Exact Variance Ratio Test with Overlapping Data

RAYMOND KAN*

First draft: January 2004
This version: January 2011

*Joseph L. Rotman School of Management, University of Toronto. I thank Robert Chen, Nikolay Gospodinov, Lynda Khalaf, Xiaolu Wang, Chu Zhang, Guofu Zhou, seminar participants at University of Southampton, and participants at the the 2005 Northern Finance Meetings and the 2006 Topics in Financial Econometrics Conference at the Federal Reserve Bank of Atlanta for helpful discussions and comments. I gratefully acknowledge financial support from the National Bank Financial of Canada and the Social Sciences and Humanities Research Council of Canada.

Exact Variance Ratio Test with Overlapping Data

ABSTRACT

Variance ratio test is popular for testing whether stock prices follow a random walk or not. However, this test is typically conducted based on its asymptotic distribution, which can be unreliable in finite samples. Under a multivariate elliptical distribution assumption with an arbitrary variance-covariance matrix, we provide the exact moments and distribution of the sample variance ratio as well as an efficient procedure to compute its exact p-value. Our results allow us to study the optimal length of multi-period return for detecting different alternatives to the random walk hypothesis of stock prices.

An important question that is often encountered in economics and finance is whether a time series follows a random walk process or not. For example, a huge literature in finance is devoted to testing whether stock prices deviate from a random walk process or not. One of the implications of the "random walk" theory of stock prices is that stock returns are uncorrelated over time. Many statistical tests were designed to test this, but the variance ratio test has gained tremendous popularity in recent years. The central idea of the variance ratio test is based on the observation that when returns are uncorrelated over time, the variance of the k-period return should be k times of the variance of the 1-period return. If a researcher finds that the ratio to be significantly different from one, then he can conclude that stock prices do not follow a random walk.

While the intuition behind the variance ratio test is rather simple, conducting a statistical inference using the variance ratio test is far from trivial. The difficulty is that the variance ratio test conducted in finance typically uses overlapping data in computing the variance of the long-horizon returns. The use of overlapping data was suggested by Lo and MacKinlay (1988) because it can potentially improve the power of the variance ratio test. However, it adds to the difficulties of analyzing the exact distribution of the sample variance ratio. Currently, virtually nothing is known about its exact distribution, and not even its moments are known. This presents a significant challenge in conducting statistical inference using the sample variance ratio. In practice, the asymptotic distribution instead of the exact one is often used for conducting statistical inference. However, there are at least three different asymptotic distributions of the sample variance ratio in the literature, and they can lead to very different conclusions. Without a good understanding of the exact distribution of the sample variance ratio, it is difficult to decide which asymptotic theory is the best one to use.

In this paper, we present the exact moments and distribution of the variance ratio test with overlapping data. The exact distribution is derived under the assumption that returns are multivariate elliptically distributed, but the results are quite robust to departures from this distributional assumption. We derive the exact distribution under the null hypothesis that returns are uncorrelated as well as under the alternatives that the returns are autocorrelated. We also present a numerically stable and fast method to compute the exact distribution and p-value of the sample variance ratio.

Traditionally, the choice of the length of the long-horizon returns (i.e., choice of k for the k-

period returns) in the variance ratio test is somewhat *ad hoc*. Under the null hypothesis that returns are uncorrelated over time, the choice of k is somewhat irrelevant. However, if returns are indeed predictable, then the choice of k could have a significant impact on the power of the variance ratio test. An important benefit of knowing the exact distribution of the sample variance ratio is that it allows us to find out what is the optimal length of multi-period return for detecting different alternatives to the random walk hypothesis.

The rest of the paper is organized as follows. Section 1 provides an overview of the variance ratio test with overlapping data and summarizes the three different asymptotic distributions that are available in the literature. Section 2 presents the exact moments and distribution of the sample variance ratio under the null hypothesis that returns are uncorrelated and have constant variance over time. Section 3 presents the exact moments and distribution of the sample variance ratio under a general variance covariance matrix of returns. It also presents the optimal length of multi-period return that maximizes the power of the variance ratio test to reject some popular alternative stock price processes. Section 4 examines the robustness of our exact distribution results to departures from our distributional assumptions. Section 5 concludes the paper. Appendix A contains proofs of all propositions and lemmata. Appendix B contains a discussion of the numerical methods in computing the exact distribution of the variance ratio test.

1. Variance Ratio Test

1.1 Definition

Let p_t be the natural log of the price of an asset at time t, and $r_t = p_t - p_{t-1}$ be the continuously compounded 1-period return of the asset at time t. Suppose we have T observations of r_t and denote r as the T-vector of returns $r = [r_1, \ldots, r_T]'$. We assume that $E[r] = \mu 1_T$ and $\mathrm{Var}[r] = \Sigma$, where 1_T is a T-vector of ones, μ is the unconditional expected return of r_t, and Σ is the variance-covariance matrix of r, which is assumed to be positive definite. When returns are uncorrelated over time, we have $\mathrm{Cov}[r_s, r_t] = 0$ for $s \neq t$ and Σ is a diagonal matrix. A popular special case is $\Sigma = \sigma^2 I_T$ for $\sigma^2 > 0$, where I_T is an identity matrix of order T, which implies returns are uncorrelated over time and they have the same mean and variance.

Multi-period returns are often used for the purpose of testing return predictability. The k-

period continuously compounded return is simply the sum of k consecutive 1-period continuously compounded returns. Denote y_t as the k-period return ending at time t, we have

$$y_t = r_{t-k+1} + \ldots + r_t. \tag{1}$$

In testing the null hypothesis of no predictability, the variance ratio test is often used (see, for examples, Cochrane (1988), Lo and MacKinlay (1988), and Poterba and Summers (1988)). Define the variance ratio of k-period return as

$$\theta(k) = \frac{\mathrm{Var}[y_t]/k}{\mathrm{Var}[r_t]} = 1 + \sum_{i=1}^{k-1} \left[\frac{2(k-i)}{k}\right] \rho_i, \tag{2}$$

where ρ_i is the i-th lag autocorrelation coefficient of $\{r_t\}$. The variance ratio test is motivated by the observation that under some mild regularity conditions, we should have $\mathrm{Var}[y_t] = k\mathrm{Var}[r_t]$ when returns are uncorrelated over time, so $\theta(k) = 1$. One can therefore think of variance ratio test as a specification test of $H_0 : \rho_1 = \cdots = \rho_k = 0$, i.e., returns are serially uncorrelated.

Due to limited sample size and the desire to improve the power of the test, variance ratio tests are often performed using overlapping long-horizon returns. With T observations of r_t, we can obtain $n \equiv T - k + 1$ observations of overlapping k-period returns from y_k to y_T. Let H be an $n \times T$ matrix with k elements of ones in each row

$$H = H(k) \equiv \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 & 0 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 1 \end{bmatrix}, \tag{3}$$

we can write the n-vector of overlapping k-period returns $y = [y_k, \ldots, y_T]'$ as $y = Hr$. There are many ways of defining the sample variance ratio, the one that we use here is based on overlapping long-horizon returns and it was advocated by Lo and MacKinlay (1988). Our analysis, however, can be easily modified to accommodate other definitions of sample variance ratio.

Using the 1-period returns r_t, the unbiased estimator of the 1-period return variance is

$$\bar{\sigma}^2(1) = \frac{1}{T-1} \sum_{t=1}^{T} (r_t - \hat{\mu})^2, \tag{4}$$

where $\hat{\mu} = \sum_{t=1}^{T} r_t/T$. For the estimator of 1-period return variance using k-period returns y_t, there are many ways to construct it. The one used by Lo and MacKinlay (1988) is based on overlapping

k-period returns and it is defined as

$$\bar{\sigma}^2(k) = \frac{1}{m}\sum_{t=k}^{T}(y_t - k\hat{\mu})^2, \tag{5}$$

where $m = kn(n-1)/T$. The value of m is chosen such that $\bar{\sigma}^2(k)$ is an unbiased estimator of the 1-period return variance when returns are uncorrelated and variance is constant over time.

The sample variance ratio is the ratio of these two different sample estimators of the variance of 1-period return, defined as[1]

$$\hat{\theta}(k) = \frac{\bar{\sigma}^2(k)}{\bar{\sigma}^2(1)}. \tag{6}$$

When $\hat{\theta}(k)$ is significantly different from one, we will reject the null hypothesis of no predictability in returns. Specifically, when $\hat{\theta}(k)$ is significantly less than one, we can conclude that the prices of the asset exhibit mean reversion behavior.

1.2 Asymptotic Distributions

In order to perform the test, we need to understand the distribution of $\hat{\theta}(k)$ under the null hypothesis. In the literature, researchers rely on the asymptotic distributions of $\hat{\theta}(k)$ under the null hypothesis to compute the p-value. Unlike most test statistics that have only one asymptotic distribution, there are at least three different asymptotic distributions of $\hat{\theta}(k)$ that are available in the literature. All three asymptotic distributions assume $T \to \infty$ but they differ on the limiting behavior of k.

The most popular asymptotic distribution of $\hat{\theta}(k)$ was presented by Lo and MacKinlay (1988) by assuming k is fixed when $T \to \infty$. They show that under the homoskedasticity assumption on the returns, we have

$$\sqrt{T}(\hat{\theta}(k) - 1) \overset{A}{\sim} N\left(0, \frac{2(2k-1)(k-1)}{3k}\right) \tag{7}$$

under the null hypothesis. However, if returns exhibit heteroskedasticity, then under some additional assumptions,[2] we have

$$\sqrt{T}(\hat{\theta}(k) - 1) \overset{A}{\sim} N\left(0, \sum_{i=1}^{k-1}\left[\frac{2(k-i)}{k}\right]^2 \tau_i\right) \tag{8}$$

[1]When non-overlapping k-period returns are used to compute $\bar{\sigma}^2(k)$, Tian, Zhang, and Huang (1999) show that the sample variance ratio has a beta distribution when r_t is i.i.d. normal. Tse, Ng, and Zhang (2004) develop a circular overlapping variance ratio and approximate its exact distribution using a beta distribution.

[2]See Lo and MacKinlay (1988, p.49) for a set of sufficient conditions.

under the null hypothesis, where

$$\tau_i = \frac{E[\epsilon_t^2 \epsilon_{t-i}^2]}{E[\epsilon_t^2]^2} = 1 + \frac{\mathrm{Cov}[\epsilon_t^2, \epsilon_{t-i}^2]}{E[\epsilon_t^2]^2} \tag{9}$$

and $\epsilon_t = r_t - \mu$. In conducting statistical inference, we often need to replace τ_i by its sample counterpart. Note that if ϵ_t^2 is serially uncorrelated, then we have $\tau_i = 1$ and the asymptotic distribution can be simplified to the one for the homoskedasticity case.

The asymptotic distribution of Lo and MacKinlay (1988) is very easy to use as it only involves computing the normal distribution. The simplicity of their asymptotic distribution probably contributes to its popular use among the researchers. However, it is not entirely clear that their asymptotic result is appropriate for all statistical inferences as researchers often use fairly large k relative to T when investigating long term predictability in stock returns. Richardson and Stock (1989) provide an alternative asymptotic analysis under the assumption that both $k \to \infty$ and $T \to \infty$ but that $k/T = \delta$ is fixed. They show that $\hat{\theta}(k)$ does not converge to one under the null hypothesis, but rather have a limiting distribution. They express this limiting distribution in terms of functionals of Brownian motion as follows:

$$\hat{\theta}(k) \xrightarrow{d} \frac{1}{\delta(1-\delta)^2} \int_{\delta}^{1} [W(\lambda) - W(\lambda - \delta) - \delta W(1)]^2 \mathrm{d}\lambda, \tag{10}$$

where $W(\lambda)$ is a standard Brownian motion.[3] Using simulations, Richardson and Stock (1989) show that this limiting distribution often provides a better approximation of the finite sample distribution than the fixed-k asymptotic distribution of Lo and MacKinlay (1988), especially when k is reasonably large. However, we need to rely on simulation to evaluate the limiting distribution in (10), so it is not as easy to use as the normal distribution for computing the p-value of the sample variance ratio.

In another attempt to provide an approximation of the finite sample distribution of $\hat{\theta}(k)$ when k is large, Campbell and Mankiw (1987) (see also Chen and Deo (2006)) present an asymptotic distribution of $\hat{\theta}(k)$ under the assumption that both $k \to \infty$ and $T \to \infty$ but $k/T = \delta \to 0$. Under this assumption, Campbell and Mankiw (1987) show that

$$\frac{1}{\sqrt{\delta}}(\hat{\theta}(k) - 1) \overset{A}{\sim} N\left(0, \frac{4}{3}\right). \tag{11}$$

[3] It should be noted that the limiting distribution presented by Richardson and Stock (1989) is for the unadjusted variance ratio (i.e., using T instead of m in computing $\bar{\sigma}^2(k)$), whereas the limiting distribution presented here is from Deo and Richardson (2003) and it is for the unbiased variance ratio of Lo and MacKinlay (1988).

Chen and Deo (2006) find that this asymptotic distribution does not provide an accurate approximation of the exact distribution of $\hat{\theta}(k)$, so they suggest using a power transformation of $\hat{\theta}(k)$ to improve the approximation.

Note that unlike the asymptotic distribution of Lo and MacKinlay (1988) which depends on whether there is conditional homoskedasticity in the returns or not, both the limiting distribution of Richardson and Stock (1989) and the asymptotic distribution of Campbell and Mankiw (1987) are not affected by conditional heteroskedasticity. This is because conditional heteroskedasticity in the single period returns has very little impact on the distribution of k-period returns when k is large. Under stationarity assumption, adjacent k-period returns would have roughly constant variance when k is large enough, so the asymptotic distributions by both Richardson and Stock (1989) and Campbell and Mankiw (1987) do not have any nuisance parameters even when there is conditional heteroskedasticity in the single period returns.

Which asymptotic distribution provides the best approximation to the exact distribution crucially depends on the values of k and T. In the absence of an algorithm to compute the exact distribution of $\hat{\theta}(k)$, one would need to rely on simulations to make such an evaluation. However, once we perform a simulation, then it would make little sense to use asymptotic distribution as we can rely on the empirical distribution from the simulation for making statistical inference. Our objective is to provide a fast and accurate algorithm for computing the exact distribution of $\hat{\theta}(k)$ so that there is no need to perform a time-consuming simulation or to decide which asymptotic distribution to use.

2. Moments and Distribution of Sample Variance Ratio under the Null Hypothesis

For the sake of deriving the exact distribution of $\hat{\theta}(k)$, we need to make an assumption on the joint distribution of r. We assume r is multivariate elliptically distributed with mean $\mu 1_T$ and a variance-covariance matrix of Σ. Although this distribution assumption is somewhat restrictive, we show in Section 4 that our finite sample results are quite robust to departure from this assumption. In order to obtain the exact distribution of $\hat{\theta}(k)$, we first need to simplify its expression. As it turns out, we can write the sample variance ratio as a ratio of two quadratic forms of standard normal random variables. It should be noted that Faust (1992) also suggests that the variance ratio can be written as a ratio of two quadratic forms of normal random variables. However, the expression that was provided by Faust (1992) is not convenient for computation purpose and as a result, its practical use has been very limited. In this paper, we take it a step further by providing a simplification of this quadratic form that will lead to a speedy computation of its distribution as well as its moments.

