# Theory and Application of an Economic Performance Measure of Risk

**Cuizhen Niu***

*School of Statistics, Beijing Normal University, Beijing*

**Xu Guo**

*School of Statistics, Beijing Normal University, Beijing*

**Michael McAleer**

*Department of Quantitative Finance, National Tsing Hua University, Taiwan*

*Discipline of Business Analytics, University of Sydney Business School*

*Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam*

*Department of Quantitative Economics, Complutense University of Madrid*

*Institute of Advanced Sciences, Yokohawa National University, Japan*

**Wing-Keung Wong**

*Department of Finance, Asia University, Taiwan*

*Department of Economics, Lingnan University, Hong Kong*

June 15, 2017

EI2017-18

*The fourth author would like to thank Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. This research has been partially supported by grants from Beijing Normal University, Erasmus University Rotterdam, Complutense University of Madrid, Asia University, Lingnan University, Research Grants Council (RGC) of Hong Kong (project numbers 12502814 and 12500915), Australian Research Council, National Science Council, Ministry of Science and Technology (MOST), Taiwan, Fundamental Research Funds for the Central Universities, Natural Science Foundation of China (11601227), China Postdoctoral Science Foundation (2016M600951, 2017M610058), and Natural Science Foundation of Jiangsu Province, China (BK20150732).

## Abstract

Homm and Pigorsch (2012a) use the Aumann and Serrano index to develop a new economic performance measure (EPM), which is well known to have advantages over other measures. In this paper, we extend the theory by constructing a one-sample confidence interval of EPM, and construct confidence intervals for the difference of EPMs for two independent samples. We also derive the asymptotic distribution for EPM and for the difference of two EPMs when the samples are independent. We conduct simulations to show the proposed theory performs well for one and two independent samples. The simulations show that the proposed approach is robust in the dependent case. The theory developed is used to construct both one-sample and two-sample confidence intervals of EPMs for Singapore and USA stock indices.



***Keywords:*** Economic performance measure; Asymptotic confidence interval; Bootstrap-based confidence interval; Method of variance estimates recovery.

**JEL**: C12, C15

# 1 Introduction

Believing that less risk averse economic agents tend to accept riskier gambles, Aumann and Serrano (2008) used the reciprocal of the absolute risk aversion (ARA) of an investor with constant ARA to develop a new economic index of riskness, namely the Aumann and Serrano (AS) index. Thereafter, Homm and Pigorsch (2012a) used the AS index to develop a new economic performance measure (EPM), which can be obtained through dividing the mean of an investment portfolio by the AS index.

The EPM has many advantages over other commonly-used risk measures, such as the Sharpe ratio. For example, EPM is strictly monotonic with respect to stochastic dominance (SD), and consistently accounts for the mean, variance and higher moments of the returns distribution. If investment returns follow a normal distribution, the EPM and Sharpe ratio have the same ranking in measuring asset performance. Thus, the EPM generalizes the Sharpe ratio with respect to non-normal distributions.

Confidence intervals are usually regarded as more informative than hypothesis tests since they can provide a range of parameter values that reflect the degree of uncertainty in estimation. The confidence interval construction of the Sharpe ratio, a common performance measure of an investment, has been investigated by many researchers. Jobson and Korkie (1981) proposed a popular tool to test the difference of Sharpe ratios of two investment strategies, where the asymptotic distributions of the estimators of the Sharpe and Treynor performance measures are derived. Memmel (2003) corrected a typographical error in the Jobson-Korkie test, without loss of any statistical properties.

The above tests are not valid when returns have tails that are heavier than the normal distribution, or are time series data. Ledoit and Wolf (2008) applied robust inference methods, suggested constructing studentized time series bootstrap confidence intervals for the difference of Sharpe ratios, and declared the two ratios as different if zero is not contained in the interval. Constructing a confidence interval for an estimator is important for studying the statistical properties. Bartlett (1953) introduced the method to construct asymptotic confidence intervals for an unknown parameter, $\theta$, with higher moments of $\partial L/\partial\theta$, especially when the sample variance is heavily skewed for moderate degrees of freedom.

Ghosh (1979) compared two confidence intervals for the binomial parameter by confidence coefficients, the lengths and Neyman shortness, which were constructed based on the extensions of Clopper-Pearson confidence intervals. Brookmeyer and Crowley (1982) constructed confidence intervals for median survival time. Efron (1987) proposed superior bootstrap confidence intervals for a single parameter in a multi-parameter family. However, to the best of our knowledge, few references focus on the construction of confidence intervals for the economic performance measure with the AS index. The present paper focuses on this issue.

We develop the statistical theory to construct one-sample confidence intervals of EPM. For one-sample confidence intervals, we recommend using three approaches, namely the asymptotic method, percentile bootstrap, and studentized bootstrap methods. The percentile bootstrap approach is the easiest approach, while the studentized bootstrap approach improves performance of the percentile bootstrap approach, and obtains more accurate results. The two bootstrap-based methods are Monte Carlo based inference approaches. van der Vaart (1998) gave a detailed introduction of the asymptotic theory, while Hall (1992), Efron (1979), Chernick (2007), Efron and Tibshirani (1993) provided information on both the percentile bootstrap and studentized bootstrap methods.

We extend the theory further by constructing confidence intervals for the difference of EPMs for two independent samples. For two-sample confidence intervals, we recommend using two methods, namely the asymptotic procedure and method of variance estimates recovery (MOVER). MOVER is a strategy that "recovers" variance estimates from the limits of individual sample parameters, and then forms approximate confidence intervals for functions of the parameters, as proposed by Zou and Donner (2008). Zou et al (2009) generalized MOVER, and established confidence limits for a linear function of binomial proportions (for further details on MOVER, see Donner and Zou (2012), Dagan et al (2010) and Newcombe (2016)). The MOVER method is an excellent and simple tool to construct confidence intervals for two independent samples.

In addition, we derive the asymptotic distribution of EPM, and the difference of two EPMs when the samples are independent. We conduct simulations to show the proposed theory performs well for one and two independent samples. The simulations also show that the proposed approach is robust in the dependent case. We apply the theory to construct both one-sample and two-sample confidence intervals of EPMs for stock indices in Singapore and USA.

The remainder of the paper is organized as follows. In Section 2, we present methods of constructing confidence intervals for EPM with one-sample, including the asymptotic method and bootstrap-based approaches. The asymptotic normality of EPM is also derived. Thereafter, we develop the theory for the construction of two-sample confidence intervals for the difference of two independent EPMs by applying both the asymptotic method and MOVER procedure. In Section 3, we conduct simulations of both the one-sample and two-sample confidence intervals for the difference in two independent EPMs. We also conduct simulations for two-sample confidence intervals for the difference in two dependent EPMs. We illustrate the theory by applying the proposed methods to real data analysis by comparing the performance of the Singapore Stock Market Index and Standard & Poor's Composite 500 Index in Section 4. Section 5 concludes the paper. Proofs of the asymptotic results are given in the Appendix.

# 2 Theory

Let $\tilde{r}$ be the stochastic return of an investment portfolio, $r^f$ be the deterministic risk-free rate, and $r = \tilde{r} - r^f$ be the excess return. The economic performance measurement (EPM) is defined as (Homm and Pigorsch, 2012a):

$$\theta(r) := \text{EPM}(r) = \frac{E(r)}{\text{AS}(r)} = \frac{E(\tilde{r}) - r^f}{\text{AS}(\tilde{r} - r^f)} \quad , \tag{2.1}$$

where $E(r)$ is the expectation of the excess return and $\text{AS}(r)$, the AS index of riskness (Aumann and Serrano, 2008) of the excess return, is the positive solution, $s > 0$, to the following equation:

$$E\left[\exp\left(-\frac{r}{s}\right)\right] = 1\ . \tag{2.2}$$

The EPM in equation (2.1) can be rewritten as $\theta(r) = \mu_r/s$, where $\mu_r = E(r)$ and the estimate of EPM is $\hat{\theta}$[1], which can be obtained from:

$$\hat{\theta} = \bar{r}/\hat{s}, \tag{2.3}$$

with $\bar{r} = \sum_{i=1}^{n} r_i/n$, in which $r_i$ $(i = 1, \cdots, n)$ is the realization of the excess return, $r$. We note that EPM in equation (2.1) may not exist as Schulze (2014) shows that the AS index of riskness may not exist. In order to ensure the existence of the EPM, we use the following assumption:

**Assumption 2.1** *A gamble/investment with returns satisfies the following conditions: (a) negative outcomes; (b) positive mean; and (c) essentially has no heavy negative tails.*

With the aid of Assumption 2.1, the following lemma is obtained.

**Lemma 2.1** *If the return, $r_i$ $(i = 1, 2, \cdots, n)$, satisfies Assumption 2.1, then the estimate $\hat{\theta}$ of the EPM defined in equation (2.3) exists and is unique.*

See Homm and Pigorsch (2012b) for further information about Lemma 2.1. In this paper, we recommend using a nonparametric approach to estimate $s$ because the distribution of the data is typically unknown. In addition, we recommend applying the method of moments (MM) (Hansen, 1982) to obtain the nonparametric estimator,

[1] We note that most of the statistics in the paper are a function of the sample size $n$. For simplicity, we omit $n$ as a subscript.

