Finance Research Letters 13 (2015) 97–104

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Testing equality of modified Sharpe ratios ☆

David Ardia [a,b,*], Kris Boudt [c,d]

[a] Département de finance, assurance et immobilier, Université Laval, Québec, Québec, Canada
[b] Centre interuniversitaire sur le risque, les politiques économiques et l'emploi, Québec, Québec, Canada
[c] Solvay Business School, Vrije Universiteit Brussel, Belgium
[d] Faculty of Economics and Business, VU University Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 27 November 2014
Accepted 26 February 2015
Available online 5 March 2015

JEL classification:
C12
C15
C22
C52

Keywords:
Bootstrap test
Hedge fund
Modified Sharpe ratio
Non-normal returns
Performance measurement

ABSTRACT

The modified Sharpe ratio is commonly used to evaluate the risk-adjusted performance of an investment with non-normal returns, such as hedge funds. In this note, a test for equality of modified Sharpe ratios of two investments is developed. A simulation study demonstrates the good size and power properties of the test. An application illustrates the complementarity between the Sharpe ratio and modified Sharpe ratio test for performance testing on hedge fund return data.

1. Introduction

Ledoit and Wolf (2008) recommend a bootstrap method to test equality of Sharpe ratios between two funds, accounting for the finite sample properties of the return distribution and for the potential autocorrelation and heteroskedasticity. If one of the funds has non-normally distributed returns,

☆ We are grateful to the Editor (Ramazan Gençay), an anonymous referee, Peter Carl, Michel Dubois, Ivan Guidotti, Lennart Hoogerheide, Simon Keel, Jeremy Kolly, Doug Martin, Enrico Schumann and participants at various conferences and seminars for useful comments. Financial support from aeris CAPITAL AG, IFM2 Montréal, and the Dutch science foundation is gratefully acknowledged. The views expressed in this paper are the sole responsibility of the authors. Any remaining errors or shortcomings are the authors' responsibility.

* Corresponding author at: Département de finance, assurance et immobilier, Université Laval, Québec, Québec, Canada.
E-mail addresses: david.ardia@fsa.ulaval.ca (D. Ardia), kris.boudt@vub.ac.be (K. Boudt).

http://dx.doi.org/10.1016/j.frl.2015.02.008
1544-6123/

comparing funds based on the Sharpe ratio is often not enough, as it ignores investors' positive preferences for odd moments and aversion to even moments. This weakness of the Sharpe ratio is well known and several alternatives have been proposed. For the analysis of hedge fund returns, the modified Sharpe ratio proposed by Favre and Galeano (2002) and Gregoriou and Gueyie (2003) is now increasingly popular. It is defined as the ratio between the average excess return of the fund and its modified Value-at-Risk. The latter is an estimator for Value-at-Risk based on the Cornish–Fisher expansion and the first four moments of the return distribution. The modified Sharpe ratio has been frequently used in research on performance of investments with non-normal returns such as hedge funds (see e.g., Amenc et al., 2003; Bali et al., 2013; Eling and Schuhmacher, 2007).

While there exist many risk-adjusted return measures to evaluate the performance of a single fund, previous research seems to have focused only on testing the equality of the Sharpe ratio of two funds (Jobson and Korkie, 1981; Memmel, 2003; Ledoit and Wolf, 2008). We develop a similar test for the modified Sharpe ratio. Compared to other performance measures taking the non-normality of the return distribution into account, the modified Sharpe ratio has the advantage of being asymptotically normal with an explicit function for its standard error. For small samples, we describe a bootstrap procedure to obtain the p-values corresponding to the studentized test statistic, accounting for the finite sample properties of the return distribution and the potential autocorrelation, heteroskedasticity and cross-dependence. The validity of the proposed methodology is verified through a Monte Carlo study. In the empirical application on the relative performance of hedge funds over the period 2008–2012 we find that, especially for hedge funds pursuing a Relative Value investment style, returns are non-normal leading to up to 22% of disagreement between the Sharpe and modified Sharpe ratio regarding equal-performance. For the other funds disagreement between the two tests ranges between 5% and 14%.

