriptive statistics of the market factors during the full evaluation period of our fund sample in Panel A and a long-term period in Panel B. This long-term period covers daily data of the market factors from 1 July 1963 to 31 March 2011 – the longest time period available to us. The mean and standard deviation of the market factors during this long-term period are used later to calculate nSRs.

Means and standard deviations of the market factors differ strongly between both panels. Our evaluation period covers the burst of the new economy bubble and the subprime mortgage crisis. This explains the lower mean and the higher standard deviations of the market index in comparison with the long-term period. In addition, the higher standard deviations of the other three factors in our evaluation period could be explained by these crises. Thus, the market climate bias of the SR arises even if long evaluation periods are studied such as the 11-year one used here.

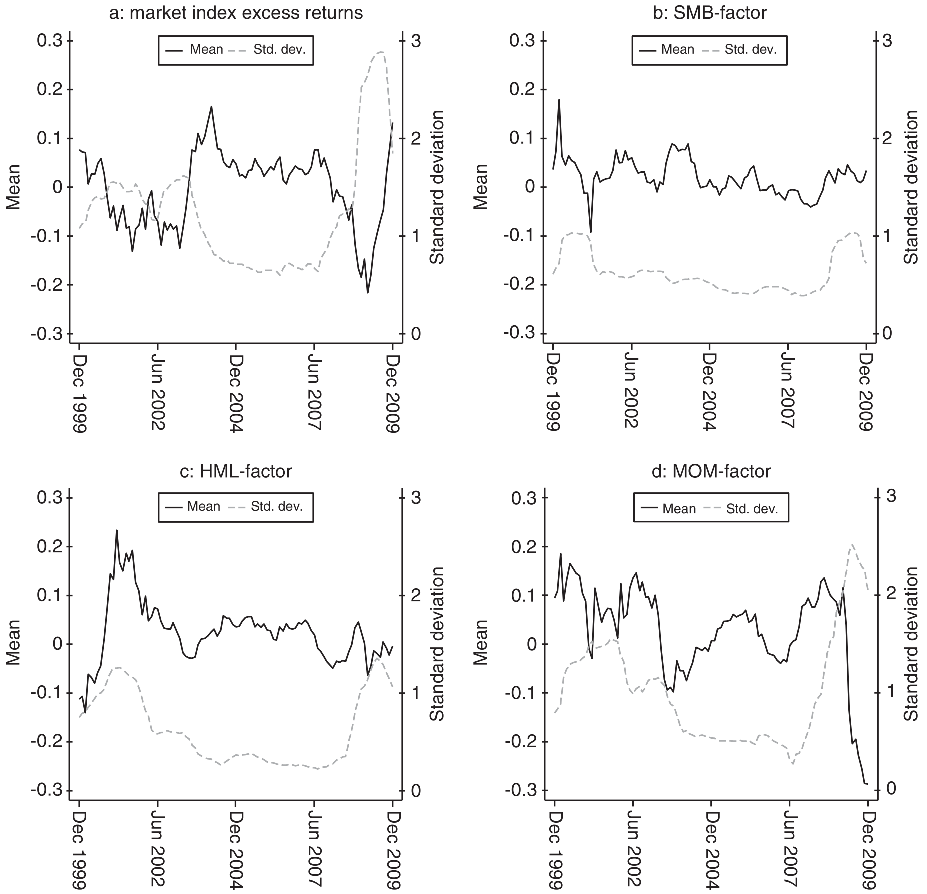

Figure 1: Market climates during the evaluation period. This figure shows the mean and the standard deviation of the value-weighted index of all NYSE, AMEX and NASDAQ stocks from CRSP (market index excess returns), the SMB-, the HML-, and the MOM-factor over 101 rolling windows. Beginning 4 January 1999, these windows iteratively roll over by 25 days with each window covering 250 consecutive trading days. The horizontal axis indicates the end dates of the respective rolling windows. All values are in per cent.

Table 1: Descriptive statistics of market factors

Market factor	*Mean*	*Median*	*5th percentile*	*95th percentile*	*Standard deviation*
Panel A: Evaluation period					
Market index	0.007	0.05	−2.12	2.01	1.379
SMB-factor	0.020	0.04	−0.97	1.00	0.649
HML-factor	0.022	0.02	−1.10	1.21	0.734
MOM-factor	0.016	0.07	−1.72	1.72	1.173
Panel B: Long-term period					
Market index	0.022	0.05	−1.49	1.44	0.983
SMB-factor	0.008	0.03	−0.78	0.74	0.508
HML-factor	0.019	0.01	−0.68	0.74	0.489
MOM-factor	0.033	0.06	−0.99	0.94	0.701

Note: This table presents the mean, median, the 5th percentile, the 95th percentile and the standard deviation of the value-weighted index of all NYSE, AMEX and NASDAQ stocks from CRSP (market index excess returns), the SMB-, the HML- and the MOM-factor. We calculate the descriptive statistics for the evaluation period from 4 January 1999 to 31 December 2009 in Panel A and for a long-term period from 1 July 1963 to 31 March 2011 in Panel B. All values are noted in per cent.

