One-Size or Tailor-Made Performance Ratios for Ranking Hedge Funds

Martin Eling, Simone Farinelli, Damiano Rossello und Luisa Tibiletti

Preprint Series: 2009-15

Fakultät für Mathematik und Wirtschaftswissenschaften
UNIVERSITÄT ULM

One-Size or Tailor-Made Performance Ratios for Ranking Hedge Funds?

Martin Eling
Institute of Insurance, Ulm University, Helmholtzstr. 22, 89081 Ulm, Germany. Email: martin.eling@uni-ulm.de

Simone Farinelli
Credit and Country Risk Control, UBS, P.O. Box, 8098 Zurich, Switzerland. Email: simone.farinelli@ubs.com

Damiano Rossello
Department of Economics and Quantitative Methods, University of Catania, Corso Italia, 55 – 95129 Catania, Italy. Email: rossello@unict.it

Luisa Tibiletti (*Corresponding Author*)
Department of Statistics and Mathematics, University of Torino, Corso Unione Sovietica, 218/bis – 10134 Torino, Italy. Email: luisa.tibiletti@unito.it

Abstract: Eling and Schuhmacher (2007) compared the Sharpe ratio with other performance measures and found virtually identical rank ordering using hedge fund data. They conclude that the choice of performance measure has no critical influence on fund evaluation and that the Sharpe ratio is generally adequate for analyzing hedge funds. Nevertheless, their analysis does not include the class of tailor-made performance ratios capable of being personalized to investment style as presented by Farinelli et al. (2009). Specifically, we deal with the Sortino-Satchell, Farinelli-Tibiletti, and Rachev ratios. Consider a large international hedge fund dataset, empirical experiments illustrate that if the ratios are tailored to a moderate investment style, they lead to rankings not too dissimilar to those found with the Sharpe ratio. But when the Rachev and Farinelli-Tibiletti ratios are used to describe aggressive investment styles rank correlations with the Sharpe ratio shrink drastically.

JEL Classification: D81, G10, G11, G23, G29

Keywords: Performance, Hedge funds, Sharpe ratio, Tailor-made performance ratios

1. Introduction

Whether the Sharpe ratio is an appropriate performance index for ranking financial products remains a controversial question among both academics and practitioners. The academic criticisms of the ratio are well known: although a trade-off ratio based on mean and variance is fully compatible with normally distributed returns (or, in general, with elliptical returns), it may lead to incorrect evaluations when returns exhibit heavy tails (see, e.g., Leland, 1999; Bernardo and Ledoit, 2000; Campbell and Kräussl, 2007). During the last two decades, numerous alternative ratios have been proposed in the literature (see, e.g., Biglova et al., 2004; Menn et al., 2005 among others); a recent review has classified more than 100 of them (see Cogneau and Hubner, 2009). Nevertheless, more sophisticated tools require more information for proper implementation, and this is not without cost, both in time and effort. Thus, a crucial question faced by many practitioners is to quantify the difference in results between sophisticated decision aid systems and the classical Sharpe ratio.

A first attempt to answer this question has been carried out by Eling and Schuhmacher (2007) using hedge fund data. They show that the rank correlations between a set of different performance ratios based on downside risk measures and the Sharpe ratio are virtually equal to 1, leading to the apparent conclusion that the Sharpe ratio is an appropriate performance measure for hedge funds. Nevertheless, the above-mentioned research does not take into account the class of tailor-made ratios, as considered in Farinelli et al. (2009). These measures are especially relevant for investors in hedge funds (i.e., sophisticated investors, such as pension funds and endowments; see Agarwal and Naik, 2004) that might seek a different risk profile compared to, for example, investors in mutual funds. Fitting a performance measure to investor preferences is exactly what tailor-made performance ratios accomplish.

The aim of the present work is to go a step further than previous research and study possible mismatching between the Sharpe ratio and tailor-made ratios. Specifically, we focus on the families of Sortino-Satchell (see Sortino and Satchell, 2001), Farinelli-Tibiletti (see Farinelli and Tibiletti, 2003, 2008), and Rachev ratios (see Biglova et al., 2004). Parameters in these ratios allow flexibility in the choice of which sector of the return distribution is focused on and create ratios tailored to the financial products under consideration and/or the investor risk profile. For the purpose of our analysis, we consider a large international database consisting of 4,048 hedge funds.

Our empirical analysis confirms Eling and Schuhmacher's (2007) results. When the tailor-made ratios describe moderate investment styles and the quantitative analysis concerns the entire return distribution (as in the case of the ratios analyzed in Eling and Schuhmacher, 2007), rankings are not too dissimilar to those established with the Sharpe ratio. However, as the parameters move to extreme values, making the ratios tailored to more aggressive investment styles, discrepancies with the Sharpe ratio ranking can be observed. As expected, the most discordant results are achieved for aggressive Rachev ratios, where only extreme tail events are taken under consideration.

