Probability in the Engineering and Informational Sciences, **35**, 2021, 404–431.
doi:10.1017/S0269964819000482

AN IMPROVED TEST OF THE SQUARED SHARPE RATIO

Wan-Yi Chiu
Department of Finance, National United University, Taiwan, Republic of China
E-mail: wychiu@nuu.edu.tw

The sample squared Sharpe ratio (SSR) is a critical statistic of the risk-return tradeoff. We show that sensitive upper-tail probabilities arise when the sample SSR is employed to test the mean-variance efficiency under different test statistics. Assuming the error's normality with a nonzero mean, we integrate the sample SSR and the arbitrage regression into a noncentral chi-square (χ^2) test. We find that the distribution of the sample SSR based on the regression error is to the left of the F-distribution when assuming the returns' normality. Compared to two benchmarks that use the noncentral F-distribution and the central F-statistic, the χ^2-statistic is more effective, competitive, significant, and locally robust when used to reject the upper-tailed mean-variance efficiency test using the usual parameters (sample size, portfolio size, and SSR).

Keywords: arbitrage regression, chi-square risk model, mean-variance efficiency test, squared sharpe ratio, upper-tail probability

1. INTRODUCTION

The squared Sharpe ratio (SSR, denoted by θ^2) is a critical component in mean-variance efficiency testing (MVET) that identifies whether the portfolio using a subset of assets is statistically close to the optimal one based on a complete set of assets. The sample SSR (denoted as $\hat{\theta}^2$) is not only one of the efficient constants used to the asset pricing but also a significant statistic in computing the power of testing a given portfolio. Explicitly speaking, the test statistics of MVET can usually be expressed as a function of $\hat{\theta}^2$ and θ^2 (e.g., the asymptotical F-statistic of Jobson and Korkie [11], the noncentral parameter of Gibbons et al. [9], the central F-statistic of Britten-Jones [3], the noncentral F-distribution of the mean-variance optimization discussed by Kan and Zhou [13]).

To apply the sample SSR to access the MVET, we need to identify the sampling distribution of $\hat{\theta}^2$ based on certain assumptions. Note that the vast literature available on the sampling distribution of the sample SSR describes all the key assumptions such as the returns' normality, independence, and constant volatility. However, widespread empirical evidence has been challenging the return's normality. For instance, Affleck-Graves and McDonald [2] indicate that the sample nonnormalities are severe, and the multivariate analysis is reasonably robust for typical levels of nonnormality. Besides, the tests conducted under weak distributional assumptions have a relatively more robust effect than the studies

 0269-9648/20 $25.00

depending on the normality assumption [16]. Zhou [25] computes the p-value of the statistic based on the other distributions. Richardson [20] presents evidence that the return's normality does not correspond to the marginal and joint distributions of returns. Moreover, Marie-Claude et al. [17] show that the MVETs reject the Gaussian error assumption.

The primary concern of the MVET is to check whether a particular portfolio, ω_τ, is statistically close to the mean-variance portfolio. In addition, investors are usually interested in their portfolio's gains 'to the right' of the specific benchmark, but they do not concern about profits to the left of the benchmark. To perform the MVET relative to this particular portfolio, we set up the upper-tailed test of the SSR as follows:

$$H_0 : \theta^2 = \theta_\tau^2 \quad \text{against} \quad H_1 : \theta^2 > \theta_\tau^2. \tag{1}$$

A significantly positive difference between $\hat{\theta}^2$ and θ_τ^2 will yield a critical opportunity cost from investing in this particular portfolio. All other things being equal (the sample size and the portfolio size), we favor the rejection of the MVET when the difference between $\hat{\theta}^2$ and θ_τ^2 increases more than a suitable amount. If not, the sample SSR is not significantly different from the defaulted SSR. It seems reasonable to establish a portfolio based on a smaller portfolio size because its management cost may be lower than the augmented portfolio. Because it is difficult to compare the performance of various statistics expressed in the different functions, the calculation of the upper-tail probability (UTP) curve provides an alternative to shape the sample SSR under different statistics. In this case, we compare the above significance test using various statistics.

Even though the return's normality problem is addressed in the literature, abandoning such an assumption does not eliminate model risk. The commonly predicated assumption about the assets' returns is that returns are independent and identically normally distributed; however, this premise may be relaxed under certain conditions (e.g., Kwon [15]). This typical assumption is broadly employed in portfolio analysis either because theoretical derivations are allowed to be made (with apparent advantages to the statistical inferences of the model itself) or because the presumption can be overcome to not cause substantial differences in the results (with the asymptotical distribution based on a larger sample). Several recent studies also employ the returns' normality assumption to estimate the SSR. For instance, Kourtis [14] provides two approximated expectations of the SSR using the first-order and the second-order Taylor series expansions under the returns' normality. Abhyankar et al. [1] assess the investor's economic gains through the difference between the conditional SSR and the unconditional SSR. Peñaranda [19] characterizes the certainty-equivalent returns between the unconditional SSR of excess returns and the maximum conditional SSR.

1.1. Benchmarks

The returns' normality assumption immediately delivers two benchmarks for the statistical inferences of the population SSR. The most obvious advantage is that the sample means and covariance matrix have a multivariate normal distribution and Wishart distribution, respectively, under the returns' normality. As a result, Kan and Zhou [13] show that the sample SSR has a noncentral F-distribution. We consider the noncentral F-distribution as the first benchmark for evaluating the sample SSR and computing the expectation of the sample SSR. For other references, see Jobson and Korkie [10], Kan and Smith [12] that obtain the same expectation of the sample SSR with the adjustment of the parameters.

Another approach to computing the sample SSR's expectation is indirectly to impose the normal assumption on the regression error in which assuming the returns' normality implies the normalized errors in linear regression. For example, Jobson and Korkie [11] regress the constant return on the assets' returns without intercept and present an asymptotical F-statistic for the MVET. Hereafter, we refer to the regression as the arbitrage regression. Also, Gibbons et al. [9] obtain the F-test by performing a multivariate regression of assets' returns conditional on the excess returns for a particular portfolio. Britten-Jones [3] expresses the sample SSR of the asset space and the restricted SSR relative to a subset of assets as a function of the central F-statistic when the arbitrage regression error is normally distributed with a zero mean. Furthermore, Britten-Jones shows that the GRS F-test for the MVET of a given portfolio weight is equivalent to perform the F-statistic using the arbitrage regression. Therefore, we consider the GRS F-test and Britten-Jones F-statistic as the second benchmark for evaluating the sample SSR.

1.2. The chi-square statistic for evaluating the SSR

One particular point between the normality and the F-statistic (either exact or asymptotical) should be highlighted.

Britten-Jones assumes multivariate normal returns on the arbitrage regression, which implies the error's normality and derives the exact F-statistic. Note that the arbitrage regression is not only without a constant but also has a nonstochastic dependent variable. Note that if variables are measured in unstandardized scores, then the intercept becomes very important. However, it is also worth noting that the arbitrage regression error is not always zero. More importantly, the nonzero mean of the regression error may violate the assumption of the F-statistic in the statistical sense.

To mitigate the impact of the zero regression mean in constructing the test statistics, we relax the assumptions used by Jobson and Korkie [11] and Britten-Jones [3]. Compared to those stronger assumptions, such as returns' normality, independence, and constant volatility, this paper assumes only that the regression errors are independent and normally distributed with a nonzero mean. Such weaker normality on the error term is typical of principles employed in regression analysis that can be applied to a broader class of risk regressions. We have several findings on the statistical inference of the SSR based on the modified arbitrage regression risk model:

- Using the weaker error's normality of the regression error with a nonzero mean, we integrate the sample SSR and the arbitrage regression into a noncentral chi-square statistic (χ^2-statistic) for testing the performance of a particular portfolio. As a result, we extend the noncentral χ^2-statistic to the MVET with probabilistic properties comparable to that of popular benchmarks based on the stronger return's normality.
- We analytically derive the sample SSR as a functional form of the noncentral χ^2-statistic that allows us to calculate an alternative expectation of the sample SSR. Consequently, we prove that the expectation of the sample SSR under the regression error with a nonzero mean is less than the expectations assuming the returns' normality for the central F-statistic and the noncentral F-distribution.
- We compare the UTP of the sample SSR of the noncentral χ^2-statistic with the two popular benchmarks mentioned above. The simulated evidence indicates that the UTP of the sample SSR under the χ^2-statistic is to the left of the UTP when assuming the returns' normality. In this case, investors are more conservative to reject the MVET using the central F-statistic and the noncentral F-distribution. As a result, modifying the error's normality assumption leads to the noncentral

χ^2-statistic being more practical to reject the MVET, unlike the central F-statistic and the noncentral F-distribution.

- Note that the sample SSR is sensitive to portfolio size and sample size under various normality assumptions. In other words, the power of the MVET (either compute analytically or numerically) is affected by sample size and portfolio size. The illustrated figures reveal that the shifted UTP of the sample SSR based on the noncentral χ^2-statistic is locally robust relative to sample size and portfolio size. As a result, the F-test may be conservative when rejecting a false null hypothesis.
- Based on the practical evidence, we suggest a parameters domain where the sample size is between 60 and 600, the portfolio size is between 5 and 100, and the SSR is from 0.01 to 1.00. We show that both the noncentral χ^2-statistic and the Britten-Jones F-statistic are superior to the noncentral F-statistic within our parameters domain. Moreover, the noncentral χ^2-statistic generally outperforms the Britten-Jones F-statistic except for in tests that probably combine a smaller portfolio size and a larger SSR.

Overall, the performance of all the sample SSR probability measures is validated using both first-order moment (the expectation) and the UTP. Our study contributes to the literature by investigating the impact of different probability measures on the sample SSR and providing an effective and a locally robust χ^2-statistic for the upper-tailed MVET based on the weaker regression error's normality and the parameters domain. The most obvious advantage is that identifying the error's normality can be easier than testing the returns' multivariate normality.

The paper proceeds as follows. In Section 2, we review the properties of the sample SSR and the arbitrage regression. In Section 3, we propose a noncentral χ^2-statistic to test the MVET relative to the particular portfolio. In addition, we validate the noncentral χ^2-statistic against the benchmarks (the noncentral F-distribution, and the central F-statistic) based on the statistical inferences of the sample SSR under different distributions. Conclusions follow in Section 4. Finally, the expectations of the sample SSR under various probability measures are derived in the Appendix. The distributions of the sample SSR using the parameters domain are also presented in the Appendix.

2. THE SQUARED SHARPE RATIO

We begin this study by reviewing the SSR and two mean-variance optimizations. Assume that the investment universe has N risky assets with $N \times 1$ excess returns vector R_t at time t. The excess return vector has the expectation μ and a positive definite covariance matrix Σ. Given the $N \times 1$ portfolio weight vector ω, the excess return on the portfolio weight vector is $R_{pt} = \omega' R_t$ at time t. The portfolio has its expected excess return and variance as $\omega'\mu$ and $\omega'\Sigma\omega$, respectively. The investor is assumed to select ω to maximize mean-variance utility

$$U(\omega) = \omega'\mu - \frac{\gamma}{2}\omega'\Sigma\omega, \tag{2}$$

where γ is the coefficient of relative risk aversion. The optimal portfolio choice $\omega^* = \Sigma^{-1}\mu/\gamma$ is proportional to the ex ante tangency portfolio. The corresponding utility is expressed as a function of the SSR and the coefficient of relative risk aversion

$$U(\omega^*) = \frac{\theta^2}{2\gamma}, \tag{3}$$

where $\theta = \sqrt{\mu'\Sigma^{-1}\mu}$ is the Sharpe ratio of the ex ante tangency portfolio of the risky assets and $\theta^2/2\gamma$ is the corresponding certainty-equivalent return. The certainty-equivalent return shows if an investor is willing to accept the portfolio ω^*. The SSR is a standard parameter of the mean-variance optimization that Kan and Zhou [13], and DeMiguel et al. [4], among others, use to evaluate the performance of the portfolio choice. In addition, Abhyankar et al. [1] assess the investor's economic gains through the difference between the conditional SSR and the unconditional SSR. Peñaranda [19] characterizes the certainty-equivalent returns in terms of the unconditional SSR of excess returns and the maximum conditional SSR.

2.1. Traditional estimation

Suppose we have T observations of $N \times 1$ excess returns vector R_t, for $t = 1, 2, \ldots, T$. The commonly used estimates of μ and Σ are given by

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t \quad \text{and} \quad \hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \hat{\mu})(R_t - \hat{\mu})'. \tag{4}$$

Under the normality assumption of returns, $\hat{\mu}$ and $\hat{\Sigma}$ are mutually independent and they have the N-dimensional normal distribution $N(\mu, \Sigma/T)$ and the Wishart distribution $W(T-1, \Sigma/T)$ with $T-1$ degrees of freedom and covariance matrix Σ/T, respectively.