EMPIRICAL RELEVANCE OF THE MARKET CLIMATE BIAS

Empirical Sharpe ratio and its components

As shown in the section 'Impact of market climates on the Sharpe ratio', the funds' PUR drive the impact of market climates on the SR. However, portfolio theory suggests that unsystematic risk and fund performance are related.[12] To avoid the bias on our empirical study resulting from such a relation, we group funds according to their PUR while controlling for funds α's. For the single-factor model, we first estimate Equation (2) for each fund and form quintiles according to α_i. We then assign the funds within each α quintile again to quintiles according to their PUR. We aggregate the resulting 25 subsamples as follows: All funds belonging to a first quintile according to PUR form the group 'low unsystematic risk' funds (henceforth 'LUR funds'). All funds belonging to a fifth quintile according to PUR form the group 'high unsystematic risk' funds (HUR funds). The remaining funds constitute the 'mid unsystematic risk' (MUR) fund group.

We repeat this process for the multi-factor model by estimating Equation (9) instead of Equation (2). For the full evaluation period, we estimate the factor models using the full data sample and apply the grouping procedure once. For the rolling analysis, we repeat the grouping procedure in each of the 101 windows using the estimation results of the respective window. Table 2 presents the mean characteristics of these fund groups based on the full evaluation period.

The fund groups clearly differ in terms of their PUR, as intended. The slight variation in their mean α is insignificant at the 10 per cent level. The multi-factor model in Panel B yields considerably lower α's and proportions of unsystematic risk. Both findings necessarily result from including the additional factors. The fund groups' exposure to the SMB and the HML factor grows with PUR, which partly applies to the exposure to the momentum factor as well.

Table 3 presents means of the original SR and its components for the fund groups defined above and based on the full evaluation period. We estimate the original SR and calculate its components according

Table 2: Descriptive statistics of fund returns and characteristics

Fund group	*Mean er_i*	*Standard deviation er_i*	α	*β against market index*	*β against SMB-factor*	*β against HML-factor*	*β against MOM-factor*	*PUR*
Panel A: Single-factor model								
LUR funds	0.012	1.372	0.0056	0.961	–	–	–	6.75
MUR funds	0.013	1.437	0.0058	0.968	–	–	–	13.52
HUR funds	0.013	1.535	0.0058	0.971	–	–	–	23.70
All funds	0.013	1.444	0.0058	0.967	–	–	–	14.28
Panel B: Multi-factor model								
LUR funds	0.008	1.391	0.0001	0.980	0.004	0.018	0.009	4.06
MUR funds	0.013	1.452	0.0007	0.986	0.197	0.050	0.019	8.19
HUR funds	0.016	1.475	0.0006	0.947	0.329	0.092	0.016	16.24
All funds	0.013	1.444	0.0006	0.977	0.186	0.052	0.017	9.03

Note: This table presents mean and standard deviation of daily fund excess returns and the mean α, β and PUR for the full evaluation period from 4 January 1999 to 31 December 2009 and for funds grouped according to their PUR. We estimate Equation (2) in Panel A and Equation (9) using market factors according to Fama and French (1993) and Carhart (1997) in Panel B for each fund. We group funds by assigning them to quintiles according to their PUR while controlling for α. The quintile of funds with the lowest PUR forms the LUR groups, the quintile with the highest PUR forms the HUR funds. The remaining funds are assigned to the MUR group. Mean and standard deviation of daily fund return, α and PUR are noted in per cent.

Table 3: Original SR and its components

Fund group	*SR*	*DSR*	*TRAP*	*CUR*	*Mean rank*
Panel A: Single-factor model					
LUR funds	0.899	0.382	0.400	−0.018	299
MUR funds	0.873	0.356	0.393	−0.036	303
HUR funds	0.861	0.345	0.411	−0.066	307
All funds	0.875	0.359	0.398	−0.039	303
Panel B: Multi-factor model					
LUR funds	0.548	−0.007	0.005	−0.013	302
MUR funds	0.895	0.009	0.048	−0.039	300
HUR funds	1.133	−0.044	0.055	−0.099	311
All funds	0.875	−0.005	0.041	−0.046	303

Note: This table presents the mean SR, DSR, TRAP, CUR and mean rank according to the DSR for the full evaluation period from 4 January 1999 to 31 December 2009 and for funds grouped according to their PUR. We estimate Equation (2) in Panel A and Equation (9) using market factors according to Fama and French (1993) and Carhart (1997) in Panel B for each fund. We group funds by assigning them to quintiles according to their PUR while controlling for α. The quintile of funds with the lowest PUR forms the LUR groups, the quintile with the highest PUR forms the HUR funds. The remaining funds are assigned to the MUR group. With the exception of mean ranks, all values are in per cent.