$\hat{s}$, from the following equation:

$$\frac{1}{n}\sum_{i=1}^{n} e^{-r_i/\hat{s}} - 1 = 0. \tag{2.4}$$

We introduce the asymptotic approach and both percentile and studentized bootstrap methods to construct a confidence interval of the EPM for one sample, and a confidence interval for the difference in two EPMs for two samples, in the following subsections.

## 2.1 One-sample confidence interval for EPM

In this section, we apply the Delta method to obtain a symmetric two-sided asymptotic confidence interval for the EPM, and thereafter discuss the bootstrap method to obtain the asymmetric two-sided asymptotic confidence interval for the EPM.

### 2.1.1 Asymptotic confidence interval

Consider the asymptotic distribution (Homm and Pigorsch, 2012a) for the estimator, $\hat{s}$, defined in equation (2.4), as follows:

$$\sqrt{n}(\hat{s} - s_0) \xrightarrow{d} N(0, V_{AS}), \tag{2.5}$$

where $s_0$ is the true value of $s$, and $V_{AS} = J/G^2$ is the asymptotic variance of the estimator, $\hat{s}$. Here, $J = E\big[(e^{-r/s_0} - 1)^2\big] = E(e^{-2r/s_0}) - 1$ and $G = E(e^{-r/s_0} r)/s_0^2$. We replace $s_0$ by $\hat{s}$ in the expressions of $J$ and $G$ to obtain the corresponding estimators $\hat{J}$ and $\hat{G}$, and obtain $\hat{V}_{AS} = \hat{J}/\hat{G}^2$.

Before we derive the symmetric two-sided asymptotic confidence interval for EPM, we derive the asymptotic distribution of $\hat{\theta}$, as given in the following theorem.

**Theorem 2.1** *Let $\{r_1, \cdots, r_n\}$ be a sample realization of returns from an investment satisfying Assumption 2.1, with the AS index of riskness defined in (2.2) for returns. Under suitable regularity conditions,*[2] *the estimate of the EPM, $\hat{\theta}$, defined in equation (2.3), satisfies:*

$$\sqrt{n}\big(\hat{\theta} - \theta\big) \xrightarrow{d} N\big(0, V_{EPM}\big), \tag{2.6}$$

*where*

$$V_{EPM} = \frac{V_r}{s_0^2} - 2\frac{\mu_r}{s_0^3}V_{AS,r} + \frac{\mu_r^2}{s_0^4}V_{AS},$$

*in which $V_r$ and $V_{AS}$ are the asymptotic variances of $r_i$ and the AS index of riskness, respectively, and $V_{r,AS} = V_{AS,r} = cov(r, -G^{-1}e^{-r/s_0})$ is the asymptotic covariance.*

The proof of Theorem 2.1 is given in the Appendix.

**Remark 2.1** *Theorem 2.1 holds under regularity conditions in different situations. For example:*

a. *for i.i.d. data, $E(r^4)$ is assumed to be finite, and*

b. *for autocorrelated time series data, under appropriate conditions, we impose a stronger assumption that $E(r^{4+\delta})$ is finite, for a small positive constant $\delta$ (see Andrews (1991) for further information).*

Based on the asymptotic result in Theorem 2.1, the two-sided symmetric $100(1-\alpha)\%$ asymptotic confidence interval for EPM can be constructed as:

$$A_n = \left(\hat{\theta} - z_{1-\alpha/2}\sqrt{\hat{V}_{EPM}/n}, \hat{\theta} + z_{1-\alpha/2}\sqrt{\hat{V}_{EPM}/n}\right), \tag{2.7}$$

where $\hat{V}_r = \sum_{i=1}^n (r_i - \bar{r})^2/n$, $\hat{V}_{AS,r} = \hat{\text{cov}}(r, -\hat{G}^{-1}e^{-r/\hat{s}})$, $\hat{V}_{EPM} = \hat{V}_r/\hat{s}^2 - 2\bar{r}\hat{V}_{AS,r}/\hat{s}^3 + \bar{r}^2\hat{J}/\hat{s}^4\hat{G}^2$, and $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

[2]See, for example, Remark 2.1 for the regularity conditions.

#### 2.1.2 Bootstrap-based confidence interval

The bootstrap method developed by Efron (1987) is a Monte Carlo based inference approach that is popular for finite samples because it is a promising tool to obtain an asymptotic variance or confidence interval of a statistic. In this paper, we propose the following two bootstrap-based methods for constructing the confidence interval for EPM, namely percentile bootstrap technique and studentized bootstrap approach. The percentile bootstrap is a simple way of obtaining a confidence interval because it uses percentiles of the bootstrap distribution, such that the confidence interval becomes $\left(\theta^\star_{\alpha/2}, \theta^\star_{1-\alpha/2}\right)$, where $\theta^\star_{\alpha/2}$ is the $\alpha/2$ percentile of the EPM computed by using the bootstrap samples. The studentized bootstrap approach (Davison and Hinkley, 1997), which is also called the bootstrap-t approach, replaces the quantiles from the normal or student approximation by the quantiles from the bootstrap distribution of the Student t-test. We state the steps to construct the bootstrap-based confidence interval, as follows:

*Step 1.* For any given returns sample $\{r_1, r_2, \cdots, r_n\}$, we estimate $r$ with $\bar{r} = \sum_{i=1}^n r_i/n$, and compute the corresponding AS index, $\hat{s}$, and asymptotic variance $\hat{V}_{EPM}$, defined in equations (2.4) and (2.7), respectively.

*Step 2.* Generate a new random sample $\{r_1^\star, r_2^\star, \cdots, r_n^\star\}$ with replacement from the given observations $\{r_1, r_2, \cdots, r_n\}$, and compute $\hat{\theta}^\star$ and the corresponding asymptotic variance, $\hat{V}^\star_{EPM}$, for the new bootstrap sample by using the approach described in Step 1.

*Step 3.* Repeat the process in Step 2 $M$ times ($M \geq 5000$) to obtain $M$ values of $\hat{\theta}^\star$ and $\hat{V}^\star_{EPM}$, and denote the $M$ values of $\hat{\theta}^\star$ and $\hat{V}^\star_{EPM}$ as $\hat{\theta}^\star_m$ and $\hat{V}^\star_{EPM,m}$ ($m = 1, \cdots, M$), respectively. Thereafter, calculate:

$$Q_m = \frac{\left|\hat{\theta}^\star_m - \hat{\theta}\right|}{\sqrt{\hat{V}^\star_{EPM,m}}} \quad (m = 1, \cdots, M) ,$$

where $|x|$ is the absolute value of $x$.

*Step 4.* **(I). Percentile Bootstrap**:

Sort $M$ values of $\hat{\theta}^\star_m$ ($m = 1, \cdots, M$) from the smallest to the largest, and obtain the corresponding order statistics $\hat{\theta}^\star_{(1)}, \cdots, \hat{\theta}^\star_{(M)}$. Thereafter, obtain the percentile bootstrap-based confidence interval:

$$B_{1,n} = \left(\hat{\theta}^\star_{([\alpha M/2])}, \hat{\theta}^\star_{([M(1-\alpha)/2])}\right) \tag{2.8}$$

at the $\alpha$ significance level, where $[\cdot]$ denotes the top integral function and $\hat{\theta}^\star_{(\cdot)}$ is the order statistic.

**(II.) Studentized Bootstrap**:

Sort $M$ values of $Q_m$ ($m = 1, \cdots, M$), and find the corresponding $(1 - \alpha)$ quantile as $q^\star_{1-\alpha} = Q_{([M(1-\alpha)])}$. Thereafter, the studentized bootstrap based confidence interval at nominal level $\alpha$ can be constructed as:

$$B_{2,n} = \left(\hat{\theta} - q^\star_{1-\alpha}\sqrt{\hat{V}_{EPM}}, \hat{\theta} + q^\star_{1-\alpha}\sqrt{\hat{V}_{EPM}}\right) . \tag{2.9}$$

One could choose $M \geq 5000$. The larger is the value of $M$, the more precise will be the constructed confidence interval. However, we need more time to compute the confidence interval for larger $M$, and suggest $M = 5000$.

## 2.2 Two-sample EPMs

In this section, we introduce two methods, namely the asymptotic approach and the method of variance estimates recovery (MOVER), to construct confidence intervals for the difference in two EPMs from two independent samples.

### 2.2.1 Asymptotic confidence interval for two independent samples

From the previous subsection, we obtain the asymptotic normal distribution of $\hat{\theta}$ for one single sample by applying Theorem 2.1. Using the property of normal distribution, we can derive the asymptotic distribution of $\hat{\Delta} = \hat{\theta}_1 - \hat{\theta}_2$. We will discuss the details in this section.

For $k = 1, 2$, consider two independent samples of returns $r_{k1}, r_{k2}, \cdots, r_{kn_k}$ and use equation (2.3) to estimate $\theta_k$. Thereafter, applying Theorem 2.1 as $n_k \to \infty$, we obtain the following asymptotic normal distribution for $\hat{\theta}_k$:

$$\sqrt{n_k}(\hat{\theta}_k - \theta_k) \xrightarrow{d} N(0, V_{EPM,k}) \ ,$$

where $V_{EPM,k}$ can be estimated by using equation (2.7) for the $k^{th}$ sample ($k = 1, 2$). Thus, we can obtain the following theorem for the asymptotic distribution of $\hat{\Delta} = \hat{\theta}_1 - \hat{\theta}_2$.

