Interpreting Sharpe Ratios – The Market Climate Bias

Hendrik Scholz
Catholic University of Eichstaett-Ingolstadt

Marco Wilkens
Catholic University of Eichstaett-Ingolstadt
and Australian Graduate School of Management

Working Paper
First Version: 2004-03-31
This Version: 2005-12-14

Key words: Sharpe ratio, performance measurement, bear market, market conditions
(JEL G11)
Interpreting Sharpe Ratios –

The Market Climate Bias

Abstract

This article adds new insights to the ongoing discussion of whether the Sharpe ratio is appropriate to assess the performance of funds in abnormal periods, e.g., when average excess returns of funds are negative. We show two main factors influencing the Sharpe ratio: first, of course, the quality of the fund management and second, the innovative aspect, the significance of the market climate during the evaluation period. Because of the influence of the market climate, a fund with constant fund-specific characteristics that outperforms the Sharpe ratio of the market index in a declining market will not necessarily have a superior Sharpe ratio in a normal market period. Moreover, based on US equity mutual fund data we discover that market climates have a considerable impact on fund rankings. In order to overcome this market climate bias we suggest using the “normalized” Sharpe ratio in performance measurement.

Key words: Sharpe ratio, performance measurement, bear market, market conditions

JEL classification: G11
1 Introduction

The Sharpe ratio, one of the most commonly cited statistics in financial analysis, has been used for almost 40 years to evaluate portfolio and mutual fund performance. Considering the worldwide declining stock prices at the beginning of this century, it is currently being discussed whether the Sharpe ratio is an appropriate performance measure even in bear markets. We broaden this question to ask whether the Sharpe ratio allows for reasonable rankings of funds in “non-normal” periods.

In general, this article deals with the influence of market climate on Sharpe ratios. The motivation is underlined by a brief overview in Section 2 of the ongoing discussion on applying the Sharpe ratio in declining markets. Based on a common factor model in performance analysis, in Section 3 we show that using the Sharpe ratio to measure mutual fund performance can – depending on a particular evaluation period – lead to different performance results compared to the market. Because of this “market climate bias” of the Sharpe ratio, investors cannot infer from an ex post measured superior Sharpe ratio of a fund what the corresponding future or expected Sharpe ratio will be. In Section 4, we confirm the practical relevance of the market climate impact on fund rankings based on US equity mutual fund data. Section 5 outlines a solution to this problem and draws the conclusion.

2 Controversy about using the Sharpe ratio in abnormal periods

The Sharpe ratio \( SR_i \) of a fund \( i \) is commonly determined by employing the mean \( (\bar{er}_i = \bar{r}_i - \bar{r}_f) \) and standard deviation \( (s_i) \) of the fund excess returns, which are calculated as difference between the total return of the fund \( r_i \) and a risk free short-term interest rate \( r_f \). Thus it measures the fund mean excess return per unit of standard deviation, upon which its original denotation in Sharpe (1966) as “reward-to-variability ratio” is to be attributed:

\[
SR_i = \frac{\bar{r}_i - \bar{r}_f}{s_i}
\]

In a return/standard deviation space, the Sharpe ratio can be interpreted as the slope of the line connecting the fund and the risk-free interest rate. For performance hypothesis testing with the Sharpe ratio, see Jobson and Korkie (1981) in conjunction with Memmel (2003).
Both, in finance practice as well as in academic literature, it is partly believed that Sharpe ratios do not offer reasonable rankings of funds in bear markets (see, e.g., Tinic and West (1979, p. 551), Jobson and Korkie (1981), and Israelsen (2005)). Often this is justified through a conventional wisdom about investments, which states that out of two funds with identical mean excess returns, the fund with the lower standard deviation exhibits a superior performance (Akeda (2003)). In declining markets, however, the Sharpe ratio leads to a reversed ranking. With identically negative excess returns, the fund with a higher total risk shows a higher (less negative) Sharpe ratio. Israelsen (2003) refers to such a ranking based on the Sharpe ratio as “The Negative ‘Excess Return’ Dilemma”. Additionally, Plantinga (1999, p. 24) states, “For negative Sharpe ratios, the ranking is consistent with a risk-seeking investor.”