2.1 Simplification of the Problem

We define an idempotent matrix M as

$$M = I_T - \frac{1}{T}1_T 1_T', \tag{12}$$

so we can write

$$\begin{aligned} \sum_{t=k}^{T}(y_t - k\hat{\mu})^2 &= r'MH'HMr, && (13)\\ \sum_{t=1}^{T}(r_t - \hat{\mu})^2 &= r'Mr, && (14) \end{aligned}$$

and the sample variance ratio as

$$\hat{\theta}(k) = \frac{T-1}{m}\left(\frac{r'MH'HMr}{r'Mr}\right). \tag{15}$$

To simplify presentation, we first assume $r \sim N(\mu 1_T, \Sigma)$ and then argue the results still hold when returns follow a multivariate elliptical distribution. Let P be a $T \times (T-1)$ orthonormal matrix such that $P'P = I_{T-1}$ and $PP' = M$. Since P is orthogonal to 1_T, we have $x = P'r \sim N(0_{T-1}, P'\Sigma P)$, where 0_{T-1} stands for a $(T-1)$-vector of zeros, and we can write the denominator as $r'Mr = x'x$ and the numerator as

$$r'MH'HMr = r'P(P'H'HP)P'r = x'P'H'HPx. \tag{16}$$

Suppose $Q\Lambda Q'$ is the spectral decomposition of $P'\Sigma P$ where Λ is a diagonal matrix of the $T-1$ eigenvalues of $P'\Sigma P$ and the columns of Q are the corresponding eigenvectors. We now define

$$z = \Lambda^{-\frac{1}{2}}Q'x \sim N(0_{T-1}, I_{T-1}) \tag{17}$$

and write the variance ratio as

$$\hat{\theta}(k) = \frac{T-1}{m}\left(\frac{x'P'H'HPx}{x'x}\right) = \frac{T-1}{m}\left(\frac{z'Az}{z'\Lambda z}\right), \tag{18}$$

where $A = \Lambda^{\frac{1}{2}}Q'(P'H'HP)Q\Lambda^{\frac{1}{2}}$. In this expression, P and H depend on T and k, but Q and Λ depend also on Σ.

In deriving (18), we assume r is multivariate normally distributed. However, (18) continues to hold when r is multivariate elliptically distributed. This is because we can write the sample variance ratio as

$$\hat{\theta}(k) = \frac{T-1}{m}\left(\frac{u'Au}{u'\Lambda u}\right), \tag{19}$$

where $u = z/(z'z)^{\frac{1}{2}}$, so the distribution of $\hat{\theta}(k)$ only depends on the distribution of u. When r is multivariate elliptically distributed with $E[r] = \mu 1_T$ and $\text{Var}[r] = \Sigma$, the same derivation allows us to show that z is also multivariate elliptically distributed with mean 0_{T-1} and variance I_{T-1}, which implies z has a spherical distribution. Since the distribution of u is the same for all spherical distributions z, we can assume z is normally distributed without any loss of generality.[4] Hence, expression (18) is still valid when r is multivariate elliptically distributed. The following proposition summarizes the results.

Proposition 1. *Suppose r is multivariate elliptically distributed with mean $\mu 1_T$ and variance-covariance matrix Σ. Let P be a $T \times (T-1)$ orthonormal matrix such that $P'1_T = 0_{T-1}$, Λ be a diagonal matrix of the eigenvalues of $P'\Sigma P$, and Q be a matrix of the corresponding eigenvectors. The distribution of the sample variance ratio $\hat{\theta}(k)$ has the same distribution as*

$$\frac{T-1}{m}\left(\frac{z'Az}{z'\Lambda z}\right), \tag{20}$$

where $z \sim N(0_{T-1}, I_{T-1})$ and $A = \Lambda^{\frac{1}{2}}Q'(P'H'HP)Q\Lambda^{\frac{1}{2}}$.

Note that the distribution of $\hat{\theta}(k)$ is not unbounded. By the Rayleigh-Ritz theorem, we have

$$\left(\frac{T-1}{m}\right)d_{T-1} \leq \hat{\theta}(k) \leq \left(\frac{T-1}{m}\right)d_1, \tag{21}$$

where d_1 and d_{T-1} are the largest and smallest eigenvalues of $\Lambda^{-\frac{1}{2}}A\Lambda^{-\frac{1}{2}}$, which are also the largest and smallest eigenvalues of $\tilde{A} \equiv P'H'HP$. Note that since H is of rank n or less, $\tilde{A}$ is not of full

[4] The distribution of u is a uniform distribution on the $(T-1)$-dimensional sphere. See, for example, Theorem 1.5.6 of Muirhead (1982).

rank and we have $d_{T-1} = 0$. Therefore, $\hat{\theta}(k)$ has a lower bound of zero and an upper bound of $(T-1)d_1/m$, and the upper bound is only a function of T and k but not a function of Σ. The fact that the distribution of $\hat{\theta}(k)$ is bounded from both above and below indicates that normal distribution may not provide a good approximation of the exact distribution.

2.2 Moments of Sample Variance Ratio when $\Sigma = \sigma^2 I_T$

When returns are uncorrelated over time, Σ is diagonal. If we further assume that $\Sigma = \sigma^2 I_T$ for some constant $\sigma^2 > 0$, we can then provide a significant simplification of the expression for $\hat{\theta}(k)$. When $\Sigma = \sigma^2 I_T$, we have $P'\Sigma P = \sigma^2 I_{T-1}$, so $Q = I_{T-1}$, $\Lambda = \sigma^2 I_{T-1}$, and we have

$$\hat{\theta}(k) = \frac{T-1}{m}\left(\frac{z'\tilde{A}z}{z'z}\right), \tag{22}$$

where $\tilde{A} = P'H'HP$, which only depends on T and k but not on μ and σ^2. In fact, this distribution only depends on $\tilde{A}$ through its eigenvalues, so we can write

$$\hat{\theta}(k) = \frac{T-1}{m}\left(\frac{z'Dz}{z'z}\right), \tag{23}$$

where D is a diagonal matrix of the eigenvalues of $\tilde{A}$ and $z \sim N(0_{T-1}, I_{T-1})$.[5] Note that although this distribution is only valid when $\Sigma = \sigma^2 I_T$, our analysis in Section 4 suggests that it still provides a very good approximation to the distribution of $\hat{\theta}(k)$ when Σ is diagonal but variance is changing over time. Before discussing the numerical method for evaluating this distribution, we present an analytical expression of the moments of $\hat{\theta}(k)$ for the special case of $\Sigma = \sigma^2 I_T$ in the following Proposition.

Proposition 2. *Suppose r is multivariate elliptically distributed with mean $\mu 1_T$ and variance-covariance matrix $\sigma^2 I_T$. The s-th moment of $\hat{\theta}(k)$ is given by*

$$E[\hat{\theta}(k)^s] = \left(\frac{T-1}{2m}\right)^s \frac{E[(z'\tilde{A}z)^s]}{\left(\frac{T-1}{2}\right)_s}, \tag{24}$$

where $(a)_r = a(a+1)\cdots(a+r-1)$ is the Pochhammer symbol and $E[(z'\tilde{A}z)^s]$ is obtained using

[5] To see this, we let GDG' be the spectral decomposition of $\tilde{A}$, then $z'\tilde{A}z/(z'z) = w'Dw/(w'w)$ where $w = G'z \sim N(0_{T-1}, I_{T-1})$ and it has the same distribution as z.

the following recursive relation[6]

$$E[(z'\tilde{A}z)^s] = (s-1)! \sum_{j=1}^{s} \frac{2^{j-1}\mathrm{tr}(\tilde{A}^j)E[(z'\tilde{A}z)^{s-j}]}{(s-j)!}. \tag{25}$$

Note that $\tilde{A}$ is a matrix that is only a function of T and k, so it is possible to express the moments of $E[\hat{\theta}(k)^s]$ as an explicit function of T and k. For the first four moments of $\hat{\theta}(k)$, we have simplified (24) to an explicit polynomial of k and T. The following lemma presents the finite sample mean and variance of $\hat{\theta}(k)$.

Lemma 1. *Suppose r follows a multivariate elliptical distribution with mean $\mu 1_T$ and variance-covariance matrix $\Sigma = \sigma^2 I_T$ for some $\sigma^2 > 0$. The finite sample mean of $\hat{\theta}(k)$ mean and variance of $\hat{\theta}(k)$ are given by*

$$E[\hat{\theta}(k)] = 1, \tag{26}$$

$$\mathrm{Var}[\hat{\theta}(k)] = \frac{2(T-1)}{T+1}\left[1 - \frac{T(n+1)}{2m} + \frac{\left[(T-k)_3 - (T-2k)_3^+\right]\left(n-k+\frac{4k^2}{T}\right)}{6m^2}\right] - \frac{2}{T+1}, \tag{27}$$

where $(a)_r^+ = \max[a, 0](a+1)\cdots(a+r-1)$.

Lemma 1 shows that $\hat{\theta}(k)$ is unbiased under the constant volatility assumption, which is a somewhat surprising result. In constructing the sample variance ratio, Lo and MacKinlay (1988) use the adjustment factor m to obtain the unbiased estimator $\bar{\sigma}^2(k)$. Although both the numerator and denominator in $\hat{\theta}(k)$ are unbiased estimators of σ^2, they are correlated with each other and also due to the Jensen's inequality, the ratio is in general not unbiased. Nevertheless, our Lemma 1 suggests that the sample variance ratio of Lo and MacKinlay (1988) still has the nice property of being unbiased.[7] Note that when $k/T > 1/2$, the expression of $\mathrm{Var}[\hat{\theta}(k)]$ is different from that for the case of $k/T \leq 1/2$. This is because with the length of the long-horizon return being longer than half of the sample period, every long-horizon return in the sample is correlated with each other even when there is no predictability in returns.[8]

[6]There exist explicit expressions of $E[(z'\tilde{A}z)^s]$ in the literature (see Magnus (1986) and Holmquist (1996)). However, as pointed out by Kan (2008), these explicit expressions are unsuitable for computational purpose except when s is very small. Recently, Hillier, Kan, and Wang (2009) provide an even faster recursive algorithm for computing $E[(z'\tilde{A}z)^s]$.

[7]The result of $E[\hat{\theta}(k)] = 1$ can be extended to the case that returns are i.i.d. Proof is available upon request. Although $\hat{\theta}(k)$ is unbiased, it should be noted that $E[1/\hat{\theta}(k)] \neq 1$, so it is indeed crucial that we define the variance ratio as $\bar{\sigma}^2(k)/\bar{\sigma}^2(1)$ but not as $\bar{\sigma}^2(1)/\bar{\sigma}^2(k)$ in order for it to be unbiased.

[8]The fact that we have different expressions for $k \leq T/2$ and $k > T/2$ is not entirely surprising. For example, Dufour and Roy (1985, 1989) show that there are also two different expressions for the finite sample variance of the

With the result in Lemma 1, it is straightforward to compute the standard error of $\hat{\theta}(k)$ under the null hypothesis of $H_0 : \Sigma = \sigma^2 I_T$ as it depends on only k and T. In addition, with our expression of finite sample variance, we can address the question as to which asymptotic distribution of $\hat{\theta}(k)$ provides a better approximation of its standard error. In order to do this comparison, we provide the limiting behavior of $\mathrm{Var}[\hat{\theta}(k)]$ under three different assumptions in the following lemma.

Lemma 2. *Suppose r follows a multivariate elliptical distribution with mean $\mu 1_T$ and variance-covariance matrix $\Sigma = \sigma^2 I_T$ for some $\sigma^2 > 0$. When k is fixed, we have*

$$\lim_{T\to\infty} T\mathrm{Var}[\hat{\theta}(k)] = \frac{2(k-1)(2k-1)}{3k}. \tag{28}$$

When both $k \to \infty$ and $T \to \infty$ but $k/T \to 0$, we have

$$\lim_{T,k\to\infty} \frac{T}{k}\mathrm{Var}[\hat{\theta}(k)] = \frac{4}{3}. \tag{29}$$

When $\delta = k/T$ is fixed, we have

$$\lim_{T\to\infty} \mathrm{Var}[\hat{\theta}(k)] = \begin{cases} \frac{\delta(6\delta^3+4\delta^2-11\delta+4)}{3(1-\delta)^4} & \text{if } \delta \leq \frac{1}{2}, \\ \frac{6\delta^2-4\delta+1}{3\delta^2} & \text{if } \delta > \frac{1}{2}. \end{cases} \tag{30}$$

Lemma 2 suggests that the limiting behavior of the sample variance ratio is drastically different under the three sets of assumptions. For the fixed-k case, the asymptotic variance agrees with (7). Using this asymptotic result, one would approximate the standard error of $\hat{\theta}(k)$ using $[2(k-1)(2k-1)/(3kT)]^{\frac{1}{2}}$. For the case that both $k \to \infty$ and $T \to \infty$ but $\delta \to 0$, the asymptotic variance is the same as given in (11). Using this asymptotic result, one would approximate the standard error of $\hat{\theta}(k)$ using $2\sqrt{\delta/3}$. Under either one of these two assumptions, the variance of $\hat{\theta}(k)$ approaches to zero as $T \to \infty$. For the fixed-δ case, Richardson and Stock (1989) provide the asymptotic analysis of $\hat{\theta}(k)$. However, they rely on simulation to evaluate the limiting distribution, so no analytical expression is available for the limiting variance. Lemma 2 complements their results by providing the limiting variance of $\hat{\theta}(k)$ under the assumption that δ is fixed. It suggests that for the fixed-δ case, the expression for the limiting variance depends on whether δ is less than or greater than 0.5. Using this asymptotic result for inference, one would approximate the standard error of $\hat{\theta}(k)$ by using $\delta^{\frac{1}{2}}(6\delta^3 + 4\delta^2 - 11\delta + 4)^{\frac{1}{2}}/[\sqrt{3}(1-\delta)^2]$ if $\delta \leq 0.5$, but using $(6\delta^2 - 4\delta + 1)^{\frac{1}{2}}/(\sqrt{3}\delta)$ if $\delta > 0.5$.

k-th lag sample autocorrelation coefficient, depending on whether $k \leq T/2$ or $k > T/2$. For the third and fourth moments of $\hat{\theta}(k)$, we have three and four different expressions, respectively, depending on k and T.

$\boxed{\text{Table I about here}}$

In Table I, we present the percentage approximation errors of the three asymptotic standard errors for various combinations of k and T. Panel A presents the results for the fixed-k asymptotic standard error. When k is small, the fixed-k asymptotic standard error performs much better than the other two asymptotic standard errors. However, when k is large, the performance of the fixed-k asymptotic standard error deteriorates and it often understates the exact standard error (when $k < T/2$). Panel B presents the results for the fixed-δ asymptotic standard error. When k is small, the fixed-δ asymptotic standard grossly overstates the exact standard error. While the overstatement of standard error continues when k increases, the approximation error is very small for large k and it often outperforms the other asymptotic standard errors. Panel C presents the approximation errors for the asymptotic standard error that assumes both k and T go to infinity but δ goes to zero. Similar to the fixed-δ standard error, the zero-δ standard error also grossly overstates the exact standard error when k is small. However, the zero-δ standard error can understate the true standard error when k increases. For a given T, there is a small range of k that the zero-δ asymptotic standard error performs the best, but once k/T is reasonably large, the zero-δ asymptotic standard error is generally dominated by the fixed-δ asymptotic standard error.

