Asset Performance Evaluation with Mean-Variance Ratio

Zhidong Bai

School of Mathematics and Statistics

North East Normal University, China

Kok Fai Phoon

School of Business

Singapore Management University, Singapore

Keyan Wang

Department of Mathematics

Shanghai Finance University, China

Wing-Keung Wong

Department of Economics

Hong Kong Baptist University, Hong Kong

April 14, 2011

Acknowledgement The fourth author would also like to thank Professors Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. The research is partially supported by the grants from North East Normal University, Singapore Management University, Shanghai Finance University, and Hong Kong Baptist University.

Asset Performance Evaluation with Mean-Variance Ratio

Abstract

Bai, et al. (2011c) develop the mean-variance-ratio (MVR) statistic to test the performance among assets for small samples. They provide theoretical reasoning to use MVR and prove that our proposed statistic is uniformly most powerful unbiased. In this paper we illustrate the superiority of our proposed test over the Sharpe ratio (SR) test by applying both tests to analyze the performance of Commodity Trading Advisors (CTAs). Our findings show that while the SR test concludes most of the CTA funds being analyzed as being indistinguishable in their performance, our proposed statistics show that some funds outperform the others. On the other hand, when we apply the SR statistic on some other funds in which the recent difference between the two funds is insignificant and even changes directions, the SR statistic indicates that one fund is significantly outperforming another fund whereas the MVR statistic could detect the change.

JEL classification: C12; G11

Keywords: Sharpe ratio; hypothesis testing; uniformly most powerful unbiased test; fund management

The pioneer work of Markowitz (1952) on the mean-variance (MV) portfolio optimization procedure has been widely used in both Economics and Finance to analyze how people make their choices concerning risky investments. The Markowitz efficient frontier also provides the basis for many important financial economics advances, including the Sharpe-Linter Capital Asset Pricing Model (Sharpe, 1964; Lintner, 1965) and the well-known optimal one-fund theorem (Tobin, 1958). Originally motivated by the MV analysis, the optimal one-fund theorem and the Sharpe-Linter Capital Asset Pricing Model, the Sharpe ratio (SR), the ratio of the excess expected return to its volatility or standard deviation, is one of the most commonly used statistics in the MV framework. The SR is now widely used in many different areas in Finance and Economics, from the evaluation of portfolio performance to market efficiency test (see, for example, Ofek and Richardson, 2003; Agarwal and Naik, 2004).

Although the SR has been widely used with a myriad of interpretations, only a few literary papers study its statistical properties. Jobson and Korkie (1981) first develop a Sharpe-ratio statistic to test for the equality of two SRs, whereby the statistic is being further modified and improved by Cadsby (1986) and Memmel (2003). On the other hand, by invoking the standard econometric methods with several different sets of assumptions imposing on the statistical behavior of the return series, Lo (2002) derives the asymptotic statistical distribution for the SR estimator and shows that confidence intervals, standard errors, and hypothesis tests can be computed for the estimated SRs in much the same way as regression coefficients such as portfolio alphas and betas are computed.

The SR test statistics developed by Jobson and Korkie (1981) and others are important as they provide a formal statistical comparison for the performances among portfolios. However, as the SR statistic possesses only the asymptotic distribution, one could only obtain its properties for large samples, but not for small samples. Nevertheless, it is important in finance to compare the performance of assets by using small samples, especially before and after markets change their directions, in which only small samples could be used to predict the assets' future performance. Also it is, sometimes, not so meaningful to measure SRs for too long periods as the means and standard deviations of the underlying assets could be empirically non-stationary and/or possessing structural breaks. The main obstacle in developing the SR test for small samples is that it is impossible to obtain a uniformly most powerful unbiased (UMPU) test to check for the equality of SRs in case of small samples. To circumvent this problem, Bai, et al. (2011c) propose to use the MV ratio (MVR) for the comparison. They also discuss the evaluation of the performance of assets for small samples by providing a theoretical framework and then invoking both one-sided and two-sided UMPU MVR tests.

To demonstrate the superiority of our proposed test over the traditional SR test, we apply both tests to analyze the performance of funds from Commodity Trading Advisors (CTAs) which involve the trading of commodity futures, financial futures and options on futures. There are many studies analyzing CTAs, in which some (Elton et al. 1987) conclude that CTAs offer neither an attractive alternative to bonds and stocks nor a profitable addition to a portfolio of bond and stocks. Whereas, others (Brorsen and Irwin 1985) conclude that commodity funds produce favorable and appropriate investment returns. We choose analyzing CTAs to illustrate the theories we developed because CTAs have become very popular with many investors, including universities; the number of universities increasingly allotting their university endowment funds to CTAs has grown significantly (Kat 2004).

Applying the traditional SR test, we fail to reject the possibility of having any significant difference among most of the CTA funds; thereby implying that most of the CTA funds being analyzed are indistinguishable in their performance. This conclusion may not necessarily be accurate as the insensitivity of the SR test is well known due to its limitation on the analysis for small samples. Thus, we invoke our proposed statistic, which is valid for small samples as well as large samples, to the analysis; the conclusion drawn from our proposed test will then be meaningful. As expected, contrary to the conclusion drawn by applying SR test, our proposed MVR test shows that the MVRs of some CTA funds are different from the others. This means that some CTA funds outperform other CTA funds in the market. Thus, the test developed in our paper provides more meaningful information in the evaluation of the portfolios' performance and enable investors to make wiser decisions in their investments.

On the other hand, when we apply the SR statistic to some other funds, we find that the statistic indicates that one fund significantly outperforms another fund even though the difference between the two funds becomes insignificantly small or even changes direction. This shows that the SR statistic may not be able to reveal the real short-run performance of the funds. On the other hand, in our analysis, we find that our proposed MVR statistic could reveal such changes. This shows the superiority of our proposed statistic in detecting short term performance, and in return, enabling the investors to make better decisions in their various investments. In addition, in this paper, we show that the values of the MVR are proportional to the corresponding investment plans in the optional MV optimization. This shows that the MVR test not only enables investors to find out which asset is superior in performance, but also enables investors to compute its corresponding investment plan over the asset. On the other hand, as the SR is not proportional to the weight of the corresponding asset, an asset with the highest Sharpe ratio does not infer that one should put highest weight on this asset whereas our MVR does. In this sense, our proposed test is superior.