The remainder of this paper is organized as follows. Section 2 provides an overview of the tailor-made performance ratios. The empirical investigation is presented in Section 3. We conclude in Section 4.

2. Tailor-made Performance Ratios

The challenge in ranking financial prospects is to choose a ratio that is not only able to discover the best return/risk tradeoff but also matches the investor goals and/or the investment style of the financial products under consideration. The use of tailor-made performance ratios just seems to hit the target. In the following we will deal with the one-size *Sharpe ratio* and tailor-made performance ratios belonging to the following three families: the *Sortino-Satchell*, the *Farinelli-Tibiletti*, and the *Rachev* ratios. We first define these various ratios as they will be used throughout this paper.

The classical *Sharpe ratio* can be calculated as (see Sharpe, 1966; see Haselmann/Herwartz, 2008, for an application in a currency hedging context):

$$\Phi_{Sharpe}\left(r;r_f\right)=\frac{E\left(r-r_f\right)}{\sigma\left(r-r_f\right)}, \qquad (1)$$

where σ denotes the standard deviation and r_f is the free-risk monthly interest rate. Using the standard deviation as a measure of risk means that upside and downside deviations to the benchmark are equally weighted. Therefore, this ratio is a good match for investors with a moderate investment style whose main concern is controlling the stability of returns around the benchmark. Its use may is questionable, however, if the investment style is more aggressive and focused on the tradeoff between large favorable/unfavorable deviations from the benchmark.

The *Sortino-Satchell ratio* is defined as:

$$\Phi_{Sortino-Satchell}\left(r;r_f\right)=\frac{E\left(r-r_f\right)}{E^{1/q}\left[\left(r-r_f\right)^{-}\right]^{q}}, \qquad (2)$$

with $q>0$. This ratio substitutes the standard deviation as a measure of risk with the left partial moment of order *q*; therefore, the only penalizing volatility is the “harmful” one below the benchmark. The original Sortino-Satchell ratio (see Sortino and Satchell, 2001) is defined for $q = 2$, then the ratio has been extended to $q \geq 1$ (see Biglova et al., 2004; Rachev et al., 2008) and, more recently, to $q > 0$ (see Farinelli and Tibiletti, 2008, and Farinelli et al., 2009).

The *Farinelli-Tibiletti ratio* (see Farinelli and Tibiletti, 2003, 2008; Menn et al., 2005, pp. 208–209) can be calculated as:

$$\Phi_{Farinelli-Tibiletti}\left(r, p, q; r_f\right) = \frac{E^{1/p}\left[\left(r - r_f\right)^{+}\right]^{p}}{E^{1/q}\left[\left(r - r_f\right)^{-}\right]^{q}}, \tag{3}$$

and $p, q > 0$. If $p = q = 1$, the index reduces to the so-called Omega index introduced in Keating and Shadwick (2002).

The parameters p and q can be balanced to match the agent's attitude toward the consequences of overperforming or underperforming. It is known (see Fishburn, 1977) that the higher p and q, the higher the agent's preference for (in the case of expected gains, parameter p) or dislike of (in the case of expected losses, parameter q) extreme events. If the agent's main concern is that the investment fund might miss the target, without particular regard to by how much, then a small value (i.e., $0 < q < 1$) for the left order is appropriate. However, if small deviations below the benchmark are relatively harmless compared to large deviations (catastrophic events), then a large value (i.e., $q>1$) for the left order is recommended. The right order p is chosen analogously and should capture the relative appreciation for outcomes above the benchmark.

Instead of measuring over- and underperformance with respect to the benchmark, *Rachev ratios* (see Biglova *et al.*, 2004) draw attention to extreme events. The ratio is defined as follows:

$$\Phi_{Rachev}\left(r, \alpha, \beta; r_f\right) = \frac{E\left[r_f - r \,\middle|\, r_f - r \geq VaR_{\alpha\%}\left(r_f - r\right)\right]}{E\left[r - r_f \,\middle|\, r - r_f \geq VaR_{\beta\%}\left(r - r_f\right)\right]}, \tag{4}$$