For a given data set, researchers often try different values of k, and we may like to find out which asymptotic standard error performs better under such situations. In Figure 1, we plot the four standard errors of $\hat{\theta}(k)$ as a function of k for different values of T. The solid lines represent the exact standard errors of $\hat{\theta}(k)$, the dotted lines represent the fixed-k asymptotic standard errors, the dashed lines represent the fixed-δ asymptotic standard errors, and the dashed-dotted lines represent the zero-δ asymptotic standard errors. From Figure 1, we can see that when k is very small, the fixed-k asymptotic standard error works better. Beyond the very small values of k, the fixed-δ asymptotic standard error works a lot better. It almost mirrors the behavior of the true standard error as k increases. When $T \geq 240$, the fixed-δ asymptotic standard error is almost indistinguishable from the exact standard error. On the contrary, the fixed-k and the zero-δ asymptotic standard errors tend to understate the true standard error when $\delta < 0.5$ and overstates the true standard error when $\delta > 0.5$. Except for very small k, the fixed-k and zero-δ asymptotic standard errors are almost identical when $T \geq 240$. Our results here echo the simulation results in Richardson and Stock (1989), where they show that the limiting distribution under the fixed-δ

assumption appears to provide an overall better approximation of the true distribution than the fixed-k asymptotic distribution. From the results in Table I and Figure 1, we can conclude that except for very small values of k, the fixed-δ asymptotic standard errors are superior to the fixed-k and zero-δ asymptotic standard errors. Nevertheless, given how easy it is to compute, there is no reason for not using the exact standard error.

Figure 1 about here

When comparing the accuracy of the three asymptotic distribution theories, one should not focus just on how well their asymptotic standard errors approximate the exact standard error. This is because the distribution of the sample variance ratio is bounded and highly non-normal, so having an accurate standard error alone is not sufficient to guarantee that one can correctly compute the p-value of the sample variance ratio. In Figure 2, we plot the coefficients of skewness (γ_1) and excess kurtosis (γ_2) of $\hat{\theta}(k)$ as a function of k for different values of T.[9] It shows that the distribution of $\hat{\theta}(k)$ is far from normal when k is large. In general, the distribution of $\hat{\theta}(k)$ is positively skewed and has very fat (right) tail, making the normal approximation highly inaccurate unless k is very small.

Figure 2 about here

In some empirical studies, we are interested in obtaining the joint moments of multiple sample variance ratios. Suppose the length of long-horizon return is equal to k_i. We use (22) to write the distribution of $\hat{\theta}(k_i)$ as

$$\hat{\theta}(k_i) = \frac{T-1}{m_i}\left(\frac{z'\tilde{A}_i z}{z'z}\right), \tag{31}$$

where $n_i = T - k_i + 1$, $m_i = k_i n_i(n_i - 1)/T$, $\tilde{A}_i = P'H_i'H_iP$ and $H_i = H(k_i)$ is an $n_i \times T$ matrix as defined in (3). This allows us to define multiple sample variance ratios and the following Proposition presents the mixed moments of p sample variance ratios.

[9] Using the results in Proposition 2 and upon further simplifications, we obtain the third and fourth moments of $\hat{\theta}(k)$ as explicit polynomials of k and T. The polynomials are somewhat lengthy, so we do not present them here but the results are available upon request. Using our results, the skewness and excess kurtosis of $\hat{\theta}(k)$ under the fixed-δ limiting distribution are easily obtained. It can be shown that as $\delta \to 0$, we have $\gamma_1 \to 0$ and $\gamma_2 \to 0$, and when $\delta \to 1$, we have $\gamma_1 \to 38/15$ and $\gamma_2 \to 1058/105$.

Proposition 3. *Suppose r is multivariate elliptically distributed with mean $\mu 1_T$ and variance-covariance matrix $\sigma^2 I_T$. The expectation of the product $\hat{\theta}(k_1)^{s_1}\cdots\hat{\theta}(k_p)^{s_p}$ is given by*

$$E\left[\prod_{i=1}^{p}\hat{\theta}(k_i)^{s_i}\right] = \left[\prod_{i=1}^{s}\left(\frac{T-1}{2m_i}\right)^{s_i}\right]\frac{1}{\left(\frac{T-1}{2}\right)_s}E\left[\prod_{i=1}^{p}(z'\tilde{A}_i z)^{s_i}\right], \tag{32}$$

where $s = s_1 + s_2 + \cdots + s_p$, and the last term can be evaluated using

$$E\left[\prod_{i=1}^{p}(z'\tilde{A}_{k_i}z)^{s_i}\right] = \frac{1}{s!}\sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_p=0}^{s_p}(-1)^{\sum_{i=1}^{p}\nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_p}{\nu_p}E[(z'B_\nu z)^s], \tag{33}$$

where $B_\nu = \left(\frac{s_1}{2}-\nu_1\right)\tilde{A}_1 + \left(\frac{s_2}{2}-\nu_2\right)\tilde{A}_2 + \cdots + \left(\frac{s_p}{2}-\nu_p\right)\tilde{A}_p$ and $E[(z'B_\nu z)^s]$ can be computed using the recursive relation in (25).

Note that in Proposition 3, we present the moments of a product of quadratic forms in normal random variables using a new formula from Kan (2008). Unlike existing expressions (e.g., Magnus (1978, 1979) and Holmquist (1996)), this expression is computationally very efficient and it allows us to compute the mixed moments of $\hat{\theta}(k_i)$ even for fairly large s.

As an example, setting $s_1 = s_2 = 1$, we obtain

$$E[\hat{\theta}(k_1)\hat{\theta}(k_2)] = \frac{(T-1)^2}{m_1 m_2}\left[\frac{\mathrm{tr}(\tilde{A}_1)\mathrm{tr}(\tilde{A}_2) + 2\mathrm{tr}(\tilde{A}_1\tilde{A}_2)}{(T-1)(T+1)}\right]. \tag{34}$$

Using this result and after simplification, we present the exact covariance between two sample variance ratios in the following lemma.

Lemma 3. *Suppose r follows a multivariate elliptical distribution with mean $\mu 1_T$ and variance-covariance matrix $\Sigma = \sigma^2 I_T$ for some $\sigma^2 > 0$. The finite sample covariance between $\hat{\theta}(k_1)$ and $\hat{\theta}(k_2)$ for $k_1 \le k_2$ is given by*

$$\frac{2(T-1)}{T+1}\left[\frac{k_1}{k_2}\left(\frac{n_2-1}{n_1-1} - \frac{T(n_2+1)}{2m_1}\right) + \frac{\left[(T-k_2)_3 - (T-k_1-k_2)_3^+\right]\left(n_2 - k_1 + \frac{4k_1k_2}{T}\right)}{6m_1m_2}\right] - \frac{2}{T+1}, \tag{35}$$

where $n_i = T - k_i + 1$ and $m_i = k_i n_i(n_i-1)/T$.

This lemma can be potentially useful for researchers who are interested in performing joint test on multiple sample variance ratios. With the exact covariance formula, the following lemma presents the limiting behavior of $\mathrm{Cov}[\hat{\theta}(k_1), \hat{\theta}(k_2)]$ under three different assumptions.

Lemma 4. *Suppose r follows a multivariate elliptical distribution with mean $\mu 1_T$ and variance-covariance matrix $\Sigma = \sigma^2 I_T$ for some $\sigma^2 > 0$. Let $k_1 \le k_2$ be the lengths of two long-horizon returns. When k_1 and k_2 is fixed, we have*

$$\lim_{T\to\infty} T\mathrm{Cov}[\hat{\theta}(k_1), \hat{\theta}(k_2)] = \frac{2(k_1-1)(3k_2-k_1-1)}{3k_2}. \tag{36}$$

When $k_1 \to \infty$, $k_2 \to \infty$ and $T \to \infty$ but with $k_1/T \to 0$, $k_2/T \to 0$ and $k_1/k_2 \to a$, where a is a positive constant, we have

$$\lim_{T,k_1,k_2\to\infty} \frac{T}{\sqrt{k_1 k_2}}\mathrm{Cov}[\hat{\theta}(k_1), \hat{\theta}(k_2)] = \frac{2\sqrt{a}(3-a)}{3}. \tag{37}$$

When $\delta_1 = k_1/T$ and $\delta_2 = k_2/T$ are fixed, we have

$$\lim_{T\to\infty} \mathrm{Cov}[\hat{\theta}(k_1), \hat{\theta}(k_2)] = \begin{cases} \frac{\delta_1[6\delta_2(1-\delta_2)^2+(4\delta_2-1)\delta_1^2-2\delta_1(1-\delta_2)(3\delta_2^2-1)]}{3\delta_2(1-\delta_1)^2(1-\delta_2)^2} & \text{if } \delta_1+\delta_2 \le 1, \\ \frac{(1-\delta_2)[6\delta_1^2(1-\delta_1)+(1-4\delta_1)(1-\delta_2)]}{3\delta_1\delta_2(1-\delta_1)^2} & \text{if } \delta_1+\delta_2 > 1. \end{cases} \tag{38}$$

Similar to Lemma 2, Lemma 4 suggests that the limiting behavior of $\mathrm{Cov}[\hat{\theta}(k_1), \hat{\theta}(k_2)]$ crucially depends on the assumptions on the limiting behavior of k_1, k_2 and T. For the case that both k_1 and k_2 are fixed, the asymptotic covariance formula in (36) was given in Richardson and Smith (1991). For the case that k_1, k_2 and T all go to infinity, but with $k_1/T \to 0$ and $k_2/T \to 0$, the limiting covariance formula in (37) was given in Chen and Deo (2006). Finally, the limiting covariance formula in (38) that assumes both δ_1 and δ_2 are fixed has not appeared in the literature. However, similar to the limiting variance case, this last formula generally provides the best approximation of the finite sample covariance between $\hat{\theta}(k_1)$ and $\hat{\theta}(k_2)$, except when both k_1 and k_2 are very small.

2.3 Evaluating the Exact Distribution of Sample Variance Ratio when $\Sigma = \sigma^2 I_T$

Under the assumption that $\Sigma = \sigma^2 I_T$, we know the exact distribution of $\hat{\theta}(k)$ is given by (23). Therefore, in order to compute the cumulative density function of $\hat{\theta}(k)$, we just need to evaluate the following probability

$$P[\hat{\theta}(k) < c] = P\left[z'Dz < \frac{cm}{T-1}z'z\right] = P\left[\sum_{i=1}^{T-1}\left(d_i - \frac{cm}{T-1}\right)z_i^2 < 0\right], \tag{39}$$

where d_i is the ith diagonal element of D. This amounts to computing the probability for a linear combination of independent χ_1^2 random variables to be less than zero. This problem has been well studied in the statistics and econometrics literature. Let $\lambda_i = d_i - cm/(T-1)$, we use the results from Gil-Pelaez (1951) and Imhof (1961) to write

$$P[\hat{\theta}(k) < c] = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin\left(\frac{1}{2}\sum_{i=1}^{T-1}\arctan(2t\lambda_i)\right)}{t\prod_{i=1}^{T-1}(1+4t^2\lambda_i^2)^{\frac{1}{4}}}\mathrm{d}t. \tag{40}$$

Numerical evaluation of this integral was studied by Imhof (1961), Davies (1980), Farebrother (1984, 1990), Ansley, Kohn, and Shively (1992), and Lu and King (2002). Our implementation follows closely to the algorithm in Lu and King (2002) and it is highly efficient and accurate.

The evaluation of this integral requires the knowledge of d_i's, the eigenvalues of $P'H'HP$. As it turns out, this is the biggest hurdle in the computation of the exact distribution of $\hat{\theta}(k)$. One may think computing eigenvalues is a trivial task which can be easily handled by programs like Matlab. However, computation of eigenvalues for a general symmetric matrix is an $O(n^3)$ operation, so it is very time consuming to compute the eigenvalues of a large matrix. More importantly, the resulting eigenvalues can be numerically inaccurate when the matrix is large. In some of our examples, we need to deal with cases with $T = 2400$, so it is impractical to rely on the standard algorithm to compute eigenvalues of $P'H'HP$. Fortunately, $P'H'HP$ is a structured matrix with its entries only depend on k and T, so we are able to develop a fast and accurate algorithm to compute its eigenvalues. However, the discussion involves numerical methods that may not interest all the readers, so we relegate its details to Appendix B.[10]

With our numerical method of computing the exact distribution of $\hat{\theta}(k)$, we plot the lower and upper 2.5 percentiles of the exact distribution of $\hat{\theta}(k)$ as a function of k in Figure 3 using the solid lines. We provide the plot for T = 60, 240, 600, and 2400, representing the typical lengths of annual, quarterly, monthly, and weekly data that we encounter in empirical studies. In Figure 3, we also plot the lower and upper 2.5 percentiles of the normal approximation using the dotted lines for comparison, where the normal approximation is based on the fixed-k asymptotic distribution of (7). As we can see from the plot, the normal approximation does a very poor job in approximating the true distribution of $\hat{\theta}(k)$. Except when k is very small, we find that the approximate normal distribution is in general to the left of the exact distribution. As a result, one can have serious size problem when testing the null hypothesis using the normal approximation, especially when using a one-tailed test.

[10] All the Matlab programs for this paper are available upon request.

$\boxed{\text{Figure 3 about here}}$

In Figure 4, we provide the actual sizes of the normal approximation test of $H_0 : \theta(k) = 1$ as a function of k under different choices of T. We look at the sizes for three different tests: a two-tailed test, a left-tailed test and a right-tailed test. All three normal approximation tests are assumed to have a nominal size of 5%, so if the tests have good size property, the actual probabilities of rejection should be close to 5% under the null hypothesis. In Figure 4, we use the solid line, dotted, and dashed lines to represent the actual sizes of the two-tailed test, the left-tailed test and the right-tailed test, respectively. We can see that the one-sided tests all behave poorly when k is reasonably large, with the left-tailed test leads to under-rejection of the null hypothesis and the right-tailed test leads to over-rejection problem. The under-rejection at the left tail is of great concerns because as pointed out by Lo and MacKinlay (1989), the cutoff point based on the normal approximation test is negative when k is large enough, whereas $\hat{\theta}(k)$ will never be negative whether it is under the null or under the alternative. As a result, the actual probability of rejection of the left-tailed test based on normal approximation will be zero when k is large enough. Unlike the one-sided tests, the size of the two-sided test is reasonably well behaved because the under-rejection at the left tail is compensated by the over-rejection at the right tail. However, one should not take comfort with the good size behavior of the two-tailed test and rely on it to test the random walk hypothesis. This is because under the alternative hypothesis that stock price has a mean reversion component, we have $\theta(k) < 1$ and we need to rely on the left-tail of the distribution of $\hat{\theta}(k)$ to provide the rejection. However, when we use the normal approximation test, whether it is one-tailed or two-tailed, such rejection will be hard to come by when k is reasonably large. This presents a significant problem of using and interpreting the normal approximation test. On the one hand, we expect mean reversion component in stock price is easier to detect with longer horizon returns, so we may like to use larger k. On the other hand, we know the normal approximation test has an under-rejection problem and it provides absolutely no rejection when k is large enough, so we have to limit its use to the case that k is very small. With the development of the exact test, we can reliably use variance ratio test for any value of k without worrying about the size problem that is in the normal approximation test.

$\boxed{\text{Figure 4 about here}}$

3. Moments and Distribution of Sample Variance Ratio under General Variance-Covariance Matrix

3.1 Moments of Sample Variance Ratio for General Σ

In order to study the power of the exact variance ratio test, we need to understand the moments and distribution of $\hat{\theta}(k)$ for the general case of Σ. The following Proposition presents an analytical expression of the moments of $\hat{\theta}(k)$ for a general variance-covariance matrix of returns.