The rest of the paper is organized as follows: Section 1 begins by providing a theoretical framework and goes on to develop the theory for both one-sided and two-sided MVR tests and studies its properties. In Section 2, we demonstrate the superiority of our proposed tests over the traditional SR tests by applying both tests to analyze the CTAs. This is then followed up by Section 3 which summarizes our conclusions and shares our insights. Technical proofs of some propositions are provided in the appendix.

1 The Theory

Let X_i and Y_i $(i = 1, 2, \cdots, n)$ be independent excess returns drawn from the corresponding normal distributions $N(\mu, \sigma^2)$ and $N(\eta, \tau^2)$ with joint density $p(x, y)$ such that

$$p(x,y) = k \times \exp(\frac{\mu}{\sigma^2}\sum x_i - \frac{1}{2\sigma^2}\sum x_i^2 + \frac{\eta}{\tau^2}\sum y_i - \frac{1}{2\tau^2}\sum y_i^2) \tag{1}$$

where $k = (2\pi\sigma^2)^{-n/2}(2\pi\tau^2)^{-n/2}\exp(-\frac{n\mu^2}{2\sigma^2})\exp(-\frac{n\eta^2}{2\tau^2})$.

To evaluate the performance of the prospects X and Y, financial practitioners and academicians are interested in testing the hypotheses

$$H_0^* : \frac{\mu}{\sigma} \leq \frac{\eta}{\tau} \quad \text{versus} \quad H_1^* : \frac{\mu}{\sigma} > \frac{\eta}{\tau} \tag{2}$$

to compare the performance of their corresponding SRs, $\frac{\mu}{\sigma}$ and $\frac{\eta}{\tau}$, the ratios of the excess expected returns to their standard deviations.

Rejecting H_0^* implies X to be the better investment prospect with larger SR that X has either larger excess mean return or smaller standard deviation or both. Jobson and Korkie (1981) and Memmel (2003) develop test statistics to test the hypotheses in (2) for large samples but their tests are not appropriate for testing small samples as the distribution of their test statistics is only valid asymptotically, but is not valid for small samples. However, it will be important in finance to test the hypotheses in (2) for small samples to provide useful investment information to investors. Furthermore, as it is impossible to obtain any UMPU test statistic to test the inequality of the SRs in (2) for small samples, Bai, et al. (2011c) propose to alter the hypothesis to test the inequality of the MVRs as shown in the following:

$$H_0 : \frac{\mu}{\sigma^2} \leq \frac{\eta}{\tau^2} \quad \text{versus} \quad H_{11} : \frac{\mu}{\sigma^2} > \frac{\eta}{\tau^2}\,. \tag{3}$$

They develop the UMPU test statistic to test the above hypotheses. Rejecting H_0 suggests X to be the better investment prospect as X possesses either smaller variance or bigger excess mean return or both. As, sometimes, investors do conduct the two-sided test to compare the MVRs, to complete the theory, they also consider the following hypotheses in this paper:

$$H_0 : \frac{\mu}{\sigma^2} = \frac{\eta}{\tau^2} \quad \text{versus} \quad H_{12} : \frac{\mu}{\sigma^2} \neq \frac{\eta}{\tau^2} \,. \tag{4}$$

One may argue that SR test is scale invariant whereas the MV ratio test is not and thus MVR test is not as good as SR test. To support the MVR test to be an alternative reasonably good choice, before developing the test statistic, we first show the theoretical justification for the use of the MV test statistic in the following remark and lemma:

Remark 1 *One may think that the MVR could be less favorable than the SR as the former is not scale invariant while the latter is. However, in some financial processes, the mean change in a short period of time is proportional to its variance change. For example, many financial processes could be characterized by the following diffusion process such for stock prices formulated as*

$$dY_t = \mu^P(Y_t)dt + \sigma(Y_t)dW_t^P,$$

(see, Cheridito et al., 2004), where μ^P *is an* N*-dimensional function,* σ *is an* $N \times N$ *matrix and* W_t^P *is an* N*-dimensional standard Brownian motion under the objective probability measure* P*. Under this model, the conditional mean of the increment* dY_t *given* Y_t *is* $\mu^P(Y_t)dt$ *and the covariance matrix is* $\sigma(Y_t)\sigma^T(Y_t)dt$*. When* $N = 1$*, the SR will be close to 0 while the MVR will be independent of* dt*. Thus, when the time period* dt *is small, it is better to consider the MVR rather than the SR.*

The above remark lends credibility to the use of the MVR tests. To give example to further support the use of MVR, in this paper we shed light on the preference of the MVR based on the Markowitz MV optimization theory as follows: suppose that there are p-branch of assets $\mathbf{S} = (s_1, ..., s_p)^T$ whose returns are denoted by $\mathbf{r} = (r_1, ..., r_p)^T$ with mean $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)^T$ and covariance matrix $\Sigma = (\sigma_{ij})$. In addition, we suppose an investor will invest capital C on the p-branch of securities $\mathbf{S}$ such that she/he wants to find out her/his optimal investment plan $\mathbf{c} = (c_1, ..., c_p)^T$ to allocate her/his investable wealth on the p-branch of securities to obtain maximize return subject to a given level of risk.

