Finance Research Letters 10 (2013) 196–208

Contents lists available at ScienceDirect

Finance Research Letters

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

Performance hypothesis testing with the Sharpe ratio: The case of hedge funds

Benjamin R. Auer *, Frank Schuhmacher

University of Leipzig, Department of Finance, Grimmaische Str. 12, 04109 Leipzig, Germany

ARTICLE INFO

Article history:
Received 20 December 2012
Accepted 13 August 2013
Available online 28 August 2013

JEL classification:
G10
G11
C12
C58

Keywords:
Sharpe ratio
Hedge funds
Robust testing
Performance measurement

ABSTRACT

As recent research highlights that the Sharpe ratio has a decision theoretic foundation even in the case of asymmetric or fat-tailed excess returns and thus is adequate even for the evaluation of hedge funds, this note provides the first Sharpe ratio based performance analysis of the hedge fund market. Furthermore, it addresses the important practical question whether the choice of hypothesis test used to statistically compare Sharpe ratios can influence an investor's hedge fund selection process. Our key findings are as follows: (i) Only a small fraction of hedge funds in our large dataset can significantly outperform passive investments in corresponding hedge fund indices. (ii) Especially in the presence of autocorrelated or skewed excess returns, the traditional test of Jobson and Korkie (1981) and Memmel (2003) tends to overstate the number of significant outperformers and thus provides potentially misleading information for investors. Decision makers are advised to use the bootstrap test of Ledoit and Wolf (2008) allowing robust and more reliable inference.

1. Introduction

The Sharpe ratio, introduced by Sharpe (1966), that measures the relationship between the mean and standard deviation of excess returns, is one of the best-known and widely used metrics to measure and compare investment performance. However, for a long time, many authors believed that hedge funds (typically characterised by asymmetric and fat-tailed excess return distributions; see Homm and Pigorsch, 2012) cannot be adequately evaluated using this classic reward-to-risk ratio

* Corresponding author. Tel.: +49 341 97 33 672; fax: +49 341 97 33 679.
E-mail address: auer@wifa.uni-leipzig.de (B.R. Auer).

1544-6123/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.frl.2013.08.001

because it was said to only have a decision theoretic justification in the case of normally distributed excess returns or quadratic investor preferences (see Brooks and Kat, 2002; Mahdavi, 2004; Sharma, 2004; Zakamouline and Koekebakker, 2009). Today, recently rediscovered insights concerning the theoretical justification of the Sharpe ratio make the measure more interesting for hedge fund evaluation. For example, Schuhmacher and Eling (2011) point out that the theoretic foundation of the Sharpe ratio is given by the widely unknown location and scale condition of Sinn (1983) and Meyer (1987). This condition states that the fund returns must be equal in distribution to one another except for their location and scale parameters. Hence, contrary to popular belief, neither asymmetry nor fat tails of excess return distributions allow a decision theoretic rejection of the measure. However, one problem with the location and scale condition is that it requires cross-sectionally identical levels of skewness and kurtosis of the funds under evaluation. Yet, this still does not justify a rejection of the Sharpe ratio because Schuhmacher and Eling (2012) note that its use can also be motivated by the generalised location and scale condition formulated by Meyer and Rasche (1992). It allows cross-sectional differences in skewness and kurtosis and thus makes the Sharpe ratio applicable to a wider range of investment assets than previously thought.

Although the Sharpe ratio can be regarded to have a solid theoretical foundation, it is prone to manipulation and has some undesirable technical properties. For example, Goetzmann et al. (2002) show that simple option-based strategies can increase the Sharpe ratio without adding value for the investor. Furthermore, Schuster and Auer (2012) and Auer (2013) formalise the high (low) return distortion of the Sharpe ratio. Specifically, they show that an exceptionally high (low) return that exceeds (deceeds) a certain limit can decrease (increase) a fund's Sharpe ratio. However, despite these unfavorable properties, practitioners do not turn their back on the measure. Financial information systems frequently publish lists where funds are ranked by their Sharpe ratios. On the basis of this information, investors are then advised to invest into funds with a high Sharpe ratio because the fund with the higher Sharpe ratio yields a higher return for the same amount of risk (see Liang, 1999). Even providers of hedge fund index data calculate Sharpe ratios to compare the index performance to the performance of alternative investment strategies (see, for example, http://www.hedgeindex.com). This popularity of the measure most likely stems from the facts that it has a long history, can be easily obtained and is simpler to interpret than most recently proposed complex performance measures.

To make inferences on whether the Sharpe ratio of a fund is significantly different from the Sharpe ratio of another investment, the literature has essentially brought forth two statistical tests: The best-known (traditional) test is the one of Jobson and Korkie (1981) and Memmel (2003) which is easy to implement but relies on the assumptions that excess returns are normally distributed and serially uncorrelated. However, it is well-known that these requirements are no typical features of asset returns making the application of the test potentially problematic in practice. The sophisticated and computationally intensive bootstrap test recently proposed by Ledoit and Wolf (2008) does not require strong distributional assumptions and can be considered a robust test taking into account the non-normality and time series nature of asset returns (including autocorrelation and heteroskedasticity). It could be considered the new standard method. However, so far this new test has received little attention in the empirical literature.[1]

In the light of this new state of knowledge (theoretical support of the Sharpe ratio and availability of an adequate testing procedure) and the popularity of the Sharpe ratio in practice, this article has two main goals: The first one is to provide the first Sharpe ratio based hedge fund performance analysis. Using the two statistical procedures mentioned above, we compare the Sharpe ratios of more than 4000 hedge funds with different investment strategies to the Sharpe ratios of passive investments in corresponding hedge fund indices in order to see whether we can detect significant outperformance of hedge fund managers. The second aim is to answer the question whether the choice of test procedure crucially influences the investment decision of a hedge fund investor, that is, whether the tests judge significant outperformance differently and thus lead to a different fund selection.[2] As

[1] Dorfleitner and Wimmer (2010), Walkshäusl and Lobe (2010), Scherer (2011), Fletcher (2011), Behr et al. (2012), Thornton and Valente (2012) and Kroencke and Schindler (2012) are notable exceptions.

[2] A question like this is typically answered in a simulation study. However, we prefer using an empirical hedge fund dataset because simulating fitted theoretical distributions only allows an approximate analysis and masks key features of real world data.