Proposition 4. *Suppose r is multivariate elliptically distributed with mean $\mu 1_T$ and variance-covariance matrix Σ. The s-th moment of $\hat{\theta}(k)$ is given by*

$$E[\hat{\theta}(k)^s] = \left(\frac{T-1}{m}\right)^s \frac{1}{(s-1)!}\int_0^\infty \frac{t^{s-1}}{|I_{T-1}+2t\Lambda|^{\frac{1}{2}}} E[(z'Bz)^s]\mathrm{d}t, \tag{41}$$

where $B = (I_{T-1}+2t\Lambda)^{-\frac{1}{2}}A(I_{T-1}+2t\Lambda)^{-\frac{1}{2}}$, where A and Λ are defined after (16).

Although the expression in Proposition 4 looks complicated, it is just a 1-dimensional integral and it is easy to numerically evaluate its value. As an illustration, we use Proposition 4 to write down the first two moments of $\hat{\theta}(k)$ as

$$\begin{aligned} E[\hat{\theta}(k)] &= \frac{T-1}{m}\int_0^\infty \frac{\mathrm{tr}(B)}{|I_{T-1}+2t\Lambda|^{\frac{1}{2}}}\mathrm{d}t, & (42)\\ E[\hat{\theta}(k)^2] &= \left(\frac{T-1}{m}\right)^2\int_0^\infty \frac{t\left[\mathrm{tr}(B)^2+2\mathrm{tr}(B^2)\right]}{|I_{T-1}+2t\Lambda|^{\frac{1}{2}}}\mathrm{d}t. & (43)\end{aligned}$$

To facilitate the numerical integration, we can use a change of variable $y = 1/(1+2t\lambda_{T-1})$ and write

$$\begin{aligned} E[\hat{\theta}(k)] &= \frac{T-1}{2\lambda_{T-1}m}\int_0^1 y^{\frac{T-3}{2}}\left(\prod_{i=1}^{T-1} c_i\right)^{-\frac{1}{2}}\left(\sum_{i=1}^{T-1}\frac{a_{ii}}{c_i}\right)\mathrm{d}y, & (44)\\ E[\hat{\theta}(k)^2] &= \left(\frac{T-1}{2\lambda_{T-1}m}\right)^2\int_0^1 (1-y)y^{\frac{T-3}{2}}\left(\prod_{i=1}^{T-1} c_i\right)^{-\frac{1}{2}}\left[\sum_{i=1}^{T-1}\sum_{j=1}^{T-1}\frac{2a_{ij}^2+a_{ii}a_{jj}}{c_ic_j}\right]\mathrm{d}y, & (45)\end{aligned}$$

where a_{ij} is the (i,j)-th element of A and $c_i = \frac{\lambda_i}{\lambda_{T-1}} - \left(\frac{\lambda_i}{\lambda_{T-1}}-1\right)y$. Ali (1984, Eq.(3.5)) provides a trick to speed up the evaluation of the integral in (45) by reducing the double summation to a single summation. Using Ali's method and after simplification, we have

$$E[\hat{\theta}(k)^2] = \left(\frac{T-1}{2\lambda_{T-1}m}\right)^2\int_0^1 (1-y)y^{\frac{T-3}{2}}\left(\prod_{i=1}^{T-1} c_i\right)^{-\frac{1}{2}}\sum_{i=1}^{T-1}\left(\frac{2h_i}{c_i}+\frac{3a_{ii}^2y}{c_i^2}\right)\mathrm{d}y, \tag{46}$$

where

$$h_i = \lambda_i \sum_{\substack{j=1 \\ j \neq i}}^{T-1} \frac{2a_{ij}^2 + a_{ii}a_{jj}}{\lambda_i - \lambda_j}. \tag{47}$$

Higher moments of $\hat{\theta}(k)$ can be numerically evaluated in a similar manner.[11]

In order to obtain the joint moments of multiple sample variance ratios, we need an expression for the expectation of a product of s sample variance ratios, which is given in the following Proposition.

Proposition 5. *Suppose r is multivariate elliptically distributed with mean $\mu 1_T$ and variance-covariance matrix Σ. Let $A_i = \Lambda^{\frac{1}{2}} Q'(P'H_i'H_iP)Q\Lambda^{\frac{1}{2}}$, where Q and Λ are the eigenvectors and the eigenvalues of $P'\Sigma P$. The expectation of the product $\hat{\theta}(k_1)^{s_1} \cdots \hat{\theta}(k_p)^{s_p}$ is given by*

$$E\left[\prod_{i=1}^{p} \hat{\theta}(k_i)^{s_i}\right] = \left[\prod_{i=1}^{p} \left(\frac{T-1}{m_i}\right)^{s_i}\right] \frac{1}{(s-1)!} \int_0^\infty \frac{t^{s-1}}{|I_{T-1} + 2t\Lambda|^{\frac{1}{2}}} E\left[\prod_{j=1}^{p} (z'B_iz)^{s_i}\right] \mathrm{d}t, \tag{48}$$

where $B_i = (I_{T-1} + 2t\Lambda)^{-\frac{1}{2}} A_i (I_{T-1} + 2t\Lambda)^{-\frac{1}{2}}$ and $s = s_1 + s_2 + \cdots + s_p$.

Following the discussion after Proposition 3, we can use Proposition 5 to obtain any joint moments of multiple sample variance ratios. For example, when $s_1 = s_2 = 1$, we have

$$E[\hat{\theta}(k_1)\hat{\theta}(k_2)] = \frac{(T-1)^2}{m_1 m_2} \int_0^\infty \frac{t[\mathrm{tr}(B_1)\mathrm{tr}(B_2) + 2\mathrm{tr}(B_1 B_2)]}{|I_{T-1} + 2t\Lambda|^{\frac{1}{2}}} \mathrm{d}t. \tag{49}$$

Similar to the case of $E[\hat{\theta}(k)^2]$, we can use a change of variable $y = 1/(1+2t\lambda_{T-1})$ and Ali's formula to express $E[\hat{\theta}(k_1)\hat{\theta}(k_2)]$ as

$$E[\hat{\theta}(k_1)\hat{\theta}(k_2)] = \frac{(T-1)^2}{4\lambda_{T-1}^2 m_1 m_2} \int_0^1 (1-y) y^{\frac{T-3}{2}} \left(\prod_{i=1}^{T-1} c_i\right)^{-\frac{1}{2}} \sum_{i=1}^{T-1} \left(\frac{h_i}{c_i} + \frac{3a_{1,ii}a_{2,ii}y}{c_i^2}\right) \mathrm{d}y, \tag{50}$$

where $c_i = \frac{\lambda_i}{\lambda_{T-1}} - \left(\frac{\lambda_i}{\lambda_{T-1}} - 1\right) y$, $a_{1,ij}$ and $a_{2,ij}$ are the (i,j)-th element of A_1 and A_2, respectively, and

$$h_i = \lambda_i \sum_{\substack{j=1 \\ j \neq i}}^{T-1} \frac{4a_{1,ij}a_{2,ij} + a_{1,ii}a_{2,jj} + a_{1,jj}a_{2,ii}}{\lambda_i - \lambda_j}. \tag{51}$$

[11] In addition to the integration approach, we can also use an alternative approach in Hillier, Kan, and Wang (2009) to evaluate the moments of $\hat{\theta}(k)$.

3.2 Evaluating the Exact Distribution of Sample Variance Ratio for General Σ

For the general Σ case, we know the exact distribution of $\hat{\theta}(k)$ is given by (18). Therefore, in order to compute the cumulative density function of $\hat{\theta}(k)$, we just need to evaluate the following probability

$$P[\hat{\theta}(k) < c] = P\left[z'Az < \frac{cm}{T-1}z'\Lambda z\right] = P\left[z'\left(A - \frac{cm}{T-1}\Lambda\right)z < 0\right]. \tag{52}$$

Let $\tilde{D}$ be a diagonal matrix of the q ($q \leq T-1$) nonzero eigenvalues of $A - (cm/(T-1))\Lambda$ and $\tilde{G}$ be a $(T-1)\times q$ matrix with its columns equal to the eigenvectors associated with the q nonzero eigenvalues. Writing $w = \tilde{G}z \sim N(0_q, I_q)$, we have

$$P[\hat{\theta}(k) < c] = P[z'\tilde{G}\tilde{D}\tilde{G}'z < 0] = P\left[\sum_{i=1}^{q}\tilde{d}_i w_i^2 < 0\right]. \tag{53}$$

Therefore, just as in the case of $\Sigma = \sigma^2 I_T$, the computation of the exact distribution under general Σ can also be reduced to computing the probability for a linear combination of independent χ_1^2 random variables to be less than zero. The difficulty in carrying out this computation is in obtaining the eigenvalues $\tilde{d}_i$. In Appendix B, we discuss the algorithm to speed up the computation of $\tilde{d}_i$, especially when Σ is a Toeplitz matrix (which is the case when returns are covariance stationary).[12]

3.3 Some Popular Alternative Hypotheses

Although we have results for general Σ, it helps to study a few special cases of Σ that are implied by some popular alternative hypotheses to the random walk theory of stock price. We consider three cases here. The first case assumes the return follows an AR(1) process as

$$(r_t - \mu) = \phi(r_{t-1} - \mu) + \epsilon_t, \tag{54}$$

where $-1 < \phi < 1$ and ϵ_t has mean zero and uncorrelated over time. This process of stock returns has some empirical support from the work of Lo and MacKinlay (1988). Under this assumption of stock return process, the i-th lag coefficient of autocorrelation of the returns is simply given by $\rho_i = \phi^i$ and the covariance matrix of the returns is proportional to

$$\Sigma = A_T(\phi), \tag{55}$$

[12] A Toeplitz matrix is a matrix which has constant entries along its diagonals. When returns are covariance stationary, $\mathrm{Cov}[r_i, r_j]$ is only a function of $|i-j|$, so Σ is a symmetric Toeplitz matrix.

where $A_T(\phi)$ is a $T \times T$ Kac, Murdoc, and Szegö (KMS) matrix with its (i,j)-th element given by $a_{ij} = \phi^{|i-j|}$. Under this alternative hypothesis, the population variance ratio is given by

$$\theta(k) = 1 + \sum_{i=1}^{k-1} \left[\frac{2(k-i)}{k} \right] \rho_i = \frac{k(1-\phi^2) - 2\phi(1-\phi^k)}{k(1-\phi)^2}, \tag{56}$$

and it monotonically approaches a limit of $(1+\phi)/(1-\phi)$ as k increases.

The second case assumes the natural log of stock price follows an AR(1) process

$$p_t - \alpha = \phi(p_{t-1} - \alpha) + \epsilon_t, \tag{57}$$

where $0 < \phi < 1$ and ϵ_t has mean zero and uncorrelated over time. This is a model of stock market fads suggested by Shiller (1981). Under this assumption of stock price process, the i-th lag autocorrelation coefficient of the returns is given by

$$\rho_i = -\left(\frac{1-\phi}{2} \right) \phi^{i-1}, \tag{58}$$

so the variance-covariance matrix of the returns is proportional to

$$\Sigma = \frac{1+\phi}{2\phi} I_T - \frac{1-\phi}{2\phi} A_T(\phi), \tag{59}$$

and the population variance ratio is given by

$$\theta(k) = 1 + \sum_{i=1}^{k-1} \left[\frac{2(k-i)}{k} \right] \rho_i = \frac{1-\phi^k}{k(1-\phi)}, \tag{60}$$

which decreases with k to a lower bound of zero.

The last case assumes the natural log of stock price is the sum of two components—a random walk component x_t and a stationary AR(1) component y_t:

$$p_t = x_t + y_t, \tag{61}$$

where

$$x_t = x_{t-1} + \eta_t, \tag{62}$$

$$y_t - \alpha = \phi(y_{t-1} - \alpha) + \epsilon_t, \tag{63}$$

where $0 < \phi < 1$ and η_t and ϵ_t have mean zero and uncorrelated with each other and over time. This specification of stock price process was used by Summers (1986), Poterba and Summers (1988), and Fama and French (1988). Under this assumption of stock price process, the returns follow an ARMA(1,1) process and the i-th lag coefficient of autocorrelation of the returns is given by

$$\rho_i = -\left(\frac{1-\phi}{2+(1+\phi)\kappa}\right)\phi^{i-1}, \tag{64}$$

where $\kappa = \mathrm{Var}[\eta_t]/\mathrm{Var}[\epsilon_t]$ is a measure of the relative importance of the random walk component to the AR(1) component in the stock price. Therefore, the variance-covariance matrix of the returns is proportional to

$$\Sigma = \frac{(1+\phi)(1+\kappa\phi)}{[2+(1+\phi)\kappa]\phi} I_T - \frac{(1-\phi)}{[2+(1+\phi)\kappa]\phi} A_T(\phi). \tag{65}$$

Under this alternative hypothesis, the population variance ratio is given by

$$\theta(k) = 1 + \sum_{i=1}^{k-1}\left[\frac{2(k-i)}{k}\right]\rho_i = 1 - \frac{2[k(1-\phi)+\phi^k - 1]}{k(1-\phi)[2+(1+\phi)\kappa]}, \tag{66}$$

and it decreases with k to a lower bound of $1 - 2/[2+(1+\phi)\kappa]$.

There is one thing common with the Σ under these three alternative hypotheses. It is that they all share the same eigenvectors with the KMS matrix $A_T(\phi)$. Let $\tilde{\lambda}_i$ be the ith eigenvalue of the KMS matrix $A_T(\phi)$ and u_i be its corresponding eigenvector, then

$$\tilde{d}_i = \frac{1+\phi-(1-\phi)\tilde{\lambda}_i}{2\phi} \tag{67}$$

is an eigenvalue for the Σ in (59) and u_i is its corresponding eigenvector. Similarly, for the third case,

$$\tilde{d}_i = \frac{(1+\phi)(1+\kappa\phi)-(1-\phi)\tilde{\lambda}_i}{[2+(1+\phi)\kappa]\phi} \tag{68}$$

is an eigenvalue of Σ in (65) and u_i is its corresponding eigenvector. Therefore, as long as we know how to compute the eigenvalues and eigenvectors of a KMS matrix, we can easily compute the eigenvalues and eigenvectors of Σ for all three cases. As it turns out, an efficient algorithm of obtaining the eigenvalues and eigenvector of a KMS matrix was given by Trench (1988) and his results are briefly summarized in Appendix B.

3.4 Distribution of Sample Variance Ratio under Alternative Hypotheses

In Figure 5, we present the distribution of $\hat{\theta}(k)$ under the assumption that return follows an AR(1) process with $\phi = 0.1$ as a function of k for different choices of T. From top to bottom, the three solid lines in Figure 5 are the upper 2.5 percentile, mean and lower 2.5 percentile of $\hat{\theta}(k)$. For comparison purpose, we also present the lower and upper 2.5 percentiles of $\hat{\theta}(k)$ under the null hypothesis using two dotted lines. When returns are positively autocorrelated, we can see that the distribution of $\hat{\theta}(k)$ shifts to the right as compared with the one under the null hypothesis, so the right-tailed test has a higher likelihood of rejecting the null hypothesis. Note that similar to the distribution under the null hypothesis, the distributions of $\hat{\theta}(k)$ across different T's under the AR(1) alternative hypothesis are largely similar when $\delta = k/T$ is fixed. This suggests that when k is reasonably large, the power of the test is largely determined by k/T. The dashed dotted line in Figure 5 shows the population variance ratio $\theta(k)$ under the AR(1) alternative. In Lemma 1, we show that $E[\hat{\theta}(k)] = 1$ under the null hypothesis, so $\hat{\theta}(k)$ is an unbiased estimator of $\theta(k)$ under the null hypothesis. However, $\hat{\theta}(k)$ is in general a biased estimator of $\theta(k)$ under the alternative hypothesis. In Figure 5, we can see that there is some small difference between $E[\hat{\theta}(k)]$ and $\theta(k)$ when $T = 60$. However, for $\phi = 0.1$, this bias is very small and becomes negligible when $T \geq 240$.