Table 4: Normalised SR and its components

Fund group	*nSR*	*nDSR*	*nTRAP*	*nCUR*	*Mean rank*
Panel A: Single-factor model					
LUR funds	2.579	0.388	0.528	−0.140	296
MUR funds	2.416	0.225	0.498	−0.273	302
HUR funds	2.233	0.042	0.504	−0.462	313
All funds	2.411	0.219	0.505	−0.285	303
Panel B: Multi-factor model					
LUR funds	2.174	−0.084	0.007	−0.091	282
MUR funds	2.391	−0.139	0.068	−0.207	298
HUR funds	2.413	−0.364	0.071	−0.436	339
All funds	2.352	−0.175	0.057	−0.231	303

Note: This table presents the mean nSR, nDSR, nTRAP, nCUR and mean rank according to the nDSR for the full evaluation period from 4 January 1999 to 31 December 2009 and for funds grouped according to their PUR. We estimate Equation (2) in Panel A and Equation (9) using market factors according to Fama and French (1993) and Carhart (1997) in Panel B for each fund. We group funds by assigning them to quintiles according to their PUR while controlling for α. The quintile of funds with the lowest PUR forms the LUR groups, the quintile with the highest PUR forms the HUR funds. The remaining funds are assigned to the MUR group. With the exception of the ranks, all values are in per cent.

to Equations (4) and (6) in the context of the single-factor model and Equations (11) and (12) in the context of the multi-factor model.

The mean SR of the fund groups varies distinctively. The variation is even stronger when funds are grouped based on the multi-factor model. The funds' SR is mostly driven by their risk exposure, as the DSRs show. The sign of TRAP in Panel A and Panel B reflects the sign of the respective α's. In both Panels, the mean TRAP is higher for HUR compared with LUR funds, while CUR is lower. The latter follows our theoretical results. Market climates especially impact funds with a higher PUR.

In Table 4, we present the corresponding analysis for the nSRs. We initially estimate fund-specific characteristics via regressions according to Equations (2) and (9). For each fund, we determine nSRs, their components and the related rankings using the long-term market parameters of Table 1. We estimate the nSR and its components as in Equation (8) for the single-factor model and Equation (14) for the multi-factor model and calculate means of each quantity for the fund groups.

Compared with the original SR, the nSR is much higher for all funds groups and in both panels. As the nDSR covers a range similar to that of the original DSR, this increase is not driven by a more positive assessment of the fund managers' skills but an improved risk-adjusted evaluation of their asset allocation. In both Panels, the LUR funds outweigh the HUR funds in terms of mean nDSR. The components of the nDSR show that these differences result from nCUR, which scatters much more over the fund groups in comparison with the CUR. The higher mean excess return of the market index in our long-term period compared with our evaluation period (Table 1) and the dominance of the market index in the asset allocation of the funds (Table 2) led to the original SR not reflecting the PUR of the funds appropriately – in other words, we find a relevant bias in the original SR, though its estimations are based on daily data over 11 years.