This rejection of the Sharpe ratio stands opposed to a justification by Sharpe (1975 and 1998). They argue that it is not the positions of funds that are directly being compared, but the risk-adjusted combinations of each fund with borrowing or lending at the risk-free interest rate. Therefore, the highest Sharpe ratio is shown for a fund with the highest average excess return attainable for any risk level even in a bear market. The intuitive assessment of fund performance – namely that if returns are the same, lower risk is more advantageous – disappears against this background in declining markets (see, e.g., Lobosco (1999) and Akeda (2003)).

Sharpe ratios of funds are usually determined using ex post-estimated distribution parameters of the fund excess returns. With respect to the choice of funds for a future period, investors are faced with the question of whether the market climate during the respective evaluation period, e.g., a bear market, would significantly distort the advantageousness of funds based on the Sharpe ratio.

3 The market climate impact on the Sharpe ratio

To answer this question, we focus on fund-specific characteristics. For this purpose, we assume the quality as well as the aggressiveness of selection activities of fund managers as stable over time. Furthermore, according to a single factor model, we presume the excess return of a fund $i$ for period $t$ ($er_{it} = r_{it} - r_{ft}$) as being due to the excess return of the market index ($er_{Mt} = r_{Mt} - r_{ft}$):
On this basis, the fund management performance is reflected in the fund-specific characteristics $J_{Ai}$, $\beta_i$, and $s_{\varepsilon_i}$, which are generally determined via regression analysis. For a discussion of factor models see, e.g., Sharpe, Alexander and Bailey (1999, pp. 256-282).

The stability of fund-specific characteristics over time represents a common assumption in performance analysis, which also underlies other traditional performance measures such as the Treynor ratio and Jensen Alpha (see Treynor (1965) and Jensen (1968)). The beta $\beta_i$ denotes the systematic risk of the fund. In this connection, we assume that the index is relatively $\mu$-$\sigma$-efficient with respect to the fund investment universes (see Grinblatt/Titman (1989)). Successful selection activities yield a positive Jensen Alpha $J_{Ai}$ and vice versa. The associated unsystematic risk is reflected in the standard deviation $s_{\varepsilon_i}$ of the fund-specific residual return. Based on the factor model, the mean and the standard deviation of fund excess returns for a particular evaluation period are:

\begin{align}
(3) \quad \bar{er}_i &= J_{Ai} + \beta_i \bar{er}_M \\
(4) \quad s_i &= \sqrt{\beta_i^2 s_M^2 + s_{\varepsilon_i}^2}
\end{align}

For given fund-specific characteristics $J_{Ai}$, $\beta_i$ as well as $s_{\varepsilon_i}$, the Sharpe ratio is determined according to:

\begin{align}
(5) \quad SR_i &= \frac{J_{Ai} + \beta_i \bar{er}_M}{\sqrt{\beta_i^2 s_M^2 + s_{\varepsilon_i}^2}}
\end{align}

Equation (5) makes it clear that the Sharpe ratio of a fund depends on the one hand on the fund-specific characteristics $J_{Ai}$, $\beta_i$ as well as $s_{\varepsilon_i}$, and on the other hand on the mean ($\bar{er}_M$) and the standard deviation ($s_M$) of the market excess returns during the evaluation period.

This impact of the market climate on Sharpe ratios carries forward to other performance measures employing the standard deviation of fund excess returns as measure of risk (for a systematic overview of basic risk-adjusted performance measures, see Scholz and Wilkens (2005)).

The following depicts the influence of the characteristics of a notional fund on its position in return/standard deviation space based on two examples of 36-month periods for the US
market. The first time frame represents the extreme “bull market” from April 1995, to March 1998; the second displays the extraordinary “bear market” from April 2000, to March 2003 (see Table 1). The index used throughout this paper is the value-weighted index of all NYSE, AMEX, and NASDAQ stocks. These two time periods are the “market climate maximum” and the “market climate minimum” for 36-month time frames from January 1994, to June 2004 (compare Table 3).