Figure 5 about here

In Figure 6, we present the distribution of $\hat{\theta}(k)$ when stock price follows an AR(1) process with $\phi = 0.975$ by using the same format as in Figure 5. Under this alternative hypothesis, the stock price is a mean-reverting process, so the returns are negatively autocorrelated. As a result, the distribution of $\hat{\theta}(k)$ shifts to the left as compared with the one under the null hypothesis, so the left-tailed test has a higher likelihood of rejecting the null hypothesis. Unlike the case of AR(1) returns, the shapes of the graphs differ significantly across different T's. For a fixed k/T, we find that the distribution of $\hat{\theta}(k)$ tightens as T increases. This suggests that even when k and T are both large, the power of the test would heavily depend on k and T, but not merely a function of k/T as in the case of AR(1) returns alternative. When stock price follows a mean-reverting AR(1) process, adjacent long-horizon returns are almost perfectly correlated when k is large, so $\theta(k)$ goes to zero as k increases. Naturally, one may like to think that using a larger k would give us better power to reject the null hypothesis. However, when T is finite, the issue becomes more complicated. First, $\hat{\theta}(k)$ is not an unbiased estimator of $\theta(k)$. Even $\theta(k)$ goes to zero as k increases, we do not find $E[\hat{\theta}(k)]$ behaves the same. In fact, we can see from Figure 6 that $E[\hat{\theta}(k)]$ exhibits a humped shape and does not decreases monotonically with k. Second, $\hat{\theta}(k)$ becomes more volatile as k increases because there are fewer independent observations of k-period returns in the sample. These two effects together suggest that larger k does not always lead to better power in finite sample.

Figure 6 about here

In Figure 7, we present the distribution of $\hat{\theta}(k)$ when stock price is a sum of a random walk and an AR(1) process with $\phi = 0.975$ by using the same format as in Figure 5. The variance of the innovation of the random walk component is assumed to be half of the variance of the innovation of the AR(1) component. As stock price has a stationary mean-reverting component, returns are still negatively autocorrelated, so the distribution of $\hat{\theta}(k)$ still shifts to the left when compared with the one under the null. However, with the presence of the random walk component in the stock price, the shift of the distribution of $\hat{\theta}(k)$ is not as strong as when stock price has only an AR(1) component, making it harder to reject the null hypothesis. Similar to Figure 6, we find that $\hat{\theta}(k)$ is biased and becomes more volatile as k increases, which suggests that larger k does not always lead to better power in rejecting the null hypothesis.

Figure 7 about here

Under the fixed-δ asymptotic that assumes both T and k go to infinity but $\delta = k/T$ is fixed, Deo and Richardson (2003) express the limiting distribution of $\hat{\theta}(k)$ under this alternative hypothesis as:

$$\hat{\theta}(k) \xrightarrow{d} \left[\frac{(1+\phi)\kappa}{2+(1+\phi)\kappa}\right] \frac{1}{\delta(1-\delta)^2} \int_{\delta}^{1} [W(\lambda) - W(\lambda-\delta) - \delta W(1)]^2 \mathrm{d}\lambda, \tag{69}$$

where $W(\lambda)$ is a standard Brownian motion. Therefore, under the fixed-δ asymptotic theory, the limiting distribution of $\hat{\theta}(k)$ under this alternative is just a rescaling of its limiting distribution under the null. However, Figure 7 shows that unlike under the null hypothesis, the fixed-δ limiting distribution of $\hat{\theta}(k)$ is a very poor approximation of the finite sample distribution under the alternative. The finite sample distribution of $\hat{\theta}(k)$ under the alternative is not just a rescaling of the null distribution, and for a fixed δ, it is still very sensitive to the choice of T.[13] In order to provide

[13]The main reason for the poor approximation of the fixed-δ limiting distribution is due to the fact that the AR(1) component still plays a significant role in determining the finite sample distribution of $\hat{\theta}(k)$ even for large k and T, especially when ϕ is close to one. The poor approximation of the fixed-δ limiting distribution under the alternative hypothesis can also be seen in Table 2 of Deo and Richardson (2003) as it shows the rejection rates are sensitive to T even when we fix δ.

a better approximation to the finite sample distribution of $\hat{\theta}(k)$ under this and other alternatives, Perron and Vodounou (2005) recently propose another fixed-δ limiting distribution of $\hat{\theta}(k)$ by assuming the span of the data is fixed (so the length of one period shrinks as T goes to infinity). As shown in Perron and Vodounou (2005), their new limiting distribution of $\hat{\theta}$ often provides a far better approximation of the finite sample distribution of $\hat{\theta}(k)$ than (69).

3.5 Power of Variance Ratio Test

In the past, the study of the power of variance ratio test was seriously hampered by the lack of finite sample distribution results. When we do not know how to evaluate the finite sample distribution of $\hat{\theta}(k)$ under the null and the alternative hypotheses, we need to rely on either asymptotic distributions or on simulations. As the use of asymptotic distributions can lead to serious size problem, the results based on asymptotic distributions are ambiguous and may not be applicable when T or k is small. While simulations can address the issue of the size problem, it is time consuming to perform, so a typical simulation study only considers a few selected values of k and T, and possibly a few selected parameters for the alternative hypothesis.[14] As a result, we do not have a complete understanding of the power properties of the variance ratio test. In implementation of the variance ratio test, one would like to know what is the optimal k to use for a given length of sample period (T) in order to maximize the power for rejecting a given class of alternative hypotheses. Without the tools to compute the exact distribution of $\hat{\theta}(k)$, definitive answer to this question is largely unavailable in the literature, so researchers are often forced to use an *ad hoc* choice of k in empirical studies.

With the ability to efficiently compute the finite sample distribution of $\hat{\theta}(k)$ under the null and the alternatives, we investigate the power of the variance ratio test against the three popular alternative hypotheses. For different values of T, we present in Figure 8 the power of the variance ratio test against the alternative hypothesis that stock return follows an AR(1) process with $\phi = 0.1$ as a function of k. The solid line and the dotted line represent the power of the 5% one-tailed and two-tailed tests, respectively. There are two patterns that are noteworthy in the graphs. The

[14] It should be noted that it is often the case that we need to use simulations to evaluate the limiting distribution of $\hat{\theta}(k)$ under the null and the alternatives (for the fixed-δ case). Therefore, asymptotic power analysis like the one conducted by Perron and Vodounou (2005) is also subject to the same computational constraints as in typical simulation studies.

first pattern is the power function is monotonically decreasing in k and the optimal power occurs at $k^* = 2$. This result is not entirely surprising. When returns follow an AR(1) process, the i-th lag autocorrelation of returns is given by $\rho_i = \phi^i$, so ρ_2 to ρ_{k-1} do not provide any additional information beyond ρ_1. As $\hat{\rho}_1$ provides the most reliable estimate of ϕ, we should not use a test that relies on higher lags of sample autocorrelation coefficient as they only provide a noisier estimate of ϕ. $\hat{\theta}(2)$ is basically $1+\hat{\rho}_1$, so it is only natural that $\hat{\theta}(2)$ has the highest power against the alternative hypothesis of return follows an AR(1) process. The second pattern is the power function for a fixed $\delta = k/T$ is largely the same across T, except when k is very small. This result can be anticipated by the graphs in Figure 5, which shows that the distribution of $\hat{\theta}(k)$ under this alternative is largely a function of δ.

Figure 8 about here

In Figure 9, we present the same results but for the alternative hypothesis that stock price follows an AR(1) process with parameter $\phi = 0.975$. We note that when $T = 60$, the alternative hypothesis is hard to be detected as the stock price is almost indistinguishable from random walk in a short sample period. With larger T, we see the power function of $\hat{\theta}(k)$ increases with k to a maximum before it comes back down. This suggests that although $\theta(k)$ decreases to zero with k under this alternative hypothesis, medium-horizon returns can actually give higher power to reject the null hypothesis than long-horizon returns. In Figure 9, we find that the optimal k typically occurs near the neighborhood of $k = T/4$, so choosing k beyond $T/4$ is inadvisable.

Figure 9 about here

In Figure 10, we present the results for the third alternative hypothesis that stock price is the sum of a random walk component and an AR(1) component with parameter $\phi = 0.975$, and the variance of the innovation of the random walk component is assumed to be half of the variance of the innovation of the AR(1) component. With the addition of the random walk component in the stock price, the predictability of returns is weakened and the power of the variance ratio test is in general lower than its counterpart when stock price has only an AR(1) component. While the general pattern in Figures 9 and 10 are the same, the optimal k is less than $T/4$ when stock price has a random walk component. This result is intuitive because with increase in k, the variance of long-horizon return is eventually dominated by the random walk component and it provides very weak evidence against the random walk hypothesis. Therefore, we would like to use a smaller k to perform the variance ratio test, especially when the random walk component is relatively important.

Figure 10 about here

So far, our analysis of optimal k^* is based on a fixed choice of the AR(1) parameter ϕ, so we would like to extend our analysis by investigating how k^* varies with ϕ. In Figure 11, we plot k^* for the one-tailed test as a function of ϕ using the dashed line (with the scale on the right hand side of the graph) when return follows an AR(1) process. As we can see from the figure, k^* is always equal to 2 regardless of the value of ϕ. In the figure, we also plot three power functions, one for $k = k^*$ (solid line), one for $k = 2$ (dashed dotted line, but invisible here as $k^* = 2$ in this case), and one for $k = T/4$. As we can see, the loss of power with the use of $k = T/4$ is very substantial, so the choice of k can have a serious impact on the outcome of the test.

Figure 11 about here

In Figure 12, we present the same results but for the alternative hypothesis that stock price follows an AR(1) process. In this case, we see that as ϕ goes up from 0.8 to 1, k^* gradually increases to $T/4$. Even when $k^* < T/4$, the power function under the optimal k^* is indistinguishable from the power function for $k = T/4$, so we can safely uses $k = T/4$ when we believe stock price follows an AR(1) process. However, if one uses $k = 2$ to conduct the variance ratio test, then there is a significant reduction in power.

Figure 12 about here

Finally, we present the results for the alternative hypothesis that stock price is the sum of a random walk component and an AR(1) component in Figure 13. Similar to the previous case, we find that k^* approaches $T/4$ when ϕ approaches one. However, when ϕ is less than one, we typically find that k^* to be far less than $T/4$. More importantly, unlike the case when there is no random walk component, there can be significant loss of power if one uses $k = T/4$ to conduct the variance ratio test. It is also inadvisable to use $k = 2$ as this choice of k leads to great reduction of the power of the variance ratio test.

Figure 13 about here

In summary, we find that when return follows an AR(1) process, the optimal k^* is always 2. However, when price follows an AR(1) process, then the optimal k^* is less than $T/4$ but using $k = T/4$ does not lead to any loss of power. When price contains both a random walk and an AR(1) component, k^* is somewhere between 2 and $T/4$ but the optimal choice of k heavily depends on T and ϕ as well as the relative importance of the two components. When ϕ is close to one or T is small, $k = T/4$ is close to optimal. However, the power of the test is very low under these conditions regardless of what k that we use, so we should not be too concerned with the choice of k. When ϕ is further away from one or T is large, then the choice of k can have significant implications on the outcome of the test.

4. Robustness of Finite Sample Results

In deriving finite sample results for the sample variance ratio, we need to make a joint elliptical distribution assumption on the returns. Whatever distribution assumption that we make is at best an approximation, so one should not take our results literally as "exact." The important issues are the robustness of our finite sample results to the departure from the elliptical distribution assumption, and whether our finite sample results provide a better approximation than the asymptotic ones.

There are two types of departure from our distributional assumption that we would like to study. The first one is to find out how well our results hold up if returns are independently and identically distributed but not normal. For example, returns may have skewness and heavy kurtosis. There are many alternative distributional assumptions that we can use to examine the robustness of our results. However, to make this exercise more relevant, we use the empirical distribution of daily returns as the alternative distribution. For a given sample size (T), we randomly draw our samples from the daily returns of the value-weighted NYSE over the period 1926/1/2–2009/12/31. We then use these resampled returns to compute the sample variance ratios for different values of k. The whole exercise is repeated 1,000,000 times. In Figure 14, we plot the empirical distribution of $\hat{\theta}(k)$ as a function of k for different choices of T. From top to bottom, the three solid lines in Figure 14 are the upper 2.5 percentile, mean and lower 2.5 percentile of $\hat{\theta}(k)$. For comparison purpose, we also present the lower and upper 2.5 percentiles of $\hat{\theta}(k)$ under the assumption that returns are jointly elliptically distributed. From Figure 14, we can see that $\hat{\theta}(k)$ is unbiased even when returns are not multivariate elliptically distributed (see the discussion in footnote 7). In addition, except for T very small, the 2.5 and 97.5 percentiles of $\hat{\theta}(k)$ are not all that different from those under the multivariate elliptical distribution.

Figure 14 about here

To further examine the robustness of the results, we report the actual sizes of two tests of $H_0 : \theta(k) = 1$ in Figure 15. The solid line represents the actual size of our exact test under the elliptical distribution assumption, and the dotted line represents the actual size of a fixed-k asymptotic test under the homoskedasticity assumption, i.e., based on the normal distribution of (7). To conserve space, we only report the results for the 5% left-tailed test in Figure 15, the results for the right-tailed and two-tailed tests are available upon request. From Figure 15, we can see that our exact test based on the elliptical distribution assumption is almost exact, with only minor over-rejection when $T = 60$ and k is large. In contrast, the fixed-k asymptotic test works well only when k is very small, but the size of the test is greatly distorted when k/T is nontrivial. We have also performed simulation experiments based on other distributional assumptions, and our experience invariably suggests that except when T is very small, our exact results on $\hat{\theta}(k)$ based on multivariate elliptical distribution are still very good approximations as long as returns are independently and identically distributed.

Figure 15 about here

The second type of departure from our distributional assumption that we study is time-varying variance. In the finance literature, there is overwhelming evidence that the variance of short-horizon return is time-varying, so it is particularly relevant for us to investigate the impact of time-varying variance on the finite sample distribution of sample variance ratio. To this end, we study two popular heteroskedastic nulls. In the first case, we assume returns are uncorrelated over time but the natural logarithm of the return variance at time t follows an AR(1) process

$$\ln \sigma_t^2 = \psi \ln \sigma_{t-1}^2 + \zeta_t, \qquad \zeta_t \sim N(0,1), \tag{70}$$

and ζ_t is independent of the returns. This heteroskedastic return process was studied by Lo and MacKinlay (1989). We assume $\psi = 0.5$ and approximate the distribution of $\hat{\theta}(k)$ using 1,000,000 simulations. In Figure 16, we present the distribution of $\hat{\theta}(k)$ as a function of k for different choices of T in the same format as Figure 14. From Figure 16, we can see that $\hat{\theta}(k)$ is almost unbiased even when return variance is time-varying. However, there are some differences in the distributions of $\hat{\theta}(k)$ under the constant variance assumption vs. the time-varying variance assumption, especially when k is small. This suggests that the standard error of $\hat{\theta}(k)$ is potentially sensitive to the time-varying variance when k is small.