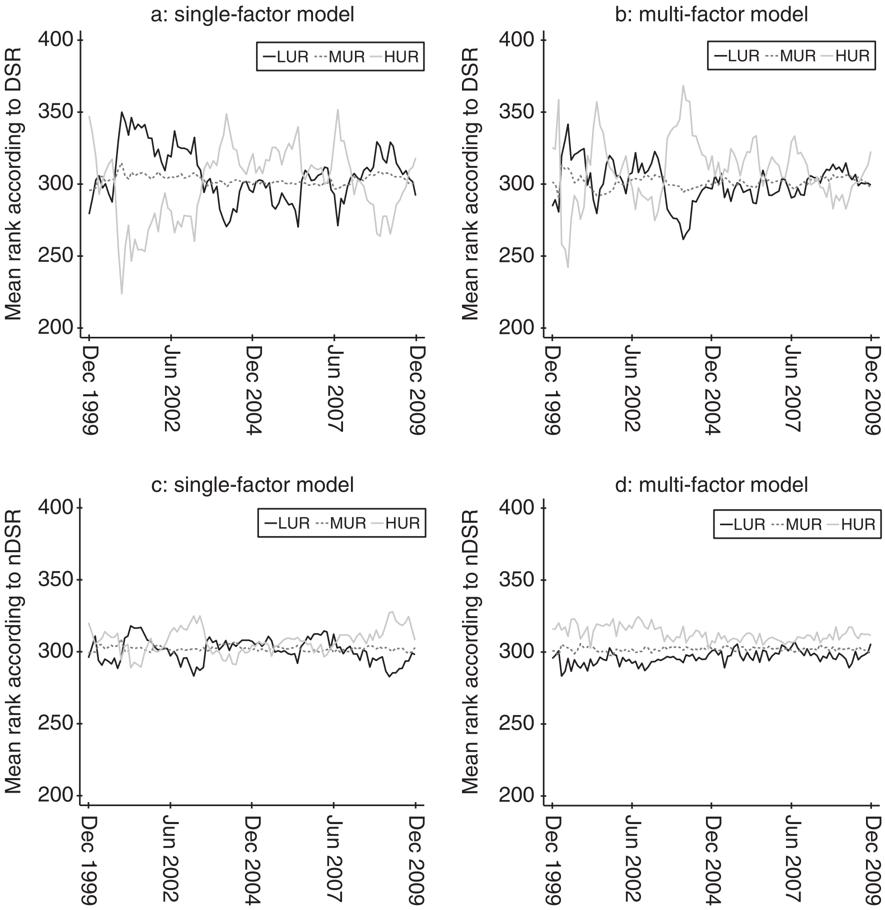

Figure 2: Mean fund ranks over time. This figure shows mean fund ranks according to the DSR and the nDSR of fund groups according to their PUR. LUR, MUR and HUR indicate the fund group with a low, medium and high PUR, respectively. SRs and fund-specific characteristics are estimated for 101 rolling windows of each 250 daily observations, which move forward by increments of 25 days. The horizontal axis indicates the end dates of the respective rolling windows. nSRs are calculated using these fund-specific characteristics and market parameters estimated for the long-term period from 1 July 1963 to 31 March 2011. We rank funds according to the DSR (nDSR) by assigning rank 1 to the fund with the highest DSR (nDSR) and rank 605 to the fund with the lowest DSR (nDSR).

Impact of market climates on Sharpe ratios

To show how market climates affect investor decisions based on the original SR and the extent to which the nSR is superior in this context, we use the rolling windows previously described. In each window, we rank funds according to the DSR and according to the nDSR. Figure 2 plots the mean ranks of the funds groups over the 101 evaluation windows.

The mean rank of LUR funds and HUR funds according to the DSR (Figures 2a and b) move in directly opposite directions. The MUR funds are nearly always between the other two groups. This applies to grouping funds according to a single-factor model as well as according to a multi-factor model and documents that investment decisions based on the original SR are highly affected by the PUR of funds. Even more, this effect strongly varies over time and shows hardly any persistence. The average ranks of LUR and HUR funds according to the nDSR (Figures 2c and d) are considerably less volatile and scatter closely around the median rank of 303. Although these two groups still show an opposing dynamic, it is less