The above maximization problem can be formulated to the following optimization problem:

$$\max R = \mathbf{c}^T \boldsymbol{\mu}, \ \text{ subject to } \ \mathbf{c}^T \Sigma \mathbf{c} \leq \sigma_0^2 \tag{5}$$

where σ_0^2 is a given risk level. We call R satisfying (5) to be optimal return and $\mathbf{c}$ to be its corresponding allocation plan. One could easily extend the separation Theorem (Cass and Stiglitz, 1970) and the mutual fund theorem (Merton, 1972) to obtain the solution of (5)[1] from the following lemma:

Lemma 1 *For the optimization setting displayed in (5), the optimal return, R, and its corresponding investment plan, c, are obtained as follows:*

$$R = \sigma_0 \sqrt{\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}}$$

and

$$\boldsymbol{c} = \frac{\sigma_0}{\sqrt{\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}}} \Sigma^{-1} \boldsymbol{\mu} \, . \tag{6}$$

From Lemma 1, we find that the investment plan, $\mathbf{c}$, is proportional to the MVR when Σ is a diagonal matrix. Hence, when the asset is concluded to be superior in performance by utilizing the MVR test, its corresponding weight could then be computed based on the corresponding MVR test value. Thus, another advantage of using the MVR test over the Sharpe ratio test is that it allows investors to compare

[1]There are several studies, for example, Ju and Pearson (1999) and Maller and Turkington (2002) that result in different solutions for settings similar to that in (5). We note that Bai et al. (2009a,b, 2011b) also use the same framework as in (5).

the performance of different portfolios as it enables investors to find out which asset they should put heavier weight and vice versa. It also enables investors to compute the corresponding allocation for the assets. On the other hand, as the SR is not proportional to the weight of the corresponding asset, an asset with the highest SR does not infer that one should put highest weight on this asset whereas our MVR does. In this sense, the test proposed by Bai, et al. (2011c) is superior.

Bai, et al. (2011c) also develop both one-sided UMPU test and two-sided UMPU test to check the equality of the MVRs for comparing the performances of different prospects with hypotheses stated in (3) and (4) respectively. We first state the one-sided UMPU test for the MVRs as follows:

Theorem 2 *Let X_i and Y_i $(i = 1, 2, \cdots, n)$ be independent random variables with joint distribution function defined in (1). For the hypotheses setup in (3), there exists a UMPU level-α test with the critical function $\phi(u, t)$ such that*

$$\phi(u,t) = \begin{cases} 1, & \text{when } u \geq C_0(t) \\ 0, & \text{when } u < C_0(t) \end{cases} \tag{7}$$

where C_0 is determined by

$$\int_{C_0}^{\infty} f^*_{n,t}(u)\ du = K_1\,; \tag{8}$$

with

$$\begin{aligned} f^*_{n,t}(u) &= (t_2 - \frac{u^2}{n})^{\frac{n-1}{2}-1}(t_3 - \frac{(t_1-u)^2}{n})^{\frac{n-1}{2}-1}\,, \\ K_1 &= \alpha \int_{\Omega} f^*_{n,t}(u)\ du\,; \end{aligned}$$

in which

$$U = \sum_{i=1}^{n} X_i\,, \quad T_1 = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i\,, \quad T_2 = \sum_{i=1}^{n} X_i^2\,, \quad T_3 = \sum_{i=1}^{n} Y_i^2\,, \quad T = (T_1, T_2, T_3)\,;$$

with $\Omega = \{u | \max(-\sqrt{nt_2}, t_1 - \sqrt{nt_3}) \le u \le \min(\sqrt{nt_2}, t_1 + \sqrt{nt_3})\}$ *to be the support of the joint density function of* (U, T).

We call the statistic U in Theorem 2 to be the one-sided MVR test statistic or simply the MVR test statistic for the hypotheses setup in (3) if no confusion occurs. In addition, Bai, et al. (2011c) introduce the two-sided UMPU test statistic as stated in the following theorem to test for the equality of the MVRs listed in (4):

Theorem 3 *Let* X_i *and* Y_i $(i = 1, 2, \cdots, n)$ *be independent random variables with joint distribution function defined in (1). Then, for the hypotheses setup in (4), there exists a UMPU level-*α *test with critical function*

$$\phi(u,t) = \begin{cases} 1, & \text{when } u \le C_1(t) \text{ or } \ge C_2(t) \\ 0, & \text{when } C_1(t) < u < C_2(t) \end{cases} \tag{9}$$

in which C_1 *and* C_2 *satisfy*

$$\begin{cases} \int_{C_1}^{C_2} f_{n,t}^*(u)\ du = K_2 \\ \int_{C_1}^{C_2} u f_{n,t}^*(u)\ du = K_3 \end{cases} \tag{10}$$

where

$$\begin{aligned} K_2 &= (1-\alpha) \int_{\Omega} f_{n,t}^*(u)\ du, \\ K_3 &= (1-\alpha) \int_{\Omega} u f_{n,t}^*(u)\ du\,. \end{aligned}$$

The terms $f^*_{n,t}(u)$, T_i $(i = 1, 2, 3)$ and T are defined in Theorem 2.

We call the statistic U in Theorem 3 to be the two-sided MVR test statistic or simply the MVR test statistic for the hypotheses setup in (4) if no confusion occurs. In this paper, Bai, et al. (2011c) propose to apply numerical methods to look for the critical values of the tests as stated in the following problem:

Problem 4 *To compute the values of the constants C_1 and C_2 in $\Omega = [I_d, I_u]$ such that*

$$\int_{C_1}^{C_2} f^*_{n,t}(u)\ du = K_2 \tag{11}$$

and

$$\int_{C_1}^{C_2} u f^*_{n,t}(u)\ du = K_3 \tag{12}$$

where

$$f^*_{n,t}(u) = (t_2 - \frac{u^2}{n})^{\frac{n-1}{2}-1}(t_3 - \frac{(t_1 - u)^2}{n})^{\frac{n-1}{2}-1},$$

$$K_2 = (1-\alpha)\int_\Omega f^*_{n,t}(u)\ du,$$

and

$$K_3 = (1-\alpha)\int_\Omega u f^*_{n,t}(u)\ du\,.$$

To solve this problem, we have to conduct the following steps:

Step 1: We first let

$$\delta_0 = (I_u - I_d)/K, \quad \text{and} \quad C_1 = I_d \tag{13}$$

where I_d and I_u are two end points of the support interval defined in Problem 4. Here, K is an integer chosen to be big enough, say for example, 400, such that $(I_u - I_d)/K$ is set to be a small increment.