Table 1: Monthly return statistics for two exemplary 36-month evaluation periods

<table>
<thead>
<tr>
<th></th>
<th>“Bull market H”</th>
<th>“Bear market B”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean excess return of the index</td>
<td>1.87 %</td>
<td>–1.64 %</td>
</tr>
<tr>
<td>Standard deviation of the index</td>
<td>3.28 %</td>
<td>5.49 %</td>
</tr>
<tr>
<td>Mean return of the one-month T-bill</td>
<td>0.43 %</td>
<td>0.28 %</td>
</tr>
</tbody>
</table>

Figure 1: The market climate bias of the Sharpe ratio

---

Index returns as well as the risk-free monthly T-bill returns are provided on Ken French’s Website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).
Initially, a notional fund A with a purely passive portfolio strategy will be considered. A corresponding strategy can be implemented through a combination of a market index and the risk-free interest rate \((J_A = 0, s_{E_i} = 0)\). In return/standard deviation space, the position of such a fund is located on the bold line connecting the index and the risk-free interest rate (see Figure 1). During the bull market, this connecting line exhibits a positive slope and during the bear market, a negative one. Position A, labeled □, results from fund A with a beta of \(\beta_A = 0.8\). Its Sharpe ratio corresponds, respectively, to the Sharpe ratio of the index. Taking exclusively systematic risk cannot lead to a superior or inferior risk-adjusted performance when compared to the index.

A second example, the passive fund B, exhibits a beta of \(\beta_B = 0.8\) as well. Moreover, the fund attains a negative Jensen Alpha of \(J_{AB} = -0.25\%\) due to management fees. This implies, according to Equation (3), a decrease in the fund average excess return of the same size. Starting from position A in both the bear market and the bull market, respectively, the negative Jensen Alpha leads to a shift in the position of the passive strategy downwards to either the □ labeled position B_H or B_B (see Figure 1). Therefore, a negative Jensen Alpha in each case implies a decrease in the Sharpe ratio of the fund, and vice versa.

An active fund C also exhibits systematic risk of \(\beta_C = 0.8\). This fund is engaged in selection activities and shows a negative Jensen Alpha of \(J_{AC} = -0.25\%\) as well as unsystematic risk of \(s_{E_C} = 4\%\). The unsystematic risk leads, according to Equation (4), to a standard deviation increase. Based on the above-mentioned positions of fund B, the higher overall risk shifts the fund position to the right, to C_H and C_B, respectively. During the bull market this shift is accompanied by an additional decrease in the Sharpe ratio. However, during the bear market a higher total risk causes a higher (less negative) Sharpe ratio. In our example, this positive effect of the unsystematic risk even compensates for the decrease of the Sharpe ratio by the negative Jensen Alpha. Fund C attains a superior Sharpe ratio compared to the market index during the bear market.

Therefore, investors cannot infer from an ex post Sharpe ratio – even when a fund has constant characteristics – how that fund would perform in the future. This holds true especially when a high Sharpe ratio for the underlying bear market can primarily be attributed to the high unsystematic risk of the fund. The influence of the market climate
during a particular evaluation period leads to the conclusion that poorly diversified funds exhibit relatively high Sharpe ratios in a declining market, and vice versa.

4 The market climate impact on mutual fund rankings

The following analysis employs monthly returns of all US equity “large funds” with a complete data history from January 1994, until June 2004 in the Morningstar data base\(^2\) (see Reichenstein (2004) for the Morningstar classification of funds).\(^3\) For each of the 532 equity mutual funds observed, there are 126 realized monthly returns. Typically, we are dealing with total returns including reinvestments of all distributions (e.g. dividends), but disregarding load charges. Linear regression analysis of monthly excess returns of funds, compared with the excess returns of the market index according to Equation (2), yield the fund-specific characteristics summarized in Table 2.

Table 2: Fund-specific characteristics of US equity “large funds” from January 1994, to June 2004

<table>
<thead>
<tr>
<th></th>
<th>Jensen Alpha</th>
<th>Beta</th>
<th>Standard deviation of term (\varepsilon)</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>0.64 %</td>
<td>1.63</td>
<td>5.75 %</td>
<td>99.43 %</td>
</tr>
<tr>
<td>2/3 quantile</td>
<td>0.05 %</td>
<td>0.96</td>
<td>2.19 %</td>
<td>88.57 %</td>
</tr>
<tr>
<td>Median</td>
<td>–0.02 %</td>
<td>0.90</td>
<td>1.89 %</td>
<td>84.47 %</td>
</tr>
<tr>
<td>1/3 quantile</td>
<td>–0.10 %</td>
<td>0.83</td>
<td>1.56 %</td>
<td>77.46 %</td>
</tr>
<tr>
<td>Minimum</td>
<td>–1.16 %</td>
<td>0.32</td>
<td>0.34 %</td>
<td>19.37 %</td>
</tr>
<tr>
<td>Mean</td>
<td>–0.03 %</td>
<td>0.92</td>
<td>1.96 %</td>
<td>80.51 %</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.20 %</td>
<td>0.20</td>
<td>0.75 %</td>
<td>13.15 %</td>
</tr>
</tbody>
</table>

In Section 3, we pointed out that stability of fund-specific characteristics is commonly assumed when the characteristics are determined according to (2). Therefore, we test if these funds perform timing activities. Successful timing activities are identified by an increase (decrease) of the systematic risk of funds in above-average positive (negative) market climates. Based on the squared-regression approach by Treynor and Mazuy (1966), at 5

\(^2\) We thank Morningstar Inc. for providing us with the mutual fund data.