**Theorem 2.2** *For $k = 1, 2$, suppose that $r_{k1}, r_{k2}, \cdots, r_{kn_k}$ are the realizations of two independent samples of investment returns satisfying Assumption 2.1. The corresponding EPMs, $\theta_k = \mu_{r,k}/s_k$, are defined in equation (2.1), and the asymptotic distribution of $\hat{\Delta} = \hat{\theta}_1 - \hat{\theta}_2$ is:*

$$\hat{\Delta} \xrightarrow{d} N\Big(\Delta, V_{\Delta}\Big), \tag{2.10}$$

*where* $V_\Delta = V_{EPM,1}/n_1 + V_{EPM,2}/n_2$, *with* $V_{EPM,k}$ *defined in equation (2.6) and the estimate* $\hat{V}_{EPM}$ *defined in equation (2.7).*

Therefore, we can construct the corresponding asymptotic $100(1-\alpha)\%$ confidence interval for $\Delta$ as:

$$D_n = \left(\hat{\Delta} - z_{1-\alpha/2}\sqrt{\hat{V}_\Delta}, \hat{\Delta} + z_{1-\alpha/2}\sqrt{\hat{V}_\Delta}\right), \tag{2.11}$$

where $\hat{V}_\Delta$ is defined in equation (2.10).

### 2.2.2 Confidence interval estimation by MOVER

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are computed based on two independent samples, an approximate two-sided $100(1-\alpha)\%$ confidence interval $(L, U)$ for $\Delta = \theta_1 - \theta_2$ is given by

$$(L, U) = \hat{\theta}_1 - \hat{\theta}_2 \mp z_{1-\alpha/2}\sqrt{\widehat{\text{var}}(\hat{\theta}_1) + \widehat{\text{var}}(\hat{\theta}_2)}, \tag{2.12}$$

where $\widehat{\text{var}}(\hat{\theta}_k)$ is an estimator of the variance of $\hat{\theta}_k$ $(k = 1, 2)$. The traditional procedure performs well when the sample sizes are large and the sampling distributions of $\hat{\theta}_k$ are close to normal. However, the procedure may not perform well when the sample sizes are not large, or the sampling distributions of $\hat{\theta}_k$ $(k = 1, 2)$ are not close to normal.

In order to improve the performance, we recommend using the method of variance estimates recovery, MOVER. This is an excellent technique to calculate confidence intervals for any linear combination, for example, a sum or a difference, of two statistics from two independent samples, especially when the sampling distributions are not asymptotically normal, or are asymmetric.

Suppose that, for $k = 1, 2$, a $100(1-\alpha)\%$ two-sided confidence interval for $\theta_k$ is given by $(l_k, u_k)$, in which case $(l_1 - u_2, u_1 - l_2)$ may be a good choice for the $100(1-\alpha)\%$ two-sided confidence interval for $\theta_1 - \theta_2$. However, this is not so because $(l_1 - u_2, u_1 - l_2)$ will cover 100% of the two-sided confidence interval of $\theta_1 - \theta_2$. In order to improve estimation, we follow the approach used by Zou (2008) to estimate the variances for both limits, not exactly at, but in the neighbourhood of $L$ and $U$, respectively.

As in Zou (2010), of all the plausible parameter values of $\theta_k$ provided by $(l_k, u_k)$, the distance between $l_1 - u_2$ and $L$ is smaller than using $\hat{\theta}_1 - \hat{\theta}_2$ and $L$, as $L$ is in the neighborhood of $l_1 - u_2$. Therefore, we can obtain the variance estimator for $L$ at $\theta_1 = l_1$ and $\theta_2 = u_2$. Similarly, the corresponding variance estimator for $U$ can be obtained by $\theta_1 = u_1$ and $\theta_2 = l_2$.

For $k = 1, 2$, to obtain a single sample confidence interval of $\theta_k$, we have:

$$(l_k, u_k) = \hat{\theta}_k \mp z_{1-\alpha/2}\sqrt{\widehat{\text{var}}(\hat{\theta}_k)}\ .$$

Similarly, to obtain a variance estimate for $\hat{\theta}_k$ at $\theta_k = l_k$, we have:

$$\widehat{\text{var}}(\hat{\theta}_k) = (\hat{\theta}_k - l_k)^2/z^2_{1-\alpha/2}\ , \tag{2.13}$$

and to obtain $\hat{\theta}_k$ at $\theta_k = u_k$, we have:

$$\widehat{\text{var}}(\hat{\theta}_k) = (u_k - \hat{\theta}_k)^2/z^2_{1-\alpha/2}\ . \tag{2.14}$$

The results in equations (2.13) and (2.14) enable construction of the confidence interval, $M_n$, by using MOVER, such that:

$$M_n = (L, U), \tag{2.15}$$

where the lower limit $L$ can be obtained by substituting the variance estimators at $\theta_1 = l_1$ from equation (2.13), and at $\theta_2 = u_2$ from equation (2.14), such that:

$$\begin{aligned} L &= \hat{\theta}_1 - \hat{\theta}_2 - z_{1-\alpha/2}\sqrt{\frac{(\hat{\theta}_1 - l_1)^2}{z^2_{1-\alpha/2}} + \frac{(u_2 - \hat{\theta}_2)^2}{z^2_{1-\alpha/2}}} \\ &= \hat{\theta}_1 - \hat{\theta}_2 - \sqrt{(\hat{\theta}_1 - l_1)^2 + (u_2 - \hat{\theta}_2)^2}. \end{aligned} \tag{2.16}$$

The upper limit $U$ can be obtained by substituting the variance estimators at $\theta_1 = u_1$ and $\theta_2 = l_2$, respectively, such that:

$$U = \hat{\theta}_1 - \hat{\theta}_2 + \sqrt{(u_1 - \hat{\theta}_1)^2 + (\hat{\theta}_2 - l_2)^2}. \tag{2.17}$$

**Remark 2.2** *A two-step approach is used to construct confidence intervals for* $\Delta = \theta_1 - \theta_2$ *by using MOVER. Construct the* $100(1-\alpha)\%$ *two-sided confidence intervals* $(l_k, u_k)$ *of* $\theta_k$ *for the independent single sample* $k$*, with* $k = 1, 2$*, by using the asymptotic approach described in Section 2.1.1, or by using the bootstrap approach described in Section 2.1.2. Thereafter, one could apply equation (2.15) to construct the confidence interval of* $\Delta = \theta_1 - \theta_2$ *by using MOVER.*

**Remark 2.3** *Using the same argument as in the above derivation, we can obtain the asymptotic confidence interval* $D_n$ *in equation (2.11) by using the MOVER method in equation (2.15), but by applying the method described in equation (2.7) to calculate individual confidence intervals for* $A_n$*. Based on the simulation results for a single sample in Section 3.1, the performance of the asymptotic confidence intervals,* $A_n$*, and studentized bootstrap-based confidence interval,* $B_{2,n}$*, are similar. Therefore, in Sections 3.2 and 3.3, we conduct simulations of two independent sample cases and two dependent sample cases, respectively. Here, we apply the percentile bootstrap confidence intervals,* $B_{1,n}$*, to compute confidence intervals* $(l_k, u_k)$*,* $k = 1, 2$*, for* $\theta_k$*,* $k = 1, 2$*, in each single sample.*

# 3 Simulations

In this section, we conduct simulations to compare the performance of the proposed methods in constructing confidence intervals for finite samples, for both the one-sample and two-sample cases. In the simulations, we assume the variables follow a normal inverse Gaussian (NIG) distribution, as NIG is one of the most commonly used and well established distributions in finance and econometrics. For example, Homm and Pigorsch (2012a) assume the NIG distribution for parametric estimation of EPM, while Zakamouline and Koekebakker (2009) use the NIG distribution when they conduct simulations for evaluating the portfolio performance of generalized Sharpe ratios.

A NIG distributed random variable, $R$, is characterized by the following density:

$$f(r;\alpha,\beta,\mu,\delta) = \frac{\alpha\delta}{\pi}\frac{K(\alpha\sqrt{\delta^2+(r-\mu)^2})}{\sqrt{\delta^2+(r-\mu)^2}}e^{\delta\gamma+\beta(r-\mu)} \ ,$$

where $\gamma = \sqrt{\alpha^2-\beta^2}$, $K(x) = (1/2)\int_0^\infty e^{-x(t+t^{-1})/2}dt$ is the modified Bessel function of the third kind with index 1, and $\delta$, $\mu$, and $\beta$ are the scale, location and asymmetry parameters, respectively, in which $\alpha \pm \beta$ determines the heaviness of the tails. Given the existence conditions of the AS index, if $R$ is a NIG-distributed random variable, with parameters $\alpha, \beta, \delta$ and $\mu$, in which $0 \leq |\beta| < \alpha$, $\delta > 0$, $\mu \in R$ and $\mu \in (-\delta\beta/\gamma, \delta(\alpha-\beta)/\gamma]$, then the AS index of $R$ exists.