2. Testing equality of modified Sharpe ratios

Gregoriou and Gueyie (2003) define the modified Sharpe ratio as the ratio between the excess return of the fund and its modified Value-at-Risk (mVaR):

$$mSR_i(\alpha) \equiv \frac{m_i - r_b}{\mathrm{mVaR}_i(\alpha)}, \tag{1}$$

where m_i is the population mean return for fund manager i and r_b is the mean return on a benchmark asset at the corresponding horizon. Modified VaR approximates the VaR under the true (unknown) distribution with the second order Cornish–Fisher expansion. Let R_i be the return of fund manager i. Denote the q-th centered portfolio moment $m_{q,i} \equiv \mathbb{E}\left[(R_i - m_i)^q\right]$, with $m_i \equiv \mathbb{E}[R_i]$. For a loss probability α, the modified VaR is:

$$\mathrm{mVaR}_i(\alpha) \equiv -m_i + \sqrt{m_{2,i}}\left(-z_\alpha - \frac{1}{6}(z_\alpha^2 - 1)s_i - \frac{1}{24}(z_\alpha^3 - 3z_\alpha)k_i + \frac{1}{36}(2z_\alpha^3 - 5z_\alpha)s_i^2\right), \tag{2}$$

with $s_i \equiv m_{3,i}/m_{2,i}^{3/2}$ the skewness, $k_i \equiv m_{4,i}/m_{2,i}^2 - 3$ the excess kurtosis, and z_α the α-percentile of the standard normal distribution. Note that, for the mSR in (1) to be correctly defined, it is required that $\mathrm{mVaR}_i(\alpha)$ is strictly positive. For loss probabilities α of 5% and 10%, this will be almost always the case for realistic return distributions. Also of interest is the skewness-kurtosis boundary derived by Jondeau and Rockinger (2001) to ensure that the corresponding Cornish–Fisher density approximation is everywhere positively valued.

Gregoriou and Gueyie (2003) set α to 5%. Favre and Galeano (2002) consider an α of 1% and 5%. Other authors, like Jaschke (December 2002) and Boudt et al. (2008), warn against the use of small values of α, based on the fact that the Cornish–Fisher approximation of the quantile function becomes less and less reliable for $\alpha \to 0$. For small estimation samples, this so-called "wrong tail behavior" of the Cornish–Fisher approximation is likely to be aggravated by the estimation error in the moments of the return distribution. We therefore follow Jaschke (December 2002) and Boudt et al. (2008) and recommend testing for equal-performance using the mSR at higher values of α, such as 5% and 10%.

We use $\alpha = 10\%$ (i.e. mVaR at the 90% risk level) in the sequel as our baseline investigation, but also explore the properties of the test for $\alpha = 5\%$.

Gregoriou and Gueyie (2003) set the benchmark return r_b in (1) to the risk-free rate. Amenc et al. (2003) note that this would only make sense if the CAPM is an appropriate model for the alternative universe. This is clearly not the case in our application on relative performance of hedge funds, as investors in hedge funds also require a premium for the funds' exposure to credit and liquidity risk, among others. In what follows, we set the benchmark return in the modified Sharpe ratio to zero implying that investors compare hedge funds based on their differences in the risk adjusted absolute returns.[1]

2.1. Pairwise test

Replacing population moments with sample ones in (1) and (2) provides us with estimates of the modified VaR and the modified Sharpe ratio. Denote the sample moments as $\hat{m}_i, \hat{m}_{2,i}, \hat{m}_{3,i}, \hat{m}_{4,i}, \hat{s}_i$ and $\hat{k}_i$, and the resulting modified VaR and Sharpe ratio as $\text{m}\hat{\text{V}}\text{aR}_i(\alpha)$ and $m\hat{S}R_i(\alpha)$. We denote by Δ_{i-j} the (true) difference of this performance measure between fund i and j ($i \neq j$) and by $\hat{\Delta}_{i-j}$ the corresponding estimate. We are interested in testing the null of equal modified Sharpe ratios between funds i and j:

$$H_0 : \Delta_{i-j} \equiv mSR_i(\alpha) - mSR_j(\alpha) = \frac{m_i}{\text{mVaR}_i(\alpha)} - \frac{m_j}{\text{mVaR}_j(\alpha)} = 0,$$

based on the estimated difference in the cross-product of the fund's return and the other fund's mVaR[2]: $\hat{\Delta}_{i-j} \equiv \hat{m}_i \text{m}\hat{\text{V}}\text{aR}_j(\alpha) - \hat{m}_j \text{m}\hat{\text{V}}\text{aR}_i(\alpha)$. The test statistic we recommend to use for evaluating the null is the ratio between $\hat{\Delta}_{i-j}$ and its standard error. Under regularity conditions, the Delta method implies that this studentized test statistic is asymptotically normally distributed.