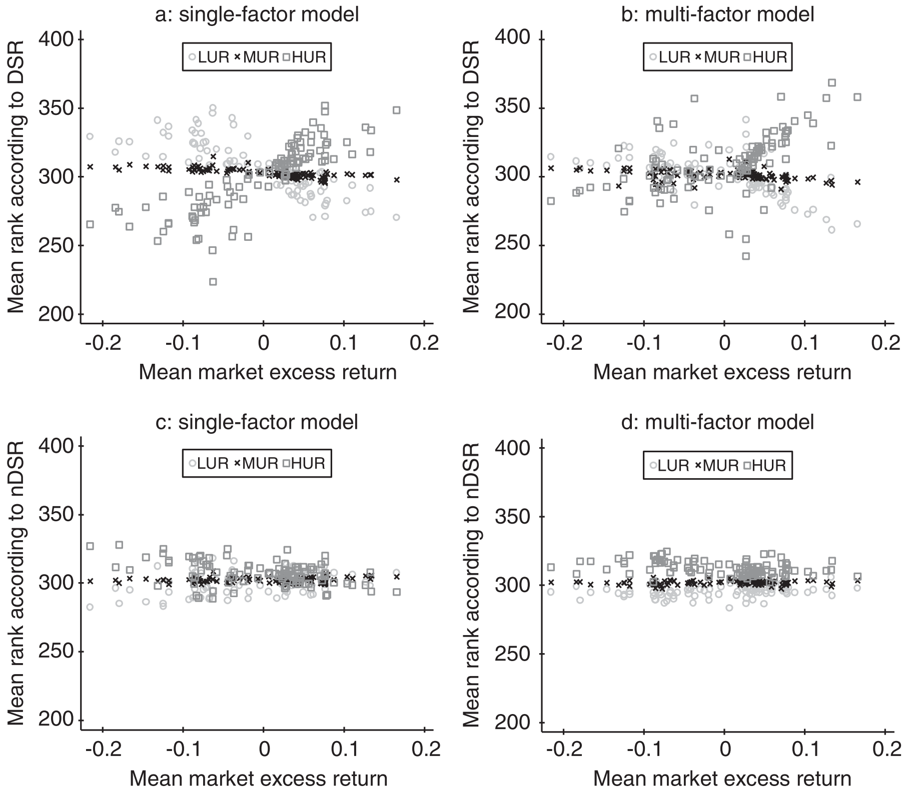

Figure 3: Mean fund ranks and the market climate.This figure shows scatter plots of mean fund ranks according to the DSR and nDSR of fund groups according to their PUR against mean market excess returns. SRs, fund-specific characteristics and mean market excess returns are estimated for 101 rolling windows of each 250 daily observations, which move forward by increments of 25 days. nSRs are calculated using these fund-specific characteristics and market parameters estimated for the long-term period from 1 July 1963 to 31 March 2011. We rank funds according to the DSR (nDSR) by assigning rank 1 to the fund with the highest DSR (nDSR) and rank 605 to the fund with the lowest DSR (nDSR).

pronounced and the differences between the two groups are now considerably smaller. The average rank of MUR funds hardly varies at all. In comparison with the original SR, the nSR yields more reliable inferences on fund performance, as it is less affected by the funds' PUR and less volatile over time.

The question remains to what extent the original SR and the nSRs are driven by market climates. To display this relation, Figure 3 depicts scatter plots of the mean rank according to the DSR and the nDSR for each group against the mean market excess return in the respective rolling window.

The mean rank according to the DSR (Figures 3a and b) is highly determined by the market climate. In bull markets, the HUR fund ranks are consistently worse and the LUR funds are consistently better. The opposite applies during bear markets, and the ranking of the MUR funds is hardly affected by market climates. Again these results hold for fund groups based on the single-factor model as well as the multi-factor model. The influence of the market climate on the mean rank of the groups mostly vanishes when funds are ranked according to their nDSR (Figures 3c and d). The scatter plots for the multi-factor model appear to be especially flat. Fund ranks according to the nSR are not biased by the contemporaneous market climate.

To further exemplify the matter, we analyse the DSR and its components during a bull, a bear and a normal market period.

Table 5: Differential SR during bull and bear markets

Fund group	*Bull market* ($\mu_m = 0.047$)			*Median market* ($\mu_m = 0.023$)			*Bear market* ($\mu_m = -0.067$)		
	DSR	*TRAP*	*CUR*	*DSR*	*TRAP*	*CUR*	*DSR*	*TRAP*	*CUR*
Panel A: Single-factor model									
LUR	−0.703	−0.533	−0.170	−0.363	−0.281	−0.082	0.293	0.196	0.097
MUR	−0.812	−0.405	−0.407	−0.383	−0.194	−0.189	0.255	0.031	0.224
HUR	−0.973	−0.162	−0.810	−0.292	0.080	−0.372	0.506	0.001	0.505
All funds	−0.823	−0.380	−0.443	−0.361	−0.155	−0.206	0.314	0.057	0.257
Panel B: Multi-factor model									
LUR	−0.338	−0.233	−0.105	0.025	0.075	−0.050	0.059	−0.001	0.060
MUR	−0.405	−0.183	−0.222	0.023	0.121	−0.098	0.057	−0.074	0.131
HUR	−0.519	0.014	−0.532	0.023	0.253	−0.230	0.056	−0.225	0.281
All funds	−0.415	−0.153	−0.263	0.023	0.139	−0.116	0.057	−0.091	0.148

Note: This table presents the mean DSR, TRAP and CUR for selected evaluation windows and for funds grouped according to their proportion of unsystematic risk while controlling for α. We select 3 out of 101 rolling windows for the analysis based on their mean market excess returns. Beginning 4 January 1999, these windows iteratively roll over by increments of 25 days with each window covering 250 consecutive trading days. 'Bull market' is the rolling window in which the 3rd quartile of the mean excess return of the market (23 December 2003 to 20 December 2004) occurs, 'Bear market' is the rolling window in which the 1st quartile of the mean excess return of the market (24 July 2007 to 18 July 2008) occurs and 'Median market' is the rolling window in which the median of the mean excess return of the market (30 January 2004 to 26 January 2005) occurs. All values are noted in per cent.