NIG distribution data can be generated easily. Assume that the random variable $X$ comes from a standard normal distribution, and $Y$ comes from the inverse Gaussian distribution, $Y \sim IG(\eta,\lambda)$, in which $\eta = \delta/\gamma = \delta/\sqrt{\alpha^2-\beta^2}$ is the mean of the inverse

Gaussian distribution and $\lambda = \delta^2$ is the shape parameter. Then $R = \mu + \beta Y + \sqrt{Y} X$ follows the NIG distribution, $R \sim NIG(\alpha, \beta, \mu, \delta)$.

### 3.1 One-sample case

Simulations are based on NIG-distributed simulation data with different parameter configurations to evaluate the finite sample performance of the proposed methods, namely $A_n$, $B_{1,n}$ and $B_{2,n}$. The coverage probabilities (CP) and average widths (AW) of two-sided 90% and 95% confidence intervals are reported in Tables 1 and 2, respectively, where 5000 samples are generated to calculate the confidence intervals.

From Table 1, we have the following observations. First, all the coverage probabilities of the three proposed confidence intervals, $A_n$, $B_{1,n}$, and $B_{2,n}$, are very close to 0.90 at the 10% significance level, and all the coverage probabilities (CP) of the three methods give similar performance in the simulations. Basically, of $A_n$, $B_{1,n}$, and $B_{2,n}$, none is superior based on the coverage probability. However, since the coverage probability of $A_n$ is further from 90% when $n$ is 70, and is closest to 90% when $n$ is 150, based on the coverage probability, $A_n$ is the worst when $n$ is small and best when $n$ is large.

Second, as expected, (a) the coverage probabilities of the three proposed confidence intervals are closer to 0.90, and (b) the average widths (AW) of the three confidence intervals decrease gradually, as $n$ increases.

Third, comparing the three methods, for average widths with the same parameter configurations, both the asymptotic confidence interval, $A_n$, and studentized bootstrap-based confidence interval, $B_{2,n}$, perform better than the percentile bootstrap-based confidence interval $B_{1,n}$, because the average AWs of both $A_n$ and $B_{2,n}$ are smaller than for $B_{1,n}$. For example, the average AW are 0.708, 0.618, 0.522 and 0.366 for $n = 70$, 80 100 and 150, respectively, for $B_{1,n}$, which are much wider than for both $A_n$ and $B_{2,n}$.

The simulated coverage probabilities (CP) and average widths (AW) of the two-sided 95% confidence intervals with the proposed three methods are given in Table 2, which suggests similar qualitative conclusions can be drawn as from Table 1.

In summary, all three proposed methods are acceptable, but the asymptotic confidence interval, $A_n$, and studentized bootstrap-based confidence interval, $B_{2,n}$, are more highly recommended.

### 3.2 Two independent samples

In this section, we conduct simulations on the finite sample performance of the proposed methods, $D_n$ and $M_n$, in equations (2.11) and (2.15), respectively, for two independent samples. Thereafter, we conduct simulations for dependent samples to check the robustness of the proposed theory.

We first discuss the simulations for the case of two independent samples. In order to compute confidence intervals with MOVER, according to Remark 2.3, we apply the percentile bootstrap approach, $B_{1,n}$, to obtain confidence intervals $(l_k, u_k),\ k = 1, 2$, for $\theta_k,\ k = 1, 2$ for a single sample. In the simulations, the first sample is from the NIG distribution, and the second sample is from the normal distribution $N(\mu, \sigma^2)$.

The coverage probabilities (CP) and average widths (AW) of two-sided 90% and 95% confidence intervals for $\Delta = \theta_1 - \theta_2$, with various parameter combinations, are given in Tables 3 and 4, respectively. The simulations are based on an average of 5000 replications.

From the tables, we obtain the following observations. First, for all parameter combinations, the coverage probabilities for both $D_n$ and $M_n$ are very close to 0.90 at the pre-specified nominal level, $\alpha = 0.10$. For example, the average coverage probabilities are 0.903, 0.903, 0.908, and 0.899 for $D_n$, and 0.911, 0.899, 0.891, 0.902 for $M_n$, when $(n_1, n_2)$ = (120,100), (150,150), (150,180), and (200,200), respectively. Second, both the coverage probabilities for $D_n$ and $M_n$ are close to the nominal levels. Third, as expected, when the sample sizes increase, the average widths for both confidence intervals fall quickly for any specified parameter configurations. Fourth, interestingly, the average widths of the MOVER method, $M_n$, are shorter than those of the asymptotic method, $D_n$, and yet the coverage probabilities of the MOVER method are higher than those of the asymptotic method, $D_n$, when $(n_1, n_2)$ = (120,100), (150,150), (150,180), and (200,200), respectively. Thus, $M_n$ is preferred to $D_n$ when $(n_1, n_2)$ = (120,100) and (200,200). Fifth, when $(n_1, n_2)$ = (150,150) and (150,180), we cannot conclude which of $D_n$ and $M_n$ is better as $D_n$ is closer to the overestimated coverage probabilities for nominal level $\alpha = 0.10$, while $M_n$ has shorter average widths, but underestimates the coverage probabilities.

The corresponding coverage probabilities and average widths for the two-sided 95% confidence intervals with the same parameter configurations are reported in

Table 4. The conclusions drawn for the pre-specified nominal level, $\alpha = 0.05$, are similar to those drawn for the pre-specified nominal level, $\alpha = 0.10$. For example, the average widths of the MOVER method, $M_n$, are remarkably shorter than those of the asymptotic method, $D_n$. The average coverage probabilities of the MOVER method are closer to the pre-specified nominal level, $\alpha = 0.05$, than those of the asymptotic method in the two-sample cases. In particular, when $(n_1, n_2)$ = (120,100), $M_n$ and $D_n$ have the same the average coverage probability, that is, 0.953, while the average width of $M_n$ is 0.483, and is much shorter than that of $D_n$, at 0.635.

When $(n_1, n_2)$ = (150,150), the average coverage probability of $M_n$ is 0.951, and is much closer to the pre-specified nominal level of $\alpha = 0.05$ than for $D_n$, at 0.958, and the average width of $M_n$ is 0.400, which is much shorter than that of $D_n$, at 0.524. When $(n_1, n_2)$ = (200, 200), the average coverage probability of $M_n$ is 0.952, which is closer to the pre-specified nominal level, $\alpha = 0.05$, than for $D_n$, at 0.953, and the average width of $M_n$ is 0.342, which is much shorter than for $D_n$, at 0.441.

In the last case, when $(n_1, n_2)$ = (150, 180), the average coverage probability of $D_n$ is 0.952, which is closer to the pre-specified nominal level than for $M_n$ at 0.941, while the average width of $M_n$ is 0.371, which is much shorter than that for $D_n$, at 0.524. For the first 3 cases, we conclude that $M_n$ performs better than $D_n$, while in the last case, we cannot conclude which of $M_n$ and $D_n$ is better. In general, $M_n$ performs better than does $D_n$.

### 3.3 Two dependent samples

We now conduct simulations to examine the robustness of the proposed approaches for simulated data of two dependent samples. The same method to compute $(l_k, u_k)$, $k = 1, 2$, is applied. The two samples are both drawn from normal distributions, where the parameters settings and correlation coefficient, $\rho$, are specified in the tables. The simulations for coverage probabilities (CP) and average widths (AW) for two-sided 90% and 95% confidence intervals of $\Delta$, for two dependent samples, are given in Tables 5 and 6, respectively.

Tables 5 and 6 lead to the following observations: 1) even though the two samples are dependent, the simulations for the asymptotic method, $D_n$, in equation (2.11) and the MOVER technique, $M_n$, in equation (2.15), are both acceptable. 2) The coverage probabilities (CP) for the two methods are very close to the pre-specified significance levels, and the average widths (AW) are shorter for increasing sample sizes. 3) AW is shorter for $D_n$ (with AW = 0.603, 0.571, 0.885, and 0.890) than for $M_n$. For example, the average AW = 0.603, 0.571, 0.885, and 0.890 for $D_n$ and = 0.657, 0.618, 0.523, and 0.496 for $M_n$ when $(n_1, n_2)$ = (100,100), (120,100), (150,150), and (150,180), respectively. However, 4) CP for $D_n$ is further from the true values, and nearly all underestimate the true CP (so that AW is shorter) with average CP = 0.891, 0.881, 0.885, and 0.890 when $(n_1, n_2)$ = (100,100), (120,100), (150,150), and (150,180), respectively. 5) CP for $M_n$ is closer to the true values. 6) In general, they overestimate the true CP with average CP=0.909, 0.902, 0.900, and 0.897, when $(n_1, n_2)$ = (100,100), (120,100), (150,150), and (150,180), respectively, for the two-sided 90% confidence intervals of $\Delta$ with $\rho$.

In order to demonstrate the effects of different values of $\rho$, we also conduct simulations for the two-sided 90% confidence intervals of $\Delta$, with $\rho$ varying from 0.0 to 0.9 for both $D_n$ and $M_n$, which are shown in Tables 7 and 8 for $D_n$ and $M_n$, respectively. The simulations show that estimation of the two-sided 90% confidence intervals for both $D_n$ and $M_n$ are robust to any values of $\rho$ from 0.1 to 0.9. Therefore, the theory works well for both independent and dependent samples.