Step 2: Thereafter, we let

$$C_1 = C_1 + k\delta_0, \quad k = 0,\ 1, ...,\ K.$$

For each C_1, we are going to solve equations in (11) and (12) to obtain two approximate values of C_2, one obtained by solving (11) and another obtained by solving (12). If the approximate values of C_2 are very close, they could be used as the approximate solutions to equations in (11) and (12). If not, we move on to let $k = k + 1$ and continue the process in Step 2 till the values of two C_2's are approximately equal. In this procedure, we can achieve the appropriate calculation precision by controlling the precision of solutions to equations in (11) and (12) respectively. We note that one could choose a very large value of K so as to get δ_0 as small as possible. However, it is not necessary to do so because any big value of K could not improve the calculation precision remarkably and thus we suggest using 400 which is large enough.

2 Illustration

In this section, we demonstrate the superiority of the MV tests developed in this paper over the traditional SR tests by illustrating the applicability of our tests to the decision making process of investing in commodity trading advisors (CTAs). For simplicity, we only demonstrate the two-sided UMPU test.[2] The data analyzed in this section are the monthly returns of 61 indices from CTAs for the sample period from January 2001 to December 2004 in which the data from Jan 2003 to Dec 2003 are used to compute the MVR in Jan 2004, while the data from Feb 2003 to Jan 2004 are used to compute the MVR in Feb 2004, and so on. However, using too short periods to compute the SRs would not be meaningful as discussed in our previous sections. Thus, we utilize a longer period from Jan 2001 to Dec 2003 to compute the SR ratio in Jan 2004, from Feb 2001 to Jan 2004 to compute the in Feb 2004, and so on.[3]

For simplicity, in our illustration we only report the comparison of three pairs of indices with the largest or smallest means, variances, or MVRs, respectively, from January 2004 and December 2004. They are: AIS Futures Fund LP (maximum mean, denoted by X_{11}) versus Beacon Currency Fund (minimum mean, X_{12}), JWH Global Financial & Energy Portfolio (maximum variance, X_{21}) versus Worldwide Financial Futures Program (minimum variance, X_{22}), Oceanus Fund Ltd (maximum MVR, X_{31}) versus Beacon Currency Fund (minimum MVR, X_{32}). Let $r_{ij,t}$ be the excess return of X_{ij} over the risk-free interest rate at time t with mean μ_{ij} and variance σ_{ij}^2

[2]The results of the one-sided test which draw a similar conclusion are available on request.

[3]We note that, actually, we should use even longer periods to compute the SRs but the earlier data are not available or do not exist at all. Also, the results for too long periods are expected to yield insignificant difference for all comparison, which is not useful to investors.

for $i = 1,2,3$ and $j = 1,2$ respectively. The 3-month Treasury bills rate obtained from Datastream is used to proxy the risk-free rate. We test the following hypotheses:

$$H_{0i} : \frac{\mu_{i1}}{\sigma_{i1}^2} = \frac{\mu_{i2}}{\sigma_{i2}^2} \quad \text{versus} \quad H_{1i} : \frac{\mu_{i1}}{\sigma_{i1}^2} \neq \frac{\mu_{i2}}{\sigma_{i2}^2} \quad \text{for } i = 1,2,3. \tag{14}$$

To test the hypotheses in (14), we first compute the values of the test function U for the MVR statistic shown in (9) for each pair of funds and display the values in Tables 1, 2 and 3 respectively. We then compute the critical values C_1 and C_2 under the test level of 0.05 for each pair of funds to test the hypotheses in (14). In addition, in order to illustrate the performance of the funds and their corresponding test results visually, we exhibit the returns of the two funds being compared and their difference for each pair of funds in Figures 1A, 2A, and 3A respectively, and display their corresponding values of U with C_1 and C_2 in Figures 1B, 2B and 3B respectively.

For comparison, we also compute the corresponding SR statistic developed by Jobson and Korkie (1981) and Memmel (2003) such that

$$z_i = \frac{\hat{\sigma}_{i2}\hat{\mu}_{i1} - \hat{\sigma}_{i1}\hat{\mu}_{i2}}{\sqrt{\hat{\theta}}} \tag{15}$$

which follows standard normal distribution asymptotically with

$$\theta = \frac{1}{T}\left[2\sigma_{i1}^2\sigma_{i2}^2 - 2\sigma_{i1}\sigma_{i2}\sigma_{i1,i2} + \frac{1}{2}\mu_{i1}^2\sigma_{i2}^2 + \frac{1}{2}\mu_{i2}^2\sigma_{i1}^2 - \frac{\mu_{i1}\mu_{i2}}{\sigma_{i1}\sigma_{i2}}\sigma_{i1,i2}^2\right]$$

to test for the equality of the SRs for the funds by setting the following hypotheses such that

$$H_{0i}^* : \frac{\mu_{i1}}{\sigma_{i1}} = \frac{\mu_{i2}}{\sigma_{i2}} \quad \text{versus} \quad H_{1i}^* : \frac{\mu_{i1}}{\sigma_{i1}} \neq \frac{\mu_{i2}}{\sigma_{i2}} \quad \text{for } i = 1,2,3. \tag{16}$$

Different from using one-year data to compute the values of our proposed statistic, we use the overlapping three-year data to compute the SR statistic for the year 2004 as discussed before. The results are also reported in Tables 1 to 3 next to the results for our proposed statistic while their plots and their critical values are depicted in Figures 1C to 3C for comparison.

We first examine the performance between the returns of AIS Futures Fund LP, the fund with the largest mean, and those of Beacon Currency Fund, the fund with the smallest mean. As shown in Table 1 and Figure 1C, we cannot detect any significant difference between their SRs, implying that the performances of these two funds are indistinguishable. We note that the three-year monthly data being used to compute the SR statistic could be too short to satisfy the asymptotic statistical properties for the test but, still, we cannot find any significant difference between the performance of these two funds. If we use any longer period, the result is expected to be insignificant as the high means in some sub-periods could be offset by the low means in other sub-periods. Thus, a possible limitation of applying the SR test is that it would usually conclude indistinguishable performances among the funds, which may not be the situation in reality. In this aspect, looking for a statistic to evaluate the performance among assets for short periods is essential. In this paper, we adopt our proposed statistic to conduct the analysis. As shown in Table 1 and Figure 1, we find that our proposed statistic does not disappoint us that it does show some significant differences in performance between these two funds in some periods. This information could be useful to investors for their decision making process.