2.1.1. Standard error of $\hat{\Delta}_{i-j}$

First we express all centered moments with uncentered ones: $\hat{g}_{q,i} \equiv \sum_{t=1}^{T} r_{i,t}^q / T$, where $r_{i,t}$ is the realized return of manager i at time t and T is the number of observations. We have $\hat{m}_{2,i} = \hat{g}_{2,i} - \hat{m}_i^2$, $\hat{m}_{3,i} = \hat{g}_{3,i} - 3\hat{m}_i \hat{g}_{2,i} + 2\hat{m}_i^3$ and $\hat{m}_{4,i} = \hat{g}_{4,i} - 4\hat{m}_i \hat{g}_{3,i} + 6\hat{m}_i^2 \hat{g}_{2,i} - 3\hat{m}_i^4$. It follows that $\hat{\Delta}_{i-j}$ is a function of the mean and second to fourth moments about the origin of the two series: $\hat{\Delta}_{i-j} \equiv f(\hat{m}_i, \hat{g}_{2,i}, \hat{g}_{3,i}, \hat{g}_{4,i}, \hat{m}_j, \hat{g}_{2,j}, \hat{g}_{3,j}, \hat{g}_{4,j})$. If $\hat{\Psi}_{i-j}$ is a consistent estimator of the asymptotic covariance matrix between these arguments, then an estimate for the standard error of $\hat{\Delta}_{i-j}$ is $s(\hat{\Delta}_{i-j}) \equiv \sqrt{\hat{\nabla}'_{i-j} \hat{\Psi}_{i-j} \hat{\nabla}_{i-j} / T}$ with the gradient of $\hat{\Delta}_{i-j}$ given by $\hat{\nabla}_{i-j} \equiv (\text{m}\hat{\text{V}}\text{aR}_j d\hat{m}_i - \hat{m}_i \text{dm}\hat{\text{V}}\text{aR}_j) - (\text{m}\hat{\text{V}}\text{aR}_i d\hat{m}_j - \hat{m}_j \text{dm}\hat{\text{V}}\text{aR}_i)$, with:

$$\begin{aligned} \text{dm}\hat{\text{V}}\text{aR}_i = & -d\hat{m}_i - z_\alpha \frac{1}{2\sqrt{\hat{m}_{2,i}}} d\hat{m}_{2,i} - \frac{1}{6}(z_\alpha^2 - 1)\left(\frac{1}{2\sqrt{\hat{m}_{2,i}}} \hat{s}_i d\hat{m}_{2,i} + \sqrt{\hat{m}_{2,i}} d\hat{s}_i\right) \\ & - \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\left(\frac{1}{2\sqrt{\hat{m}_{2,i}}} \hat{k}_i d\hat{m}_{2,i} + \sqrt{\hat{m}_{2,i}} d\hat{k}_i\right) + \frac{1}{36}(2z_\alpha^3 - 5z_\alpha)\left(\frac{1}{2\sqrt{\hat{m}_{2,i}}} \hat{s}_i^2 d\hat{m}_{2,i} + 2\sqrt{\hat{m}_{2,i}} \hat{s}_i d\hat{s}_i\right), \end{aligned}$$

[1] As mentioned by a referee, the proposed equal *mSR* performance test could also be of interest for other risk-free rate definitions. If the risk free rate is constant over the period, this is straightforward. When the risk free rate is stochastic, it would imply that the distribution of the test statistic not only depends on the (co) moments of the hedge fund returns, but also on those of the risk free rate process. This comes at the price of making the test statistic less tractable. As a solution, we recommend to apply the proposed test directly on the excess returns. This leads to a small approximation error in the calculation of the mVaR, which, as shown in Daníelsson (2011, Chapter 5) is negligible compared to the statistical uncertainty in the volatility, skewness and kurtosis.