The simulations for the two-sided 95% confidence intervals of $\Delta$, with $\rho$ varying from 0.0 to 0.9 are similar. We also conducted simulations to check the robustness of the non-normal distribution. The simulations show that the proposed theory is robust to non-normal distributions.

## 4 Empirical application

In order to illustrate the theory, in this section we construct confidence intervals for the EPMs for Singapore and USA stock markets, and their differences, by using weekly returns data of the Singapore Stock Market Index (STI) and Standard & Poor's Composite 500 Index (S&P500), from January 1, 2000 to December 31, 2015. The STI is a capitalisation-weighted stock market index that is regarded as the benchmark index of the Singapore stock market to track the performance of the top 30 companies listed on the Singapore Exchange. The S&P500 Index is a stock market index based on the market capitalizations of 500 large companies having common stock listed on the NYSE or NASDAQ. It is one of the most commonly followed equity indices, and many consider it one of the best representations of the USA stock market and a bellwether for the USA economy. The time series plots from January 1, 2000 to December 31, 2015 for STI and S&P500 are given in Figure 1.

We first apply the three proposed methods, namely the asymptotic method, $A_n$, percentile bootstrap-based approach, $B_{1,n}$ and studentized bootstrap-based procedure $B_{2,n}$, to construct one-sample confidence intervals for EPMs at a confidence level 95% for both STI and S&P500. The results are given in Table 9, which show that the average widths of the percentile bootstrap-based confidence intervals are the longest for all sub-periods, while the other two methods have similar performance, which is consistent with the simulation results. Table 9 also shows that the EPM of STI and S&P500 for each sub-period have both positive and negative values, implying that we do not reject the EPM as zero which, in turn, implies that the average returns of both STI and S&P500 could be zero for each sub-period.

Before we construct the two-sample confidence intervals for the differences in EPMs for the returns of STI ($r_{STI}$) and S&P500 ($r_{SP}$) by using the proposed methods, we first test whether $r_{STI}$ and $r_{SP}$ are independent. We use the Kendall $\tau$ test to examine whether $r_{STI}$ and $r_{SP}$ are correlated, with the test results given in Table 10. It can be concluded that the correlation between $r_{STI}$ and $r_{SP}$ is rejected as zero, so the samples are dependent. The simulations show that the theory developed in the paper is robust to dependent samples. Thus, we can apply the proposed methods to dependent samples in the empirical illustration.

We apply both the asymptotic method, $D_n$, in equation (2.11), and the MOVER procedure, $M_n$, in equation (2.15), to construct confidence intervals for the difference, $\Delta = \theta_{STI} - \theta_{SP}$, in EPMs between $r_{STI}$ and $r_{SP}$, with the results given in Table 11. We can see that zero is included in both the confidence intervals of $D_n$ and $M_n$ for each sub-period. Thus, we do not reject the null hypothesis that the two EPMs, $\theta_{STI}$ and $\theta_{SP}$, for Singapore and USA stock markets, respectively, are the same.

# 5 Conclusion

In this paper, the confidence intervals for EPM using the AS index with one-sample, and the difference of EPMs with two samples, were constructed. For the single sample case, three approaches were considered, namely the asymptotic method, $A_n$, percentile bootstrap method, $B_{1,n}$ and studentized bootstrap method, $B_{2,n}$. The simulations indicated that all three methods were acceptable, but $A_n$ and $B_{2,n}$ were more highly recommended, with both presenting higher coverage probabilities and shorter average widths than $B_{1,n}$.

For the two-sample case, in the case of both independent and dependent samples, the asymptotic procedure, $D_n$, and method of variance estimates recovery (MOVER), $M_n$, were used. Simulations for the two-sample situations were conducted, where the results showed that, for two independent samples, $M_n$ performed better than $D_n$, which had similar coverage probabilities, but the average widths for $M_n$ were shorter. For two dependent samples, both methods were reasonable, which indicated that the proposed methods were robust.

The returns data of the Singapore Stock Market Index (STI) and USA Stock

Market Standard & Poor's Composite 500 Index (S&P500) from January 1, 2000 to December 31, 2015 confirmed the veracity of the proposed methods. The empirical results showed that the two indices were not statistically different.

# Appendix

### Proof of Theorem 2.1

Before proving Theorem 2.1, we derive the asymptotic joint distribution of $(\bar{r}, \hat{s})$, as shown in the following lemma.

**Lemma 5.1** *Assume $\{r_1, \cdots, r_n\}$ is a sample realization of returns from a portfolio investment satisfying Assumption 2.1, and the AS index of riskness can be calculated based on (2.2). Then we have:*

$$\sqrt{n}\begin{pmatrix} \bar{r}-\mu_r \\ \hat{s}-s_0 \end{pmatrix} \xrightarrow{d} N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_r & V_{r,AS} \\ V_{AS,r} & V_{AS} \end{pmatrix}\right],$$

*where $V_r$ and $V_{AS}$ are the asymptotic variances of $r_i$ and the AS index of riskness, respectively, and $V_{r,AS} = V_{AS,r} = cov(r, -G^{-1}e^{-r/s_0})$ is the asymptotic covariance.*

### Proof of Lemma 5.1

Let $f(r,s) = e^{-r/s} - 1$. Based on equation (2.4), and taking a Taylor expansion, it can be shown that:

$$\begin{aligned} 0 &= \frac{1}{n}\sum_{i=1}^{n}(e^{-\frac{r_i}{\hat{s}}} - 1) \\ &= \frac{1}{n}\sum_{i=1}^{n}(e^{-\frac{r_i}{s_0}} - 1) + \frac{1}{n}\sum_{i=1}^{n}\frac{\partial f(r_i, s)}{\partial s}\Big|_{s=\tilde{s}}(\hat{s} - s_0), \end{aligned} \tag{5.1}$$

where $\tilde{s}$ is between $s$ and $s_0$. From equation (5.1), we have:

$$\begin{aligned} \sqrt{n}(\hat{s} - s_0) &= \Big\{-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial f(r_i, s_0)}{\partial s}\Big\}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(e^{-\frac{r_i}{s_0}} - 1) + o_p(1), \\ &= G^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(1 - e^{-\frac{r_i}{s_0}}) + o_p(1). \end{aligned}$$

According to the Central Limit Theorem (CLT), when $n \to \infty$:

$$\sqrt{n}\begin{pmatrix} \bar{r}-\mu_r \\ \hat{s}-s_0 \end{pmatrix} = \frac{1}{\sqrt{n}}\begin{pmatrix} \sum_{i=1}^{n}(r_i-\mu_r) \\ \sum_{i=1}^{n}\{G^{-1}(1-e^{-\frac{r_i}{s_0}})-s_0\} \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_r & V_{r,AS} \\ V_{AS,r} & V_{AS} \end{pmatrix}\right).$$

Thus, Lemma 5.1 holds. □

Now we will prove Theorem 2.1. Note that $\theta(r) := \text{EPM}(r) = \mu_r/s$, so that combining the asymptotic joint distribution in Lemma 5.1 and the Delta method, the proof of Theorem 2.1 is obtained. □

## References

[1] Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59**, 817-858.

[2] Aumann, R. and Serrano, R. (2008). An economic index of riskiness. *Journal of Political Economy*, **116**, 810-836.

[3] Bartlett, M.S., (1953). Approximate confidence intervals. *Biometrika*, **40**, 12-19.

[4] Brookmeyer, R. and Crowley, J. (1982). A confidence interval for the median survival time. *Biometrics*, **38**, 29-41.

[5] Chernick, M.R. (2007). *Bootstrap Methods: A Guide for Practitioners and Researchers*. Wiley, New Jersey.

[6] Dagan, R., Poolman, J. and Siegrist, C.A. (2010). Glycoconjugate vaccines and immune interference: A review. *Vaccine*, **28**, 5513-5523.

[7] Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge.

[8] Donner, A. and Zou, G.Y. (2012). Closed-form confidence intervals for functions of the normal mean and standard deviation. *Statistical Methods in Medical Research*, **21**, 347-359.

[9] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, **7**, 1-26.

[10] Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association*, **82**, 171-185.

[11] Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, London.

[12] Ghosh, B.K. (1979). A comparison of some approximate confidence intervals for the binomial parameter. *Journal of the American Statistical Association*, **74**, 894-900.

[13] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer, New York.

[14] Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, **50**, 1029-1054.

[15] Homm, U. and Pigorsch, C. (2012a). Beyond the Sharpe ratio: An application of the Aumann-Serrano index to performance measurement. *Journal of Banking & Finance*, **36**, 2274-2284.

[16] Homm, U. and Pigorsch, C. (2012b). An operational interpretation and existence of the Aumann-Serrano index of riskiness. *Economics Letters*, **114**, 265-267.

[17] Jobson, J.D. and Korkie, B.M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, **36**, 889-908.

[18] Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, **15**, 850-859.

[19] Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, **1**, 21-23.

[20] Newcombe, R.G. (2016). MOVER-R confidence intervals for ratios and products of two independently estimated quantities. *Statistical Methods in Medical Research*, *25*, 1774-1778.