Similar conclusion could also be drawn for the comparison between JWH Global Financial & Energy Portfolio and Worldwide Financial Futures Program; the former is the fund possessing the maximum variance while the latter attains the minimum variance. Again, applying the SR test concludes that the performance between these two funds is indistinguishable while invoking our proposed statistic enables us to detect some significant differences.

Then, we turn to investigate the performance between Oceanus Fund Ltd and Beacon Currency Fund in 2004, with the former possessing the maximum MVR while the latter attaining the minimum MVR. From Table 3, we find that the differences between these two funds become very small after June 2004 and even turn positive to negative in September 2004. However, the SR test cannot detect such change and indicates that Oceanus Fund Ltd performs significantly better than Beacon Currency Fund in the entire 2004. In applying our proposed MVR test, this test reveals that the change in its value has become insignificant after June 2004. The information that is derived from our proposed test is thus useful for investors who takes their decision making with regard to their investment seriously.

3 Concluding Remarks

In this paper, to evaluate the performance among the assets for small samples, we propose to apply the MV test statistics developed by Bai, et al. (2011c). We illustrate the superiority of the proposed test over the traditional SR test by applying both tests to analyze the performance of funds from Commodity Trading Advisors. Our findings show that while the traditional SR test concludes most of the CTA funds being analyzed as being indistinguishable in their performance, our proposed statistic shows that some funds outperform the others. In addition, when we apply the SR statistic on some other funds, we find that the statistic indicates that one fund is significantly outperforming another fund even though the difference between the two funds become insignificantly small or even changes directions. However, when our proposed MVR statistic is applied, we could detect such changes. This shows the superiority of our proposed statistic in revealing short term performance and in return enables the investors to make better decision about their investments.

We note that although in some situations, data could be transformed to be normally-distributed as discussed in our theory section whereas, in some other situations, this transformation may not be possible. Thus, further research could include applying the approaches by Dufour et al. (2003) and others to extend our proposed MV test to relax the normality assumption. However, we note that the price to relax normality assumption is that the test may no longer be uniformly most powerful unbiased as shown in our paper. Nonetheless, our proposed test statistic will still have some merits over the statistics relaxing the normality assumption. Further research could also include conducting simulation to study the robustness of our proposed MVR test. If our proposed MVR test is found to be robust to non-normality, the MVR test will then be a good test for non-normality data as well as normality data. Another direction of further research is to develop confidence interval for the MVR which could shed new light on asset investments.

There are two basic approaches to the problem of portfolio selection under uncertainty. One approach is based on the concept of utility theory, see for example, Wong and Li (1999), Post and Levy (2005), Wong and Chan (2008), Wong and Ma (2008), Wong et al. (2006, 2008), and Sriboonchita et al. (2009) for more information. Davidson and Duclos (2000), Barrett and Donald (2003), Linton et al. (2005), Bai, et al. (2011a) and others have developed several stochastic dominance (SD) test statistics using this approach. This approach offers a mathematically rigorous treatment for portfolio selection but it is not popular among investors since few investors like to specify their utility functions and choose a distributional assumption for the returns before their investment decision making.

The other is the mean-risk (MR) analysis as discussed in this paper. In this approach, the portfolio choice is made with respect to two measures – the expected portfolio mean return and portfolio risk. A portfolio is preferred if it has higher expected return and smaller risk. There are convenient computational recipes and geometric interpretations of the trade-off between the two measures. A disadvantage of the latter approach is that it is derived by assuming the Von Neumann-Morgenstern quadratic utility function and that returns are normally distributed (Feldstein, 1969; Hanoch and Levy, 1969). Thus, it cannot capture the richness of the former. Among the MR analysis, the most popular measure is the SR introduced by Sharpe (1966). As the SR requires strong assumptions that the assets being analyzed have to be iid, various measures for MR analysis have been developed to improve the SR, including the Sortino ratio (Sortino and van der Meer, 1991), the conditional SR (Agarwal and Naik, 2004), the modified SR (Gregoriou and Gueyie, 2003), Value-at-Risk (Christoffersen, 2004; Chen, 2005; Kuester, et al., 2006), Expected Shortfall (Chen, 2008) and others. However, most empirical studies, see for example, Eling and Schuhmacher (2007), find that the conclusions drawn by using these ratios are basically the same as that drawn by the SR. Nonetheless, recently Leung and Wong (2008) develop a multiple SR statistic and find that the results drawn from the multiple Sharpe ratio statistic could be different from its counterpart pair-wise SR statistic comparison, indicating that there are some relationships among the assets that have not being revealed from the pair-wise SR statistics.

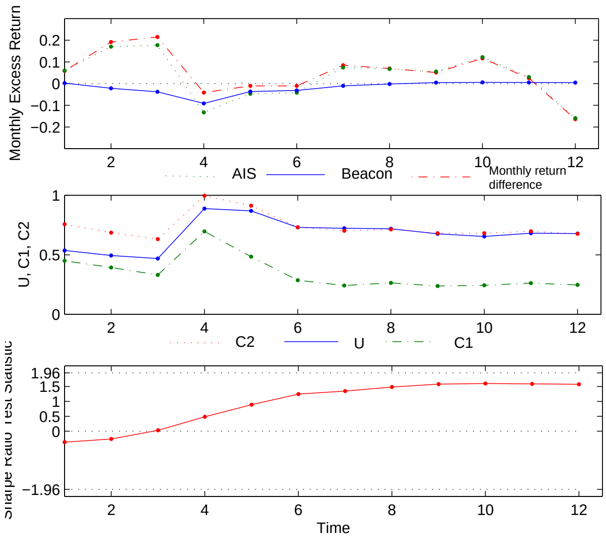
The major limitation of these pair-wise MR statistics is that up to now academics can only develop their asymptotic distributions, but not their distribution for small samples. Need not to say about their performance in small samples, even for large sample, investors do not know how large the sample size should be to make these distributions valid for testing purpose. As their testing results are only valid asymptotically, they may not be valid in small samples nor samples with not too big sizes. For very large samples, the results of these tests could be valid. However, as discussed in our introduction section, too large sample could result in extreme positive difference canceling out the extreme negative difference and the conditions of the market may not be the same over long period of time. Thus, we are not surprised that in most empirical studies these MR statistics draw similar conclusions in their testing.