We define these periods according to the distribution of the mean market excess return over the 101 evaluation windows. The 'bull market' is the 250-trading-day window in which the 3rd quartile of the mean market excess return (23 December 2003 to 20 December 2004) occurs, the 'bear market' denotes the window covering the 1st quartile of the market excess return (24 July 2007 to 18 July 2008) and 'median market' describes the window containing the median value of the mean market excess returns (30 January 2004 to 26 January 2005). Table 5 presents the mean DSR and its components for these three windows and the fund groups according to the PUR.

In Panel A, the funds' mean DSR develop contrarily in the bull and the bear market. This reversal is caused by the CUR. In the bull market period, the HUR funds are highly punished for their unsystematic risk, whereas in the bear market, the HUR funds are rewarded for their unsystematic risk. This impact of the market climate outweighs their differential skills. In the bull market, the HUR group achieves the lowest mean DSR though it possesses the highest mean TRAP. In the median market, the differences in DSR are small, as the TRAP and the CUR mostly outweigh each other in each group. The results in Panel B are quite similar; however, here TRAP and CUR mostly outweigh each other in the median and the bear market. Fund rankings based on the original SR have little reliability even during modest market climates as 50 per cent of the 250-day windows in our rolling analysis were more bullish or bearish than the periods analysed here. We repeat the analysis for the same market periods using the nDSR. Table 6 shows the mean nDSR and its components for the fund groups.

In Panel A, a ranking based on the fund groups' mean nDSR reverses from the bull market to the bear market. This reversal is driven by the differences in skill of the groups. As nCUR is negative and similarly decreases with PUR in all three market climates, funds achieve a higher nDSR if the skill of the management as measured by nTRAP is sufficiently high to compensate for nCUR. The latter is the case in the bull and median market in Panel A. In the bear market, both components of the nDSR decrease with PUR. In Panel B, the LUR funds are on average superior according to

Table 6: Normalised differential SR during bull and bear markets

Fund group	Bull market ($\mu_m = 0.047$)			Median market ($\mu_m = 0.023$)			Bear market ($\mu_m = -0.067$)		
	nDSR	*nTRAP*	*nCUR*	*nDSR*	*nTRAP*	*nCUR*	*nDSR*	*nTRAP*	*nCUR*
Panel A: Single-factor model									
LUR	−0.417	−0.386	−0.030	−0.233	−0.203	−0.030	0.177	0.244	−0.067
MUR	−0.372	−0.298	−0.074	−0.214	−0.144	−0.070	−0.118	0.032	−0.150
HUR	−0.273	−0.117	−0.156	−0.081	0.063	−0.144	−0.303	0.016	−0.319
All funds	−0.361	−0.278	−0.083	−0.190	−0.113	−0.077	−0.097	0.071	−0.168
Panel B: Multi-factor model									
LUR	−0.129	−0.101	−0.027	0.136	0.163	−0.027	−0.054	−0.015	−0.040
MUR	−0.173	−0.107	−0.066	0.082	0.146	−0.064	−0.212	−0.116	−0.096
HUR	−0.185	−0.013	−0.171	0.035	0.193	−0.158	−0.518	−0.304	−0.215
All funds	−0.166	−0.087	−0.080	0.083	0.159	−0.076	−0.244	−0.135	−0.109

Note: This table presents the mean nDSR, nTRAP and nCUR for selected evaluation windows and for funds grouped according to their proportion of unsystematic risk while controlling for α. We select 3 out of 101 rolling windows for the analysis based on their mean market excess returns. Beginning 4 January 1999, these windows iteratively roll over by increments of 25 days with each window covering 250 consecutive trading days. 'Bull market' is the rolling window in which the 3rd quartile of the mean excess return of the market (23 December 2003 to 20 December 2004) occurs, 'Bear market' is the rolling window in which the first quartile of the mean excess return of the market (24 July 2007 to 18 July 2008) occurs and 'Median market' is the rolling window in which the median of the mean excess return of the market (30 January 2004 to 26 January 2005) occurs. All values are noted in per cent.

the nDSR in all market climates. Fund groups with higher unsystematic risk are unable to compensate for the associated punishment with a higher nTRAP in the bull and in the median market and achieve the least nTRAP in the bear market. However, the unsystematic risk hurts the funds' performance on a similar scale in all market climates. The nDSR is thus free from a market climate bias. While fund managers are punished for taking unsystematic risk (which follows elemental capital market theory as the market offers no premium for it), the magnitude of this punishment is irrespective of the market climate.

CONCLUSION

This study examines whether and to what extent market climates impact the performance of funds measured by the original SR. Defining fund-specific characteristics based on standard linear factor models, we analyse theoretically the impact of market climate on the SR. In particular, we found that the proportion of unsystematic risk of funds has a reverse impact on SRs in bearish and bullish market periods. Thus, the SR of a fund not only reflects the skill of the fund management, but also the market climate. Ranking funds according to the SR is subject to a market climate bias, even when the specific characteristics of the funds are stable over time.