[21] Schreiber, A. (2014). Economic indices of absolute and relative riskiness. *Economic Theory*, **56**, 309-331.

[22] Schulze, K. (2014). Existence and computation of the Aumann-Serrano index of riskness and its extension. *Journal of Mathematical Economics*, **50**, 219-224.

[23] van der Vaart A.W. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.

[24] Zakamouline, V. and Koekebakker, S. (2009). Portfolio performance evaluation with generalized Sharpe ratio: Beyond the mean and variance. *Journal of Banking & Finance*, **33**, 1242-1254.

[25] Zou, G.Y. (2008). On the estimation of additive interaction by use of the four-by-two table and beyond. *American Journal of Epidemiology*, **168**, 212-224.

[26] Zou, G.Y. (2010). Confidence interval estimation under inverse sampling. *Computational Statistics and Data Analysis*, **54**, 55-64.

[27] Zou, G.Y. and Donner, A. (2008). Construction of confidence limits about effect measures: A general approach. *Statistics in Medicine*, **27**, 1693-1702.

[28] Zou, G.Y., Huang, W.Y. and Zhang, X. H. (2009). A note on confidence interval estimation for a linear function of binomial proportions. *Computational Statistics and Data Analysis*. **53**, 1080-1085.

Table 1: Coverage probabilities (CP) and average widths (AW) of two-sided 90% confidence intervals for one-sample EPM with NIG distribution

| $(\alpha,\beta,\mu,\delta)$ | $n$ | $A_n$ | | $B_{1,n}$ | | $B_{2,n}$ | |
|---|---|---|---|---|---|---|---|
| | | CP | AW | CP | AW | CP | AW |
| (2.5,-2,1.8,1) | 70 | 0.861 | 0.612 | 0.860 | 0.885 | 0.901 | 0.592 |
| | 80 | 0.872 | 0.530 | 0.875 | 0.734 | 0.881 | 0.511 |
| | 100 | 0.911 | 0.461 | 0.878 | 0.622 | 0.899 | 0.452 |
| | 150 | 0.914 | 0.321 | 0.898 | 0.405 | 0.915 | 0.313 |
| (3,-0.8,0.5,1) | 70 | 0.885 | 0.604 | 0.883 | 0.691 | 0.877 | 0.576 |
| | 80 | 0.886 | 0.545 | 0.903 | 0.634 | 0.882 | 0.540 |
| | 100 | 0.890 | 0.477 | 0.893 | 0.559 | 0.883 | 0.493 |
| | 150 | 0.903 | 0.392 | 0.894 | 0.419 | 0.905 | 0.386 |
| (2,-0.5,0.3,0.5) | 70 | 0.870 | 0.561 | 0.875 | 0.693 | 0.882 | 0.555 |
| | 80 | 0.881 | 0.503 | 0.877 | 0.606 | 0.879 | 0.508 |
| | 100 | 0.878 | 0.439 | 0.889 | 0.531 | 0.876 | 0.438 |
| | 150 | 0.896 | 0.340 | 0.882 | 0.389 | 0.898 | 0.349 |
| (1.5,-0.4,0.4,0.8) | 70 | 0.884 | 0.381 | 0.914 | 0.489 | 0.877 | 0.386 |
| | 80 | 0.869 | 0.371 | 0.894 | 0.435 | 0.871 | 0.352 |
| | 100 | 0.874 | 0.313 | 0.886 | 0.355 | 0.870 | 0.296 |
| | 150 | 0.888 | 0.248 | 0.891 | 0.264 | 0.855 | 0.236 |
| (3,-0.8,1,2) | 70 | 0.882 | 0.652 | 0.888 | 0.783 | 0.888 | 0.661 |
| | 80 | 0.913 | 0.592 | 0.907 | 0.680 | 0.917 | 0.598 |
| | 100 | 0.898 | 0.408 | 0.872 | 0.543 | 0.877 | 0.415 |
| | 150 | 0.903 | 0.330 | 0.913 | 0.355 | 0.905 | 0.321 |
| average | 70 | 0.876 | 0.562 | 0.884 | 0.708 | 0.885 | 0.554 |
| | 80 | 0.884 | 0.508 | 0.891 | 0.618 | 0.886 | 0.502 |
| | 100 | 0.890 | 0.420 | 0.884 | 0.522 | 0.881 | 0.419 |
| | 150 | 0.901 | 0.3262 | 0.896 | 0.366 | 0.896 | 0.321 |

Table 2: Coverage probabilities (CP) and average widths (AW) of two-sided 95% confidence intervals for one-sample EPM with NIG distribution

| $(\alpha, \beta, \mu, \delta)$ | $n$ | $A_n$ | | $B_{1,n}$ | | $B_{2,n}$ | |
|---|---|---|---|---|---|---|---|
| | | CP | AW | CP | AW | CP | AW |
| (2.5,-2,1.8,1) | 70 | 0.929 | 0.786 | 0.925 | 0.979 | 0.929 | 0.769 |
| | 80 | 0.932 | 0.676 | 0.927 | 0.861 | 0.939 | 0.670 |
| | 100 | 0.948 | 0.510 | 0.932 | 0.712 | 0.941 | 0.506 |
| | 150 | 0.945 | 0.387 | 0.943 | 0.407 | 0.949 | 0.382 |
| (3,-0.8,0.5,1) | 70 | 0.923 | 0.689 | 0.931 | 0.863 | 0.926 | 0.687 |
| | 80 | 0.938 | 0.634 | 0.933 | 0.794 | 0.937 | 0.635 |
| | 100 | 0.939 | 0.565 | 0.948 | 0.669 | 0.936 | 0.561 |
| | 150 | 0.942 | 0.438 | 0.945 | 0.464 | 0.945 | 0.443 |
| (2,-0.5,0.3,0.5) | 70 | 0.927 | 0.653 | 0.929 | 0.879 | 0.930 | 0.649 |
| | 80 | 0.938 | 0.595 | 0.939 | 0.691 | 0.936 | 0.581 |
| | 100 | 0.944 | 0.512 | 0.939 | 0.626 | 0.939 | 0.508 |
| | 150 | 0.944 | 0.406 | 0.942 | 0.468 | 0.947 | 0.407 |
| (1.5,-0.4,0.4,0.8) | 70 | 0.921 | 0.472 | 0.939 | 0.613 | 0.922 | 0.453 |
| | 80 | 0.930 | 0.421 | 0.937 | 0.539 | 0.937 | 0.411 |
| | 100 | 0.938 | 0.366 | 0.937 | 0.437 | 0.939 | 0.346 |
| | 150 | 0.938 | 0.293 | 0.949 | 0.326 | 0.942 | 0.279 |
| (3,-0.8,1,2) | 70 | 0.935 | 0.795 | 0.940 | 0.873 | 0.938 | 0.786 |
| | 80 | 0.934 | 0.681 | 0.941 | 0.755 | 0.934 | 0.697 |
| | 100 | 0.942 | 0.548 | 0.953 | 0.651 | 0.937 | 0.561 |
| | 150 | 0.949 | 0.392 | 0.943 | 0.440 | 0.942 | 0.406 |
| average | 70 | 0.927 | 0.679 | 0.932 | 0.841 | 0.929 | 0.669 |
| | 80 | 0.934 | 0.601 | 0.935 | 0.728 | 0.937 | 0.599 |
| | 100 | 0.942 | 0.500 | 0.942 | 0.619 | 0.938 | 0.496 |
| | 150 | 0.944 | 0.383 | 0.944 | 0.421 | 0.945 | 0.383 |

Table 3: Coverage probabilities (CP) and average widths (AW) of two-sided 90% confidence intervals for $\Delta$ with two independent samples

| $(\alpha, \beta, \nu, \delta)$ | $(\mu, \sigma^2)$ | $(n_1, n_2)$ | $D_n$ | | $M_n$ | |
|---|---|---|---|---|---|---|
| | | | CP | AW | CP | AW |
| (2.5,-2,1.8,1) | (0.3,1) | (120,100) | 0.913 | 0.574 | 0.918 | 0.466 |
| | | (150,150) | 0.905 | 0.488 | 0.901 | 0.384 |
| | | (150,180) | 0.896 | 0.458 | 0.896 | 0.354 |
| | | (200,200) | 0.888 | 0.415 | 0.902 | 0.327 |
| (2.5,-2,1.8,1) | (1,3.2) | (120,100) | 0.886 | 0.593 | 0.905 | 0.483 |
| | | (150,150) | 0.905 | 0.498 | 0.910 | 0.394 |
| | | (150,180) | 0.910 | 0.472 | 0.904 | 0.368 |
| | | (200,200) | 0.904 | 0.412 | 0.905 | 0.343 |
| (2,-1.5,1.2,0.8) | (0.17,0.8) | (120,100) | 0.913 | 0.459 | 0.908 | 0.331 |
| | | (150,150) | 0.883 | 0.368 | 0.910 | 0.276 |
| | | (150,180) | 0.898 | 0.350 | 0.904 | 0.256 |
| | | (200,200) | 0.906 | 0.307 | 0.908 | 0.236 |
| (1.5,-1,1.5,1.2) | (0.3,1.2) | (120,100) | 0.913 | 0.516 | 0.914 | 0.398 |
| | | (150,150) | 0.909 | 0.427 | 0.889 | 0.330 |
| | | (150,180) | 0.905 | 0.408 | 0.866 | 0.301 |
| | | (200,200) | 0.896 | 0.361 | 0.896 | 0.278 |
| (1.5,-1,1.5,1.2) | (0.5,2) | (120,100) | 0.890 | 0.527 | 0.911 | 0.399 |
| | | (150,150) | 0.915 | 0.430 | 0.887 | 0.333 |
| | | (150,180) | 0.929 | 0.406 | 0.886 | 0.305 |
| | | (200,200) | 0.902 | 0.359 | 0.899 | 0.283 |
| average | | (120,100) | 0.903 | 0.534 | 0.911 | 0.415 |
| | | (150,150) | 0.903 | 0.442 | 0.899 | 0.343 |
| | | (150,180) | 0.908 | 0.419 | 0.891 | 0.317 |
| | | (200,200) | 0.899 | 0.371 | 0.902 | 0.293 |