The SD test statistics could be superior to the MR test statistics as the conclusions drawn by these SD test statistics between the assets being examined could be used by investors to compare their expected utility on these assets since they do not require investors to possess quadratic utility function nor any form of the distribution for the assets being analyzed.

So far, in the literature of the development of test statistics for portfolio selection, academics have developed the SD test statistics and the MR test statistics to examine the preferences of different assets or portfolios. However, all the SD test statistics and the MR test statistics are valid only asymptotically, and thus the conclusion drawn by these statistics may not be valid if one applies these tests to small samples. Nonetheless, as discussed in our introduction, the comparison of asset performance for small samples is very important but so far, in the literature, such test is not available. Thus, the test developed in this paper sets a milestone in the literature of financial economics. The test developed in our paper is the first test making such comparison possible.

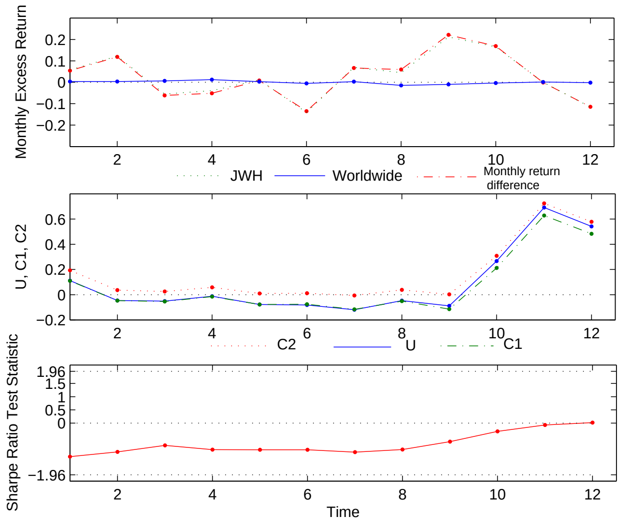
One may claim that the limitation of our proposed MV statistic is that it could only draw conclusion for investors with quadratic utility functions and for normal-distributed assets. Our answer is that it may not be. Meyer (1987), Wong (2007), and Wong and Ma (2008) have shown that the conclusion drawn from the MR comparison is equivalent to the comparison of expected utility maximization for any risk-averse investor, not necessarily with only quadratic utility function, and for assets with any distribution, not necessarily normal distribution, if the assets being examined belong to the same location-scale family. In addition, one could also apply Theorem 10 in Li and Wong (1999) to generalize the result so that it can be valid for any risk-averse investor and for portfolios with any distribution if the portfolios being examined belong to the same convex combinations of (same or different) location-scale families. The location-scale family can be very large, containing normal distributions as well as t-distributions, gamma distributions, etc. The stock returns could be expressed as convex combinations of normal distributions, t-distributions and other location-scale families, see for example, Fama (1963), Clark (1973), Fielitz and Rozelle (1983), and Kon (1984). Thus, the conclusions drawn from the test statistics developed in this paper are valid for most of the stationary data including most, if not all, of the returns of different portfolios.

Figure 1: Plots of the monthly excess returns for AIS Futures Fund LP and Beacon Currency Fund and corresponding Mean-Variance Ratio test U and Sharpe ratio test statistic Z



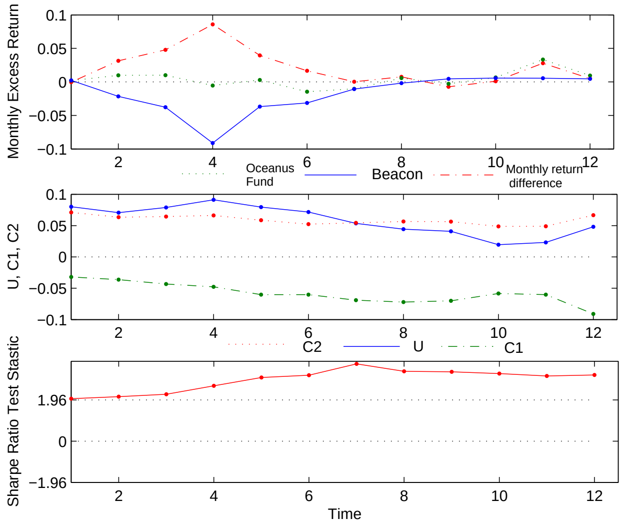
Note: The Mean-Variance Ratio test U is defined in Theorem 2 with C_1 and C_2 defined in (10) and the Sharpe ratio test statistic Z is defined in (15).

Figure 2: Plots of Monthly excess returns of JWH Global Financial & Energy Portfolio and Worldwide Financial Futures Program and corresponding Mean-Variance Ratio test U and Sharpe ratio test statistic Z



Note: The Mean-Variance Ratio test U is defined in Theorem 2 with C_1 and C_2 defined in (10) and the Sharpe ratio test statistic Z is defined in (15).

Figure 3: Plots of the monthly excess returns for Oceanus Fund Ltd versus Beacon Currency Fund and corresponding Mean-Variance Ratio test U and Sharpe ratio test statistic Z



Note: The Mean-Variance Ratio test U is defined in Theorem 2 with C_1 and C_2 defined in (10) and the Sharpe ratio test statistic Z is defined in (15).