Incorporating the sample counterparts $\hat{\mu}$ and $\hat{\Sigma}$ into the SSR, the sample SSR is defined as $\hat{\theta}^2 = \hat{\mu}'\hat{\Sigma}^{-1}\hat{\mu}$. There are a few ways to obtain the expected value and the variance of $\hat{\theta}^2$. Kan and Zhou [13, p. 637] show that the sample SSR has the following noncentral F-distribution with the noncentrality $T\theta^2$:

$$\hat{\theta}^2 \sim \frac{N}{T-N} F_{N, T-N, T\theta^2}. \tag{5}$$

In this case, the first two moments are denoted as $E_{\mathrm{NF}}(\hat{\theta}^2)$ and $V_{\mathrm{NF}}(\hat{\theta}^2)$, and the derivations are straightforward.

Lemma 1: *The expectation and the variance of $\hat{\theta}^2$ are*

$$E_{\mathrm{NF}}(\hat{\theta}^2) = \frac{N + T\theta^2}{T-N-2}, \quad V_{\mathrm{NF}}(\hat{\theta}^2) = \frac{2T^2\theta^4 + 2(T-2)(N + 2T\theta^2)}{(T-N-2)^2(T-N-4)}. \tag{6}$$

There are other ways to compute the statistical moments of the sample SSR. With the appropriate scaling adjustment of the factors, one possibility is how Kan and Smith [12] rewrite the results in Jobson and Korkie [10] as Eq. (6). Other possibilities for deriving the same results by conditioning the joint distribution of the efficiency set constants may refer to Okhrin and Schmid [18]. All the derivations assume the normality in the assets' returns.

2.2. Inferences based on the arbitrage regression

Alternatively, Jobson and Korkie [11] regress the constant return on the assets' returns without an intercept. In the empirical studies, the arbitrage regression takes the form:

$$\ell = R\omega + \epsilon, \tag{7}$$

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

where ℓ denotes the nonstochastic arbitrage returns (dependent variable) with $T \times 1$ vector of one; the $T \times N$ matrix R only contains the excess returns of the N risky assets without the intercept terms; the $T \times 1$ vector ϵ is the regression error.

Specifically, Britten-Jones [3] provides regression analyses to infer the mean-variance spanning based on linear restriction on the portfolio weights. Instead of citing the well-known results, we explicitly outline some regression procedures. Thus, our discussion may be a self-contained exposition of the properties of the sample SSR.

Jobson and Korkie derive the OLS estimator of the regression coefficient as:

$$\hat{\omega} = (R'R)^{-1}R'\ell = \frac{\hat{\Sigma}^{-1}\hat{\mu}}{1+\hat{\theta}^2}. \tag{8}$$

Hereafter, we refer to the regression coefficients as the mean-variance efficient portfolio (MVEP). Note that the regression coefficients may not sum to one. The scaling MVEP will result in the sample's tangency portfolio. It is evident that both the MVEP and the sample's tangency portfolio have the same SSR. Jobson and Korkie and Britten-Jones show that the unrestricted sum of the squared errors from regression (7) is:

$$\mathrm{SSE}_u = (\ell - R\hat{\omega})'(\ell - R\hat{\omega}) = \frac{T}{1+\hat{\theta}^2}. \tag{9}$$

Expression (9) shows that the mean squared error (MSE), $\hat{\sigma}^2 = \mathrm{SSE}_u/T$, is inversely proportional to the sample SSR. Geometrically, a steeper tangent line for the efficient frontier produces a larger Sharpe ratio, which implies a significantly smaller MSE. Thus, the MVEP has a good fit with the arbitrage regression and produces the largest SSR.

Britten-Jones also suggests that the restricted sum of the squared errors (SSE_r) for testing a particular portfolio ω_τ can be obtained using the following univariate regression with an appropriate parameter adjustment:

$$\ell = (R\omega_\tau)\beta + \epsilon. \tag{10}$$

Thus, the estimators of the coefficients β and SSE_r can be further expressed as:

$$\begin{aligned} \hat{\beta} &= \frac{(R\omega_\tau)'\ell}{(R\omega_\tau)'(R\omega_\tau)}, \\ \mathrm{SSE}_r &= \left(\ell - (R\omega_\tau)\hat{\beta}\right)'\left(\ell - (R\omega_\tau)\hat{\beta}\right) = \frac{T}{1+\theta_\tau^2}, \end{aligned} \tag{11}$$

where θ_τ^2 is the SSR relative to the particular portfolio ω_τ. In addition, Britten-Jones extends the GRS F-statistic to an OLS F-statistic for use in testing a particular portfolio where the null hypotheses are expressed in terms of the linear restrictions on the efficient portfolio weights. In this study, we consider the portfolio ω_τ examined in the MVET (1). When normal multivariate returns are obtained, the regression's error is normally distributed. Britten-Jones integrates the features of the return's normality and the arbitrage regression into the following central F-statistic:

Lemma 2: *The exact central F-statistic for testing a particular portfolio ω_τ of Eq. (1) is expressed as:*

$$F = \frac{(\mathrm{SSE}_r - \mathrm{SSE}_u)/N}{\mathrm{SSE}_u/(T-N)} = \frac{T-N}{N} \times \frac{\hat{\theta}^2 - \theta_\tau^2}{1+\theta_\tau^2} \sim F_{N,T-N}. \tag{12}$$

Consequently, the large values of the F-statistic (values of SSE_r larger than SSE_u by a suitable amount) favor the rejection of the mean-variance efficiency between the MVEP and *the portfolio ω_τ. With an appropriate scaling adjustment of the factors, Eq.* (12) *is similar to the asymptotical F-statistic of Jobson and Korkie* [11, Eq. (32) on p. 195].

3. MAIN RESULTS

Using the weaker error's normality of the regression error with a nonzero mean, we integrate the sample SSR and the arbitrage regression into a noncentral χ^2-statistic for evaluating the performance of a particular portfolio. To validate the noncentral χ^2-statistic, we discuss our findings for the moments, the significance tests of the SSR, and the robustness of the statistics with the two popular benchmarks mentioned above.

3.1. Moment of the sample SSR under the χ^2 distributions

The arbitrage regression can be written as:

$$1 = \omega' R_t + \epsilon_t. \tag{13}$$

The MVEP of minimizing the expected MSE is $\omega = \Sigma^{-1}\mu/(1+\theta^2)$. Moreover, the expected MSE relative to the MVEP is:

$$E\left[1 - \frac{\mu'\Sigma^{-1}R_t}{1+\theta^2}\right]^2 = \frac{1}{1+\theta^2}. \tag{14}$$

Conditional on the returns' normality, the commonly used manner of characterizing inferences in a linear regression is to assume that the regression error is normally distributed with a zero mean and a finite variance, $\epsilon_t \sim N(0, \sigma^2)$. In the present situation, Seber [21, p. 97] indicates that the ratio of SSE_u to the error's variance σ^2 takes the form:

$$\frac{\mathrm{SSE}_u}{\sigma^2} \sim \chi^2_{T-N}. \tag{15}$$

One particular point between the normality and the central F-statistic should be highlighted. Britten-Jones notes that "the lack of an intercept in the arbitrage regression may result in errors need not sum to zero," (page 659). In fact, the error's expectation is definitely not zero since:

$$E(\epsilon_t) = E\left[1 - \frac{\mu'\Sigma^{-1}R_t}{1+\theta^2}\right] = \frac{1}{1+\theta^2}. \tag{16}$$

In addition, the error's variance in (13) is given by:

$$\sigma^2 = E\left[1 - \frac{\mu'\Sigma^{-1}R_t}{1+\theta^2} - \frac{1}{1+\theta^2}\right]^2 = \frac{\theta^2}{(1+\theta^2)^2}. \tag{17}$$

Thus, it is more appropriate for us to assume that the regression error is normally distributed as $\epsilon_t \sim N(1/(1+\theta^2), \sigma^2)$. We modify the χ^2-distribution as follows.

PROPOSITION 1: *Conditional on the normal returns and on the i.i.d. errors, the ratio of* SSE_u *to the error's variance* σ^2 *is of the form:*

$$\frac{\text{SSE}_u}{\sigma^2} = \frac{\epsilon'\epsilon}{\sigma^2} = \frac{T(1+\theta^2)^2}{\theta^2(1+\hat{\theta}^2)} \sim \chi^2_{T-N,\delta}, \tag{18}$$

where $\chi^2_{T-N,\delta}$ *is the noncentral* χ^2*-distribution with* $(T-N)$ *degrees of freedom and has the noncentrality parameter:*

$$\delta = \sum_{t=1}^{T} (E(\epsilon_t/\sigma))^2 = \frac{T}{(1+\theta^2)^2} \times \frac{(1+\theta^2)^2}{\theta^2} = \frac{T}{\theta^2}. \tag{19}$$

At this point, the sample SSR is a function of the χ^2-distribution:

$$\hat{\theta}^2 = \frac{T}{\sigma^2} \times \frac{1}{\text{SSE}_u/\sigma^2} - 1 = \frac{T(1+\theta^2)^2}{\theta^2 \chi^2_{T-N,T/\theta^2}} - 1. \tag{20}$$

A significant concern when using the noncentral χ^2-statistic in portfolio analysis is whether to consider competitors when comparing moments. The following proposition shapes the locations and measures the dispersions of the sample SSR among various distributions. A notable feature is that the distribution of the sample SSR under the regression error with a nonzero mean is to the left of the distribution when assuming the returns' normality.

PROPOSITION 2: *Confining the SSR to the condition* $\theta^2 \le 1$ *and ignoring the higher orders of* δ*, the moments of the sample SSR are approximated as:*

$$E_{\text{NC}}(\hat{\theta}^2) \simeq \frac{T(T-N)\theta^2 + (2T - N^2 - 2N)}{(T-N-2)(T-N)} \tag{21}$$

and

$$V_{\text{NC}}(\hat{\theta}^2) \simeq \frac{2T^2(T-N)\theta^4 + 4T^2(T-2N)\theta^2 + 4T^2(-T-3N)}{(T-N)(T-N-2)^2(T-N-4)}. \tag{22}$$

Moreover, the moments of the sample SSR among the different distributions are:

$$E_{\text{NC}}(\hat{\theta}^2) < E_{BJ}(\hat{\theta}^2) < E_{\text{NF}}(\hat{\theta}^2) \tag{23}$$

and

$$\left\{V_{\text{NC}}(\hat{\theta}^2), V_{\text{BJ}}(\hat{\theta}^2)\right\} < V_{\text{NF}}(\hat{\theta}^2), \tag{24}$$

where the subscripts "BJ," "NF," and "NC" denote the moments based on the Britten-Jones F-statistic, the noncentral F-statistic, and the noncentral χ^2*-statistic, respectively. (See the Appendix.)*

Using the normalization at the same confidence level, Proposition 2 shows that both the concentral χ^2-statistic and the Britten-Jones F-statistic have the lower critical values than the noncentral F-statistic as T and N approach infinite. However, the comparison between the concentral χ^2-statistic and the Britten-Jones F-statistic should be validated case by case. Note that the sensitive ranking of moments under different distributions depends on the parameters (T, N, θ^2). In particular, the determining factor of comparing those moments is the SSR's default value. There is one major reason for employing the restricted domain

TABLE 1. Parameters domain in previous studies. This table contains the usual sample size and portfolio size used in empirical studies and the realized Sharpe ratios across different portfolios. Some important points follow: (1) DeMiguel et al. [4] assess a total of 84 portfolios using the Sharpe ratio. Only the top two portfolios produce Sharpe ratios of 0.510 and 0.536, respectively. The remaining 82 portfolios' Sharpe ratios are all between −0.035 and 0.385. (2) Frazzini et al. [7] report that the 2872 funds in CRSP data with at least 10-year history during the period 1976–2011 have the median of 0.39, a 95th percentile of 0.62, and a maximum of 0.99. In addition, the impressive Sharpe ratio of Warren Buffett's Berkshire Hathaway is 0.76; it is ranked in the top 11 and the 99.7% percentile. Note that the Berkshire Hathaway portfolio holdings of 2019 Q2 consists of 47 leading stocks, such as AAPL, BAC, KO, etc. (3) Gibbons et al. [9] report that the Sharpe ratio is 0.172 over the time period of 1926–1982. They also present 4 portfolios with a subperiod of 10 years in which the Sharpe ratios range from 0.286 to 0.604. (4) Sharpe [22] also evaluates 34 mutual funds between 1954 and 1963 in which the Sharpe ratios range from 0.431 to 0.778. (5) Tang theoretically uses 100 stocks to diversify a portfolio up to 99% of total diversifiable risk based on a universe with infinite assets.