with $\alpha, \beta \in (0,1)$ and $VaR_c(x) := -\inf\{z | P(x \le z) > c\}$ interpreted as the smallest value to be added to the random profit and loss x to avoid negative results with probability at least $1-c$. Formula (4) is related to the expected shortfall $\mathrm{ES}_c(x) = -E\left[x | x \le -VaR_{c\%}(x)\right]$ also known as tail conditional expectation or Conditional *VaR* (*CVaR*) (see Acerbi and Tasche, 2002): it measures the expected value of profit and loss, given that the *VaR* has been exceeded. By changing the sign in the *ES*, the Rachev ratio can be interpreted as the ratio of the expected tail return above a certain level, i.e., the $VaR_{\alpha\%}$ divided by the expected tail loss below a certain level, i.e., the $VaR_{\beta\%}$. In other words, this ratio awards extreme returns adjusted for extreme losses. The STARR ratio (also called *CVaR* ratio, see Favre and Galeano, 2002; Martin et al., 2003) is a special case of the Rachev ratio. For example, STARR(5%) = Rachev ratio with (α, β): = (1, 0.05). We analyze the Rachev ratio for different parameters α and β; the lower they are, the more the focus is concentrated on the extreme tails.

In conclusion, by properly balancing parameters p, q, α, and β, we can tailor the ratios to investor style and/or capture different features of the financial products under consideration. As the parameters tend toward the extreme, the correspondent ratios shift to describe a more "extreme" investment style. Specifically, if our goal is to focus on extreme events at the tails (high stakes/huge losses), thus needing an aggressive ratio, parameters p and q in the Farinelli-Tibiletti ratios are fixed at high values, whereas parameters α and β in the Rachev ratios are fixed at low values.

3. Empirical Analysis

3.1. Data and Methodology

We consider hedge fund data provided by the Center for International Securities and Derivatives Markets (CISDM). We decided not to employ the hedge fund data that Eling and Schuhmacher (2007) used in their analysis because the CISDM database is larger and its use more wide-spread.[1]

The database contains 4,048 hedge funds reporting monthly returns, net of fees, for the time period of January 1996 to December 2005. Table 1 contains descriptive statistics on the return distributions of the hedge funds. On the basis of the Jarque-Bera test, the assumption of normally distributed hedge fund returns must be rejected for 37.67% (43.60%) of the funds at the 1% (5%) significance level.

Fund	Mean	Median	Standard deviation	Minimum	Maximum
Mean value (%)	0.97	0.86	1.48	-18.96	19.58
Standard deviation (%)	4.37	3.01	4.32	0.03	49.50
Skewness	0.01	0.00	1.15	-9.21	6.23
Excess kurtosis	2.45	0.91	6.13	-4.71	95.00

Table 1: Descriptive statistics for 4,048 hedge fund return distributions

[1] The CISDM database has been the subject of many academic studies; see, e.g., Capocci and Hübner, (2004); Ding and Shawky (2007). We also conducted the empirical analysis with the ehedge database used in Eling and Schuhmacher (2007) and found that our results are robust with regard to a variation of the dataset (results are available under request). Eling and Schuhmacher (2007) also used the CISDM data (analyzed in this paper) as a robustness test in their study in order to see whether their finding is driven by the dataset used; they found robust results as well which confirms our findings. Note that the standard deviation of the monhtly returns is on average higher in the CISDM database (4.37%) compared to the ehedge database (3.18%; see Eling and Schuhmacher, 2007, p. 2638). Possible explanations for this might be differences in database composition and investigation period.

The findings reported in the following Section were generated by first using the measures presented in Section 2 to determine hedge fund performance. To produce results comparable to those of Eling and Schuhmacher (2007), we chose a minimal acceptable return equal to the risk-free monthly interest rate (r_f) of 0.35%. Next, for each performance measure, the funds were ranked on the basis of the measured values. Finally, the rank correlations between the performance measures were calculated. This research design is of high relevance, as the performance of funds is regularly ranked on basis of risk-adjusted performance measures in order to benchmark the success of the fund compared with that of other funds and to serve as the basis for investment decisions.

A large number of different parameter combinations were included in the analysis: For the Sortino-Satchell ratio, the parameter q is varied between 0.01 and 10. For the Farinelli-Tibiletti ratio, the parameters p and q are both varied between 0.01 and 10. For the Rachev ratio, the parameters α and β are varied between 0.1% and 90%.

3.2. Findings

Figure 1 presents the rank correlation between the ranking resulting from the Sharpe ratio and that of the Sortino-Satchell ratio for different parameters q.