The results of our empirical analysis confirm the practical relevance of this market climate bias. Initially, we show that fund rankings are systematically driven by market climates. This effect is robust against the factor model used and dominates the cross-section of fund ranks. We find poorly diversified funds to show superior ranks in declining markets, and vice versa. Subsequently, we repeat the analysis for rolling windows and find that in each window, the cross-section of the SR is highly driven by the proportion of unsystematic risk. Thus, we consider investor decisions based on the SR as unjustified. Investors should not, as it is currently the case, rely on the original SR in order to assess the performance of funds. Instead, we introduce the normalised SR, as this measure reflects the pure performance of fund management and is not affected by a market climate bias.

The market climate bias of the SR raises several interesting questions warranting additional research. The SR is not only biased for equity funds but generally for assets whose returns are the outcome of a linear factor model. Clearly, further studies of all applications of the SR are justified. Moreover, the results of existing empirical studies based on the original SR should be interpreted anew, taking the market climate bias into consideration. The normalised SR also provides new possibilities for forecasting future SRs of funds. In this context, one could empirically evaluate which form of normalisation would produce appropriate estimators for the future performance of funds. To do this, different underlying time frames will have to be evaluated, as well as models or methods to estimate the specific characteristics of funds and the distribution parameters of the market factors. Lastly, the conclusions presented here can be applied to other issues, such as merit-based reward for fund managers or risk-adjusted performance measurement of business units, for example, using RORAC or RAROC concepts in the banking industry.

ACKNOWLEDGEMENTS

Parts of this research were done while Marco Wilkens was visiting the Australian Graduate School of Management, University of New South Wales. He thanks Timothy Devinney and the academic and administrative staff for their hospitality and support. We are grateful to participants at research seminars at the University of Eichstätt-Ingolstadt; the University of Essex; the University of Melbourne; the University of New South Wales; the University of New Zealand; the University of Regensburg; the University of Tasmania; the University of Western Australia; the 46th Southern Finance Association Annual Meeting 2006, Destin; the 36th Annual Meeting of the Financial Management Association International 2006, Salt Lake City; the 15th European Financial Management Association Annual Meeting 2006, Madrid; the 68th Association of University Professors of Management Annual Meeting 2006, Dresden; the 9th Conference of the Swiss Society for Financial Market Research 2006, Zurich; the International Scientific Annual Conference Operations Research 2005, Bremen; and, especially, to Vikas Agarwal, Jerry Coakley, Klaus Eberl, Oliver Entrop, Norris L. Larrymore, Christoph Memmel, Stewart Myers and Marco Pagani for helpful comments and suggestions on earlier drafts of this article.

NOTES

1. We assume that the market is relatively μ-σ-efficient with respect to the fund's investment universe; see Grinblatt and Titman (1989).
2. This also applies to the funds we study empirically in Section 'Empirical relevance of the market climate bias'. Each fund has a positive market risk exposure, in the total evaluation period as well as in each of the rolling windows we analyse.
3. Instead of long-term estimates, an investor could also use the mean and standard deviation of market excess returns she expects during her planned investment period. This does not affect further results.
4. As asset allocations of funds differ, rankings according to the SR and the DSR are usually not identical in the multi-factor case.
5. These are the codes CA (capital appreciation), EI (equity income), G (growth), GI (growth income), MC (mid-cap) and SG (small growth).
6. The survivorship bias is studied in detail by Brown and Goetzmann (1995), Elton *et al* (1996), Carhart *et al* (2002) and Rohleder *et al* (2011).
7. All return observations that are larger than 100 per cent in absolute value were excluded.
8. Next to the analysis of equity funds, these approaches are also used to measure the market timing of hedge funds. See, for example, Gregoriou *et al* (2002), Gupta *et al* (2003) and Chen and Liang (2007).
9. The tests are based on heteroscedasticity and autocorrelation consistent standard errors according to Newey and West (1987).
10. We also tested for simultaneous timing activities in several market factors along the lines of Chan *et al* (2002) and Comer (2006) and find no evidence of such activities.
11. We thank Ken French for providing this data online under www.mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
12. This follows from the necessity of an active portfolio to deviate from the efficient market portfolio in order to be able to disclose performance; see Treynor and Black (1973).

REFERENCES

Brown, S.J. and Goetzmann, W.N. (1995) Performance persistence. *Journal of Finance* 50(2): 679–698.

Carhart, M.M. (1997) On persistence in mutual fund performance. *Journal of Finance* 52(1): 57–82.