Literature	Sample size (T)	Portfolio size (N)	Sharpe ratio
DeMiguel et al. [4, pp. 1918 & 1931]	264, 379, 497	3, 9, 11, 21, 24	(1)
	497	21	0.510
	497	24	0.536 (maximum)
Frazzini et al. [7, pp. 35 & 37]	432	47	0.790 (Buffett's)
	120–480	2872 Funds (2)	0.990 (maximum)
Gibbons et al. [9, pp. 1141 & 1132]	60, 120, 240	10, 20	N/A
	684	10	0.172
	120	10	(3)
	84	10	0.538
Green and Hollifield [8, p. 1803]	60	10, 30, 50	N/A
Sharpe [22, pp. 125 & 136]	120	DOW30	0.667
	120	34 Funds (4)	0.778 (maximum)
Tang [24]	∞ (5)	8–40, 20, 100	N/A
Parameters domain used in this paper	60, 120, 240 360, 480, 600	5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0

($\theta^2 \leq 1$) instead of considering all possible values of this SSR: it is difficult to derive a closed-form comparison between the noncentral χ^2-statistic and the F-statistic. Alternatively, the comparisons are best examined numerically using an operational domain of parameters. Table 1 summarizes the usual sample size, the portfolio size used in empirical studies, and the realized Sharpe ratios in different portfolios from several popular articles. This parameters domain is somewhat consistent with the practical SSR operation used by academic researchers and financial practitioners.

For the sample size (time period), DeMiguel et al. [4, p. 1919] indicate that models are typically estimated using only 60 or 120 months of data. However, Column 2 of Table 1 shows that sample sizes ranging from 60 months to a long horizon of 684 months have been previously used. This may be because investors are concerned about the realized performance of their portfolio in the most recent 5–10 years. However, researchers are interested in precisely estimating the return and evaluating the portfolio performance from the long run perspective. In sum, this paper sets the sample size between 60 and 600.

For portfolio size, Column 2 of Table 1 indicates that empirical studies usually use a number of assets between 3 and 100. Notably, Tang [24] summarizes the optimal number

TABLE 2. The test statistic of SSR under various distributions. Given a particular critical value (C), we compute the upper-tailed probabilities under various distributions. Conversely, we can compute the critical values under various distributions at the same significance level (α).

Test statistic	Critical value	Upper-tailed probability
Noncentral $\chi^2_{T-N,T/\theta^2_\tau}$	$C_{\mathrm{NC}} = \dfrac{T(1+\theta^2_\tau)^2}{\theta^2_\tau \times \chi^2_{T-N,T/\theta^2_\tau;\alpha}} - 1$	$P_{\mathrm{NC}}\left(\chi^2_{T-N,T/\theta^2_\tau} \leq \dfrac{T(1+\theta^2_\tau)^2}{\theta^2_\tau(1+C)}\right)$
Britten-Jones $F_{N,T-N}$	$C_{\mathrm{BJ}} = \dfrac{N(1+\theta^2_\tau)F_{N,T-N;\alpha}}{T-N} + \theta^2_\tau$	$P_{\mathrm{BJ}}\left(F_{N,T-N} \geq \dfrac{(T-N)(C-\theta^2_\tau)}{N(1+\theta^2_\tau)}\right)$
Noncentral $F_{N,T-N,T\theta^2_\tau}$	$C_{\mathrm{NF}} = \dfrac{N \times F_{N,T-N,T\theta^2_\tau;\alpha}}{T-N}$	$P_{\mathrm{NF}}\left(F_{N,T-N,T\theta^2_\tau} \geq \dfrac{C(T-N)}{N}\right)$

of assets for diversification purposes drawing on 10 investment textbooks and 10 financial management textbooks. Across the 20 books, the smallest number is 8, and the largest is approximately 40. Most of these textbooks' reviews are based on papers from Evans and Archer [6], Elton and Gruber [5], and Statman [23]. However, given an universe of infinite stocks, Tang also analytically shows that at least 20 stocks are required to eliminate 95% of the diversifiable risk on average. However, investors need a portfolio size of 100 to diversify up to 99% of total diversifiable risk. Note that a portfolio size between 8 and 40 is primarily influenced by the portfolio's management cost, and a larger portfolio size, such as 100 stocks, indicates a well-diversified portfolio. In sum, this paper set the portfolio size as a range from 5 to 100.

For the Sharpe ratios, Frazzini et al. [7] present a comprehensive investigation of mutual funds' Sharpe ratios where all 2872 funds in the CRSP data with at least 10-year history in the period 1976–2011 have a median of 0.39, a 95th percentile of 0.62, and a maximum of 0.99. In particular, the impressive Sharpe ratio of Warren Buffett's Berkshire Hathaway is 0.76, which is ranked in the top 11 and the 99.7% percentile. Of course, some mutual funds with the higher Sharpe ratios over one possibly exist, but they often survive for only short time. Table 1 indicates that only a tiny number of mutual funds could outperform the Berkshire Hathaway portfolio holdings over the long horizon. Moreover, there are many miles to go in searching for better tests for the cases $\theta^2 > 1$ since the sample SSR using all the three distributions is poorly behaved for these cases. For example, based on the calculations of Table 2, the noncentral χ^2, the Britten-Jones F, and the noncentral F require $\mathrm{CV_{NC}} = 23.292$, $\mathrm{CV_{BJ}} = 51.664$, and $\mathrm{CV_{NF}} = 54.046$ to reject $H_0 : \theta^2 = 2.00$ in the case $(T, N) = (120, 100)$. The justification is unacceptable due to the extremely unrealistic critical values. Therefore, we summarize the above meaningful information and focus on the reasonable Sharpe ratios such as $\theta = 0.1, 0.2, \ldots, 1.0$ (i.e., $\theta^2 \in \Theta^2_\tau = \{0.01, 0.04, 0.09, \ldots, 0.81, 1.00\}$).

3.2. Significance test of the SSR

Another major concern of the mean-variance efficiency analysis is whether a particular portfolio, ω_τ, is statistically close to the MVEP. In a situation such as this, we may compare the portfolio ω_τ with the MVEP based on the arbitrage regression. Thus, we set up the null hypothesis (1) that the SSR equals the estimated SSR of the portfolio ω_τ ($H_0 : \theta^2 = \theta^2_\tau$ against $H_1 : \theta^2 > \theta^2_\tau$). We hope to reject the hypothesis by obtaining a sample SSR of the MVEP that is statistically larger than θ^2_τ. According to the noncentral χ^2-statistic (18), we propose the following test statistic.

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

Table 3. The upper-tail probabilities of the sample SSR under various distributions for $T = 120$ and $N = 30$. Taking the case $\theta_\tau^2 = 0.09$ at the UTP $= 0.002$ as an example, this means that $\pi_{\mathrm{NC}}(0.09) = P_{\mathrm{NC}}(\hat{\theta}^2 \geq 0.30|\theta_\tau^2 = 0.09) = 0.002$. In addition, $\pi_{\mathrm{BJ}}(0.09) = P_{\mathrm{BJ}}(\hat{\theta}^2 \geq 0.90|\theta_\tau^2 = 0.09) = P_{\mathrm{BJ}}(\hat{\theta}^2 \geq 1.00|\theta_\tau^2 = 0.09) = \pi_{\mathrm{NF}}(0.09) = 0.002$.

	The default SSR under ($H_0 : \theta^2 = \theta_\tau^2$)									
Critical value	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
Panel A: Noncentral χ^2-statistic										
0.05	0.024	0.505	0.874	0.982	0.999	1.000	1.000	1.000	1.000	1.000
0.10	0.000	0.101	0.594	0.913	0.991	0.999	1.000	1.000	1.000	1.000
0.20	0.000	0.000	0.080	0.514	0.883	0.987	0.999	1.000	1.000	1.000
0.30	0.000	0.000	**0.002**	0.128	0.568	0.902	0.990	1.000	1.000	1.000
0.40	0.000	0.000	0.000	0.015	0.230	0.677	0.937	0.995	1.000	1.000
0.50	0.000	0.000	0.000	0.001	0.060	0.384	0.794	0.968	0.998	1.000
0.60	0.000	0.000	0.000	0.000	0.011	0.164	0.568	0.889	0.987	0.999
0.70	0.000	0.000	0.000	0.000	0.001	0.055	0.337	0.741	0.950	0.996
0.80	0.000	0.000	0.000	0.000	0.000	0.015	0.168	0.547	0.870	0.982
0.90	0.000	0.000	0.000	0.000	0.000	0.003	0.072	0.356	0.740	0.946
1.00	0.000	0.000	0.000	0.000	0.000	0.001	0.027	0.205	0.577	0.876
1.10	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.107	0.412	0.770
1.20	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.051	0.271	0.637
1.30	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.023	0.165	0.494
1.40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.094	0.360
Panel B: Britten-Jones F-statistic										
0.05	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.20	0.961	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.30	0.671	0.812	0.955	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.40	0.292	0.430	0.682	0.930	0.999	1.000	1.000	1.000	1.000	1.000
0.50	0.090	0.155	0.324	0.646	0.943	1.000	1.000	1.000	1.000	1.000
0.60	0.023	0.044	0.113	0.314	0.700	0.975	1.000	1.000	1.000	1.000
0.70	0.005	0.011	0.032	0.116	0.379	0.812	0.995	1.000	1.000	1.000
0.80	0.001	0.002	0.008	0.036	0.159	0.520	0.928	1.000	1.000	1.000
0.90	0.000	0.001	**0.002**	0.010	0.056	0.260	0.719	0.988	1.000	1.000
1.00	0.000	0.000	0.000	0.003	0.018	0.109	0.445	0.902	1.000	1.000
1.10	0.000	0.000	0.000	0.001	0.005	0.040	0.227	0.698	0.987	1.000
1.20	0.000	0.000	0.000	0.000	0.002	0.014	0.101	0.448	0.912	1.000
1.30	0.000	0.000	0.000	0.000	0.000	0.004	0.041	0.246	0.736	0.992
1.40	0.000	0.000	0.000	0.000	0.000	0.001	0.015	0.119	0.510	0.943
Panel C: Noncentral F-statistic										
0.05	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.20	0.955	0.979	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.30	0.665	0.784	0.906	0.976	0.997	1.000	1.000	1.000	1.000	1.000
0.40	0.297	0.436	0.648	0.850	0.962	0.995	1.000	1.000	1.000	1.000
0.50	0.095	0.176	0.351	0.609	0.843	0.963	0.995	1.000	1.000	1.000
0.60	0.025	0.057	0.152	0.356	0.639	0.871	0.973	0.997	1.000	1.000
0.70	0.006	0.016	0.056	0.176	0.416	0.710	0.911	0.985	0.999	1.000
0.80	0.001	0.004	0.019	0.077	0.238	0.517	0.796	0.949	0.993	0.999
0.90	0.000	0.001	0.006	0.031	0.122	0.338	0.639	0.875	0.975	0.997

(continued)

TABLE 3. Continued

	The default SSR under ($H_0 : \theta^2 = \theta_\tau^2$)									
Critical value	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
1.00	0.000	0.000	**0.002**	0.011	0.058	0.201	0.472	0.762	0.935	0.990
1.10	0.000	0.000	0.001	0.004	0.026	0.111	0.322	0.623	0.864	0.971
1.20	0.000	0.000	0.000	0.001	0.011	0.058	0.205	0.477	0.764	0.934
1.30	0.000	0.000	0.000	0.000	0.004	0.029	0.124	0.344	0.643	0.873
1.40	0.000	0.000	0.000	0.000	0.002	0.014	0.071	0.236	0.515	0.789

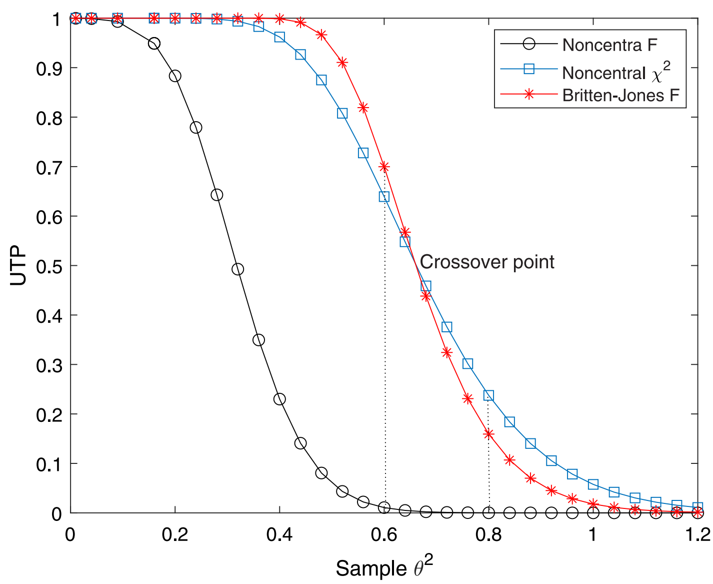

FIGURE 1. The UTPs of the sample SSR under various statistics. This graph depicts the UTP curves of the sample SSR for the case that is generated by the χ^2-statistic versus when the UTP curves of the sample SSR is plotted under the central and noncentral F distributions, keeping the parameters $(T,\ N,\ \theta^2) = (120$ observations, 30 stocks, 0.25 SSR) to be the same in three distributions.