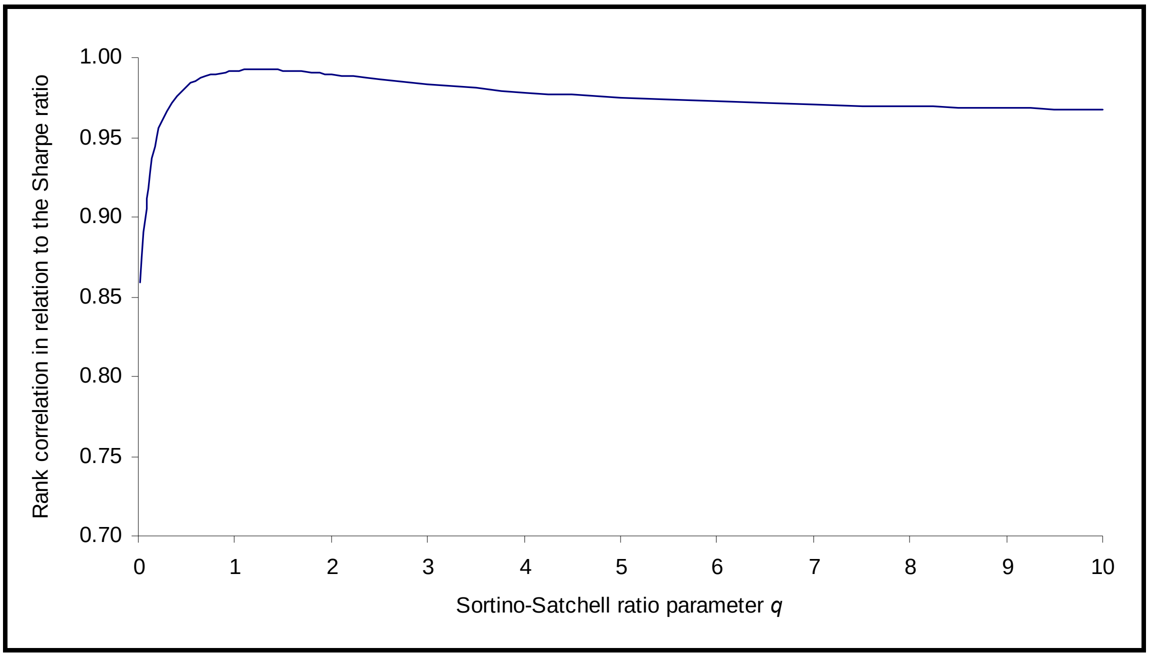

Figure 1: Sortino-Satchell ratio

The value assigned to parameter q appears to have little effect on the hedge fund ranking. In fact, the rank correlation is relatively close to 1 and a kind of lower bound with high values of q seems to exist with a rank correlation about 0.965 (this lower bound is confirmed by an analysis of higher values for q that is available upon request). For the original Sortino-Satchell ratio (q = 2), the rank correlation is 0.98, which confirms the high rank correlation found by Eling and Schuhmacher (2007) for this measure. Note that it is not common to consider values for q much lower than 1. For example, Fishburn (1977) reports, that in practice, values for q range from slightly less than 1 to 4, while Farinelli et al. (2009) use a value of q = 0.8 to describe an aggressive investor, and a value of q = 2.5 for a conservative investor. For all these values of q, the rank correlations are very close to 1. This is convincing evidence that the Sortino-Satchell and the Sharpe ratios lead to similar rankings.

Next, the Farinelli-Tibiletti ratio is analyzed. The upper part of Figure 2 presents the rank correlation between the Sharpe ratio and the Farinelli-Tibiletti ratio depending on the parameter p

(with q = 1); the lower part of the figure shows the rank correlation depending on the parameter q (with p = 1).