[2] An alternative way to test the difference in modified Sharpe ratios is to consider directly the estimated difference in modified Sharpe ratios $\hat{\Delta}_{i-j} \equiv m\hat{S}R_i(\alpha) - m\hat{S}R_j(\alpha)$. The apparent advantage of this alternative test statistic is that its construction is similar to the traditional one used by Jobson and Korkie (1981), Memmel (2003) and Ledoit and Wolf (2008) to test for the equality of two Sharpe ratios. The disadvantage however is that both the numerator and denominator depend on the inverse of the modified VaR. For funds with positive sample skewness, the modified Sharpe ratio can be close to zero, and extremely fat tailed. By considering the test statistic that cross-multiplies the mean and mVaR of the two funds, we obtain a statistic that is analytically more simple and better behaved. We verified this statement in a simulation study (results are available from the authors upon request).

where $d\hat{s}_i = (\hat{m}_{2,i}^{3/2} d\hat{m}_{3,i} - 1.5\hat{m}_{3,i}\hat{m}_{2,i}^{1/2} d\hat{m}_{2,i})/\hat{m}_{2,i}^3$, $d\hat{k}_i = (\hat{m}_{2,i}^2 d\hat{m}_{4,i} - 2\hat{m}_{4,i}\hat{m}_{2,i} d\hat{m}_{2,i})/\hat{m}_{2,i}^4$ and similarly for j, and:

$$
\begin{aligned}
d\hat{m}_i &= (1,0,0,0,0,0,0,0)' \\
d\hat{m}_j &= (0,0,0,0,1,0,0,0)' \\
d\hat{m}_{2,i} &= (-2\hat{m}_i,1,0,0,0,0,0,0)' \\
d\hat{m}_{2,j} &= (0,0,0,0,-2\hat{m}_j,1,0,0)' \\
d\hat{m}_{3,i} &= (-3\hat{g}_{2,i}+6\hat{m}_i^2,-3\hat{m}_i,1,0,0,0,0,0)' \\
d\hat{m}_{3,j} &= (0,0,0,0,-3\hat{g}_{2,j}+6\hat{m}_j^2,-3\hat{m}_j,1,0)' \\
d\hat{m}_{4,i} &= (-4\hat{g}_{3,i}+12\hat{m}_i\hat{g}_{2,i}-12\hat{m}_i^3,6\hat{m}_i^2,-4\hat{m}_i,1,0,0,0,0)' \\
d\hat{m}_{4,j} &= (0,0,0,0,-4\hat{g}_{3,i}+12\hat{m}_i\hat{g}_{2,i}-12\hat{m}_i^3,6\hat{m}_i^2,-4\hat{m}_i,1)'.
\end{aligned}
$$

To account for the autocorrelation and heteroscedasticity in hedge fund returns, we recommend to estimate $\hat{\Psi}_{i-j}$ using the heteroscedasticity and autocorrelation robust (HAC) kernel estimators of Andrews (1991) and Andrews and Monahan (1992).

2.1.2. Calculation of p-values

From the central limit theorem, it follows that, under standard regularity conditions, the absolute value of the studentized test statistic $\hat{\tau}_{i-j} \equiv |\hat{\Delta}_{i-j}|/s(\hat{\Delta}_{i-j})$ is asymptotically distributed as the absolute value of a standard normal random variable. The corresponding p-value is computed as $\hat{p}_{i-j} \equiv 2\Phi(-\hat{\tau}_{i-j})$ where Φ denotes the cumulative distribution of a standard normal random variable.

Because of the small sample size in our empirical application, we also consider a bootstrap method to compute the p-value. To generate bootstrap data, we resample with replacement either individual pairs, as in the *i.i.d.* bootstrap of Efron (1979), or blocks of fixed size $l \geqslant 1$, following the circular block-bootstrap of Politis and Romano (1992). Based on the B bootstrap pairs $(r_{t,i}^{*b}, r_{t,j}^{*b})$, the bootstrap test statistics are computed as $\hat{\tau}_{i-j}^{*b} \equiv (\hat{\Delta}_{i-j}^{*b} - \hat{\Delta}_{i-j})/s(\hat{\Delta}_{i-j}^{*b})$, where $*b$ denotes the estimators computed on the b-th bootstrap data set. The correction for the non-zero performance differential in the sample ensures that in the bootstrap population the null hypothesis of equal-performance is satisfied. Under the bootstrap approach, we compute the p-value as $\hat{p}_{i-j} \equiv (\sum_{b=1}^{B} I\{|\hat{\tau}_{i-j}^{*b}| \geqslant |\hat{\tau}_{i-j}|\} + 1)/(B+1)$ where $I\{\cdot\}$ is the indicator function (Davison, 1997, p. 141). We set $B \equiv 499$ in the sequel.