Figure 16 about here

When return is heteroskedastic, one may like to use the heteroksedastic consistent standard error in (8) developed by Lo and MacKinlay (1988) to conduct the statistical test. In Figure 17, we report the actual sizes of two tests of $H_0 : \theta(k) = 1$. Both tests are left-tailed tests with nominal size of 5%. The solid line represents the actual size of our exact test under the elliptical distribution assumption, and the dotted line represents the actual size of Lo and MacKinlay's heteroskedastic consistent test. From Figure 17, we can see that our exact test over-rejects the null hypothesis when k is small. However, when k is moderately large, our exact test has almost the perfect size. This result is expected because when k is reasonably large, the long-horizon returns would behave like i.i.d. normal, so it is not surprising to find that heteroskedasticity in the single period returns has little impact on the distribution of $\hat{\theta}(k)$ when k is moderately large. In contrast, Lo and MacKinlay (1988) heteroskedastic consistent test works very well when k is very small, but the size of the test is greatly distorted when k/T is nontrivial. In fact, Lo and MacKinlay's test shows no power in rejecting the null hypothesis when $k > T/4$. This is because $\hat{\theta}(k)$ is bounded below by zero, but yet the fixed-k asymptotic theory assumes it is normally distributed. When k is large enough, the left tail cutoff point from the normal approximation falls below zero and hence it is unattainable.

Figure 17 about here

In the second case, we consider a case that the returns follow a GARCH(1,1) model proposed by Bollerslev (1986). Under this model, returns are uncorrelated over time but their conditional variance changes over time. The GARCH(1,1) process assumes $r_t = \mu + \epsilon_t$, where $\epsilon_t \sim N(0, \sigma_t^2)$

and

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2. \tag{71}$$

We assume $\alpha = 0.1171$ and $\beta = 0.8575$ (the choice of ω does not matter as long as $\omega > 0$). These are the same parameters used by Chen and Deo (2006) in their study of variance ratio test. We approximate the distribution of $\hat{\theta}(k)$ using 1,000,000 simulations. In Figure 18, we present the distribution of $\hat{\theta}(k)$ as a function of k for different choices of T in the same format as Figure 14. Similar to Figure 14, we find that $\hat{\theta}(k)$ is almost unbiased when return follows a GARCH(1,1) model. However, there are some differences in the distributions of $\hat{\theta}(k)$ under the constant variance assumption vs. the GARCH(1,1) assumption, especially when k is small.

Figure 18 about here

In Figure 19, we report the actual sizes of two tests of $H_0 : \theta(k) = 1$ in the same format as in Figure 17. Similar to Figure 17, we can see that our exact test over-rejects the null hypothesis when k is small but it is almost perfect when k is moderately large. In contrast, Lo and MacKinlay (1988) heteroskedastic consistent test works very well when k is very small, but the size of the test is greatly distorted when k/T is nontrivial.

Figure 19 about here

In summary, we find that our finite sample results based on joint elliptical distribution assumption are robust to the distribution assumption when k is reasonably large. The results for small k are still robust if returns are independently and identically distributed. However, for returns that exhibit time-varying variance, our exact distribution understates the standard error of $\hat{\theta}(k)$ when k is small and that leads to over-rejection. In comparison, Lo and MacKinlay's heteroskedastic consistent test does a good job in approximating the standard error of $\hat{\theta}(k)$ for small k, but it is grossly inappropriate for moderately large k. Therefore, we recommend researchers to use Lo and MacKinlay's test for very small k especially if one believes variance is time-varying, but our exact test should be used for moderately large k, regardless of whether there is time-varying variance or not.

5. Conclusion

In this paper, we provide a finite sample analysis of variance ratio tests with overlapping observations. Under the assumption that returns are multivariate elliptically distributed, we provide analytical formulas for the moments of the sample variance ratio for an arbitrary variance-covariance matrix of the returns. In addition, we provide an efficient numerical algorithm for evaluating the cumulative density function of the sample variance ratio under the null and the alternatives. This allows us to conduct exact inference using variance ratio test as well as studying its power against popular alternative hypotheses to the "random walk" theory of the stock prices.

We find that the fixed-k asymptotic distribution that is typically used in empirical studies is grossly inappropriate except when k is very small. However, when stock price contains a mean-reverting component, one may like to use a larger k for conducting the variance ratio test. Our exact variance ratio test allows us to make reliable inference for any values of k and T. In addition, it allows us to investigate the optimal choice of k for a given class of alternative hypothesis. Due to the inability to compute the exact distribution, this question has been largely unanswered in the literature and we provide the tools for researchers to address this question.

We study the power property of the variance ratio test against three popular alternative hypotheses. For the case that return follows an AR(1) process, we find that the optimal k is always 2. For the case that stock price follows an AR(1) process, we find that $k = T/4$ is close to optimal. Finally, when stock price contains both a random walk and an AR(1) component, the optimal k depends on T, the AR(1) parameter ϕ as well as the relative importance of the two components.

For future work, we would like to extend our finite sample analysis to other predictability tests that use long-horizon returns. For example, we would like to study the long-horizon regression test of Fama and French (1988) and other autocorrelation based tests. This would give us a better understanding of the finite sample properties of these tests, which in turn will provide us with better ways of detecting departures from the random walk hypothesis.

Appendix A

Proof of Propositions 2 and 3: As Proposition 2 is a special case of Proposition 3, we only provide the proof of Proposition 3 here. It is well known that when z has a spherical distribution, $u = z/(z'z)^{\frac{1}{2}}$ and $z'z$ are independent of each other (see, for example, Theorem 1.5.6 of Muirhead (1982)). Therefore,

$$\begin{aligned} E\left[\prod_{i=1}^{p}(z'\tilde{A}_i z)^{s_i}\right] &= E\left[(z'z)^s\prod_{i=1}^{p}\left(\frac{z'\tilde{A}_i z}{z'z}\right)^{s_i}\right] \\ &= E\left[(z'z)^s\prod_{i=1}^{p}(u'\tilde{A}_i u)^{s_i}\right] \\ &= E[(z'z)^s]E\left[\prod_{i=1}^{s}(u'\tilde{A}_i u)^{s_i}\right] \\ &= 2^s\left(\frac{T-1}{2}\right)_s E\left[\prod_{i=1}^{p}\left(\frac{z'\tilde{A}_i z}{z'z}\right)^{s_i}\right] \end{aligned} \qquad \text{(A1)}$$

The third equality follows from the independence of u and $z'z$. The last equality follows because $z'z \sim \chi^2_{T-1}$, so its s-th moment is given by $2^s\left(\frac{T-1}{2}\right)_s$ (see, for example, Johnson, Kotz, and Balakrishnan (1994, Eq. 18.8)).

The recursive relation in (25) is based on a recursive relation between moments and cumulants, and can be found in Mathai and Provost (1992, Eq. 3.2b.8). Finally, (33) is obtained from Proposition 4 of Kan (2008). This completes the proof.

Proof of Lemmeta 1 and 3: Using the expression in Proposition 2 and the fact that $\mathrm{tr}(\tilde{A}) = \mathrm{tr}(P'HH'P) = \mathrm{tr}(H'PP'H) = \mathrm{tr}(HMH)$, we have

$$E[\hat{\theta}(k)] = \frac{\mathrm{tr}(\tilde{A})}{m} = \frac{\mathrm{tr}(HMH')}{m}. \qquad \text{(A2)}$$

It is straightforward to show that HH' is a Toeplitz matix with its (i,j)-th element given by $\max[k-|i-j|,0]$, and $HMH' = HH' - \frac{k^2}{T}1_n1_n'$ is also a Toeplitz matrix with its (i,j)-th element given by $\max[k-|i-j|,0] - k^2/T$. Therefore,

$$\mathrm{tr}(HMH') = n\left(k - \frac{k^2}{T}\right) = m, \qquad \text{(A3)}$$

so we have $E[\hat{\theta}(k)] = 1$.

For the covariance formula, we first simplify (34) as

$$E[\hat{\theta}(k_1)\hat{\theta}(k_2)] = \frac{(T-1)^2}{m_1 m_2}\left[\frac{\mathrm{tr}(\tilde{A}_1)\mathrm{tr}(\tilde{A}_2)+2\mathrm{tr}(\tilde{A}_1\tilde{A}_2)}{(T-1)(T+1)}\right] = \frac{2(T-1)\mathrm{tr}(\tilde{A}_1\tilde{A}_2)}{(T+1)m_1m_2}+\frac{T-1}{T+1}, \tag{A4}$$

where the last equality follows because $\mathrm{tr}(\tilde{A}_i) = \mathrm{tr}(H_i M H_i') = m_i$. Using the fact that $E[\hat{\theta}(k_i)] = 1$,

$$\mathrm{Cov}[\hat{\theta}(k_1), \hat{\theta}(k_2)] = E[\hat{\theta}(k_1)\hat{\theta}(k_2)] - 1 = \frac{2(T-1)\mathrm{tr}(\tilde{A}_1\tilde{A}_2)}{(T+1)m_1m_2} - \frac{2}{T+1}. \tag{A5}$$

The only term that needs to be evaluated is $\mathrm{tr}(\tilde{A}_1\tilde{A}_2)$, which can be written as

$$\mathrm{tr}(\tilde{A}_1\tilde{A}_2) = \mathrm{tr}(P'H_1'H_1PP'H_2'H_2P) = \mathrm{tr}(H_2MH_1'H_1MH_2') = \mathrm{tr}(C'C) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} c_{ij}^2, \tag{A6}$$

where $C = H_1MH_2' = H_1H_2' - \frac{k_1k_2}{T}1_{n_1}1_{n_2}'$. When $k_1 \leq k_2$, it is straightforward to show that H_1H_2' is a Toeplitz matrix with its (i,j)-th element given by $\max[k_1-(j-i),0]$ if $i-j<0$, $\max[k_2-(i-j),0]$ if $i-j>k_2-k_1$, and k_1 if $0 \leq i-j \leq k_2-k_1$. It follows that C is also a Toeplitz matrix with its (i,j)-th element given by

$$c_{ij} = \begin{cases} \max[k_1-(j-i),0] - \frac{k_1k_2}{T} & \text{if} \quad i-j<0, \\ k_1 - \frac{k_1k_2}{T} & \text{if} \quad 0 \leq i-j \leq k_2-k_1, \\ \max[k_2-(i-j),0] - \frac{k_1k_2}{T} & \text{if} \quad i-j > k_2-k_1. \end{cases} \tag{A7}$$

Since C is a Toeplitz matrix, we can obtain $\mathrm{tr}(C'C)$ by summing c_{ij}^2 along the diagonals. Counting downward from the principal diagonal of C, there are altogether k_2-k_1+1 diagonals, each has n_2 elements of $k_1-(k_1k_2)/T$. For the i-th superdiagonal of C, it has n_2-i elements of $\max[0,k_1-i]-(k_1k_2)/T$. By symmetry, there is a corresponding subdiagonal of C that has the same elements. Using this pattern, we can express $\mathrm{tr}(C'C)/(m_1m_2)$ as

$$\begin{aligned} \frac{\mathrm{tr}(C'C)}{m_1m_2} &= \frac{1}{m_1m_2}\left[(k_2-k_1+1)n_2\left(k_1-\frac{k_1k_2}{T}\right)^2 + 2\sum_{i=1}^{n_2-1}(n_2-i)\left(\max[k_1-i,0]-\frac{k_1k_2}{T}\right)^2\right] \\ &= \frac{k_1}{k_2}\left(\frac{n_2-1}{n_1-1}-\frac{T(n_2+1)}{2m_1}\right) + \frac{[(T-k_2)_3-(T-k_1-k_2)_3^+]\left(n_2-k_1+\frac{4k_1k_2}{T}\right)}{6m_1m_2}. \end{aligned} \tag{A8}$$

Plugging (A8) in (A5) gives us the expressions for $\mathrm{Cov}[\hat{\theta}(k_1), \hat{\theta}(k_2)]$. The formula for $\mathrm{Var}[\hat{\theta}(k)]$ is obtained by setting both k_1 and k_2 in the covariance formula to k. This completes the proof.

Proof of Lemmeta 2 and 4: Since Lemma 2 is a special case of Lemma 4, we only provide the proof for Lemma 4 here. When $\delta_1 \to 0$ and $\delta_2 \to 0$, we must have $k_1 + k_2 \leq T$, so the appropriate expression of finite sample covariance is obtained by dropping the + sign in $(T-k_1-k_2)_3^+$, and the limiting variance in (36) and (37) can be obtained by taking the appropriate limit. When δ_1 and δ_2 are fixed, we need to consider two cases. When $\delta_1+\delta_2 \le 1$, we have $k_1+k_2 \le T$, so the appropriate expression of finite sample covariance is obtained by dropping the + sign in $(T-k_1-k_2)_3^+$. Setting $k_1=\delta_1 T$ and $k_2=\delta_2 T$ in this expression and letting $T\to\infty$, we obtain the first expression in (38). When $\delta_1+\delta_2>1$, we have $k_1+k_2>T$ and the appropriate expression of finite sample covariance is obtained by setting $(T-k_1-k_2)_3^+$ to zero. Setting $k_1=\delta_1 T$ and $k_2=\delta_2 T$ in this expression and letting $T\to\infty$ gives us the second expression in (38). This completes the proof.

Proof of Propositions 4 and 5: Since Propositions 4 is a special case of Propositions 5, we only provide the proof of Proposition 5 here. Let $X_i = z'A_i z$ and $Y = z'\Lambda z$. Using Lemma 5 of Magnus (1986), the joint moment generating function of $X_1, X_2, \cdots, X_p, Y$ is given by

$$\begin{aligned}\phi(u_1,\cdots,u_p,-t) &= E[\exp(u_1X_1+\cdots u_pX_p-tY)]\\ &= |I_{T-1}-2(u_1A_1+\cdots+u_pA_p-t\Lambda)|^{-\frac{1}{2}}\\ &= |I_{T-1}+2t\Lambda|^{-\frac{1}{2}}|I_{T-1}-2C|^{-\frac{1}{2}},\end{aligned}\tag{A9}$$

where $C=u_1B_1+\cdots u_pB_p$ and $B_i=(I_{T-1}+2t\Lambda)^{-\frac{1}{2}}A_i(I_{T-1}+2t\Lambda)^{-\frac{1}{2}}$. Using Theorem 1 of Meng (2005), the expectation of the product of s sample variance ratios is given by

$$\begin{aligned}E\left[\prod_{i=1}^p \hat{\theta}(k_i)^{s_i}\right] &= \left[\prod_{i=1}^p\left(\frac{T-1}{m_i}\right)^{s_i}\right]E\left[\frac{X_1^{s_1}\cdots X_p^{s_p}}{Y^s}\right]\\ &= \left[\prod_{i=1}^p\left(\frac{T-1}{m_i}\right)^{s_i}\right]\frac{1}{(s-1)!}\int_0^\infty t^{s-1}\left.\frac{\partial^s\phi(u_1,\cdots,u_s,-t)}{\partial u_1^{s_1}\cdots\partial u_p^{s_p}}\right|_{u_1=\cdots u_p=0}\mathrm{d}t\\ &= \left[\prod_{i=1}^p\left(\frac{T-1}{m_i}\right)^{s_i}\right]\frac{1}{(s-1)!}\int_0^\infty \frac{t^{s-1}}{|I_{T-1}+2t\Lambda|^{\frac{1}{2}}}\left.\frac{\partial^s|I_{T-1}-2C|^{-\frac{1}{2}}}{\partial u_1^{s_1}\cdots\partial u_p^{s_p}}\right|_{u_1=\cdots u_p=0}\mathrm{d}t\\ &= \left[\prod_{i=1}^p\left(\frac{T-1}{m_i}\right)^{s_i}\right]\frac{1}{(s-1)!}\int_0^\infty \frac{t^{s-1}}{|I_{T-1}+2t\Lambda|^{\frac{1}{2}}}\left.\frac{\partial^s|I_{T-1}-2C|^{-\frac{1}{2}}}{\partial u_1^{s_1}\cdots\partial u_p^{s_p}}\right|_{u_1=\cdots u_p=0}\mathrm{d}t\\ &= \left[\prod_{i=1}^p\left(\frac{T-1}{m_i}\right)^{s_i}\right]\frac{1}{(s-1)!}\int_0^\infty \frac{t^{s-1}}{|I_{T-1}+2t\Lambda|^{\frac{1}{2}}}E\left[\prod_{i=1}^p(z'B_iz)^{s_i}\right]\mathrm{d}t.\end{aligned}\tag{A10}$$

The last equality follows because $|I_{T-1}-2C|^{-\frac{1}{2}}$ is the joint moment generating function of $z'B_1z,\cdots$, $z'B_pz$, so the expectation of the product can be obtained by differentiating the joint moment generating function. This completes the proof.