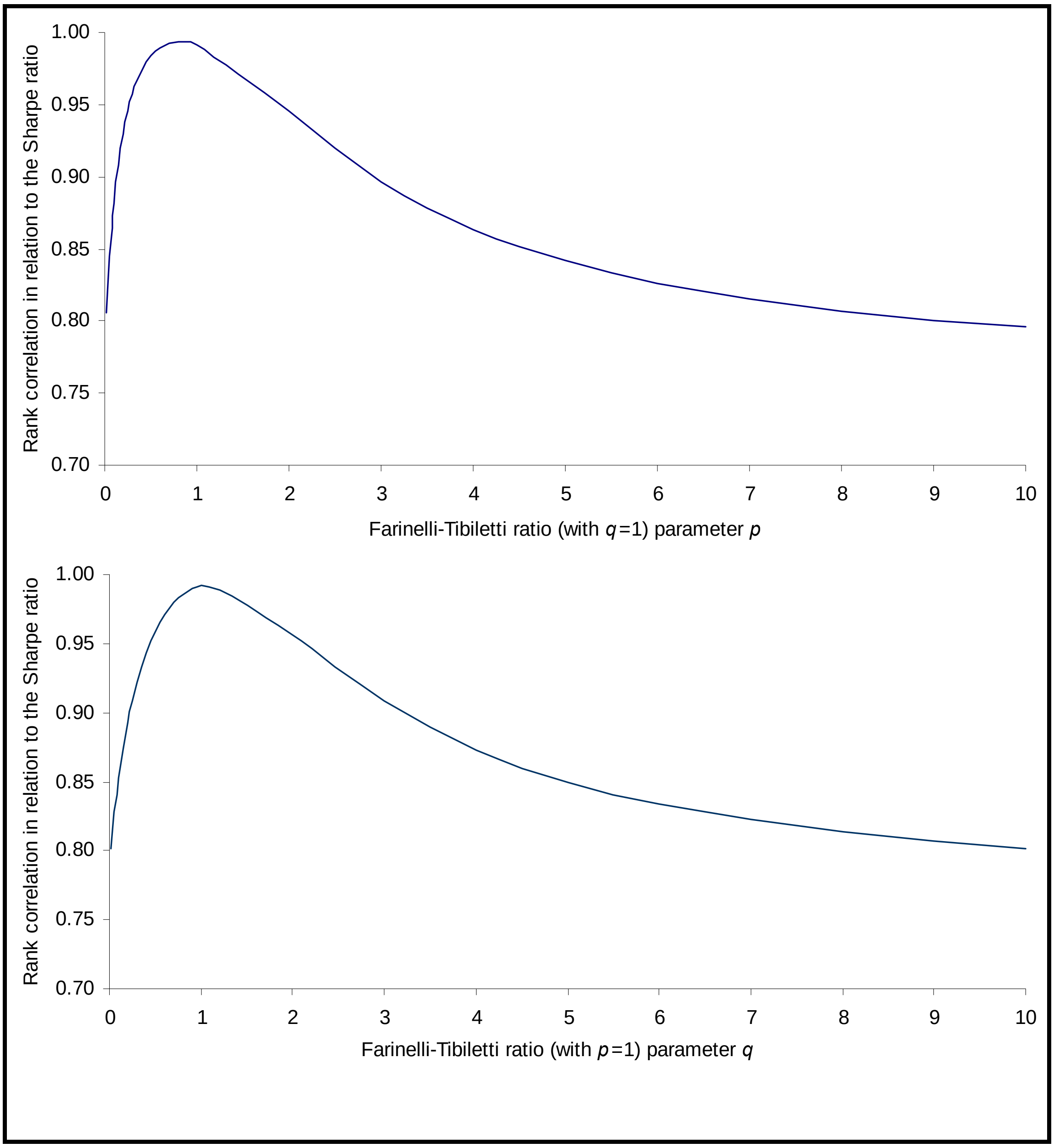

Figure 2: Farinelli-Tibiletti ratio (upper part: $0.01 < p < 10$, with $q = 1$; lower part: $0.01 < q < 10$, with $p = 1$)

Again, our results are in line with conjectures deriving from the study of the influence of the parameters (see Fishburn, 1977). As expected, the highest rank correlations occur for values of p

close to 1. For $p = q = 1$, the Farinelli-Tibiletti ratio coincides with the Omega index, which Eling and Schuhmacher (2007) showed to produce rankings similar to those derived by the Sharpe ratio.

Figure 3 presents rank correlations between the Sharpe ratio and the Farinelli-Tibiletti ratio for different combinations of p and q (the kink at $p, q = 1$ is due to the different scaling between $0.01 < p, q < 1$ and $1 < p, q < 10$).