3. Monte Carlo study

3.1. Size

To test the size properties of the proposed test (at the 90% and 95% risk levels, i.e. $\alpha = 10\%$ and $\alpha = 5\%$) of two funds, i and j, we generate the returns of fund i and j from an identical marginal return process, that are joined together in a bivariate distribution using a Normal copula with correlation of 0.5. The marginal return process has an unconditional mean and variance equal to one. They either (i) are *i.i.d.*, or (ii) follow an AR(1) model with auto-regressive parameter equal to 0.2. The conditional marginal innovations are either: (i) normal, (ii) standardized Student-t with six degrees of freedom or (iii) standardized skewed Student-t with six degrees of freedom and asymmetry parameter $\xi = 0.75$ (negative skewness); see Fernández and Steel (1998). Overall, this leads to six data generating processes (DGPs). The size of the simulated DPGs is set to $T = 60$ and $T = 1000$ to study the small size and asymptotic properties of the estimators.

For each DGP, we report in Fig. 1 the 95% confidence bands of the rejection rate (over five hundred Monte Carlo replications) using the asymptotic (solid bars) and block-bootstrap (dashed bars) based p-values at the 5% significance level. For the block-bootstrap test, the block-size is set to the $\lceil T^{1/3} \rceil$

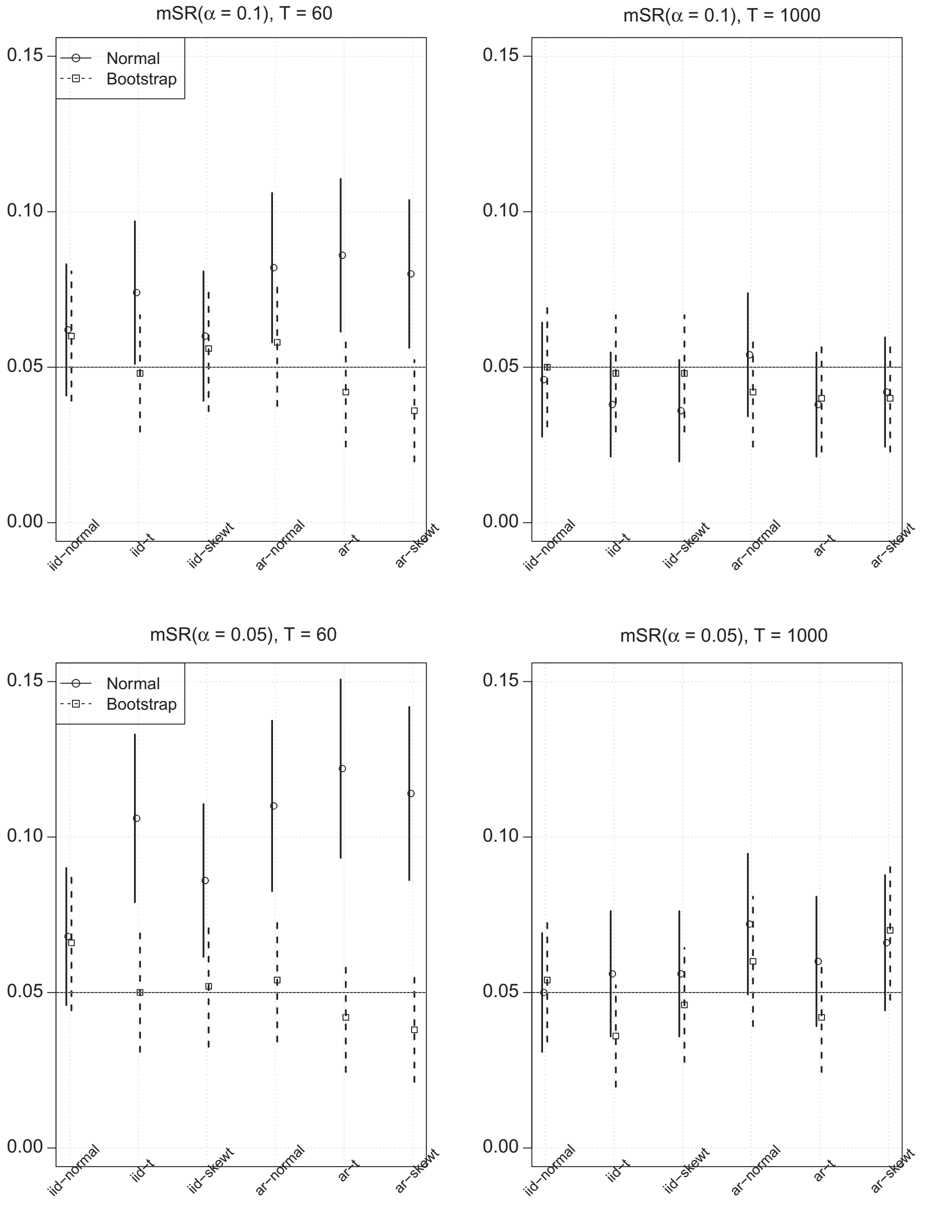

Fig. 1. Size results for the modified Sharpe ratio at the 90% (top plots) and 95% (bottom plots) risk level when *p*-values are computed under the asymptotic Normal approximation (full line) or bootstrap approach (dashed line). The vertical bars represent the 95% confidence bands of the frequency of rejecting the null hypothesis at the 5% significance level.

rule-of-thumb. The top and bottom panels correspond to testing for equal performance using the 90% and 95% risk level in the modified VaR calculation, respectively.[3] We see that in the presence of AR dynamics in the return, the test using the asymptotic normal distribution is oversized in small samples, while the block-bootstrap test still has correct size properties. The same observation holds when testing

[3] To save space, we report size results at the 5% level. Results at the 10% significance level are qualitatively similar and are available from the authors upon request.