PROPOSITION 3: *For testing the hypothesis* (1), *the test statistic is given by:*

$$\chi^2 = \frac{T(1+\theta_\tau^2)^2}{\theta_\tau^2(1+\hat{\theta}^2)} \sim \chi^2_{T-N,T/\theta_\tau^2}. \quad \mathbf{(25)}$$

Moreover, given the (*critical value, sample size, portfolio size*) *is* (C, T, N), *the upper-tail probability* (*UTP*) *of the sample SSR under the null hypothesis is defined as:*

$$\pi_{\mathrm{NC}}(\theta_\tau^2) = P_{\mathrm{NC}}(\hat{\theta}^2 \geq C|\theta_\tau^2, T, N) = P\left(\chi^2_{T-N,T/\theta_\tau^2} \leq \frac{T(1+\theta_\tau^2)^2}{\theta_\tau^2(1+C)}\right). \quad \mathbf{(26)}$$

Note that the small values of the χ^2-statistic (values of $\hat{\theta}^2$ larger than θ_τ^2 by a suitable amount) favor rejection of the mean-variance efficiency between the MVEP and the portfolio ω_τ. It is not easy to compare the probability density functions of various statistics expressed in the forms of totally different functions. To characterize the impact of three

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

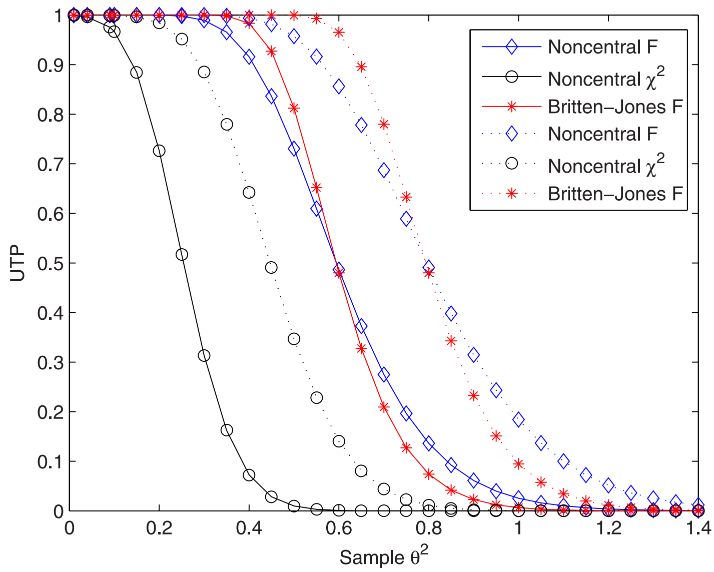

FIGURE **2.** The shifted UTPs of the sample SSR with respect to the SSR. This graph depicts that the UTP curves of the sample SSR generated using $(T, N, \theta^2) = (120, 30, 0.36)$ move toward the right side of the UTPs based on $(T, N, \theta^2) = (120, 30, 0.25)$. Keeping the sample size and the portfolio size to be the same in both cases, the solid curves indicate the UTPs of various statistics based on the SSR $\theta^2 = 0.25$. The dotted curves indicate the UTPs of various statistics using the SSR $\theta^2 = 0.36$ for the cases where (T, N) is fixed to (120, 30).

alternative distributions on the sample SSR, the UTP curve provides another way to shape the sample SSR under different statistics. Therefore, we also compute the UTPs of the sample SSR using the noncentral F-statistic and the Britten-Jones F-statistic. For the case of the noncentral F-statistic (5), the sample SSR under the null hypothesis (1) has the following UTP:

$$\pi_{\mathrm{NF}}(\theta_\tau^2) = P_{\mathrm{NF}}(\hat{\theta}^2 \geq C|\theta_\tau^2, T, N) = P\left(F_{N,T-N,T\theta_\tau^2} \geq \frac{C(T-N)}{N}\right). \tag{27}$$

Alternatively, under the null hypothesis (1), the Britten-Jones F-statistic (12) has the following UTP for the sample SSR:

$$\pi_{\mathrm{BJ}}(\theta_\tau^2) = P_{\mathrm{BJ}}(\hat{\theta}^2 \geq C|\theta_\tau^2, T, N) = P\left(F_{N,T-N} \geq \frac{(T-N)(C-\theta_\tau^2)}{N(1+\theta_\tau^2)}\right). \tag{28}$$

For comparison, Table 2 summarizes the formulas that simultaneously compute the critical value for rejecting the upper-tailed MVET (1) for satisfying a particular UTP (or the significance level) under various distribution. As an example, Table 3 summarizes the UTPs of the sample SSR for the parameters $T = 120$, $N = 30$, and $\theta_\tau^2 \in \Theta_\tau^2$. To better illustrate the impact of (T, N, θ_τ^2) on the UTPs using Eqs. (26)–(28), Figure 1 depicts the UTPs in relation to the sample SSR for the noncentral χ^2-statistic, the noncentral F-statistic, and the Britten-Jones F-statistic. In Figure 1, we vary the sample SSR on the horizontal axis and report the corresponding UTP on the vertical axis given $\theta^2 = 0.25$. Each curve presents the value of the sample SSR and the corresponding UTP. The crossover point is the intersection

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

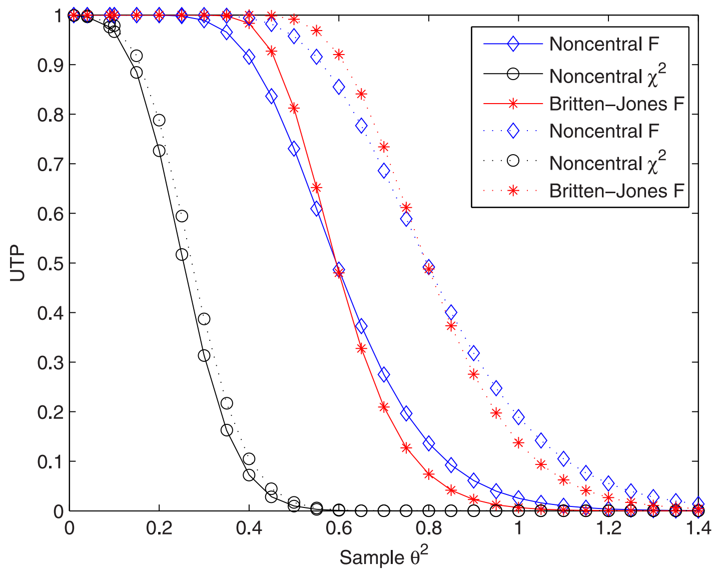

FIGURE 3. The shifted UTPs of the sample SSR with respect to the portfolio size. This graph depicts that the UTP curves of the sample SSR generated using $(T, N, \theta^2) = (120, 40, 0.25)$ move toward the right side of the UTPs based on $(T, N, \theta^2) = (120, 30, 0.25)$. Keeping the sample size and the SSR to be the same in both cases, the solid curves indicate the UTPs of various statistics based on the portfolio size N = 30. The dotted curves indicate the UTPs of various statistics using the portfolio size N = 40 for the cases where (T, θ^2) is fixed to (120, 0.25).

of the UTPs between the noncentral F-statistic and the Britten-Jones F-statistic. We see that the crossover point is near the point $(\hat{\theta}^2, \pi) = (0.67, 0.51)$ in the case of (120, 30, 0.25). We also observe that the horizontal coordinate of the crossover point is quite near to the average of $E_{\mathrm{NF}}(\hat{\theta}^2) = 0.682$ and $E_{\mathrm{BJ}}(\hat{\theta}^2) = 0.676$. If the critical value is at the left (or right) side of the crossover point such as $C = 0.6$ (or $C = 0.8$), Figure 1 and Table 3 indicate that $\pi_{\mathrm{NC}}(0.25) < \pi_{\mathrm{NF}}(0.25) < \pi_{\mathrm{BJ}}(0.25)$ (or $\pi_{\mathrm{NC}}(0.25) < \pi_{\mathrm{BJ}}(0.25) < \pi_{\mathrm{NF}}(0.25)$).

Figure 1 also shows how much the statistics shift the significance levels for rejecting the MVET. Taking the particular SSR $\theta^2 = 0.09$ as an example, Column 4 of Table 3 reports that the UTP is 0.002 for $\pi_{\mathrm{NC}}(0.09) = P_{\mathrm{NC}}(\hat{\theta}^2 \geq 0.30|\theta^2_\tau = 0.09)$, $\pi_{\mathrm{BJ}}(0.09) = P_{\mathrm{BJ}}(\hat{\theta}^2 \geq 0.90|\theta^2_\tau = 0.09)$, and $\pi_{\mathrm{NF}}(0.09) = P_{\mathrm{BJ}}(\hat{\theta}^2 \geq 1.00|\theta^2_\tau = 0.09)$. This implies that the noncentral and central F statistics will be more conservative in regard to rejecting the MVET at the 5% significance level Compared to the noncentral χ^2-statistic.

To illustrate the impact of sample size and portfolio size on the critical values using various statistics, we plot the UTPs relative to sample size, portfolio size, and default SSR. Figure 2 depicts the shifted UTPs with respect to the values of θ^2_τ, and we plot the UTPs for the cases where (T, N) is fixed to (120, 30), and the value of θ^2_τ changes from 0.25 to 0.36. We observe that the UTPs using $\theta^2_\tau = 0.36$ move toward the right side of the UTPs using $\theta^2_\tau = 0.25$. This is evident because all the expectations $E_{\mathrm{BJ}}(\hat{\theta}^2)$, $E_{\mathrm{NF}}(\hat{\theta}^2)$, and $E_{\mathrm{NC}}(\hat{\theta}^2)$ increase as the value of θ^2_τ increases from 0.25 to 0.36. Figure 3 depicts the shifted UTPs with respect to the number of assets, and we plot the UTPs for the case where (N, θ^2_τ) is fixed to (120, 0.25), but where the portfolio size changes from 30 to 40. We observe that the UTPs using $N = 40$ are located to the right of the UTPs using $N = 30$. When the portfolio size is small and the sample size is fixed, e.g., $N = 30$, the degrees of freedom,

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

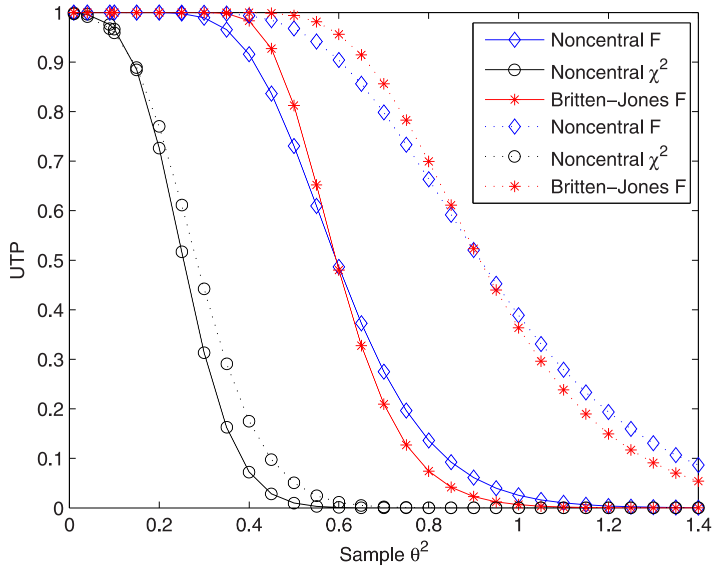

FIGURE 4. The shifted UTPs of the sample SSR with respect to the sample size. This graph depicts that the UTP curves of the sample SSR generated using $(T, N, \theta^2) = (60, 30, 0.25)$ move toward the right side of the UTPs based on $(T, N, \theta^2) = (120, 30, 0.25)$. Keeping the portfolio size and the SSR to be the same in both cases, the solid curves indicate the UTPs of various statistics based on the sample size $T = 120$. The dotted curves indicate the UTPs of various statistics using the sample size $T = 60$ for the cases where (N, θ^2) is fixed to (30, 0.25).

$T - N$, become very large. Thus, the probability density function of the sample SSR is more positively skewed than the case of $N = 40$. Figure 4 depicts the shifted UTPs with respect to the number of observations, and we plot the UTPs for the case where (N, θ_τ^2) is fixed to (120, 0.25), but where the sample size changes from 60 to 120. We observe that the UTPs using $T = 120$ move toward the left side of the UTPs using $T = 60$. The reason is similar to that of Figure 3. When the sample size is large and the portfolio size is fixed, e.g., $T = 120$, the degrees of freedom, $T - N$, become very large.

3.3. Effective test compared to the F-statistics

Analyzing Figures 3 and 4 more closely reveals the local robustness of the noncentral χ^2-statistic. We observe that the UTP's larger change in the F-statistics increases with sample size (holding portfolio size and SSR constant) as well as with portfolio size (holding sample size and SSR constant). For a comprehensive examination of methods at the 5% significance level, we calculate six distributions for the sample SSR based on our parameters domain.[1] The most apparent advantage of using distributions of the sample SSR is that the investor is more directly interested in knowing whether the sample SSR is greater than the hypothesized SSR and is not indirectly interested in the sample SSR's UTP using different critical values in Table 3. For the sake of brevity, we do not include similar results using the 1% and 10% significance levels. We consider the case $\theta = 0.25$ and set up the following

[1] To prevent the tables of distributions from taking up space and affecting the complete comparison, these six tables are numbered Tables 5–10 and include them in the Appendix.