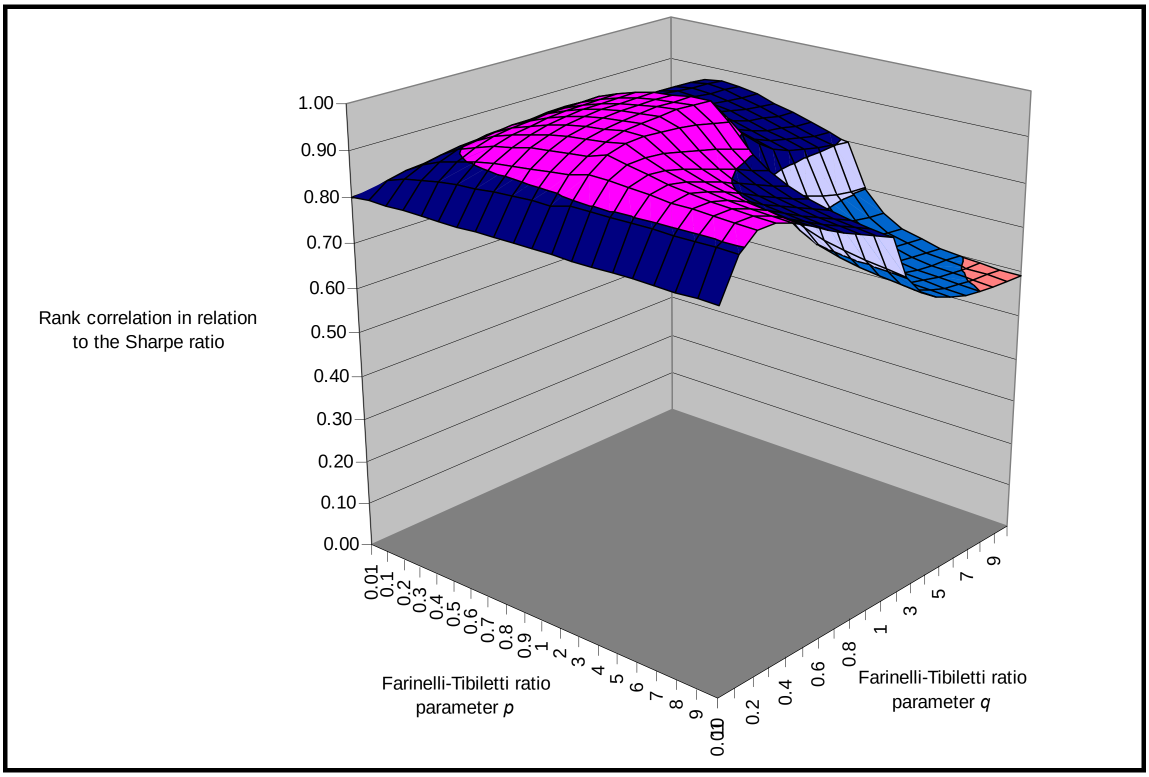

Figure 3: Farinelli-Tibiletti ratio ($0.01 < p, q < 10$)

The Farinelli-Tibiletti ratio is more sensitive to rank correlations than the Sortino-Satchell ratio, but still provides relatively high values, especially for reasonable values of p and q. For example, Farinelli et al. (2009) use values of $p = 2.8$ and $q = 0.8$ to describe an aggressive investor, which here results in a rank correlation of 0.92 to the Sharpe ratio. A conservative investor is described by $p = 0.8$ and $q = 2.5$, which gives a rank correlation of 0.95.[2] In both cases, the pa-

[2] These values are in the range of the high rank correlations found by Eling and Schuhmacher (2007); this finding therefore again confirms the results presented by them.

rameters are chosen according to Fishburn (1977) and expected utility theory, i.e., conservative ($p < 1$, $q > 1$) and aggressive ($p > 1$, $q < 1$). If p (<1) tends toward 0, the ratio assumes a conservative investor most interested in gaining small returns rather than seeking high stakes. According to Fishburn (1977), a conservative ratio should express aversion to high losses, so the parameter shaping an attitude toward negative returns should be $q > 1$. Conversely, as $p > 1$ increases, the ratio describes a more aggressive investor hoping to profit from a high-stakes strategy. Therefore, an aggressive ratio with $p > 1$ should show indifference to high losses, thus $q < 1$.

However, the Farinelli-Tibiletti ratio is a flexible tool that can be used in various ways. The parameters can be chosen so that the ratio can be read as the tradeoff between moderate gain/moderate risk or between high stakes/huge losses. In such a case, p and q go hand in hand , i.e., $p < 1$ goes with $q < 1$, and $p > 1$ goes with $q > 1$. The ratio can then be interpreted as the price of one unit of return for one unit of loss, where returns and losses are weighted by p and q. As the ratio moves to extreme investment styles, rank correlations with the moderate Sharpe ratio decrease. This is most evident for p and q close to 10, where the rank correlation falls to 0.59. It is worth noting that this occurs in correspondence with the case where the ratio detects the tradeoff between high stakes/huge losses.

Finally, we consider the Rachev ratio. Figure 4 shows the rank correlation between the Sharpe ratio and the Rachev ratio for different combinations of the parameters α and β.