\(^3\) The data set points to a survivorship bias, which leads to a biased average performance of funds compared with the market. Since we are specifically analyzing the changes of the Sharpe ratios of individual funds and the rankings of funds, our analyses are not sensitive to survivorship issues. See, for example, Brown and Goetzmann (1995) and Elton, Gruber and Blake (1996) for survivorship bias.
percent level, we can verify that only 10 funds – or 1.88 percent – show significant timing activities. Using the dummy variable regression approach by Henriksson and Merton (1981), only 6 funds, or 1.13 percent, lead to the same outcome. The two-tailed t-tests are based on standard errors corrected for heteroskedasticity and autocorrelation according to Newey and West (1987). As a result, timing activities cannot be verified for most funds in our data set. Therefore, timing activities of funds should not be a serious problem for our empirical study.

Employing fund-specific characteristics, we determine the Sharpe ratios according to (5) for five time frames, these being representative of the entire time period starting January 1994, to June 2004. Presuming constant characteristics over time makes it possible to separately consider the impact of market climates on the Sharpe ratios of funds. This assumption would be unrealistic if timing activities of funds were determined in the tests mentioned above.

In order to select five representative periods for the overall evaluation period, we calculate average monthly excess returns of the market index for 91 time frames. Starting January 1994, these time frames are defined as 36-month periods which are rolled over monthly ending December 1996, to June 2004. The market excess returns represent – as shown in Section 3 – a strong market climate effect on the Sharpe ratios of funds. With the aim of displaying the entire spectrum of market climates, we order the 91 time frames and select the ones representing the maximum, median, minimum as well as the 2/3 quantile, and 1/3 quantile of the average market excess returns. The corresponding values for the average and the standard deviation of excess returns as well as the Sharpe ratios of the market index are reproduced in Table 3.

### Table 3: Representative market climates for 36-month time frames starting from January 1994, to June 2004

<table>
<thead>
<tr>
<th>Parameters of the market index</th>
<th>“Market climate maximum”</th>
<th>“Market climate 2/3 quantile”</th>
<th>“Market climate median”</th>
<th>“Market climate 1/3 quantile”</th>
<th>“Market climate minimum”</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{e}_M )</td>
<td>1.87 %</td>
<td>1.40 %</td>
<td>0.94 %</td>
<td>–0.05 %</td>
<td>–1.64 %</td>
</tr>
<tr>
<td>( s_M )</td>
<td>3.28 %</td>
<td>4.91 %</td>
<td>5.20 %</td>
<td>4.79 %</td>
<td>5.49 %</td>
</tr>
<tr>
<td>( SR_M )</td>
<td>56.98 %</td>
<td>28.49 %</td>
<td>18.12 %</td>
<td>–1.02 %</td>
<td>–29.84 %</td>
</tr>
</tbody>
</table>

Using fund-specific characteristics, according to (5), we calculate Sharpe ratios for each fund in our data sample for these five market climates. Subsequently, we determine the resulting fund ranking for each market climate. Since the unchanged fund-specific characteristics for
all market climates suggest identical rankings, all differences can be attributed solely to the influence of the market climate.

Table 4 shows the Bravais-Pearson correlation coefficients between the Sharpe ratios of the funds as well as the corresponding Spearman rank correlation coefficients. Evidently, market climates have a considerable impact on the fund rankings in practice. Thus, between the “maximum” and “minimum” phase, the Spearman rank and the Bravais-Pearson correlation coefficients are only 0.05 and 0.15, respectively. Additionally, between the “1/3 quantile” and the “2/3 quantile” market climate, the Sharpe ratios and the fund rankings vary from one another by a correlation coefficient of 0.84 and 0.85, respectively.