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

Table 4. Noncentral χ^2's performance relative to F-statistics. This table summarizes the cases where one approach outperforms the other methods. For the case of $T = 120$ in Panel A based on the critical values in Table 6, we observe that 94 cases of the noncentral χ^2-statistic are superior to the F-statistics for rejecting H_0. Conversely, only 6 cases of the Britten-Jones F-statistic can defeat the χ^2-statistic, and none of the noncentral F-statistics are better than the other two statistics. Similarly, for the case $N = 50$, summarized from Tables 5–10, we see that 57 cases of the noncentral χ^2-statistic outperform the other statistics. Only 3 cases of the Britten-Jones F-statistic are better than the noncentral χ^2-statistic. For the case $\theta = 0.16$ summarized in Tables 5–10, we see that 51 cases of the noncentral χ^2-statistic outperform the other statistics. Only 4 cases of Britten-Jones F-statistic are better than the noncentral χ^2-statistic. The numbers in parentheses denote the fraction of cases with superior performance to the other methods. The range is the difference between the maximum and the minimum of critical values given in the parameters (the sample size or portfolio size or SSR).

	Noncentral χ^2-statistic			Britten-Jones F-statistic			Noncentral F-statistic		
Attributes	Number	Ratio	Range	Number	Ratio	Range	Number	Ratio	Range
Panel A: Categorized using the sample size (T)									
60	49	(98%)	4.250	1	(2%)	27.142	0	(0%)	28.450
120	94	(94%)	3.563	6	(6%)	19.880	0	(0%)	20.575
240	88	(88%)	1.934	12	(12%)	2.839	0	(0%)	3.051
360	84	(84%)	1.596	16	(16%)	1.939	0	(0%)	2.100
480	80	(80%)	1.448	20	(20%)	1.626	0	(0%)	1.763
600	76	(76%)	1.363	24	(24%)	1.468	0	(0%)	1.591
Panel B: Categorized using the portfolio size (N)									
10	24	(40%)	1.929	36	(60%)	1.768	0	(0%)	2.250
20	44	(73%)	2.299	16	(27%)	2.774	0	(0%)	3.272
30	46	(77%)	2.774	14	(23%)	4.593	0	(0%)	5.160
40	54	(90%)	3.404	6	(10%)	8.863	0	(0%)	9.595
50	57	(95%)	4.281	3	(5%)	27.235	0	(0%)	28.552
60	49	(98%)	2.371	1	(2%)	3.908	0	(0%)	4.260
70	50	(100%)	2.610	0	(0%)	5.177	0	(0%)	5.561
80	50	(100%)	2.885	0	(0%)	7.219	0	(0%)	7.654
90	50	(100%)	3.205	0	(0%)	10.977	0	(0%)	11.499
100	50	(100%)	3.580	0	(0%)	19.798	0	(0%)	20.500
Panel C: Categorized using the SSR (θ_τ^2)									
0.01	55	(100%)	0.039	0	(0%)	13.285	0	(0%)	13.296
0.04	55	(100%)	0.104	0	(0%)	13.680	0	(0%)	13.726
0.09	53	(96%)	0.201	2	(4%)	14.338	0	(0%)	14.449
0.16	51	(93%)	0.343	4	(7%)	15.259	0	(0%)	15.459
0.25	49	(89%)	0.544	6	(11%)	16.443	0	(0%)	16.758
0.36	45	(82%)	0.819	10	(18%)	17.889	0	(0%)	18.339
0.49	42	(76%)	1.190	13	(24%)	19.600	0	(0%)	20.203
0.64	41	(75%)	1.679	14	(25%)	21.573	0	(0%)	22.348
0.81	39	(71%)	2.310	16	(29%)	23.809	0	(0%)	24.773
1.00	33	(60%)	3.109	22	(40%)	26.308	0	(0%)	27.478

TABLE 5. Distributions of the sample SSR using 60 monthly returns at the 5% significance level. This table reports the critical value of rejecting $H_0 : \theta^2 = \theta_\tau^2$ against $H_1 : \theta^2 > \theta_\tau^2$ at the 5% significance level. Taking the case $(N, \theta_\tau^2) = (30, 0.25)$ at the $UTP = 0.05$ into consideration, this means that $\pi_{\text{NC}}(0.25) = P_{\text{NC}}(\hat{\theta}^2 \geq 0.711 | \theta_\tau^2 = 0.25) = 0.05$. In addition, $P_{\text{BJ}}(\hat{\theta}^2 \geq 2.551 | \theta_\tau^2 = 0.25) = P_{\text{BJ}}(\hat{\theta}^2 \geq 2.728 | \theta_\tau^2 = 0.25) = 0.05$. The boldface number indicates that the statistic is the best among three statistics. The more the boldface area is, the better the statistic is. Thus, in this table, the noncentral χ^2 outperforms the F-statistics except for the seven cases of $N = 5, \theta_\tau^2 = 0.25, 0.36, 0.49, 0.64, 0.81, 1.00$, and $(N, \theta_\tau^2) = (10, 1.00)$.

	The default value of SSR under the hypothesis $H_0 : \theta^2 = \theta_\tau^2$									
Size (N)	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
Panel A: Noncentral χ^2-statistic										
5	**0.055**	**0.136**	**0.244**	**0.380**	0.544	0.737	0.958	1.209	1.488	1.796
10	**0.056**	**0.140**	**0.254**	**0.398**	**0.575**	**0.784**	**1.026**	**1.302**	**1.610**	1.953
15	**0.057**	**0.144**	**0.263**	**0.417**	**0.607**	**0.834**	**1.099**	**1.403**	**1.745**	**2.127**
20	**0.058**	**0.148**	**0.273**	**0.436**	**0.640**	**0.886**	**1.177**	**1.513**	**1.895**	**2.323**
25	**0.059**	**0.152**	**0.283**	**0.456**	**0.675**	**0.942**	**1.261**	**1.634**	**2.061**	**2.545**
30	**0.060**	**0.156**	**0.293**	**0.476**	**0.711**	**1.001**	**1.352**	**1.766**	**2.248**	**2.798**
35	**0.060**	**0.160**	**0.303**	**0.497**	**0.749**	**1.064**	**1.450**	**1.913**	**2.458**	**3.089**
40	**0.061**	**0.164**	**0.314**	**0.519**	**0.788**	**1.131**	**1.557**	**2.076**	**2.697**	**3.428**
45	**0.062**	**0.168**	**0.325**	**0.541**	**0.829**	**1.202**	**1.673**	**2.258**	**2.971**	**3.828**
50	**0.063**	**0.172**	**0.335**	**0.564**	**0.873**	**1.278**	**1.801**	**2.463**	**3.289**	**4.305**
Panel B: Britten-Jones F-statistic										
5	0.229	0.265	0.326	0.411	**0.521**	**0.655**	**0.813**	**0.995**	**1.202**	**1.433**
10	0.419	0.461	0.532	0.630	0.757	0.911	1.094	1.305	1.543	**1.810**
15	0.648	0.697	0.778	0.893	1.040	1.219	1.431	1.676	1.953	2.263
20	0.939	0.996	1.092	1.227	1.399	1.610	1.860	2.148	2.474	2.839
25	1.326	1.395	1.510	1.671	1.878	2.132	2.431	2.777	3.168	3.606
30	1.869	1.955	2.097	2.295	2.551	2.864	3.233	3.659	4.142	4.682
35	2.686	2.795	2.978	3.233	3.562	3.963	4.438	4.985	5.605	6.299
40	4.038	4.187	4.437	4.786	5.235	5.783	6.432	7.180	8.028	8.975
45	6.645	6.872	7.250	7.780	8.462	9.294	10.278	11.414	12.700	14.138
50	13.327	13.753	14.462	15.455	16.732	18.292	20.137	22.264	24.676	27.371
Panel C: Noncentral F-statistic										
5	0.242	0.311	0.414	0.546	0.707	0.896	1.113	1.359	1.634	1.937
10	0.429	0.498	0.607	0.750	0.927	1.136	1.378	1.652	1.959	2.298
15	0.657	0.730	0.848	1.007	1.204	1.439	1.712	2.021	2.368	2.753
20	0.947	1.028	1.160	1.339	1.563	1.832	2.145	2.500	2.900	3.342
25	1.334	1.426	1.578	1.785	2.047	2.361	2.727	3.145	3.614	4.135
30	1.878	1.987	2.167	2.415	2.728	3.107	3.548	4.053	4.621	5.253
35	2.695	2.830	3.053	3.362	3.755	4.229	4.785	5.420	6.137	6.933
40	4.047	4.226	4.521	4.932	5.454	6.087	6.830	7.682	8.642	9.711
45	6.657	6.919	7.353	7.957	8.729	9.667	10.769	12.035	13.464	15.055
50	13.344	13.817	14.605	15.702	17.108	18.819	20.834	23.151	25.771	28.692

right-tailed hypothesis (MVET) at the 5% significance level (in terms of 5% UTP) as an illustration.

$$H_0 : \theta^2 = 0.25 \quad \text{against} \quad H_1 : \theta^2 > 0.25.$$

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

TABLE 6. Distributions of the sample SSR using 120 monthly returns at the 5% significance level.

	The default value of SSR under the hypothesis $H_0 : \theta^2 = \theta_\tau^2$									
Size (N)	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
Panel A: Noncentral χ^2-statistic										
10	**0.042**	**0.107**	**0.198**	**0.313**	**0.455**	0.622	0.815	1.034	1.280	1.552
20	**0.043**	**0.111**	**0.207**	**0.330**	**0.482**	**0.664**	**0.874**	**1.115**	**1.386**	1.686
30	**0.044**	**0.115**	**0.216**	**0.347**	**0.511**	**0.708**	**0.938**	**1.202**	**1.501**	**1.834**
40	**0.044**	**0.119**	**0.225**	**0.365**	**0.541**	**0.754**	**1.006**	**1.297**	**1.628**	**2.000**
50	**0.045**	**0.123**	**0.234**	**0.383**	**0.572**	**0.803**	**1.079**	**1.400**	**1.769**	**2.186**
60	**0.046**	**0.126**	**0.244**	**0.402**	**0.605**	**0.855**	**1.157**	**1.513**	**1.925**	**2.396**
70	**0.047**	**0.130**	**0.253**	**0.421**	**0.638**	**0.910**	**1.242**	**1.637**	**2.100**	**2.635**
80	**0.048**	**0.134**	**0.263**	**0.441**	**0.674**	**0.968**	**1.333**	**1.774**	**2.297**	**2.910**
90	**0.049**	**0.138**	**0.273**	**0.461**	**0.710**	**1.030**	**1.432**	**1.925**	**2.521**	**3.230**
100	**0.050**	**0.142**	**0.283**	**0.482**	**0.749**	**1.096**	**1.539**	**2.094**	**2.777**	**3.605**
Panel B: Britten-Jones F-statistic										
10	0.186	0.221	0.280	0.362	0.468	**0.597**	**0.750**	**0.926**	**1.126**	**1.349**
20	0.349	0.389	0.455	0.549	0.669	0.816	0.990	1.190	1.417	**1.671**
30	0.544	0.590	0.666	0.773	0.911	1.079	1.278	1.507	1.767	2.057
40	0.790	0.843	0.932	1.056	1.216	1.411	1.641	1.907	2.208	2.545
50	1.114	1.177	1.281	1.428	1.616	1.846	2.118	2.432	2.788	3.186
60	1.560	1.636	1.762	1.940	2.168	2.447	2.776	3.156	3.587	4.069
70	2.213	2.309	2.468	2.690	2.977	3.327	3.740	4.217	4.758	5.363
80	3.257	3.384	3.595	3.890	4.269	4.733	5.281	5.913	6.630	7.431
90	5.169	5.352	5.658	6.085	6.635	7.307	8.101	9.017	10.055	11.216
100	9.638	9.954	10.481	11.218	12.166	13.325	14.694	16.274	18.064	20.066
Panel C: Noncentral F-statistic										
10	0.195	0.251	0.338	0.451	0.591	0.756	0.948	1.165	1.409	1.679
20	0.355	0.413	0.505	0.628	0.781	0.963	1.175	1.417	1.688	1.988
30	0.550	0.612	0.712	0.847	1.018	1.222	1.460	1.731	2.035	2.373
40	0.796	0.864	0.976	1.129	1.321	1.553	1.824	2.133	2.481	2.867
50	1.119	1.197	1.325	1.501	1.724	1.992	2.307	2.666	3.071	3.520
60	1.565	1.656	1.807	2.015	2.280	2.600	2.974	3.403	3.886	4.424
70	2.219	2.330	2.515	2.770	3.096	3.491	3.954	4.485	5.084	5.750
80	3.264	3.408	3.646	3.978	4.402	4.916	5.521	6.214	6.997	7.869
90	5.176	5.380	5.718	6.189	6.791	7.524	8.385	9.376	10.495	11.741
100	9.647	9.990	10.559	11.354	12.371	13.611	15.071	16.751	18.651	20.770

Table 6 reports the critical values for the case $(T, N) = (120, 30)$ at the 5% significance level as follows:

$$C_{\text{NC}} = 0.511 < C_{\text{BJ}} = 0.911 < C_{\text{NF}} = 1.018$$

For the above upper tail test of the SSR, it is evident that not only is the noncentral χ^2 best when compared to F-statistics, but both the Britten-Jones' F, and noncentral F are conservative in rejecting the MVET at the same significance level. However, the greatest limitation with the F-statistics is the smaller sample size and the larger portfolio size. For example, given the 0.05 UTP for the case $(T, N) = (60, 50)$ in Table 5, it is notable that the noncentral χ^2, Britten-Jones' F, and noncentral F will require $\hat{\theta}^2 > 0.873$, $\hat{\theta}^2 > 16.732$, and $\hat{\theta}^2 > 17.108$ to reject the null hypothesis $H_0 : \theta^2 = 0.25$, respectively. Similarly, for the case $(T, N) = (120, 80)$ in Table 6, note that the noncentral χ^2, Britten-Jones' F, and

Table 7. Distributions of the sample SSR using 240 monthly returns at the 5% significance level.