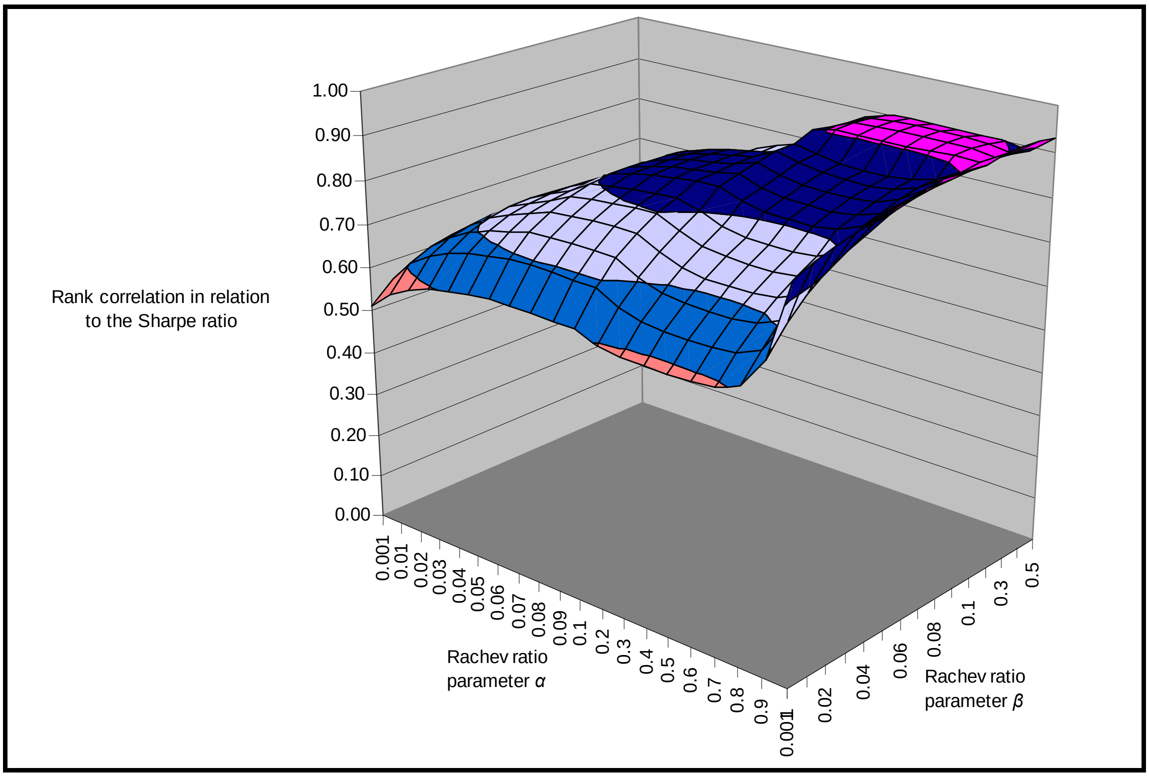

Figure 4: Rachev ratio

Among the tailor-made ratios, Rachev ratios are the most different from the Sharpe ratio. In fact, in all previous analyses, the entire return distribution is taken into account, although with a different emphasis given to the tails (according to parameters p and q, i.e., both for return and risk). In contrast, Rachev ratios with α and β less than 0.5 ignore even a portion of upside variability in the evaluation of return and equally ignore even a portion of downside variability in the evaluation of risk. Remember that the Rachev ratio can be interpreted as the tradeoff between the expected return above the $VaR_{\alpha\%}$, i.e., the $CVaR_{\alpha\%}$, and the expected loss below the $VaR_{\beta\%}$, i.e., the $CVaR_{\beta\%}$.

Again, our expectations are confirmed: the highest rank correlation is achieved for values of α and β close to 0.5, which is just the same as the case of a moderate ratio achieving the tradeoff between the expected returns above and below the median. In this situation, the Rachev ratio acts similarly to the Omega index (note that it collapses into the Omega if the distribution is symmetrical). Vice versa, as α and β decrease, central data are removed from the analysis of return and risk. The ratio becomes more aggressive, providing only the tradeoff between the expected high stakes and the expected huge losses. In such circumstances, the Rachev ratio is focused merely on the tails, whereas the Sharpe ratio is focused on the "stability" around central values. Therefore, the two ratios show the biggest divergence in the way they capture information from the data and, as expected, their rank correlations shrink to 0.51. Moreover, when β tends toward 1, the denominator tends toward the mean (given for β = 1), clearly failing to be an accurate measure of risk and meaning that the ratio itself is no longer a valid return/risk tradeoff; $0.5 < \beta < 1$ is thus not relevant. In conclusion, when Rachev ratios are tailored to moderate investment styles (i.e., for α and β close to 0.5), the rank correlation is about 0.90, whereas when they are fitted to more aggressive investment styles (i.e., for α and β close to 0), the rank correlation falls to 0.51.

We can conclude that among the three families of tailor-made ratios analyzed here, the Sortino-Satchell is the one that behaves most like the Sharpe ratio. There may be two reasons for this: first, the Sortino-Satchell ratio captures the attitude toward gains with the mean, as does the Sharpe ratio; second the choice of q varies in accordance with Fishburn's (1977) approach, i.e., the greater the aversion to huge losses, the higher $q > 1$ and the less the aversion to huge losses, the lower $q < 1$, which is compatible with expected utility theory. The biggest discrepancies with the Sharpe ratio are found for Farinelli-Tibiletti and Rachev ratios fitting "extreme" investment styles. Specifically, the worst mismatch is achieved when the ratio is built to act as a *tradeoff* between moderate gains/moderate losses or between high stakes/huge losses, so that the parameter regulating aversion to huge losses no longer follows the Fishburn (1977) paradigm. Since by definition, the Rachev ratio is the tradeoff between gains and losses, its largest discrepancy from the Sharpe ratio occurs when it is set up for the most aggressive investor style, that is, small α and β. In this case, the rank correlation between the two measures falls as low as 0.51.