Significance tests of correlation coefficients (see, e.g., Altman (1991, pp. 295-296)) specified in Table 4 show that at 0.1 percent level almost all coefficients are significant, i.e. not equal to zero (see the *** marked coefficients in Table 4). Only the Bravais-Pearson correlation coefficient between the “maximum” and “minimum” phase yields a p-value of 0.21, thus we cannot reject the null hypothesis of this correlation coefficient being equal to zero.

Table 4: Spearman rank correlation coefficients (upper triangular matrix) and Bravais-Pearson correlation coefficients (lower triangular matrix) between the Sharpe ratios of funds for different market climates

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>1/3 quantile</th>
<th>Median</th>
<th>2/3 quantile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>1</td>
<td>0.80 ***</td>
<td>0.66 ***</td>
<td>0.41 ***</td>
<td>0.15 ***</td>
</tr>
<tr>
<td>1/3 quantile</td>
<td>0.82 ***</td>
<td>1</td>
<td>0.97 ***</td>
<td>0.85 ***</td>
<td>0.67 ***</td>
</tr>
<tr>
<td>Median</td>
<td>0.67 ***</td>
<td>0.97 ***</td>
<td>1</td>
<td>0.94 ***</td>
<td>0.80 ***</td>
</tr>
<tr>
<td>2/3 quantile</td>
<td>0.39 ***</td>
<td>0.84 ***</td>
<td>0.94 ***</td>
<td>1</td>
<td>0.95 ***</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.05 ***</td>
<td>0.61 ***</td>
<td>0.77 ***</td>
<td>0.94 ***</td>
<td>1</td>
</tr>
</tbody>
</table>

H₀: coefficient = 0. Values significant at the 0.1 percent level are marked ***.
H₀: coefficient ≥ 0.85. At the 1 percent or 0.1 percent level significant values are marked ++ and +++, respectively.

Evidence of the market climate impact on Sharpe ratios and fund rankings is already given when the correlation coefficients are significantly smaller than one. Therefore, as an example, we test the null hypothesis of correlation coefficients greater than or equal to 0.85. This test is based upon Fisher’s z transformation of correlation coefficients. See, for example, Altman (1991, pp. 293-295).

Table 4 shows coefficients significantly smaller than 0.85, which are, at the 0.1 percent and 1 percent level, denoted with +++ and ++, respectively. Apparently, diverse coefficients are
significant. This outcome verifies that there is a considerable market climate influence on Sharpe ratios and fund rankings, especially during varying market climates.

These findings make it clear that the market climate impact on Sharpe Ratios is eminently relevant in abnormal market climates, such as during exceptional bear and bull markets. In such climates, one cannot draw meaningful conclusions about the management performance of funds derived from original Sharpe ratios and resulting rankings owing to a potentially dominant market climate bias.

5 Conclusion and outlook

The key contribution of this article is our finding that the Sharpe ratio is determined not only by the efficiency of fund managers but also by the respective market climate. The first influence is well known, of course; the second one is an innovative aspect that could have a large impact on performance measurement both in theoretical and empirical research. The general idea was shown in a theoretical analysis based on a single factor model. We show how a fund with constant fund-specific characteristics over time compared to the market index exhibits a superior Sharpe ratio in a declining market and an inferior Sharpe ratio in a rising market.

The cause for this is the influence of the unsystematic risk of funds on their Sharpe ratios subject to the market climate during a particular evaluation period. As shown, poorly diversified funds have relatively high Sharpe ratios in bear markets and vice versa. Based on this market climate bias, an appraisal of funds according to their Sharpe ratios, e.g., during a declining market, will lead to a biased evaluation of the “true” fund management performance. The analysis in Section 4 confirmed the practical relevance of the market climate bias for fund rankings. Based on fund-specific characteristics of 532 US equity mutual funds, we found that different market climates can lead to significantly diverging rankings of funds.

However, to assess the “pure” fund management performance, the original Sharpe ratio needs to be adjusted for the market climate. Therefore, we suggest a separate calculation of the specific characteristics of funds and the distribution parameters of the market excess returns. While the characteristics of funds can be still estimated based on relatively short time frames, one should use longer evaluation periods for estimating the distribution parameter of the
market excess returns in order to overcome the market climate bias. Based on this data, a “normalized” Sharpe ratio can be determined according to Equation (5). This performance figure measures the pure performance of fund management. Therefore, it should be employed in corresponding (scientific) inquiries in the future.
References


