Journal of Empirical Finance 15 (2008) 850–859

Contents lists available at ScienceDirect

Journal of Empirical Finance

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

Robust performance hypothesis testing with the Sharpe ratio

Oliver Ledoit [a,1], Michael Wolf [b,*,2]

[a] Equity Proprietary Trading, Credit Suisse, London E14 4QJ, UK
[b] Institute for Empirical Research in Economics, University of Zurich, Switzerland

ARTICLE INFO

Article history:
Received 7 May 2007
Received in revised form 16 March 2008
Accepted 18 March 2008
Available online 26 March 2008

JEL classification:
C12
C14
C22

Keywords:
Bootstrap
HAC inference
Sharpe ratio

ABSTRACT

Applied researchers often test for the difference of the Sharpe ratios of two investment strategies. A very popular tool to this end is the test of Jobson and Korkie [Jobson, J.D. and Korkie, B.M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. Journal of Finance, 36:889–908], which has been corrected by Memmel [Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. Finance Letters, 1:21–23]. Unfortunately, this test is not valid when returns have tails heavier than the normal distribution or are of time series nature. Instead, we propose the use of robust inference methods. In particular, we suggest to construct a studentized time series bootstrap confidence interval for the difference of the Sharpe ratios and to declare the two ratios different if zero is not contained in the obtained interval. This approach has the advantage that one can simply resample from the observed data as opposed to some null-restricted data. A simulation study demonstrates the improved finite sample performance compared to existing methods. In addition, two applications to real data are provided.

1. Introduction

Many applications of financial performance analysis are concerned with the comparison of the Sharpe ratios of two investment strategies (such as stocks, portfolios, mutual funds, hedge funds, or technical trading rules). Since the true quantities are not observable, the Sharpe ratios have to be estimated from historical return data and the comparison has to be based on statistical inference, such as hypothesis tests or confidence intervals.

It appears that the *status quo* in the applied literature is the test of Jobson and Korkie (1981) and its corrected version by Memmel (2003); for example, see DeMiguel et al. (in press), DeMiguel and Nogales (2007), and Gasbarro et al. (2007), among others. Unfortunately, this test is not valid when returns have tails heavier than the normal distribution or are of time series nature. The former is a quite common, and by now well-known, property of financial returns. As far as the latter is concerned, serial correlation of the actual returns is, arguably, only a minor concern for stocks and mutual funds, but it is certainly relevant to hedge funds; for example, see Brooks and Kat (2002) and Malkiel and Saha (2005). However, even stocks and mutual funds often exhibit correlation of the squared returns, that is, volatility clustering; for example, see Campbell et al. (1997, Chapter 12) and Alexander (2001, Chapter 4).

In this paper, we discuss inference methods that are more generally valid. One possibility is to compute a HAC standard error[3] for the difference of the estimated Sharpe ratios by the methods of Andrews (1991) and Andrews and Monahan (1992), say. Such an approach works asymptotically but does not always have satisfactory properties in finite samples. As an improved alternative, we suggest a studentized time series bootstrap.

* Corresponding author.
E-mail address: mwolf@iew.uzh.ch (M. Wolf).

[1] The views expressed herein are those of the author and do not necessarily reflect or represent those of Credit Suisse.

[2] The research has been supported by the University Research Priority Program "Finance and Financial Markets", University of Zurich, and by the Spanish Ministry of Science and Technology and FEDER, grant MTM2006-05650.

[3] In this paper, a standard error of an estimator denotes an estimate of the true standard deviation of the estimator.

0927-5398/$ – see front matter
doi:10.1016/j.jempfin.2008.03.002