4. Conclusion

Whether using the Sharpe ratio to rank funds is advisable remains an open question in academia and among practitioners. The empirical analysis carried out here confirms the results of Eling and Schuhmacher (2007); as long as tailor-made ratios describe moderate and conservative investment styles, the rank correlation with the Sharpe ratio ranking is close to 1. However, if ratios such as Farinelli-Tibiletti or Rachev are tailored to describe more aggressive investment styles, the rank correlation is drastically reduced and the use of the Sharpe ratio becomes questionable.

References

Acerbi, C., Tasche, D., 2002. On the Coherence of Expected Shortfall. Journal of Banking & Finance 26, 1487–1503.

Agarwal, V., Naik, N.Y., 2004. Risk and Portfolio Decisions Involving Hedge Funds. Review of Financial Studies 17, 63–98.

Bernardo, A., Ledoit, O., 2000. Gain, Loss and Asset Pricing. Journal of Political Economy 108, 144–172.

Biglova, A., Ortobelli, S., Rachev, S. T., Stoyanov, S., 2004. Different Approaches to Risk Estimation in Portfolio Theory. Journal of Portfolio Management 31, 103–112.

Campbell, A., Kräussl, R., 2007. Revisiting the Home Bias Puzzle: Downside Equity Risk. Journal of International Money and Finance 26, 1239–1260.

Capocci, D., Hübner, G., 2004. Analysis of Hedge Fund Performance. Journal of Empirical Finance 11, 55–89.

Cogneau, P., Hubner, G., 2009. The 101 Ways to Measure Portfolio Performance(January 11, 2009). Available at SSRN: http://ssrn.com/abstract=1326076

Ding, B., Shawky, H. A., 2007. The Performance of Hedge Fund Strategies and the Asymmetry of Return Distributions. European Financial Management 13, 309–331.

Eling, M., Schuhmacher, F., 2007. Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds? Journal of Banking & Finance 31, 2632–2647.

Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M., Tibiletti, L., 2008. Beyond Sharpe Ratio: Optimal Asset Allocation Using Different Performance Ratios. Journal of Banking & Finance 32, 2057–2063.

Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M., Tibiletti, L., 2009. Optimal asset allocation aid system: From “one-size” vs “tailor-made” performance ratio. European Journal of Operational Research 192, 209–215.

Farinelli, S., Tibiletti, L., 2003. Sharpe Thinking with Asymmetrical Preferences, Technical Report presented at European Bond Commission; Winter Meeting; Frankfurt.

Farinelli, S., Tibiletti, L., 2008. Sharpe Thinking in Asset Ranking with One-Sided Measures. European Journal of Operational Research 185, 1542–1547.

Favre, L., Galeano, J. A., 2002. Mean-Modified Value at Risk Optimization with Hedge Funds. Journal of Alternative Investments 5, 21–25.

Fishburn, P.C., 1977. Mean-Risk Analysis with Risk Associated with Below-Target Returns. American Economic Review 66, 116–126.

Haselmann, R., Herwartz, H., 2008. Portfolio Performance and the Euro: Prospects for New Potential EMU members. Journal of International Money and Finance 27, 314–330.

Keating, C., Shadwick, W., 2002. A Universal Performance Measure. Journal of Performance Measurement 6(3), 59–84.

Leland, H. E., 1999. Beyond Mean-Variance: Performance Measurement in a Non-Symmetrical World. Financial Analysts Journal 55, 27–36.

Martin, D., Rachev, S. T., Siboulet, F., 2003. Phi-Alpha Optimal Portfolios and Extreme Risk Management. Wilmott Magazine of Finance November, 70–83.

Menn, C., Fabozzi, F. J., Rachev, S. T., 2005. Fat-Tailed and Skewed Asset Return Distributions: Implications for Risk Management, Portfolio Selection, and Option Pricing. John Wiley & Sons: Hoboken, NJ.

Rachev, S. T., Ortobelli, S., Stoyanov, S., Fabozzi, F., Biglova, A., 2008. Desirable properties of an ideal risk measure in portfolio theory. International Journal of Theoretical and Applied Finance 11**,** 19–54

Sharpe, W. F., 1966. Mutual Fund Performance. Journal of Business 39, 119–138.

Sortino, F. A., Satchell, S., 2001. Managing Downside Risk in Financial Markets. Butterworth Heinemann: Oxford.