Table 4: Coverage probabilities (CP) and average widths (AW) of two-sided 95% confidence intervals for $\Delta$ with two independent samples

| $(\alpha,\beta,\nu,\delta)$ | $(\mu,\sigma^2)$ | $(n_1,n_2)$ | $D_n$ | | $M_n$ | |
|---|---|---|---|---|---|---|
| | | | CP | AW | CP | AW |
| (2.5,-2,1.8,1) | (0.3,1) | (120,100) | 0.943 | 0.688 | 0.945 | 0.514 |
| | | (150,150) | 0.949 | 0.577 | 0.948 | 0.449 |
| | | (150,180) | 0.956 | 0.553 | 0.940 | 0.415 |
| | | (200,200) | 0.944 | 0.483 | 0.953 | 0.384 |
| (2.5,-2,1.8,1) | (1,3.2) | (120,100) | 0.956 | 0.721 | 0.957 | 0.571 |
| | | (150,150) | 0.955 | 0.591 | 0.946 | 0.467 |
| | | (150,180) | 0.957 | 0.552 | 0.937 | 0.431 |
| | | (200,200) | 0.956 | 0.495 | 0.957 | 0.400 |
| (2,-1.5,1.2,0.8) | (0.17,0.8) | (120,100) | 0.951 | 0.533 | 0.952 | 0.395 |
| | | (150,150) | 0.960 | 0.439 | 0.952 | 0.324 |
| | | (150,180) | 0.955 | 0.423 | 0.941 | 0.299 |
| | | (200,200) | 0.957 | 0.362 | 0.954 | 0.275 |
| (1.5,-1,1.5,1.2) | (0.3,1.2) | (120,100) | 0.953 | 0.615 | 0.957 | 0.469 |
| | | (150,150) | 0.967 | 0.499 | 0.953 | 0.380 |
| | | (150,180) | 0.950 | 0.489 | 0.951 | 0.353 |
| | | (200,200) | 0.958 | 0.429 | 0.947 | 0.325 |
| (1.5,-1,1.5,1.2) | (0.5,2) | (120,100) | 0.962 | 0.616 | 0.953 | 0.468 |
| | | (150,150) | 0.961 | 0.512 | 0.954 | 0.382 |
| | | (150,180) | 0.952 | 0.492 | 0.935 | 0.355 |
| | | (200,200) | 0.948 | 0.436 | 0.948 | 0.327 |
| average | | (120,100) | 0.953 | 0.635 | 0.953 | 0.483 |
| | | (150,150) | 0.958 | 0.524 | 0.951 | 0.400 |
| | | (150,180) | 0.954 | 0.502 | 0.941 | 0.371 |
| | | (200,200) | 0.953 | 0.441 | 0.952 | 0.342 |

Table 5: Coverage probabilities (CP) and average widths (AW) of two-sided 90% confidence intervals for $\Delta$ with two dependent samples

| $(\mu_1, \sigma_1^2)$ | $(\mu_2, \sigma_2^2)$ | $\rho$ | $(n_1, n_2)$ | $D_n$ | | $M_n$ | |
|---|---|---|---|---|---|---|---|
| | | | | CP | AW | CP | AW |
| (0.2,1) | (0.4,0.9) | 0.3 | (100,100) | 0.876 | 0.604 | 0.921 | 0.703 |
| | | | (120,100) | 0.872 | 0.565 | 0.905 | 0.643 |
| | | | (150,150) | 0.875 | 0.481 | 0.912 | 0.555 |
| | | | (150,180) | 0.883 | 0.467 | 0.900 | 0.512 |
| (0.2,1) | (0.4,0.9) | 0.4 | (100,100) | 0.879 | 0.600 | 0.910 | 0.637 |
| | | | (120,100) | 0.875 | 0.558 | 0.905 | 0.604 |
| | | | (150,150) | 0.885 | 0.478 | 0.913 | 0.523 |
| | | | (150,180) | 0.887 | 0.466 | 0.904 | 0.502 |
| (0.5,4) | (0.7,4) | 0.3 | (100,100) | 0.900 | 0.583 | 0.904 | 0.643 |
| | | | (120,100) | 0.879 | 0.553 | 0.890 | 0.606 |
| | | | (150,150) | 0.886 | 0.466 | 0.899 | 0.509 |
| | | | (150,180) | 0.895 | 0.447 | 0.896 | 0.494 |
| (0.3,1) | (0.4,1.6) | 0.24 | (100,100) | 0.894 | 0.609 | 0.911 | 0.655 |
| | | | (120,100) | 0.886 | 0.595 | 0.904 | 0.619 |
| | | | (150,150) | 0.882 | 0.497 | 0.886 | 0.513 |
| | | | (150,180) | 0.891 | 0.468 | 0.896 | 0.486 |
| (0.3,1) | (0.4,1.6) | 0.32 | (100,100) | 0.905 | 0.619 | 0.901 | 0.648 |
| | | | (120,100) | 0.895 | 0.584 | 0.905 | 0.616 |
| | | | (150,150) | 0.897 | 0.495 | 0.892 | 0.516 |
| | | | (150,180) | 0.893 | 0.465 | 0.888 | 0.488 |
| average | | | (100,100) | 0.891 | 0.603 | 0.909 | 0.657 |
| | | | (120,100) | 0.881 | 0.571 | 0.902 | 0.618 |
| | | | (150,150) | 0.885 | 0.483 | 0.900 | 0.523 |
| | | | (150,180) | 0.890 | 0.463 | 0.897 | 0.496 |

Table 6: Coverage probabilities (CP) and average widths (AW) of two-sided 95% confidence intervals for $\Delta$ with two dependent samples

| $(\mu_1, \sigma_1^2)$ | $(\mu_2, \sigma_2^2)$ | $\rho$ | $(n_1, n_2)$ | $D_n$ | | $M_n$ | |
|---|---|---|---|---|---|---|---|
| | | | | CP | AW | CP | AW |
| (0.2,1) | (0.4,0.9) | 0.3 | (100,100) | 0.922 | 0.715 | 0.961 | 0.860 |
| | | | (120,100) | 0.929 | 0.666 | 0.958 | 0.788 |
| | | | (150,150) | 0.922 | 0.569 | 0.964 | 0.671 |
| | | | (150,180) | 0.929 | 0.564 | 0.954 | 0.655 |
| (0.2,1) | (0.4,0.9) | 0.4 | (100,100) | 0.926 | 0.715 | 0.962 | 0.842 |
| | | | (120,100) | 0.925 | 0.668 | 0.956 | 0.780 |
| | | | (150,150) | 0.933 | 0.578 | 0.952 | 0.667 |
| | | | (150,180) | 0.922 | 0.562 | 0.953 | 0.651 |
| (0.5,4) | (0.7,4) | 0.3 | (100,100) | 0.956 | 0.694 | 0.957 | 0.777 |
| | | | (120,100) | 0.936 | 0.664 | 0.952 | 0.732 |
| | | | (150,150) | 0.930 | 0.554 | 0.947 | 0.615 |
| | | | (150,180) | 0.923 | 0.541 | 0.945 | 0.592 |
| (0.3,1) | (0.4,1.6) | 0.24 | (100,100) | 0.952 | 0.739 | 0.961 | 0.796 |
| | | | (120,100) | 0.935 | 0.699 | 0.948 | 0.748 |
| | | | (150,150) | 0.929 | 0.591 | 0.942 | 0.622 |
| | | | (150,180) | 0.949 | 0.565 | 0.945 | 0.595 |
| (0.3,1) | (0.4,1.6) | 0.32 | (100,100) | 0.951 | 0.736 | 0.953 | 0.801 |
| | | | (120,100) | 0.939 | 0.695 | 0.951 | 0.758 |
| | | | (150,150) | 0.938 | 0.591 | 0.947 | 0.620 |
| | | | (150,180) | 0.942 | 0.556 | 0.946 | 0.597 |
| average | | | (100,100) | 0.941 | 0.720 | 0.959 | 0.815 |
| | | | (120,100) | 0.933 | 0.678 | 0.953 | 0.761 |
| | | | (150,150) | 0.930 | 0.577 | 0.950 | 0.639 |
| | | | (150,180) | 0.933 | 0.558 | 0.947 | 0.618 |