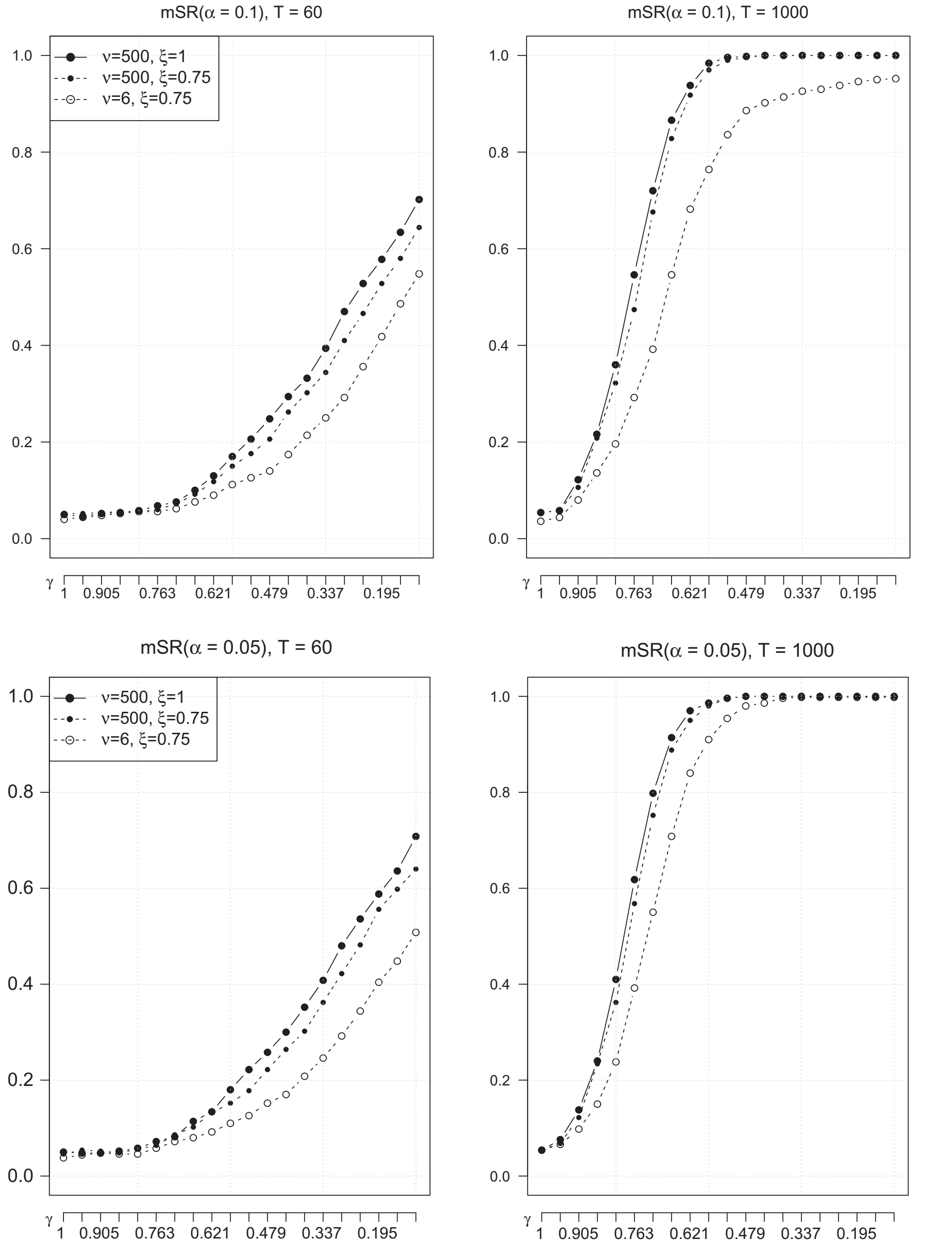

Fig. 2. Power results for the modified Sharpe ratio at the 90% (top plots) and 95% (bottom plots) risk level when varying the mean parameter μ_j between μ_i ($\gamma = 1$: equal-performance) and $\mu_i/10$ ($\gamma = 1/10$: under-performance of fund *j*). The plot displays the frequency of rejections at the 5% significance level of the null hypothesis of equal modified Sharpe ratio for γ ranging from 1 (equal-performance) to 1/10 (under-performance of fund *j*).

at the 95% risk level (bottom plots) in the presence of fat-tailed return distributions: in contrast with the bootstrap test having correct size, the asymptotic test is oversized. For large samples, the null hypothesis of correct size cannot be rejected for both the asymptotic and bootstrap based tests.

Table 1
Left: Average mean and standard deviation (in percent) together with average skewness and kurtosis of monthly hedge fund returns over the period 2008–2012. Middle and right parts report percentages of (non-) rejections (at the 5% significance level) of the Sharpe ratio and modified Sharpe ratio (at the 90% risk level for the middle part, and at the 95% risk level for the right part) equality tests. R-R: both rejections; NR-R: rejection of modified Sharpe test only; R-NR: rejection of Sharpe ratio test only; NR-NR: non-rejection of both tests.

Style	Summary statistics				Equal SR – mSR($\alpha = 0.1$) tests				Equal SR – mSR($\alpha = 0.05$) tests			
	Mean	Std	Skew	Kurt	R-R	NR-R	R-NR	NR-NR	R-R	NR-R	R-NR	NR-NR
All	0.44	4.71	−0.33	5.82	6.57	0.70	8.52	84.22	5.63	0.29	9.74	84.33
Equity hedge	0.34	5.28	−0.27	4.83	4.44	0.34	8.05	87.17	4.06	0.21	8.65	87.07
Event-driven	0.52	3.89	−0.77	6.85	5.30	0.51	12.19	82.00	4.32	0.18	14.07	81.42