	The default value of SSR under the hypothesis $H_0 : \theta^2 = \theta_\tau^2$									
Size (N)	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
Panel A: Noncentral χ^2-statistic										
10	**0.032**	**0.086**	**0.163**	0.262	0.384	0.529	0.698	0.890	1.105	1.344
20	**0.033**	**0.088**	**0.167**	**0.270**	**0.397**	**0.548**	0.724	0.924	1.150	1.400
30	**0.033**	**0.090**	**0.171**	**0.278**	**0.409**	**0.567**	**0.750**	**0.960**	**1.196**	1.459
40	**0.033**	**0.092**	**0.175**	**0.286**	**0.422**	**0.586**	**0.778**	**0.998**	**1.245**	**1.520**
50	**0.034**	**0.093**	**0.180**	**0.294**	**0.436**	**0.607**	**0.807**	**1.036**	**1.296**	**1.585**
60	**0.034**	**0.095**	**0.184**	**0.302**	**0.449**	**0.627**	**0.836**	**1.077**	**1.349**	**1.653**
70	**0.035**	**0.097**	**0.188**	**0.310**	**0.463**	**0.648**	**0.867**	**1.119**	**1.405**	**1.725**
80	**0.035**	**0.099**	**0.193**	**0.318**	**0.477**	**0.670**	**0.898**	**1.162**	**1.463**	**1.801**
90	**0.036**	**0.101**	**0.197**	**0.327**	**0.491**	**0.692**	**0.930**	**1.208**	**1.524**	**1.881**
100	**0.036**	**0.103**	**0.202**	**0.335**	**0.506**	**0.715**	**0.964**	**1.255**	**1.589**	**1.966**
Panel B: Britten-Jones F-statistic										
10	0.092	0.125	0.179	**0.254**	**0.352**	**0.471**	**0.611**	**0.773**	**0.957**	**1.163**
20	0.159	0.193	0.250	0.331	0.434	0.560	**0.709**	**0.881**	**1.076**	**1.294**
30	0.228	0.265	0.326	0.411	0.520	0.654	0.812	0.995	1.201	**1.432**
40	0.304	0.343	0.407	0.498	0.614	0.756	0.924	1.117	1.337	1.582
50	0.387	0.428	0.497	0.593	0.716	0.867	1.046	1.252	1.485	1.746
60	0.479	0.523	0.596	0.699	0.830	0.991	1.182	1.401	1.650	1.929
70	0.582	0.629	0.707	0.817	0.958	1.130	1.334	1.569	1.835	2.133
80	0.699	0.749	0.833	0.951	1.102	1.287	1.506	1.758	2.044	2.364
90	0.832	0.886	0.977	1.104	1.267	1.466	1.702	1.974	2.283	2.627
100	0.985	1.044	1.142	1.280	1.457	1.673	1.929	2.223	2.557	2.931
Panel C: Noncentral F-statistic										
10	0.100	0.149	0.224	0.322	0.443	0.587	0.754	0.945	1.159	1.397
20	0.164	0.213	0.289	0.391	0.517	0.668	0.843	1.043	1.267	1.517
30	0.233	0.283	0.361	0.467	0.599	0.757	0.940	1.150	1.387	1.649
40	0.308	0.359	0.440	0.551	0.689	0.855	1.049	1.270	1.518	1.795
50	0.391	0.443	0.528	0.644	0.790	0.965	1.169	1.402	1.665	1.957
60	0.483	0.537	0.626	0.748	0.902	1.087	1.304	1.551	1.829	2.139
70	0.586	0.643	0.737	0.866	1.029	1.225	1.455	1.718	2.015	2.344
80	0.702	0.763	0.862	0.999	1.173	1.382	1.628	1.908	2.225	2.577
90	0.835	0.900	1.006	1.152	1.338	1.562	1.825	2.126	2.465	2.843
100	0.989	1.058	1.171	1.328	1.528	1.770	2.053	2.377	2.743	3.151

noncentral F will require $\hat{\theta}^2 > 0.674$, $\hat{\theta}^2 > 4.269$, and $\hat{\theta}^2 > 4.402$, respectively, to reject the null hypothesis $H_0 : \theta^2 = 0.25$. Compared to the noncentral χ^2 approach, we observe that the critical values of the sample SSR using the F-statistic methods are overly conservative and extremely unrealistic for rejecting the null hypothesis $H_0 : \theta^2 = 0.25$ when checking the cases using the smaller sample size and the larger portfolio size.

Collectively, these six distributions of the sample SSR summarized in Table 4 that presents the noncentral χ^2's performance relative to F-statistics at the 5% significance level. We have several observations. First, the boldface areas in Tables 5–10 and the winning ratios in Table 4 show that both the noncentral χ^2-statistic and the Britten-Jones F-statistic are definitely superior to the noncentral F-statistic within our parameters domain. Second, holding constant the portfolio size and the SSR, the critical value decreases as the sample

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

TABLE 8. Distributions of the sample SSR using 360 monthly returns at the 5% significance level.

	The default value of SSR under the hypothesis $H_0 : \theta^2 = \theta_\tau^2$									
Size (N)	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
Panel A: Noncentral χ^2-statistic										
10	**0.028**	**0.077**	**0.148**	0.241	0.356	0.493	0.652	0.834	1.039	1.266
20	**0.028**	**0.078**	**0.151**	**0.246**	**0.364**	0.505	0.669	0.856	1.067	1.301
30	**0.028**	**0.080**	**0.154**	**0.251**	**0.372**	**0.517**	0.686	0.879	1.096	1.337
40	**0.029**	**0.081**	**0.156**	**0.256**	**0.380**	**0.529**	**0.703**	**0.901**	**1.125**	1.374
50	**0.029**	**0.082**	**0.159**	**0.261**	**0.389**	**0.542**	**0.720**	**0.925**	**1.156**	**1.412**
60	**0.029**	**0.083**	**0.162**	**0.267**	**0.397**	**0.554**	**0.738**	**0.949**	**1.187**	**1.452**
70	**0.030**	**0.084**	**0.165**	**0.272**	**0.406**	**0.567**	**0.756**	**0.973**	**1.219**	**1.493**
80	**0.030**	**0.086**	**0.168**	**0.277**	**0.414**	**0.580**	**0.775**	**0.999**	**1.252**	**1.535**
90	**0.030**	**0.087**	**0.170**	**0.282**	**0.423**	**0.593**	**0.794**	**1.025**	**1.286**	**1.579**
100	**0.031**	**0.088**	**0.173**	**0.288**	**0.432**	**0.607**	**0.813**	**1.051**	**1.321**	**1.624**
Panel B: Britten-Jones F-statistic										
10	0.064	0.095	0.148	**0.222**	**0.316**	**0.432**	**0.569**	**0.727**	**0.906**	**1.106**
20	0.105	0.138	0.193	0.269	0.368	**0.488**	**0.630**	**0.795**	**0.981**	**1.188**
30	0.147	0.181	0.238	0.318	0.420	0.545	**0.692**	**0.863**	**1.056**	**1.272**
40	0.191	0.226	0.285	0.368	0.474	0.603	0.757	0.934	1.134	**1.358**
50	0.237	0.273	0.335	0.420	0.531	0.665	0.825	1.008	1.216	1.449
60	0.285	0.324	0.387	0.476	0.591	0.731	0.896	1.087	1.304	1.545
70	0.337	0.377	0.443	0.536	0.655	0.801	0.973	1.171	1.396	1.648
80	0.393	0.434	0.503	0.599	0.724	0.875	1.055	1.261	1.496	1.758
90	0.452	0.495	0.567	0.668	0.797	0.955	1.142	1.358	1.602	1.876
100	0.516	0.561	0.636	0.742	0.877	1.042	1.237	1.462	1.717	2.003
Panel C: Noncentral F-statistic										
10	0.071	0.117	0.187	0.279	0.393	0.529	0.688	0.869	1.073	1.299
20	0.111	0.156	0.227	0.321	0.438	0.578	0.742	0.929	1.139	1.372
30	0.152	0.197	0.269	0.366	0.486	0.631	0.799	0.992	1.209	1.450
40	0.195	0.241	0.314	0.413	0.538	0.687	0.861	1.060	1.283	1.532
50	0.241	0.287	0.362	0.464	0.592	0.746	0.926	1.132	1.363	1.621
60	0.289	0.337	0.414	0.519	0.651	0.810	0.996	1.209	1.449	1.715
70	0.341	0.389	0.469	0.577	0.714	0.879	1.071	1.292	1.540	1.817
80	0.396	0.446	0.528	0.640	0.782	0.952	1.152	1.381	1.639	1.926
90	0.455	0.507	0.591	0.708	0.855	1.032	1.239	1.477	1.745	2.043
100	0.519	0.573	0.660	0.781	0.933	1.118	1.334	1.581	1.860	2.171

size increases for all the statistics considered. Similarly, holding constant the sample size and the SSR, the critical value increases as the portfolio size increases. Holding constant the sample and portfolio sizes, the larger the SSR, the greater the critical value. Third, another comparison of advantages and disadvantages between methods can be observed through the critical value's range. The range is the difference in critical values between the maximum and the minimum relative to the sample size, the portfolio size, and the SSR based on different tables. For example, the noncentral χ^2's range (2.371) of $N = 60$ in Panel A of Table 4 is computed as the maximum (2.396) from Panel A in Table 6 deducting the minimum (0.025) from Panel A in Table 10. Table 4 indicates that the noncentral χ^2's ranges are from 1.363 to 4.250, from 1.929 to 3.580, from 0.039 to 3.109 using T, N, and θ_τ^2, respectively. Table 4 similarly reports that the Britten-Jones F's ranges are between 1.468

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

Table 9. Distributions of the sample SSR using 480 monthly returns at the 5% significance level.

	The default value of SSR under the hypothesis $H_0 : \theta^2 = \theta_\tau^2$									
Size (N)	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
Panel A: Noncentral χ^2-statistic										
10	**0.025**	**0.072**	0.140	0.229	0.340	0.473	0.627	0.804	1.002	1.223
20	**0.026**	**0.073**	**0.142**	**0.233**	0.346	0.481	0.639	0.820	1.023	1.249
30	**0.026**	**0.074**	**0.144**	**0.236**	**0.352**	**0.490**	0.651	0.836	1.043	1.274
40	**0.026**	**0.075**	**0.146**	**0.240**	**0.358**	**0.499**	**0.664**	**0.852**	1.064	1.301
50	**0.026**	**0.075**	**0.148**	**0.244**	**0.364**	**0.508**	**0.676**	**0.869**	**1.086**	1.328
60	**0.027**	**0.076**	**0.150**	**0.248**	**0.370**	**0.517**	**0.689**	**0.886**	**1.108**	**1.355**
70	**0.027**	**0.077**	**0.152**	**0.252**	**0.376**	**0.526**	**0.702**	**0.903**	**1.130**	**1.384**
80	**0.027**	**0.078**	**0.154**	**0.255**	**0.382**	**0.535**	**0.715**	**0.921**	**1.153**	**1.413**
90	**0.027**	**0.079**	**0.156**	**0.259**	**0.389**	**0.545**	**0.728**	**0.939**	**1.177**	**1.442**
100	**0.027**	**0.080**	**0.158**	**0.263**	**0.395**	**0.554**	**0.742**	**0.957**	**1.201**	**1.473**
Panel B: Britten-Jones F-statistic										
10	0.050	0.081	**0.133**	**0.206**	**0.299**	**0.414**	**0.549**	**0.705**	**0.881**	**1.079**
20	0.080	0.112	0.166	0.240	**0.337**	**0.454**	**0.593**	**0.754**	**0.935**	**1.139**
30	0.110	0.143	0.198	0.275	0.374	0.495	**0.637**	**0.802**	**0.989**	**1.198**
40	0.141	0.174	0.231	0.310	0.412	0.536	0.683	0.852	**1.044**	**1.259**
50	0.172	0.207	0.265	0.346	0.451	0.578	0.729	0.903	1.101	**1.321**
60	0.205	0.241	0.300	0.384	0.491	0.622	0.777	0.956	1.159	1.386
70	0.239	0.276	0.337	0.423	0.533	0.668	0.828	1.012	1.220	1.454
80	0.275	0.313	0.376	0.464	0.578	0.716	0.881	1.070	1.284	1.524
90	0.312	0.351	0.416	0.507	0.624	0.767	0.936	1.131	1.352	1.598
100	0.352	0.392	0.459	0.552	0.673	0.820	0.994	1.195	1.422	1.676
Panel C: Noncentral F-statistic										
10	0.057	0.101	0.168	0.256	0.367	0.499	0.653	0.829	1.027	1.247
20	0.085	0.129	0.196	0.287	0.399	0.534	0.692	0.871	1.074	1.300
30	0.114	0.158	0.226	0.318	0.433	0.571	0.732	0.916	1.123	1.354
40	0.145	0.188	0.258	0.351	0.469	0.610	0.774	0.963	1.175	1.411
50	0.176	0.220	0.290	0.386	0.506	0.650	0.819	1.012	1.229	1.471
60	0.208	0.253	0.325	0.422	0.545	0.693	0.865	1.063	1.286	1.534
70	0.242	0.287	0.361	0.460	0.586	0.737	0.914	1.117	1.346	1.600
80	0.278	0.324	0.398	0.500	0.629	0.784	0.966	1.174	1.408	1.669
90	0.315	0.362	0.438	0.543	0.675	0.834	1.020	1.234	1.475	1.743
100	0.354	0.402	0.480	0.587	0.723	0.886	1.078	1.297	1.545	1.820

and 27.142, between 1.768 and 19.798 as well as between 13.285 and 26.308. By focusing on the parameters domain considered, we observe that the range of the noncentral χ^2-statistic is locally robust compared to the F-statistics.