Appendix B

Numerical Method for Computing Eigenvalues of $P'H'HP$: We take a number of steps to simplify the computation of the eigenvalues of $P'H'HP$. The first step is to realize that $P'H'HP$ and $HPP'H' = HMH'$ share the same nonzero eigenvalues, so we know $P'H'HP$ has $k-2$ eigenvalues of zero and the rest of the eigenvalues are the same as the eigenvalues of HMH'. So instead of computing the eigenvalues of a $(T-1)\times(T-1)$ matrix, we only need to compute the eigenvalues of a smaller $n \times n$ matrix.

The second step is to write HMH' as

$$HMH' = H\left(I_T - \frac{1_T 1_T'}{T}\right)H' = HH' - \frac{k^2}{T}1_n 1_n'. \tag{B1}$$

It is straightforward to show that HH' is a symmetric Toeplitz matrix with its (i,j)-th element given by $\max[k-|i-j|,0]$, so HMH' is a symmetric Toeplitz matrix with its (i,j)-th element given by $\max[k-|i-j|,0]-k^2/T$. For a Toeplitz matrix, the following facts due to Cantoni and Butler (1976) allow us to greatly reduce the computation time of its eigenvalues.

Suppose A is an $n\times n$ symmetric Toeplitz matrix. Let $n_1 = [n/2]$ be the integral part of $n/2$ and $n_2 = n - n_1$. Denote J as a "flip" matrix that has ones along the southwest-northeast diagonal and zeros elsewhere. We have

1. A has n_1 odd eigenvectors (i.e., an eigenvector u such that $Ju = -u$) and n_2 even eigenvectors (i.e., $Ju = u$).

2. When n is even, we write
$$A = \begin{bmatrix} A_1 & JA_2J \\ A_2 & JA_1J \end{bmatrix}, \tag{B2}$$
where A_1 and A_2 are $n_1 \times n_1$ submatrices of A. The odd eigenvalues of A are the same as the eigenvalues of $A_1 - JA_2$. The even eigenvalues of A are the same as the eigenvalues of $A_1 + JA_2$.

3. When n is odd, we write
$$A = \begin{bmatrix} A_1 & x & JA_2J \\ x' & q & x'J \\ A_2 & Jx & JA_1J \end{bmatrix}. \tag{B3}$$

The odd eigenvalues of A are the same as the eigenvalues of $A_1 - JA_2$. The even eigenvalues of A are the same as the eigenvalues of

$$\begin{bmatrix} A_1 + JA_2 & \sqrt{2}x \\ \sqrt{2}x' & q \end{bmatrix}. \tag{B4}$$

These two steps allow us to reduce the problem of finding the eigenvalues of a $(T-1) \times (T-1)$ matrix to a problem of finding the eigenvalues of two symmetric matrices with dimensions n_1 and n_2 respectively. Since computing eigenvalues is an $O(n^3)$ operation, these two steps would cut down the computation time by at least 75% (but often substantially more).

The final step is to realize that HMH' differs from HH' by a matrix of rank one. Suppose $Q\Lambda Q'$ is the spectral decomposition of HH', where Λ is a diagonal matrix of the eigenvalues of HH', and Q is a matrix of the corresponding eigenvectors. Since $HMH' = HMH'QQ'$ and $Q'HMH'Q$ share the same eigenvalues, the eigenvalues of HMH' are the same as the eigenvalues of

$$Q'HMH'Q = Q'\left(HH' - \frac{k^2}{T}1_n 1_n'\right)Q = \Lambda - \frac{k^2}{T}qq', \tag{B5}$$

where $q = Q'1_n$. Since HH' is Toeplitz, we can partition its eigenvalues into odd and even. Writing

$$\Lambda = \begin{bmatrix} \Lambda_o & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & \Lambda_e \end{bmatrix}, \tag{B6}$$

where Λ_o is a diagonal matrix of the odd eigenvalues and Λ_e is a diagonal matrix of the even eigenvalues. Since for an odd eigenvalue, the sum of the elements of its eigenvector is equal to zero, so we have $q = [0_{n_1}', \ q_e']'$, where q_e is the sum of the elements of the even eigenvectors. With this observations, we can write

$$\Lambda - \frac{k^2}{T}qq' = \begin{bmatrix} \Lambda_o & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & \Lambda_e - \frac{k^2}{T}q_e q_e' \end{bmatrix}, \tag{B7}$$

so the odd eigenvalues of HMH' are the same as Λ_o, and the even eigenvalues of HMH' are the same as the eigenvalues of $\Lambda_e - \frac{k^2}{T}q_e q_e'$, which differs from Λ_e by a matrix of rank one. This problem is known as the rank-one update problem in numerical matrix algebra, and fast and stable algorithm is well developed (see Li (1993) and Gu and Eisenstat (1994) and references therein).

Note that we should only use the third step when Q and Λ for HH' are easy to obtain. This would be the case when k is small or when k is large. When k is small, HH' has only a small band of nonzero elements and this type of matrix is called the banded, symmetric Toeplitz matrix. There are a number of fast algorithms that one can use to obtain the eigenvalues and eigenvectors of a banded, symmetric Toeplitz matrix (see Trench (1985), Arbenz (1991), and Handy and Barlow (1994)).

When $k \geq T/2$, the (i,j)-th element of HH' is simply $k - |i-j|$ and we can write

$$HMH' = S + \left[\frac{(T-k)(2k-T)}{2T} - 1\right] 1_n 1_n', \tag{B8}$$

where S is a Toeplitz matrix with its (i,j)-th element given by $(n+1)/2 - |i-j|$. The eigenvalues and eigenvectors of S can be solved analytically.[15] The odd eigenvalues of S are $1/[1-\cos((2i-1)\pi/n)]$, $i = 1, \ldots, n_1$, and they are also the odd eigenvalues of HMH'. The even eigenvalues of S are $1/[1-\cos((2i-1)\pi/(n+1)]$, $i = 1, \ldots, n_2$, and the sum of the elements of the eigenvectors are $\sqrt{2/(n+1)}\cot((2i-1)\pi/(2n+2))$, $i = 1, \ldots, n_2$. The even eigenvalues of HMH' can then be obtained using a rank-one update.

Computation of Eigenvalues and Eigenvectors of KMS Matrix: This section presents a fast algorithm of computing the eigenvalues and eigenvectors of a KMS matrix that is due to Trench (1988). Let $A_T(\phi)$ be a $T \times T$ KMS matrix. Following Trench (1988), we let

$$S_T(\gamma) = (1 - \phi\cos\gamma)\sin\left(\frac{(T+1)\gamma}{2}\right) - \phi\sin\gamma\cos\left(\frac{(T+1)\gamma}{2}\right). \tag{B9}$$

$$C_T(\gamma) = (1 - \phi\cos\gamma)\cos\left(\frac{(T+1)\gamma}{2}\right) - \phi\sin\gamma\sin\left(\frac{(T+1)\gamma}{2}\right), \tag{B10}$$

The KMS matrix has $T_1 = [T/2]$ odd eigenvalues and $T_2 = T - T_1$ even eigenvalues. For (B9), there are T_1 solutions. If $0 < \phi < 1$, we have the i-th solution is in the following range

$$\frac{(2i-1)\pi}{T+1} < \gamma_i < \frac{2i\pi}{T+1}, \qquad i = 1, \ldots, T_1. \tag{B11}$$

If $-1 < \phi < 0$, the range of γ_i is given by

$$\frac{2i\pi}{T+1} < \gamma_i < \frac{(2i+1)\pi}{T+1}, \qquad i = 1, \ldots, T_1. \tag{B12}$$

The odd eigenvalues of $A_T(\phi)$ are given by

$$\tilde{\lambda}_i = \frac{(1-\phi)^2}{1 - 2\phi\cos\gamma_i + \phi^2}, \qquad i = 1, \ldots, T_1. \tag{B13}$$

[15] Proof is available upon request.

The odd eigenvector associated with $\tilde{\lambda}_i$ is proportional to u_i, where

$$u_{ij} = \sin\left(\frac{(T-2j+1)\gamma_i}{2}\right), \qquad j = 1, \dots, T. \tag{B14}$$

Note that $\sum_{j=1}^{T} u_{ij} = 0$ and $\sum_{j=1}^{T} u_{ij}^2 = [T - \sin(T\gamma_i)/\sin(\gamma_i)]/2$.

For (B10), there are T_2 solutions. If $0 < \phi < 1$, we have the j-th solution is in the following range

$$\frac{(2i-2)\pi}{T+1} < \gamma_i < \frac{(2i-1)\pi}{T+1}, \qquad i = 1, \dots, T_2. \tag{B15}$$

If $-1 < \phi < 0$, the range of γ_i is given by

$$\frac{(2i-1)\pi}{T+1} < \gamma_i < \frac{2i\pi}{T+1}, \qquad i = 1, \dots, T_2. \tag{B16}$$

The even eigenvalues of $A_T(\phi)$ are given by

$$\tilde{\lambda}_i = \frac{(1-\phi)^2}{1 - 2\phi\cos\gamma_i + \phi^2}, \qquad i = 1, \dots, T_2. \tag{B17}$$

The symmetric eigenvector associated with $\tilde{\lambda}_i$ is proportional to u_i, where

$$u_{ij} = \cos\left(\frac{(T-2j+1)\gamma_i}{2}\right), \qquad j = 1, \dots, T. \tag{B18}$$

Note that $\sum_{j=1}^{T} u_{ij} = \sin(T\gamma_i/2)/\sin(\gamma_i/2)$ and $\sum_{j=1}^{T} u_{ij}^2 = [T + \sin(T\gamma_i)/\sin(\gamma_i)]/2$. With this, we can show that the sum of the elements of the even eigenvector is

$$\frac{\sum_{j=1}^{T} u_{ij}}{\left[\sum_{j=1}^{T} u_{ij}^2\right]^{\frac{1}{2}}} = \frac{\sqrt{2}\csc(\gamma_i/2)\sin(T\gamma_i/2)}{[T + \csc(\gamma_i)\sin(T\gamma_i)]^{\frac{1}{2}}}. \tag{B19}$$

Numerical Method for Computing the Eigenvalues of $P'\Sigma P$ when Σ is a Symmetric Toeplitz Matrix: Suppose $\tilde{Q}\tilde{\Lambda}\tilde{Q}'$ is the spectral decomposition of Σ. Denote $T_1 = [T/2]$ and $T_2 = T - T_1$. Since Σ is a symmetric Toeplitz matrix, we can write

$$\tilde{\Lambda} = \begin{bmatrix} \tilde{\Lambda}_o & 0_{T_1 \times T_2} \\ 0_{T_2 \times T_1} & \tilde{\Lambda}_e \end{bmatrix}, \tag{B20}$$

where $\tilde{\Lambda}_o$ and Λ_e are diagonal matrices of the odd and even eigenvalues of Σ, respectively. In addition, we write $\tilde{Q}$ as $\tilde{Q} = [\tilde{Q}_o,\ \tilde{Q}_e]$ where $\tilde{Q}_o$ and $\tilde{Q}_e$ are the matrices of odd and even eigenvectors of Σ, respectively.

Let $x = \tilde{\Lambda}^{\frac{1}{2}} Q' 1_T$. Suppose Λ is a diagonal matrix of the $T-1$ nonzero eigenvalues of $\tilde{\Lambda} - \frac{1}{T}xx'$ and Q_a is a matrix of the corresponding eigenvectors. It is then straightforward to show that (1) Λ is also the matrix of eigenvalues of $P'\Sigma P$ and, (2) $Q = P'\tilde{Q}\tilde{\Lambda}^{\frac{1}{2}} Q_a \Lambda^{-\frac{1}{2}}$ is the matrix of eigenvectors of $P'\Sigma P$. Note that for the odd eigenvectors of Σ, we have $\tilde{Q}_o' 1_T = 0_{T_1}$ so

$$\tilde{\Lambda} - \frac{1}{T}xx' = \begin{bmatrix} \tilde{\Lambda}_o & 0_{T_1 \times T_2} \\ 0_{T_2 \times T_1} & \tilde{\Lambda}_e - \frac{1}{T}x_e x_e' \end{bmatrix}, \tag{B21}$$

where $x_e = \tilde{\Lambda}_e^{\frac{1}{2}} \tilde{Q}_e' 1_T$. It follows that its eigenvalues Λ and eigenvectors Q_a are given by

$$\Lambda = \begin{bmatrix} \tilde{\Lambda}_o & 0_{T_1 \times (T_2-1)} \\ 0_{(T_2-1) \times T_1} & \Lambda_e \end{bmatrix}, \qquad Q_a = \begin{bmatrix} I_{T_1} & 0_{T_1 \times (T_2-1)} \\ 0_{(T_2-1) \times T_1} & Q_b \end{bmatrix}, \tag{B22}$$

where Λ_e are the nonzero eigenvalues of $\tilde{\Lambda}_e - \frac{1}{T}x_e x_e'$ and Q_b is the matrix of the corresponding eigenvectors. Therefore, once $\tilde{Q}$ and $\tilde{\Lambda}$ of Σ are available, we only need to do a rank-one update on $\tilde{\Lambda}_e$ to obtain the eigenvalues and eigenvectors of $P'\Sigma P$.

After we obtain Λ_e and Q_b, we can then write $A = \Lambda^{\frac{1}{2}} Q' P' H H' P Q \Lambda^{\frac{1}{2}}$ as

$$A = Q_a' \tilde{\Lambda}^{\frac{1}{2}} \tilde{Q}' M H H' M \tilde{Q} \tilde{\Lambda}^{\frac{1}{2}} Q_a = \begin{bmatrix} \tilde{\Lambda}_o^{\frac{1}{2}} \tilde{Q}_o' H H' \tilde{Q}_o \tilde{\Lambda}_o^{\frac{1}{2}} & 0_{T_1 \times (T_2-1)} \\ 0_{(T_2-1) \times T_1} & Q_b' \Lambda_e^{\frac{1}{2}} \tilde{Q}_e' M H H' M \tilde{Q}_e \Lambda_e^{\frac{1}{2}} Q_b \end{bmatrix}. \tag{B23}$$

The first block is obtained using the fact that $\tilde{Q}_o' M = \tilde{Q}_o$. The off diagonal blocks are zero matrices because the columns of $A_1 \equiv H' \tilde{Q}_o \tilde{\Lambda}_o^{\frac{1}{2}}$ are antisymmetric (i.e., $JA_1 = -A_1$) and the columns of $A_2 \equiv H' M \tilde{Q}_e \Lambda_e^{\frac{1}{2}} Q_b$ are symmetric (i.e., $JA_2 = A_2$). It follows that $A_1' A_2 = A_1' J' J A_2 = -A_1' A_2$, so $A_1' A_2 = 0_{T_1 \times (T_2-1)}$. Since A is block diagonal, $A - cm\Lambda/(T-1)$ is also block diagonal and the computation time of its eigenvalues is reduced by at least 75%.