Table 1: The Results of the Mean-Variance Ratio Test and Sharpe ratio Test for AIS Futures Fund LP versus Beacon Currency Fund in 2004

Time	$X_{11} - X_{12}$	Mean-Variance Ratio Test			Sharpe ratio Test	
		U	C_1	C_2	Z	p-value
Jan	0.0580	0.5368	0.4501	0.7574	-0.3717	0.71
Feb	0.1923	0.4943	0.3938	0.6860	-0.2634	0.79
Mar	0.2153	0.4692	0.3311	0.6315	0.0291	0.98
Apr	-0.0412	0.8881	0.6969	0.9963	0.4823	0.63
May	-0.0104	0.8691	0.4846	0.9127	0.8899	0.37
Jun	-0.0107	0.7300	0.2861	0.7310	1.2447	0.21
Jul	0.0851	0.7234*	0.2424	0.7013	1.3469	0.18
Aug	0.0697	0.7190*	0.2647	0.7130	1.4847	0.14
Sep	0.0513	0.6762	0.2381	0.6817	1.5838	0.11
Oct	0.1166	0.6545	0.2441	0.6812	1.6034	0.11
Nov	0.0251	0.6813	0.2618	0.6971	1.5911	0.11
Dec	-0.1639	0.6784*	0.2471	0.6760	1.5783	0.11

* $p < 5\%$, the Mean-Variance Ratio Test U is defined in (9) while the Sharpe ratio Test Z is defined in (15).

Table 2: The Results of the Mean-Variance Ratio Test and Sharpe ratio Test for JWH Global Financial & Energy Portfolio versus Worldwide Financial Futures Program in 2004

		Mean-Variance Ratio Test			Sharpe ratio Test	
Time	$X_{21} - X_{22}$	U	C_1	C_2	Z	p-value
Jan	0.0542	0.1125	0.1103	0.1932	-1.2710	0.20
Feb	0.1188	-0.0455	-0.0465	0.0366	-1.0933	0.27
Mar	-0.0616	-0.0507	-0.0538	0.0253	-0.8457	0.40
Apr	-0.0514	-0.0115	-0.0153	0.0581	-1.0057	0.31
May	0.0082	-0.0773	-0.0775	0.0101	-1.0128	0.31
Jun	-0.1350	-0.0814*	-0.0757	0.0116	-1.0132	0.31
Jul	0.0664	-0.1188*	-0.1157	-0.0062	-1.1002	0.27
Aug	0.0597	-0.0464	-0.0504	0.0381	-1.0017	0.32
Sep	0.2215	-0.0879	-0.1143	0.0032	-0.7042	0.48
Oct	0.1690	0.2671	0.2124	0.3086	-0.3144	0.75
Nov	-0.0011	0.6918	0.6276	0.7242	-0.0742	0.94
Dec	-0.1143	0.5415	0.4828	0.5787	0.0165	0.99

* $p < 5\%$, the Mean-Variance Ratio Test U is defined in (9) while the Sharpe ratio Test Z is defined in (15).

Table 3: The Results of the Mean-Variance Ratio Test and Sharpe ratio Test for Oceanus Fund Ltd versus Beacon Currency Fund in 2004

		Mean-Variance Ratio Test			Sharpe ratio Test	
Time	$X_{31} - X_{32}$	U	C_1	C_2	Z	p-value
Jan	-0.0003	0.0801*	-0.0320	0.0711	2.0180	0.04*
Feb	0.0315	0.0707*	-0.0362	0.0635	2.1148	0.03*
Mar	0.0477	0.0789*	-0.0434	0.0644	2.2311	0.03*
Apr	0.0858	0.0913*	-0.0478	0.0662	2.6265	0.01*
May	0.0396	0.0795*	-0.0602	0.0586	3.0249	0.00*
Jun	0.0166	0.0718*	-0.0602	0.0524	3.1306	0.00*
Jul	0.0002	0.0535	-0.0690	0.0543	3.6708	0.00*
Aug	0.0075	0.0444	-0.0722	0.0566	3.3177	0.00*
Sep	-0.0074	0.0410	-0.0701	0.0563	3.2926	0.00*
Oct	0.0010	0.0195	-0.0584	0.0488	3.2088	0.00*
Nov	0.0278	0.0231	-0.0602	0.0490	3.0957	0.00*
Dec	0.0049	0.0481	-0.0911	0.0666	3.1418	0.00*

* $p < 5\%$, the Mean-Variance Ratio Test U is defined in (9) while the Sharpe ratio Test Z is defined in (15).

References

Agarwal, V., Naik, N.Y., 2004. Risk and portfolios decisions involving hedge funds. Review of Financial Studies 17(1), 63-98.

Bai, Z.D., H. Li, H.X. Liu, W.K. Wong. 2011a. Test Statistics for Prospect and Markowitz Stochastic Dominances with Applications, Econometrics Journal, (forthcoming).

Bai, Z.D., H.X. Liu, W.K. Wong. 2009a. Enhancement of the Applicability of Markowitz's Portfolio Optimization by Utilizing Random Matrix Theory. Mathematical Finance 19(4) 639-667.

Bai, Z.D., Liu, H.X., Wong, W.K., 2009b. On the Markowitz Mean-Variance Analysis of Self-Financing Portfolios, Risk and Decision analysis, 1(1), 35-42.

Bai, Z.D., H.X. Liu, W.K. Wong. 2011b. Asymptotic Properties of Eigenmatrices of A Large Sample Covariance Matrix. Annals of Applied Probability (forthcoming).

Bai, Z.D., K.Y. Wang, W.K. Wong. 2011c. Mean-Variance Ratio Test, A Complement to Coefficient of Variation Test and Sharpe Ratio Test, Statistics and Probability Letters, (forthcoming).

Bai, Z.D., W.K. Wong, B. Zhang. 2010. Multivariate Linear and Non-Linear Causality Tests. Mathematics and Computers in Simulation 81 5-17.

Barrett, G.F., Donald, S.G., 2003. Consistent tests for stochastic dominance. Econometrica 71, 71-104.

Brorsen, B.W., Irwin, S.H., 1985. Examination of commodity fund performance. Review of Futures Market 4, 84-94.

Cadsby C.B., 1986. Performance hypothesis testing with the Sharpe and Treynor measures: a comment. Journal of Finance 41, 1175-1176.

Cass, D., Stiglitz, J.E., 1970. The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds. Journal of Economic Theory 2(2), 122-160.

