**Referee report** on "Comparing and selecting performance measures using rank correlations", by M. Caporin and F. Lisi (MS 530).

**General comment.** The authors address the important question of how to select a non-redundant set of asset performance measures. They suggest a rank correlation approach which is applied in a variety of different conditions. My opinion is that the proposed methodology constitutes a useful contribution to the research on asset selection criteria.

## Point remarks

- For redundant performance measures, the mean of the asymptotic distribution of the sample rank correlation is an increasing function of the length of the series, T. In light of this, one should adopt a T-dependent redundancy threshold. As an extreme example, consider 1000 series of iid zero-mean gaussian returns of length T = 4. The expected Omega/Sharp rank correlation is .65, with a standard deviation equal to .035; for T = 10, it is .99, with a standard deviation equal to .006 (values computed via MC methods). For a threshold set to .8, if T = 4 one would always accept the null hypothesis of redundant performance measures, whereas, if T = 10, one would always reject it. My feeling, indeed, is that for the sample sizes considered by the authors the choice of a common threshold should result in a minor problem. It would be nice, however, to substantiate such a choice by means of some *ad hoc* empirical and/or theoretical argument.
- I cannot understand the sentence "Eling and Schuhmacher (2007) tested the null hypothesis  $R_S \leq p$ , and determined the value of p associated with a rejection of the null for all assets" (see p. 11). Please, restate this sentence in a more explicit fashion.
- In table 2 and table 7, "Conditional Sharp" should be replaced by "STARR ratio".
- Maybe the instability of the Omega/Sharp rank correlation in Fig. 1 appears to be persistent as a fictitious effect of the autocorrelation induced by the rolling window.