Finally, another expression used to demonstrate the local robustness of the noncentral χ^2-statistic is graphing the tradeoff between sample size (X-axis), portfolio size (Y-axis), and critical values (Z-axis). The table-based Figure 2 shows how the critical values of the sample SSR are affected by the sample size and the portfolio size in rejecting the hypothesis $H_0 : \theta^2 = 0.25$ at the 5% significance level. As shown in Figure 5, the upper (steeper and colorful) plane indicates the critical values of the sample SSR using the noncentral F-statistic over the sample size and the portfolio size provided in Tables 5–10, and then we plot the result as a surface. The lower (smoother) plane indicates corresponding results based

TABLE 10. Distributions of the sample SSR using 600 monthly returns at the 5% significance level.

	The default value of SSR under the hypothesis $H_0 : \theta^2 = \theta_\tau^2$									
Size (N)	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
Panel A: Noncentral χ^2-statistic										
10	**0.024**	**0.068**	0.134	0.221	0.329	0.459	0.611	0.784	0.979	1.196
20	**0.024**	**0.069**	**0.136**	**0.224**	0.334	0.466	0.620	0.796	0.995	1.215
30	**0.024**	**0.070**	**0.137**	**0.227**	**0.339**	0.473	0.630	0.809	1.011	1.235
40	**0.024**	**0.071**	**0.139**	**0.230**	**0.343**	**0.480**	**0.639**	0.821	1.027	1.256
50	**0.024**	**0.071**	**0.141**	**0.233**	**0.348**	**0.487**	**0.649**	**0.834**	1.044	1.277
60	**0.025**	**0.072**	**0.142**	**0.236**	**0.353**	**0.494**	**0.659**	**0.848**	**1.061**	1.298
70	**0.025**	**0.073**	**0.144**	**0.239**	**0.358**	**0.501**	**0.669**	**0.861**	**1.078**	**1.319**
80	**0.025**	**0.073**	**0.146**	**0.242**	**0.363**	**0.508**	**0.679**	**0.874**	**1.095**	**1.341**
90	**0.025**	**0.074**	**0.147**	**0.245**	**0.368**	**0.515**	**0.689**	**0.888**	**1.113**	**1.364**
100	**0.025**	**0.075**	**0.149**	**0.248**	**0.372**	**0.523**	**0.699**	**0.902**	**1.131**	**1.387**
Panel B: Britten-Jones F-statistic										
10	0.042	0.073	**0.124**	**0.196**	**0.289**	**0.403**	**0.537**	**0.691**	**0.867**	**1.063**
20	0.065	0.097	0.150	0.224	**0.318**	**0.435**	**0.572**	**0.730**	**0.909**	**1.110**
30	0.089	0.121	0.175	0.250	0.347	**0.466**	**0.606**	**0.768**	**0.951**	**1.156**
40	0.112	0.145	0.200	0.277	0.376	0.498	0.641	**0.806**	**0.993**	**1.202**
50	0.136	0.170	0.226	0.305	0.406	0.530	0.676	0.845	**1.036**	**1.250**
60	0.161	0.195	0.253	0.333	0.437	0.563	0.712	0.885	1.080	**1.299**
70	0.186	0.221	0.280	0.362	0.468	0.597	0.750	0.926	1.126	1.349
80	0.212	0.248	0.308	0.392	0.500	0.632	0.789	0.969	1.173	1.401
90	0.239	0.276	0.338	0.424	0.534	0.669	0.829	1.013	1.221	1.454
100	0.268	0.305	0.368	0.456	0.569	0.707	0.870	1.058	1.272	1.510
Panel C: Noncentral F-statistic										
10	0.048	0.091	0.156	0.243	0.350	0.480	0.631	0.803	0.998	1.214
20	0.070	0.113	0.178	0.266	0.376	0.507	0.661	0.836	1.034	1.255
30	0.093	0.135	0.201	0.290	0.402	0.535	0.692	0.871	1.072	1.297
40	0.116	0.158	0.225	0.315	0.429	0.565	0.724	0.906	1.112	1.340
50	0.140	0.182	0.250	0.342	0.457	0.595	0.758	0.943	1.152	1.385
60	0.164	0.207	0.275	0.369	0.486	0.627	0.792	0.981	1.195	1.432
70	0.189	0.232	0.302	0.397	0.516	0.660	0.828	1.021	1.239	1.481
80	0.215	0.259	0.330	0.426	0.548	0.694	0.866	1.063	1.284	1.531
90	0.242	0.286	0.358	0.456	0.580	0.730	0.905	1.106	1.332	1.584
100	0.270	0.315	0.388	0.488	0.615	0.767	0.946	1.151	1.382	1.639

on the noncentral χ^2-statistic. Additionally, the difference between two surfaces steeply increases in the cases $T = 120, 240, 360, 480, 600$, and $N = 10, 20, 30, \ldots, 90, 100$. In these cases, the lower surface is smooth and robust, but the upper surface shows the abrupt UTP changes.

Overall, the noncentral χ^2-statistic reveals the consistency relative to portfolio size, sample size, and SSR compared to the F statistics. From the perspective of rejecting the upper-tailed hypothesis of SSR, Table 4 and Figure 5 show that our noncentral χ^2 approach is more competitive, significant, and locally robust when compared to the parameters domain suggested. In addition, note that the Britten-Jones' F-statistic could likely apply to the test combining the smaller portfolio and the larger SSR.

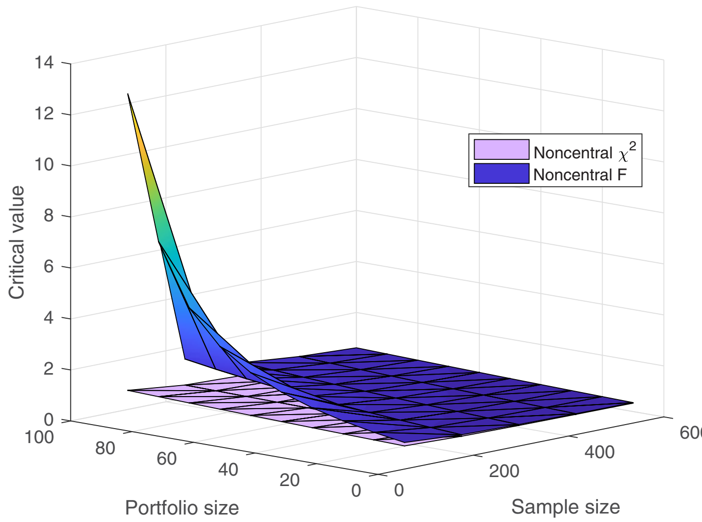

FIGURE **5.** Critical values with respect to the sample size and the portfolio size. This table-based graph plots how the critical values of the sample SSR are affected by the sample size and the portfolio size in rejecting the right-tailed hypothesis at the 5% significance level. The surfaces are generated using $\theta^2 = 0.25$, $T = 120, 240, \ldots, 600$, and $N = 10, 20, \ldots, 100$ under the noncentral F- and χ^2-statistics. The upper (colorful) plane indicates the critical values of the noncentral F-statistic over the sample and portfolio sizes provided and then plot the result as a surface. The lower plane indicates the critical values of the noncentral χ^2-statistic in terms of the same sample size and the same portfolio size.

4. CONCLUSION

The traditional MVET (either the noncentral F-distribution or the central F-statistic) based on the sample SSR employs the returns' normality, independence, and constant volatility. However, the rejection regions of the sample SSR using F tests are not only affected by the returns' normality but also the sample and portfolio sizes. Note that the upper-tailed MVET of a particular portfolio may result in entirely different conclusions under various normality assumptions. Thus, investors have the critical opportunity cost using the inappropriate sampling distribution of the sample SSR.

This paper proposes a new sampling distribution for identifying the upper-tailed MVET using the noncentral χ^2-statistic of the sample SSR. We compare our method against two popular methods (the noncentral F-distribution and the Britten-Jones F-statistic). Under the error's normality, we integrate the regression error with a nonzero mean and the arbitrage regression into a noncentral χ^2-statistic. In this framework, the evidence shows that the sampling distribution of the sample SSR is to the left of the sampling distributions of the noncentral F-distribution and the Britten-Jones F-statistic. Compared to these two benchmarks when using stronger returns' normality assumptions, the noncentral χ^2-statistic is more effective, significant, and locally robust in rejecting the upper-tailed MVET for a particular portfolio with the parameters domain employed as usual.

This study also finds that when using the sample SSR for the MVET, the following principles should be remarked:

- Both the noncentral χ^2-statistic and the Britten-Jones F-statistic are superior to the noncentral F-statistic within our parameters domain. Moreover, the noncentral

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

χ^2-statistic generally outperforms the Britten-Jones F-statistic except for in tests that probably combine a smaller portfolio size and a larger SSR.

- If returns have the multivariate normal distribution, the sample SSR theoretically follows the noncentral F-distribution. To employ the exact noncentral F-distribution for the MVET, we should first perform the returns' multivariate normality test.
- If the F-statistic in the arbitrage regression is used to implement the MVET, the regression error's normality (including the zero mean) should be tested. Note that the UTP's difference between the noncentral F-distribution and the central F-statistic is small.
- When the regression error's zero mean is not tenable; the noncentral χ^2-statistic of the sample SSR could be a locally robust test for implementing the MVET.

Acknowledgments

The author is thankful to the Editor and two anonymous reviewers for helpful suggestions which have improved the presentation in the paper. The author also gratefully acknowledges partial financial support from the National United University through project 108NUUPRJ-03.

References

1. Abhyankar, A., Basu, D., & Stremme, A. (2012). The optimal use of return predictability: an empirical study. *Journal of Financial and Quantitative* 47(5): 973–1001.
2. Affleck-Graves, J. & McDonald, B. (1989). Nonnormalities and tests of asset pricing theories. *Journal of Finance* 44(4): 889–908.
3. Britten-Jones, M. (1999). The sampling error in estimating of mean-variance efficient portfolio weights. *Journal of Finance* 54(2): 655–671.
4. DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. *Review of Financial Studies* 22(5): 1915–1953.
5. Elton, E.J. & Gruber, M.J. (1977). Risk reduction and portfolio size: an analytic solution. *Journal of Business* 50(4): 415–437.
6. Evans, J.L. & Archer, S.H. (1968). Diversification and the reduction of dispersion: an empirical analysis. *Journal of Finance* 23(5): 761–767.
7. Frazzini, A., Kabiller, D., & Pedersen, L.H. (2018). Buffett's alpha. *Financial Analysts Journal* 74(4): 35–55.
8. Green, R.C. & Hollifield, B. (1992). When will mean-variance efficient portfolios be well diversified?. *Journal of Finance* 47(5): 1785–1809.
9. Gibbons, M.R., Ross, S.A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica* 57(5): 1121–1152.
10. Jobson, J.D. & Korkie, B. (1980). Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association* 75(371): 544–554.
11. Jobson, J.D. & Korkie, B. (1983). Statistical inference in two-parameter portfolio theory with multiple regression software. *Journal of Financial and Quantitative Analysis* 18(2): 189–197.
12. Kan, R. & Smith, D.R. (2007). The distribution of the sample minimum-variance frontier. *Management Science* 54(7): 1364–1380.
13. Kan, R. & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative* 42(3): 621–656.
14. Kourtis, A. (2016). The Sharpe ratio of estimated efficient portfolios. *Finance Research Letters* 17: 72–78.
15. Kwon, Y.K. (1985). Derivation of the capital asset pricing model without normality or quadratic preference: a note. *Journal of Finance* 40(5): 1505–1509.
16. MacKinlay, A.C. & Richardson, M. (1991). Using generalized method of moments to test mean-variance efficiency. *Journal of Finance* 46(2): 511–527.
17. Marie-Claude, D., Jean-Marie, B. & Lynda, K. (2007). Multivariate tests of mean-variance efficiency with possibly non-Gaussian errors: an exact simulation-based approach. *Journal of Business and Economic Statistics* 25: 398–410.