Macro	0.45	4.39	0.16	5.10	4.68	0.66	4.46	90.21	4.16	0.30	5.03	90.50
Relative value	0.67	3.94	−0.95	9.34	4.75	0.57	20.24	74.43	3.75	0.26	21.78	74.22

3.2. Power

In Fig. 2, we study the power of the test when the difference in modified Sharpe ratio between the two funds increases. As a reference point, we take for fund i the returns from a skewed Student-t distribution with mean $\mu = 0.445\%$, standard deviation $\sigma = 1.07\%$, skewness parameter $\xi = 1$ (no skewness) and degrees of freedom parameter $\nu = 500$ (close to normal tails). This parameter setup corresponds to one of the outperforming funds in our empirical database. We consider the following variations to test the power of the test statistics: (i) approximative normal: $\mu_i = 0.445\%, \sigma_i = \sigma_j = 1.07\%, \xi_i = \xi_j = 1, \nu_i = \nu_j = 500$, (ii) skewness: $\mu_i = 0.445\%, \sigma_i = \sigma_j = 1.07\%, \xi_i = 1, \xi_j = 0.75, \nu_i = \nu_j = 500$ and (iii) skewness and fat tails: $\mu_i = 0.445\%, \sigma_i = \sigma_j = 1.07\%, \xi_i = 1, \xi_j = 0.75, \nu_i = 500, \nu_j = 6$. In both cases, for several values of γ between 1 and 1/10, μ_j is set such that $\mu_j = \gamma\mu_i$. For each setup, we generate *i.i.d.* data and analyze the power of the modified Sharpe ratio bootstrap test.[4]

The power is computed as the average frequency of rejections of the null hypothesis at the 5% level. We see that, in all three scenarios, the test has correct size for $\gamma = 1$ and, when increasing the difference between the two modified Sharpe ratios, the power increases relatively slowly for $T = 60$, but the slope becomes substantially more steep for $T = 1000$. The presence of skewness and fat tails in the data seems to reduce slightly the power of the test.