Table 7: Coverage probabilities (CP) and average widths (AW) of two-sided 90% confidence intervals for $\Delta$ with two dependent samples for different $\rho$

| $(n_1, n_2)$ | $\rho$ | $D_n$ | | $\rho$ | $D_n$ | |
|---|---|---|---|---|---|---|
| | | CP | AW | | CP | AW |
| (100,100) | 0 | 0.923 | 0.594 | 0.1 | 0.868 | 0.605 |
| (120,100) | | 0.921 | 0.557 | | 0.870 | 0.559 |
| (150,150) | | 0.928 | 0.475 | | 0.858 | 0.486 |
| (150,180) | | 0.914 | 0.468 | | 0.846 | 0.469 |
| (100,100) | 0.2 | 0.867 | 0.607 | 0.3 | 0.855 | 0.598 |
| (120,100) | | 0.854 | 0.564 | | 0.864 | 0.564 |
| (150,150) | | 0.861 | 0.481 | | 0.865 | 0.483 |
| (150,180) | | 0.842 | 0.471 | | 0.866 | 0.471 |
| (100,100) | 0.4 | 0.863 | 0.601 | 0.5 | 0.857 | 0.608 |
| (120,100) | | 0.875 | 0.555 | | 0.868 | 0.560 |
| (150,150) | | 0.882 | 0.481 | | 0.853 | 0.478 |
| (150,180) | | 0.856 | 0.468 | | 0.843 | 0.475 |
| (100,100) | 0.6 | 0.899 | 0.592 | 0.7 | 0.904 | 0.594 |
| (120,100) | | 0.896 | 0.554 | | 0.909 | 0.643 |
| (150,150) | | 0.901 | 0.480 | | 0.913 | 0.481 |
| (150,180) | | 0.873 | 0.470 | | 0.905 | 0.464 |
| (100,100) | 0.8 | 0.914 | 0.594 | 0.9 | 0.902 | 0.585 |
| (120,100) | | 0.927 | 0.565 | | 0.937 | 0.558 |
| (150,150) | | 0.917 | 0.478 | | 0.914 | 0.473 |
| (150,180) | | 0.915 | 0.469 | | 0.907 | 0.467 |
| (100,100) | average | | | | 0.885 | 0.598 |
| (120,100) | average | | | | 0.892 | 0.568 |
| (150,150) | average | | | | 0.889 | 0.480 |
| (150,180) | average | | | | 0.877 | 0.469 |

Note: We use the first pair in Table 5 for different $\rho$ for $D_n$.

Table 8: Coverage probabilities (CP) and average widths (AW) of two-sided 90% confidence intervals for $\Delta$ with two dependent samples for different $\rho$

| $(n_1, n_2)$ | $\rho$ | $M_n$ | | $\rho$ | $M_n$ | |
|---|---|---|---|---|---|---|
| | | CP | AW | | CP | AW |
| (100,100) | 0 | 0.881 | 0.703 | 0.1 | 0.890 | 0.702 |
| (120,100) | | 0.887 | 0.641 | | 0.891 | 0.652 |
| (150,150) | | 0.887 | 0.555 | | 0.893 | 0.553 |
| (150,180) | | 0.903 | 0.543 | | 0.894 | 0.547 |
| (100,100) | 0.2 | 0.888 | 0.701 | 0.3 | 0.921 | 0.703 |
| (120,100) | | 0.902 | 0.648 | | 0.905 | 0.643 |
| (150,150) | | 0.909 | 0.554 | | 0.912 | 0.555 |
| (150,180) | | 0.909 | 0.540 | | 0.900 | 0.512 |
| (100,100) | 0.4 | 0.910 | 0.637 | 0.5 | 0.930 | 0.700 |
| (120,100) | | 0.905 | 0.604 | | 0.924 | 0.646 |
| (150,150) | | 0.913 | 0.523 | | 0.920 | 0.556 |
| (150,180) | | 0.904 | 0.502 | | 0.915 | 0.541 |
| (100,100) | 0.6 | 0.942 | 0.689 | 0.7 | 0.945 | 0.687 |
| (120,100) | | 0.932 | 0.641 | | 0.945 | 0.641 |
| (150,150) | | 0.943 | 0.558 | | 0.939 | 0.552 |
| (150,180) | | 0.939 | 0.538 | | 0.934 | 0.535 |
| (100,100) | 0.8 | 0.953 | 0.691 | 0.9 | 0.948 | 0.692 |
| (120,100) | | 0.947 | 0.649 | | 0.956 | 0.630 |
| (150,150) | | 0.962 | 0.551 | | 0.958 | 0.543 |
| (150,180) | | 0.951 | 0.542 | | 0.960 | 0.540 |
| (100,100) | average | | | | 0.920 | 0.690 |
| (120,100) | average | | | | 0.919 | 0.639 |
| (150,150) | average | | | | 0.923 | 0.550 |
| (150,180) | average | | | | 0.921 | 0.534 |

Note: We use the first pair in Table 5 for different $\rho$ for $M_n$.


Figure 1: Time series plots of the standardized Stock Market Index of Singapore (STI) and Standard & Poor's 500 Index (S&P500).

Table 9: Confidence intervals for returns of one-sample with confidence level 95%

| Date(mm/dd/yy) | STI | | |
|---|---|---|---|
| | $A_n$ | $B_{1,n}$ | $B_{2,n}$ |
| 01/07/2000-12/28/2001 | (-0.072,0.162) | (-0.002,0.2646) | (-0.043,0.132) |
| 01/04/2002-12/26/2003 | (-0.013,0.015) | (-0.001,0.111) | (-0.008,0.009) |
| 01/02/2004-12/30/2005 | (-0.072,0.253) | (-0.001,0.348) | (-0.059,0.240) |
| 01/06/2006-12/28/2007 | (-0.072,0.222) | (-0.001,0.284) | (-0.052,0.202) |
| 01/04/2008-12/25/2009 | (-0.034,0.045) | (-0.002,0.132) | (-0.020,0.031) |
| 01/01/2010-12/30/2011 | (-0.020,0.023) | (-0.001,0.117) | (-0.011,0.015) |
| 01/06/2012-12/27/2013 | (-0.061,0.111) | (-0.001,0.236) | (-0.040,0.090) |
| 01/03/2014-12/25/2015 | (-0.043,0.062) | (-0.001,0.128) | (-0.022,0.040) |

| Date(mm/dd/yy) | S&P500 | | |
|---|---|---|---|
| | $A_n$ | $B_{1,n}$ | $B_{2,n}$ |
| 01/07/2000-12/28/2001 | (-0.058,0.098) | (-0.001,0.173) | (-0.032,0.072) |
| 01/04/2002-12/26/2003 | (-0.015,0.017) | (-0.003,0.268) | (-0.043,0.133) |
| 01/02/2004-12/30/2005 | (-0.063,0.120) | (-0.001,0.197) | (-0.034,0.090) |
| 01/06/2006-12/28/2007 | (-0.065,0.132) | (-0.001,0.224) | (-0.039,0.106) |
| 01/04/2008-12/25/2009 | (-0.053,0.083) | (-0.002,0.161) | (-0.028,0.058) |
| 01/01/2010-12/30/2011 | (-0.035,0.046) | (-0.002,0.120) | (-0.020,0.031) |
| 01/06/2012-12/27/2013 | (-0.058,0.384) | (-0.005,0.477) | (-0.059,0.385) |
| 01/03/2014-12/25/2015 | (-0.045,0.066) | (-0.001,0.139) | (-0.025,0.046) |

Table 10: Test for independence of returns of STI and S&P500

| Date(mm/dd/yy) | $p$-values of the Correlation Test |
|---|---|
| 01/07/2000-12/28/2001 | 0.001 |
| 01/04/2002-12/26/2003 | 0 |
| 01/02/2004-12/30/2005 | 0 |
| 01/06/2006-12/28/2007 | 0.001 |
| 01/04/2008-12/25/2009 | 0.002 |
| 01/01/2010-12/30/2011 | 0.001 |
| 01/06/2012-12/27/2013 | 0 |
| 01/03/2014-12/25/2015 | 0.001 |

Table 11: Confidence intervals for difference of EPM between returns of STI and S&P 500 with confidence level 95%

| Date(mm/dd/yy) | $D_n$ | $M_n$ |
|---|---|---|
| 01/07/2000-12/28/2001 | (-0.116,0.165) | (-0.043,0.082) |
| 01/04/2002-12/26/2003 | (-0.021,0.021) | (-0.012,0.010) |
| 01/02/2004-12/30/2005 | (-0.124,0.248) | (-0.034,0.114) |
| 01/06/2006-12/28/2007 | (-0.135,0.219) | (-0.041,0.137) |
| 01/04/2008-12/25/2009 | (-0.088,0.069) | (-0.061,0.043) |
| 01/01/2010-12/30/2011 | (-0.049,0.042) | (-0.035,0.020) |
| 01/06/2012-12/27/2013 | (-0.375,0.100) | (-0.272,0.053) |
| 01/03/2014-12/25/2015 | (-0.078,0.075) | (-0.053,0.028) |