18. Okhrin, Y. & Schmid, W. (2006). Distributional properties of portfolio weights. *Journal of Econometrics* 134(1): 235–256.
19. Peñaranda, F. (2016). Understanding portfolio efficiency with conditioning information. *Journal of Financial and Quantitative* 51(3): 985–1011.
20. Richardson, M. (1993). A test for multivariate normality in stock returns. *Journal of Business* 66(2): 295–321.
21. Seber, G.A.F. (1977). *Linear regression analysis*. New York: John Wiley & Sons.
22. Sharpe, W.F. (1966). Mutual fund performance. *Journal of Business* 39: 119–138.
23. Statman, M. (1987). How many stocks make a diversified portfolio?. *Journal of Financial and Quantitative Analysis* 22(3): 353–363.
24. Tang, G.Y.N. (2004). How efficient is naive portfolio diversification? An educational note. *Omega* 32: 155–160.
25. Zhou, G. (1993). Asset-pricing tests under alternative distributions. *Journal of Finance* 48(5): 1927–1942.

APPENDIX A.

In this appendix, we provide the derivations of Eqs. (21)–(24). We first compare the noncentral F-distribution to the Britten-Jones' F-statistic. Using Eq. (20), the moments of $\hat{\theta}^2$ under the Britten-Jones' model are:

$$E_{\mathrm{BJ}}(\hat{\theta}^2) = \frac{N(1+\theta^2)}{(T-N)} E(F_{N,T-N}) + \theta^2 = \frac{N+(T-2)\theta^2}{T-N-2} < E_{\mathrm{NF}}(\hat{\theta}^2) \qquad \textbf{(A.1)}$$

and

$$\begin{aligned} V_{\mathrm{BJ}}(\hat{\theta}^2) &= \frac{N^2(1+\theta^2)^2}{(T-N)^2} V(F_{N,T-N}) \\ &= \frac{2(T-2)N\theta^4 + 2(T-2)(N+2N\theta^2)}{(T-N-2)^2(T-N-4)} \\ &< \frac{2T^2\theta^4 + 2(T-2)(N+2T\theta^2)}{(T-N-2)^2(T-N-4)} \\ &= V_{\mathrm{NF}}(\hat{\theta}^2). \end{aligned} \qquad \textbf{(A.2)}$$

To derive the moments of a noncentral χ^2, for simplicity, let $X \sim \chi^2_{h,\delta}$ denote a noncentral χ^2-distribution with h degrees of freedom and with the noncentrality parameter δ. Thus, the pdf of $\chi^2_{h,\delta}$ is given by:

$$f_{\chi^2_h}(x;\delta) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} f_{\chi^2_{h+2i}}(x;0),$$

where $f_{\chi^2_{h+2i}}(x;0)$ is the pdf of $\chi^2_{h+2i}(0)$.

Assume that the power series has a positive radius of convergence. Changing the order of summation and the integration, the expectation of $1/X$ is:

$$\begin{aligned} E\left(\frac{1}{X}\right) &= \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} \int_0^{\infty} \frac{x^{[(h+2i)/2]-2} e^{-x/2}}{\Gamma\left(\frac{h+2i}{2}\right) 2^{(h+2i)/2}} dx \\ &= \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} \times \frac{1}{h+2i-2} \\ &= \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} \times \int_0^1 x^{h+2i-3} dx \end{aligned}$$

$$= e^{-\delta/2} \int_0^1 x^{h-3} e^{(x^2\delta/2)} dx$$

$$= \frac{e^{-\delta/2}}{2} \int_0^1 u^{(h-4)/2} e^{u\delta/2} du$$

Integrating by parts sequentially gives:

$$E\left(\frac{1}{X}\right) = \frac{1}{h-2} - \frac{\delta}{(h-2)h} + \frac{\delta^2}{(h-2)h(h+2)} - \cdots = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}\delta^{m-1}}{\Pi_{n=1}^{m}(h+2n-4)}.$$

The procedures previously stated also apply to the expectation of $1/X^2$.

$$\begin{aligned}
E\left(\frac{1}{X^2}\right) &= \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} \times \frac{1}{(h+2i-2)(h+2i-4)} \\
&= \sum_{i=0}^{\infty} \left[\frac{e^{-\delta/2}(\delta/2)^i}{i!} \times \int_0^1 \int_0^1 s^{h+2i-3} t^{h+2i-5} ds dt\right] \\
&= e^{-\delta/2} \int_0^1 \left[\int_0^1 \sum_{i=0}^{\infty} \frac{(s^2t^2\delta/2)^i}{i!} \times s^{h-3} ds\right] t^{h-5} dt \\
&= e^{-\delta/2} \int_0^1 \left[\int_0^1 e^{(s^2t^2\delta/2)} s^{h-3} ds\right] t^{h-5} dt \\
&= \frac{e^{-\delta/2}}{2} \int_0^1 \left[\int_0^1 e^{(ut^2\delta/2)} u^{(h-4)/2} du\right] t^{h-5} dt \\
&\simeq \frac{e^{-\delta/2}}{2} \int_0^1 \left[e^{(t^2\delta/2)} \frac{2}{h-2} - e^{(t^2\delta/2)} \left(\frac{t^2\delta}{2}\right) \frac{2^2}{(h-2)h} \right. \\
&\quad \left. + e^{(t^2\delta/2)} \left(\frac{t^2\delta}{2}\right)^2 \frac{2^3}{(h-2)h(h+2)}\right] t^{h-5} dt \\
&= e^{-\delta/2} \int_0^1 \left[\frac{e^{(t^2\delta/2)} t^{h-5}}{h-2} - \frac{e^{(t^2\delta/2)} \delta t^{h-3}}{(h-2)h} + \frac{e^{(t^2\delta/2)} \delta^2 t^{h-1}}{(h-2)h(h+2)}\right] dt \\
&= \frac{e^{-\delta/2}}{2} \int_0^1 \left[\frac{e^{(v\delta/2)} v^{(h-6)/2}}{h-2} - \frac{e^{(v\delta/2)} \delta v^{(h-4)/2}}{(h-2)h} + \frac{e^{(v\delta/2)} \delta^2 v^{(h-2)/2}}{(h-2)h(h+2)}\right] dv \\
&\simeq \frac{1}{h-2} \left[\frac{1}{h-4} - \frac{\delta}{(h-4)(h-2)} + \frac{\delta^2}{(h-4)(h-2)h}\right] \\
&\quad - \frac{\delta}{(h-2)h} \left[\frac{1}{h-2} - \frac{\delta}{(h-2)h}\right] + \frac{\delta^2}{(h-2)h(h+2)} \left[\frac{1}{h}\right] \\
&= \frac{1}{(h-2)(h-4)} - \frac{2\delta}{h(h-2)(h-4)} + \frac{3\delta^2}{(h+2)h(h-2)(h-4)}
\end{aligned}$$

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press

As a result, we can express the variance of $1/X$ as

$$\begin{aligned}
V\left(\frac{1}{X^2}\right) &= \left[\frac{1}{(h-2)(h-4)} - \frac{1}{(h-2)^2}\right] + \left[-\frac{2\delta}{h(h-2)(h-4)} + \frac{2\delta}{h(h-2)^2}\right] \\
&\quad + \left[\frac{3\delta^2}{(h+2)h(h-2)(h-4)} - \frac{\delta^2}{(h-2)^2h^2} - \frac{2\delta^2}{(h-2)^2h(h+2)}\right] + \cdots \\
&= \frac{2}{(h-2)^2(h-4)} - \frac{4\delta}{h(h-2)^2(h-4)} + \frac{4\delta^2}{h^2(h-2)^2(h-4)} \\
&\quad - \frac{4\delta^3}{(h+2)h^2(h-2)^2(h-4)} + \cdots
\end{aligned} \tag{A.3}$$

By replacing the random variance X with the random variable $\chi^2_{T-N,\delta}$, we obtain the expectation of the sample SSR.

$$\begin{aligned}
E_{\mathrm{NC}}(\hat{\theta}^2) &= \frac{T}{\sigma^2}E\left[\frac{1}{\chi^2_{T-N,\delta}}\right] - 1 \\
&= \left[\frac{T(1+\theta^2)^2}{(T-N-2)\theta^2} - \frac{T^2(1+\theta^2)^2}{(T-N-2)(T-N)\theta^4}\right. \\
&\quad \left. + \frac{T^3(1+\theta^2)^2}{(T-N-2)(T-N)(T-N+2)\theta^6} - \cdots\right] - 1
\end{aligned}$$

Ignoring the higher orders of θ under the convergent condition, we can approximate $E_{\mathrm{NC}}(\hat{\theta}^2)$ as follows.

$$\begin{aligned}
E_{\mathrm{NC}}(\hat{\theta}^2) &\simeq \frac{T(1+\theta^2)^2}{(T-N-2)\theta^2} - \frac{T^2(1+\theta^2)^2}{(T-N-2)(T-N)\theta^4} - 1 \\
&\simeq \frac{T(2+\theta^2)}{(T-N-2)} - \frac{T^2}{(T-N-2)(T-N)} - 1 \\
&= \frac{T\theta^2}{(T-N-2)} + \frac{2T(T-N) - T^2 - (T-N-2)(T-N)}{(T-N-2)(T-N)} \\
&= \frac{T\theta^2}{(T-N-2)} + \frac{2T - N^2 - 2N}{(T-N-2)(T-N)} \\
&= \frac{(T-2)\theta^2 + N}{(T-N-2)} + \frac{2\theta^2 - N}{(T-N-2)} + \frac{2T - N^2 - 2N}{(T-N-2)(T-N)} \\
&= \frac{(T-2)\theta^2 + N}{(T-N-2)} + \frac{2\theta^2(T-N) - N(T-N) + (2T - N^2 - 2N)}{(T-N-2)(T-N)} \\
&< \frac{(T-2)\theta^2 + N}{(T-N-2)} + \frac{2\theta^2(T-N) - N(T-N) + (2T - N^2 - 2N)}{(T-N-2)(T-N)} \\
&= \frac{(T-2)\theta^2 + N}{(T-N-2)} + \frac{2(T-N) - N(T-N) + (2T - N^2 - 2N)}{(T-N-2)(T-N)} \\
&= E_{\mathrm{BJ}}(\hat{\theta}^2) - \frac{T(N-4) + N^2 + 2N + 2}{(T-N-2)(T-N)} \\
&< E_{\mathrm{BJ}}(\hat{\theta}^2)
\end{aligned} \tag{A.4}$$

Based on Eq. (A.3), we obtain that

$$\begin{aligned}
V_{\mathrm{NC}}(\hat{\theta}^2) &= \frac{T^2}{\sigma^4} V\left[\frac{1}{\chi^2_{T-N,\delta}}\right] \\
&\simeq \frac{T^2}{\sigma^4}\left[\frac{2}{(T-N-2)^2(T-N-4)} - \frac{4\delta}{(T-N)(T-N-2)^2(T-N-4)}\right] \\
&= \frac{2T^2(1+\theta^2)^4}{(T-N-2)^2(T-N-4)\theta^4} - \frac{4T^3(1+\theta^2)^4}{(T-N)(T-N-2)^2(T-N-4)\theta^6} \\
&\simeq \frac{2T^2(\theta^4+4\theta^2+6)}{(T-N-2)^2(T-N-4)} - \frac{4T^3(\theta^2+4)}{(T-N)(T-N-2)^2(T-N-4)} \\
&= \frac{2T^2(T-N)\theta^4 + 4T^2(2T-2N-T)\theta^2 + 4T^2(3T-3N-4T)}{(T-N)(T-N-2)^2(T-N-4)} \\
&= \frac{2T^2(T-N)\theta^4 + 4T^2(T-2N)\theta^2 + 4T^2(-T-3N)}{(T-N)(T-N-2)^2(T-N-4)} \\
&< \frac{2T^2(T-N)\theta^4 + 4T(T-2)(T-N)\theta^2 + 2(T-2)N(T-N)}{(T-N)(T-N-2)^2(T-N-4)} \\
&= \frac{(2T^2\theta^4 + 2(T-2)(N+2T\theta^2))(T-N)}{(T-N)(T-N-2)^2(T-N-4)} \\
&= V_{\mathrm{NF}}(\hat{\theta}^2)
\end{aligned} \qquad \textbf{(A.5)}$$

https://doi.org/10.1017/S0269964819000482 Published online by Cambridge University Press