Clark, P.K., 1973. A subordinated stochastic process model with finite variance for speculative prices. Econometrica 37, 135-155.

Chen, S.X., 2005. Nonparametric inference of value-at-risk for dependent financial returns. Journal of Financial Econometrics 3(2), 227-255.

Chen, S.X., 2008. Nonparametric estimation of expected shortfall. Journal of Financial Econometrics 6, 87-107.

Cheridito, P., Filipović, D., Kimmel, R.L., 2004. Market price of risk specifications for affine models: theory and evidence. Econometric Society 2004 North American Winter Meetings 536.

Christoffersen, P., 2004. Backtesting value-at-risk: a duration-based approach. Journal of Financial Econometrics 2(1), 84-108.

Davidson, R.. Duclos, J.Y., 2000. Statistical inference for stochastic dominance and for the measurement of poverty and inequality. Econometrica 68, 1435-1464.

Dufour, J.M., Khalaf, L., Beaulieu, M.C., 2003. Exact skewness-kurtosis tests for multivariate normality and goodness-of-fit in multivariate regressions with application to asset pricing models. Oxford Bulletin of Economics and Statistics 65, 891-906.

Eling, M., Schuhmacher, F., 2007. Does the choice of performance measure influence the evaluation of hedge funds? Journal of Banking and Finance 31, 2632-2647.

Elton, E.J., Gruber, M.J., Rentzler, J., 1987. Professionally managed, publicly traded commodity funds. Journal of Business 60, 175-199.

Fama, E.F., 1963. Mandelbrot and the stable Paretian hypothesis. Journal of Business 36, 420-429.

Feldstein, M.S., 1969. Mean variance analysis in the theory of liquidity preference and portfolio selection. Review of Economics Studies 36, 5-12.

Fielitz, B.D., Rozelle, J.P., 1983. Stable distributions and mixtures of distributions hypotheses for common stock return. Journal of American Statistical Association 78, 28-36.

Gregoriou, G.N., Gueyie, J.-P., 2003. Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio. Journal of Wealth Management 6 (Winter), 77-83.

Hanoch, G., Levy, H., 1969. The efficiency analysis of choices involving risk. Review of Economic Studies 36, 335-346.

Jobson, J.D., Korkie, B., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. Journal of Finance 36, 889-908.

Ju, X., Pearson, N., 1999. Using value-at-risk to control risk taking: how wrong can you be? Journal of Risk 1(2), 5-36.

Kat, H.M., 2004. In search of the optimal fund of hedge funds. Journal of Wealth Management 6(4), 43-51.

Kon, S.J., 1984. Models of stock returns – a comparison. Journal of Finance 39, 147-165.

Kuester, K., Mittnik, S., Paolella, M.S., 2006. Value-at-risk prediction: a comparison of alternative strategies. Journal of Financial Econometrics 4(1), 53-89.

Lehmann, E.L., 1986. Testing statistical hypotheses. John Wiley & Sons, University of California, Berkeley.

Leung, P.L., Wong, W.K., 2008. On testing the equality of the multiple Sharpe Ratios, with application on the evaluation of iShares. Journal of Risk 10(3), 1-16.

Li, C.K., Wong, W.K., 1999. Extension of stochastic dominance theory to random variables. RAIRO Recherche Opérationnelle 33(4), 509-524.

Linton, O., Maasoumi, E., Whang, Y.J., 2005. Consistent testing for stochastic dominance under general sampling schemes. Review of Economic Studies 72, 735-765.

Lintner, J., 1965. The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets. Review of Economics and Statistics 47, 13-37.

Lo, A., 2002. The statistics of Sharpe ratios. Financial Analysis Journal 58, 36-52.

Maller, R.A., Turkington, D.A., 2002. New light on the portfolio allocation problem. Mathematical Methods of Operations Research 56(3), 501-511.

Markowitz, H.M., 1952. Portfolio selection. Journal of Finance 7, 77-91.

Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. Finance Letters 1, 21-23.

Merton, R.C., 1972. An analytic derivation of the efficient portfolio frontier. Journal of Financial and Quantitative Analysis 7, 1851-1872.

Meyer, J., 1987. Two-Moment decision models and expected utility maximization. American Economic Review 77(3), 421-30.

Ofek, E., Richardson, M., 2003. A survey of market efficiency in the internet sector. Journal of Finance 58, 1113-1138.

Post, T., Levy, H., 2005. Does risk seeking drive stock prices? a stochastic dominance of aggregate investor preferences and beliefs. Review of Financial Studies 18, 925-953.

Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance 19, 425-442.

Sharpe, W.F., 1966. Mutual funds performance. Journal of Business 39(1), 119-138.

Sortino, F.A., van der Meer, R., 1991. Downside risk. Journal of Portfolio Management 17 (Spring), 27-31.

Sriboonchita, S., W.K. Wong, D. Dhompongsa, H.T. Nguyen. 2009. Stochastic Dominance and Applications to Finance, Risk and Economics. Chapman and Hall/CRC, Boca Raton, Florida.

Tobin, J., 1958. Liquidity preference as behavior towards risk. Review of Economic Studies 25, 65-86.

Wong, W.K., 2007. Stochastic dominance and mean-variance measures of profit and loss for business planning and investment. European Journal of Operational Research 182, 829-843.

Wong, W.K., Chan, R., 2008. Markowitz and prospect stochastic dominances. Annals of Finance 4(1), 105-129.

Wong, W.K., Li, C.K., 1999. A note on convex stochastic dominance theory. Economics Letters 62, 293-300.

Wong, W.K., Ma, C., 2008. Preferences over location-scale family. Economic Theory 37(1), 119-146.

Wong, W.K., K.F. Phoon, H.H. Lean. 2008. Stochastic dominance analysis of Asian hedge funds. Pacific-Basin Finance Journal 16(3), 204-223.

Wong, W.K., H.E. Thompson, S. Wei, Y.F. Chow. 2006. Do Winners perform better than Losers? A Stochastic Dominance Approach. Advances in Quantitative Analysis of Finance and Accounting 4, 219-254.