4. Empirical illustration

The proposed modified Sharpe ratio will be useful when it leads to different decisions on equal–performance than the test using the Sharpe ratio as performance measure. We investigate now the (dis) agreement regarding equal–performance of the Sharpe ratio and modified Sharpe ratio test for the 1552 funds in the Hedge Funds Research database over the period ranging from January 2008 to December 2012 (60 monthly observations). We perform the tests of equal-performance of Sharpe and the modified Sharpe ratios for the 1,203,576 pairs using the block-bootstrap approach with five hundred replications and block-size set to four; alternative specifications lead to similar results.[5]

In Table 1 we report the percentage of (non-)rejections (at the 5% significance level) of the Sharpe ratio and modified Sharpe ratio (at the 90% and 95% risk levels) equality tests for all funds, and per investment style. We see that especially for the Relative Value investment style and for the mVaR at the 90% risk level, there is a large proportion of disagreement between the test; for 0.57% (resp. 20.24%) of the pairs the modified Sharpe ratio (resp. Sharpe ratio) rejects equal-performance, while the other test does not. For the mVaR at the 95% risk level, the level of disagreement for Relative

[4] Similar power plots are obtained when considering a reference fund with a higher return and volatility or when keeping $\mu_i = \mu_j$ but allowing for differences in volatility between funds i and j. To save space, these power plots are not reported, but they are available from the authors upon request.

[5] All computations are performed in the R statistical computing language with the package **PeerPerformance** (Ardia and Boudt, 2014). Computer code is available from the authors upon request.

Value funds is even higher (around 22%). For the other funds, the disagreement ranges from 5% to 14%. The high degree of disagreement for the Relative Value style can be explained by the more extreme values for the skewness and kurtosis of the funds pursuing the relative value strategy, compared to the other investment styles.

References

Amenc, N., Martellini, L., Vaissie, M., 2003. Benefits and risk of alternative investment strategies. J. Asset Manage. 4, 96–118.
Andrews, D., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.
Andrews, D., Monahan, J., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 60, 953–966.
Ardia, D., Boudt, K., 2014. PeerPerformance: peer performance analysis in R. Version 2.0.6 <https://r-forge.r-project.org/projects/peerperformance/>.
Bali, T.G., Brown, S.J., Demirtas, K.O., 2013. Do hedge funds outperform stocks and bonds? Manage. Sci. 59, 1887–1903.
Boudt, K., Peterson, B.G., Croux, C., 2008. Estimation and decomposition of downside risk for portfolios with non-normal returns. J. Risk 11, 79–103.
Daníelsson, J., 2011. Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab, vol. 587. John Wiley & Sons.
Davison, A.C., 1997. Bootstrap Methods and their Application. Cambridge University Press.
Efron, B., 1979. Bootstrap methods: another look at the jackknife. Ann. Statist. 7, 1–26.
Eling, M., Schuhmacher, F., 2007. Does the choice of performance measure influence the evaluation of hedge funds? J. Bank. Finan. 31, 2632–2647.
Favre, L., Galeano, J., 2002. Mean-modified value-at-risk with hedge funds. J. Alter. Invest. 5 (2), 21–25.
Fernández, C., Steel, M.F., 1998. On Bayesian modelling of fat tails and skewness. J. Am. Statist. Assoc. 93, 359–371.
Gregoriou, G.N., Gueyie, J.-P., 2003. Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio. J. Wealth Manage. 6, 77–83.
Jaschke, S.R., December 2002. The Cornish–Fisher-expansion in the context of Delta–Gamma–Normal approximations. J. Risk 4 (4), 33–52.
Jobson, J., Korkie, B., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. J. Finan. 36, 889–908.
Jondeau, E., Rockinger, M., 2001. Gram–Charlier densities. J. Econ. Dynamics Control 25 (10), 1457–1483.
Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. J. Emp. Finan. 15, 850–859.
Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. Finan. Lett. 1, 21–23.
Politis, D., Romano, J., 1992. Exploring the limits of bootstrap. John Wiley, New York (Ch. A circular block-resampling procedure for stationary data, pp. 263–270).