Carhart, M.M., Carpenter, J.N., Lynch, A.W. and Musto, D.K. (2002) Mutual fund survivorship. *Review of Financial Studies* 15(5): 1439–1143.

Chan, L.K.C., Chen, H.L. and Lakonishok, J. (2002) On mutual fund investment styles. *Review of Financial Studies* 15(2): 1407–1437.

Chen, W.K., Chen, Y.J. and Chen, T.C. (2008) Using efficiency ratio to measure fund performance. *Journal of Asset Management* 8(6): 352–360.

Chen, Y. and Liang, B. (2007) Do market timing hedge funds time the market? *Journal of Financial and Quantitative Analysis* 42(4): 827–856.

Comer, G. (2006) Hybrid mutual funds and market timing performance. *Journal of Business* 79(2): 771–797.

Elton, E.J., Gruber, M.J. and Blake, C.R. (1996) Survivorship bias and mutual fund performance. *Review of Financial Studies* 9(4): 1097–1120.

Fama, E.F. and French, K.R. (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1): 3–56.

Ferruz, L.A. and Sarto, J.L.M. (2004) An analysis of Spanish investment fund performance: Some considerations concerning Sharpe's ratio. *Omega – The International Journal of Management Science* 32(4): 273–284.

Gregoriou, G.N., Rouah, F. and Sedzro, K. (2002) Market timing and security selection: The case of hedge funds. *Derivatives Use, Trading & Regulation* 8(2): 140–158.

Grinblatt, M. and Titman, S. (1989) Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies* 2(3): 393–421.

Gupta, B., Cerrahoglu, B. and Daglioglu, A. (2003) Evaluating hedge fund performance: Traditional versus conditional approaches. *Journal of Alternative Investments* 6(3): 7–24.

Henriksson, R.D. and Merton, R.C. (1981) On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. *Journal of Business* 54(4): 513–533.

Israelsen, C.L. (2003) Sharpening the Sharpe ratio. *Financial Planning* 33(1): 49–51.

Israelsen, C.L. (2005) A refinement to the Sharpe ratio and information ratio. *Journal of Asset Management* 5(6): 423–427.

Israelsen, C.L. (2009) Refining the Sharpe ratio. *Journal of Performance Measurement* 13(3): 23–27.

Jensen, M.C. (1968) The performance of mutual funds in the period 1945–1964. *Journal of Finance* 23(2): 389–416.

Jobson, J.D. and Korkie, B.M. (1981) Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance* 36(4): 889–908.

Lo, A.W. (2002) The statistics of Sharpe ratios. *Financial Analysts Journal* 58(4): 36–52.

Lobosco, A. (1999) Style/Risk-adjusted performance. *Journal of Portfolio Management* 25(3): 65–68.

McLeod, W. and van Vuuren, B. (2004) Interpreting the Sharpe ratio when excess returns are negative. *Investment Analysts Journal* 59(1): 15–20.

Modigliani, F. and Modigliani, L. (1997) Risk-adjusted performance: How to measure it and why. *Journal of Portfolio Management* 23(2): 45–54.

Newey, W.K. and West, K.D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3): 703–708.

Pástor, L. and Stambaugh, R.F. (2002) Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics* 63(3): 315–349.

Rohleder, M., Scholz, H. and Wilkens, M. (2011) Survivorship bias and mutual fund performance: Relevance, significance, and methodical differences. *Review of Finance* 15(2): 441–474.

Scholz, H. (2007) Refinements to the Sharpe ratio: Comparing alternatives for bear markets. *Journal of Asset Management* 7(5): 347–357.

Sharpe, W.F. (1966) Mutual fund performance. *Journal of Business* 39(1): 119–138.

Sharpe, W.F. (1975) Adjusting for risk in portfolio performance measurement. *Journal of Portfolio Management* 1(2): 29–34.

Sharpe, W.F. (1998) Morningstar's risk-adjusted ratings. *Financial Analysts Journal* 54(4): 21–33.

Tinic, S.M. and West, R.R. (1979) *Investing in Securities: An Efficient Markets Approach*. Reading, MA: Addison-Wesley.

Treynor, J.L. and Black, F. (1973) How to use security analysis to improve portfolio selection. *Journal of Business* 46(1): 66–86.

Treynor, J.L. and Mazuy, K.K. (1966) Can mutual funds outguess the market? *Harvard Business Review* 44(4): 131–136.

Vinod, H.D. and Morey, M.R. (2000) Confidence intervals and hypothesis testing for the Sharpe and Treynor performance measures. In: Y.S. Abu-Mostafe, B. LeBaron, A.W. Lo and A.S. Weigend (eds.) *Computational Finance 1999*. Cambridge and London: MIT Press, pp. 25–39.