References

Ali, Mukhtar M., 1984, Distributions of the sample autocorrelations when observations are from a stationary autoregressive-moving-average process, *Journal of Business and Economic Statistics* 2, 271–278.

Ansley, Craig F., Robert Kohn, and Thomas S. Shively, 1992, Computing p-values of the generalized Durbin-Watson and other invariant test statistics, *Journal of Econometrics* 54, 277–300.

Arbenz, Peter, 1991, Computing eigenvalues of banded symmetric Toeplitz matrices, *SIAM Journal on Scientific Computing* 12, 743–754.

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307–327.

Campbell, John Y., and N. Gregory Mankiw, 1987, Permanent and transitory components in marcoeconomic fluctuations, *American Economic Review* 77, 111–117.

Cantoni, A., and P. Butler, 1976, Eigenvalues and eigenvectors of symmetric centrosymmetric matrices, *Linear Algebra and its Applications* 13, 275–288.

Chen, Willa W., and Rohit S. Deo, 2006, The variance ratio statistic at large horizons, *Econometric Theory* 22, 206–234.

Cochrane, John, 1988, How big is the random walk in GNP? *Journal of Political Economy* 96, 893–920.

Davies, Robert B., 1980, The distribution of a linear combination of χ^2 random variables, *Applied Statistics* 29, 323–333.

Deo, Rohit S., and Matthew Richardson, 2003, On the asymptotic power of the variance ratio test, *Econometric Theory* 19, 231–239.

Dufour, Jean-Marie, and Roch Roy, 1985. Some robust exact results on sample autocorrelations and tests of randomness, *Journal of Econometrics* 29, 257–273.

Dufour, Jean-Marie, and Roch Roy, 1989. Corrigendum, *Journal of Econometrics* 41, 279–281.

Fama, Eugene F., and Kenneth F. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

Farebrother, R. W., 1984, The distribution of a linear combination of χ^2 random variables, *Applied Statistics* 33, 366–369.

Farebrother, R. W., 1990, The distribution of a quadratic form in normal variables, *Applied Statistics* 39, 294–309.

Faust, Jon, 1992, When are variance ratio tests for serial dependence optimal? *Econometrica* 60, 1215–1226.

Gil-Pelzes, J., 1951, Note on the inversion theorem, *Biometrika* 38, 481–482.

Gu, Ming, and Stanley C. Eisenstat, 1994, A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem, *SIAM Journal on Matrix Analysis and Applications* 15, 1266–1276.

Handy, L. Susan, and Jesse L. Barlow, 1994, Numerical solution of the eigenproblem for banded, symmetric Toeplitz matrices, *SIAM Journal on Matrix Analysis and Applications* 15, 205–214.

Hillier, Grant, Raymond Kan, and Xiaolu Wang, 2009, Computationally efficient recursions for top-order invariant polynomials with applications, *Econometric Theory*, 25, 211–242.

Holmquist, Björn, 1996. Expectations of products of quadratic forms in normal variables, *Stochastic Analysis and Applications* 14, 149–164.

Imhof, J. P., 1961, Computing the distribution of quadratic forms in normal variables, *Biometrika* 48, 419–426.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan, 1995, *Continuous Univariate Distributions, Volume 2* (Wiley, New York).

Kan, Raymond, 2008. From moments of sum to moments to product. *Journal of Multivariate Analysis* 99, 542–554.

Li, Ren-Cang, 1993, Solving secular equations stably and efficiently, LAPACK working notes No. 89.

Lo, Andrew W., and Craig A. MacKinlay, 1988, Stock prices do not follow random walks, *Review of Financial Studies* 1, 41–66.

Lo, Andrew W., and Craig A. MacKinlay, 1989, The size and power of the variance ratio test in finite samples, *Journal of Econometrics* 40, 203–238.

Lu, Zeng-Hua, and Maxwell L. King, 2002, Improving the numerical technique for computing the accumlated distribution of a quadratic form in normal variables, *Econometric Reviews* 21, 149–165.

Magnus, Jan R., 1978, The moments of products of quadratic forms in normal variables, *Statistica Neerlandica* 32, 201–210.

Magnus, Jan R., 1979, The expectation of products of quadratic forms in normal variables: the practice, *Statistica Neerlandica* 33, 131–136.

Magnus, Jan R., 1986, The exact moments of a ratio of quadratic forms in normal variables, *Annales d'Économie et de Statistique* 4, 95–109.

Magnus, Jan R., 1990, On certain moments relating to ratios of quadratic forms in normal variables: further results, *Sankhyā: The Indian Journal of Statistics* 52, Series B, Part 1, 1–13.

Mathai, A. M., Serge B. Provost, 1992, *Quadratic Forms in Random Variables: Theory and Applications* (Marcel Dekker, New York).

Meng, Xiao-Li, 2005, From unit root to Stein's estimator to Fisher's k statistics: If you have a moment, I can tell you more, *Statistical Science* 20, 141–162.

Muirhead, Robb J., 1982, *Aspects of multivariate statistical theory* (Wiley, New York).

Perron, Pierre, and Cosme Vodounou, 2005, The variance ratio test: an analysis of size and power based on a continuous-time asymptotic framework, *Econometric Theory* 21, 562–592.

Poterba, James M., and Lawrence H. Summers, 1988, Mean reversion in stock prices — evidence and implications, *Journal of Financial Economics* 22, 27–59.

Richardson, Matthew, and James Stock, 1989, Drawing inferences from statistics based on multi-year asset returns, *Journal of Financial Economics* 25, 323–348.

Richardson, Matthew, and Tom Smith, 1991, Tests of financial models in the presence of overlapping observations, *Review of Financial Studies* 4, 227–254.

Shiller, Robert J., 1981, The use of volatility measures in assessing market efficiency, *Journal of Finance* 36, 291–304.

Summers, Lawrence H., 1986, Does the stock market rationally reflect fundamental values? *Journal of Finance* 41, 591–600.

Tian, G., Y. Zhang, and W. Huang, 1999, A note on the exact distributions of variance ratio statistics, working paper, Beijing University.

Tse, Y. K., K. W. Ng, and Xibin Zhang, 2004, A small sample overlapping variance-ratio test, *Journal of Time Series Analysis* 25, 127–135.

Trench, William F., 1985, On the eigenvalue problem for Toeplitz band matrices, *Linear Algebra and its Applications* 64, 199–214.

Trench, William F., 1988, Numerical solution of the eigenvalue problem for symmetric rationally generated Toeplitz matrices, *SIAM Journal on Matrix Analysis and Applications* 9, 291–303.

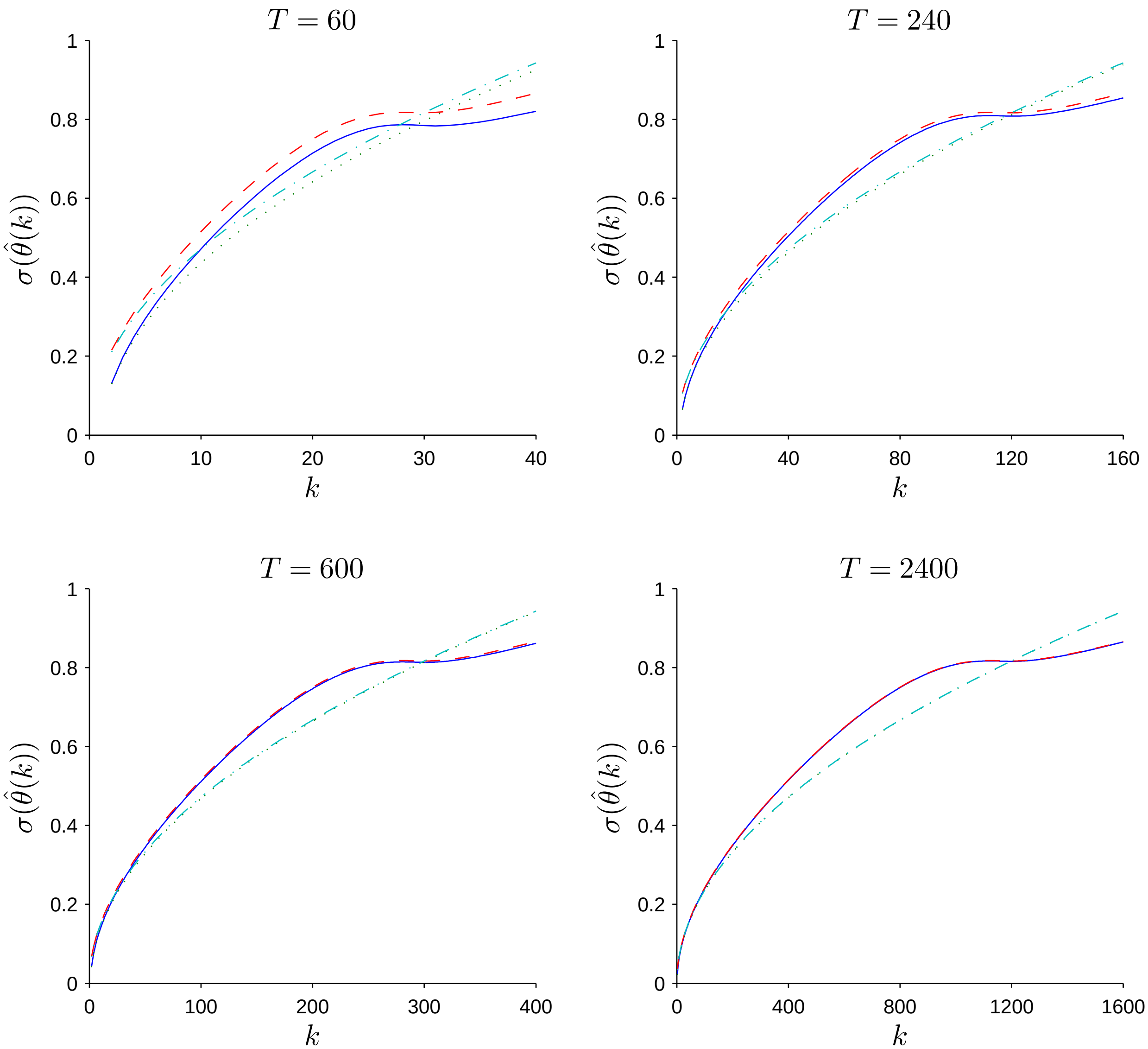

Figure 1: Plots of Exact and Asymptotic Standard Errors of Sample Variance Ratio as a Function of the Length of the Multi-period Return

The figure presents the plots of standard errors of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are uncorrelated and jointly elliptically distributed with constant mean and variance. The solid line represents the exact finite sample standard error. The dotted line represents the standard error based on the fixed-k asymptotic theory. The dashed line represents the standard error based on the fixed-k/T asymptotic theory. The dashed-dotted line represents the standard error based on the zero-k/T asymptotic theory.

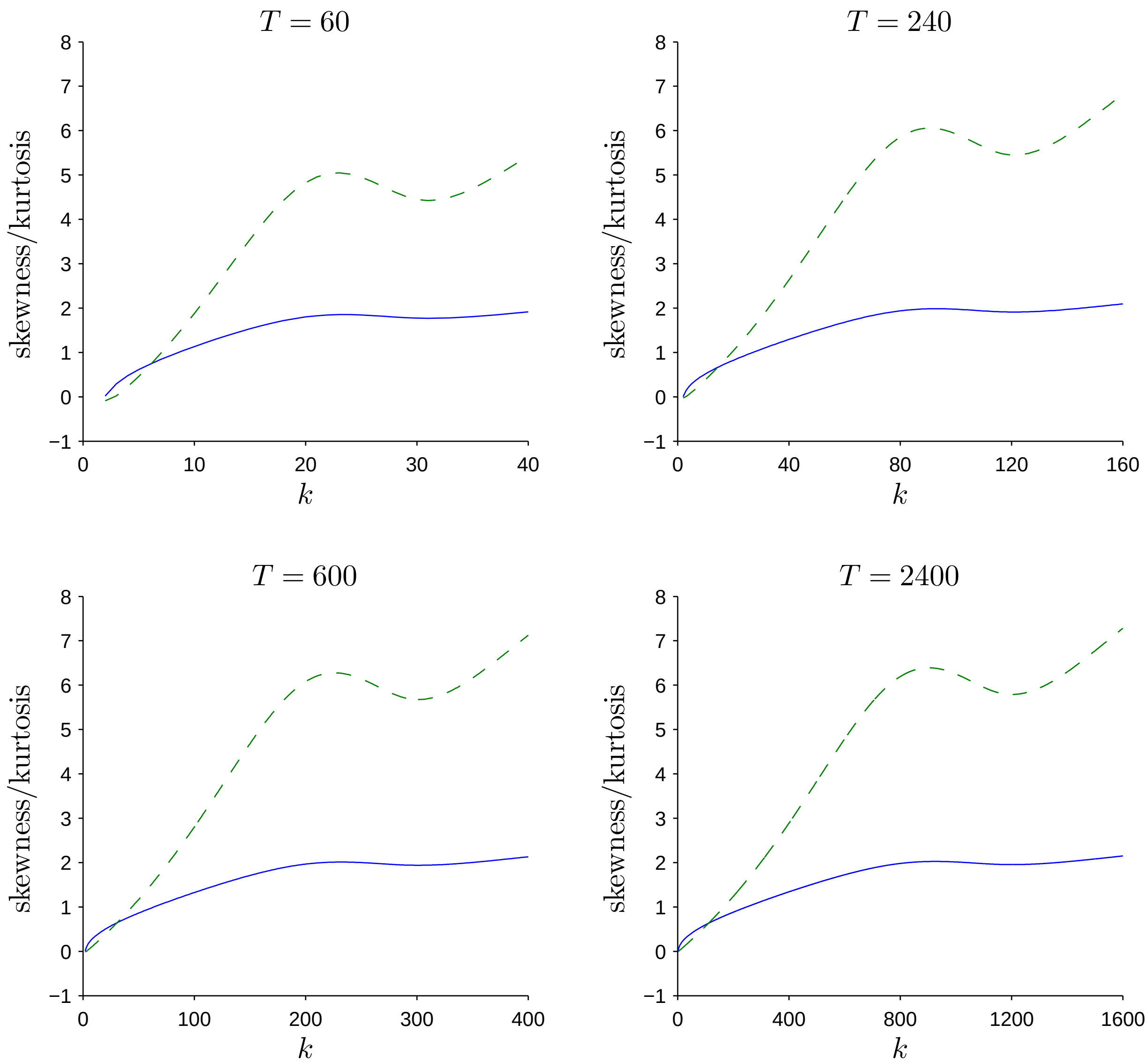

Figure 2: Plots of Coefficients of Skewness and Excess Kurtosis of Sample Variance Ratio as a Function of the Length of the Multi-period Return

The figure presents the plots of the coefficients of skewness $(E[(\hat{\theta}(k)-1)^3]/\mathrm{Var}[\hat{\theta}(k)]^{\frac{3}{2}})$ and excess kurtosis $(E[(\hat{\theta}(k)-1)^4]/\mathrm{Var}[\hat{\theta}(k)]^2-3)$ of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are uncorrelated and jointly elliptically distributed with constant mean and variance. The solid line represents the coefficient of skewness and the dashed line represents the coefficient of excess kurtosis.

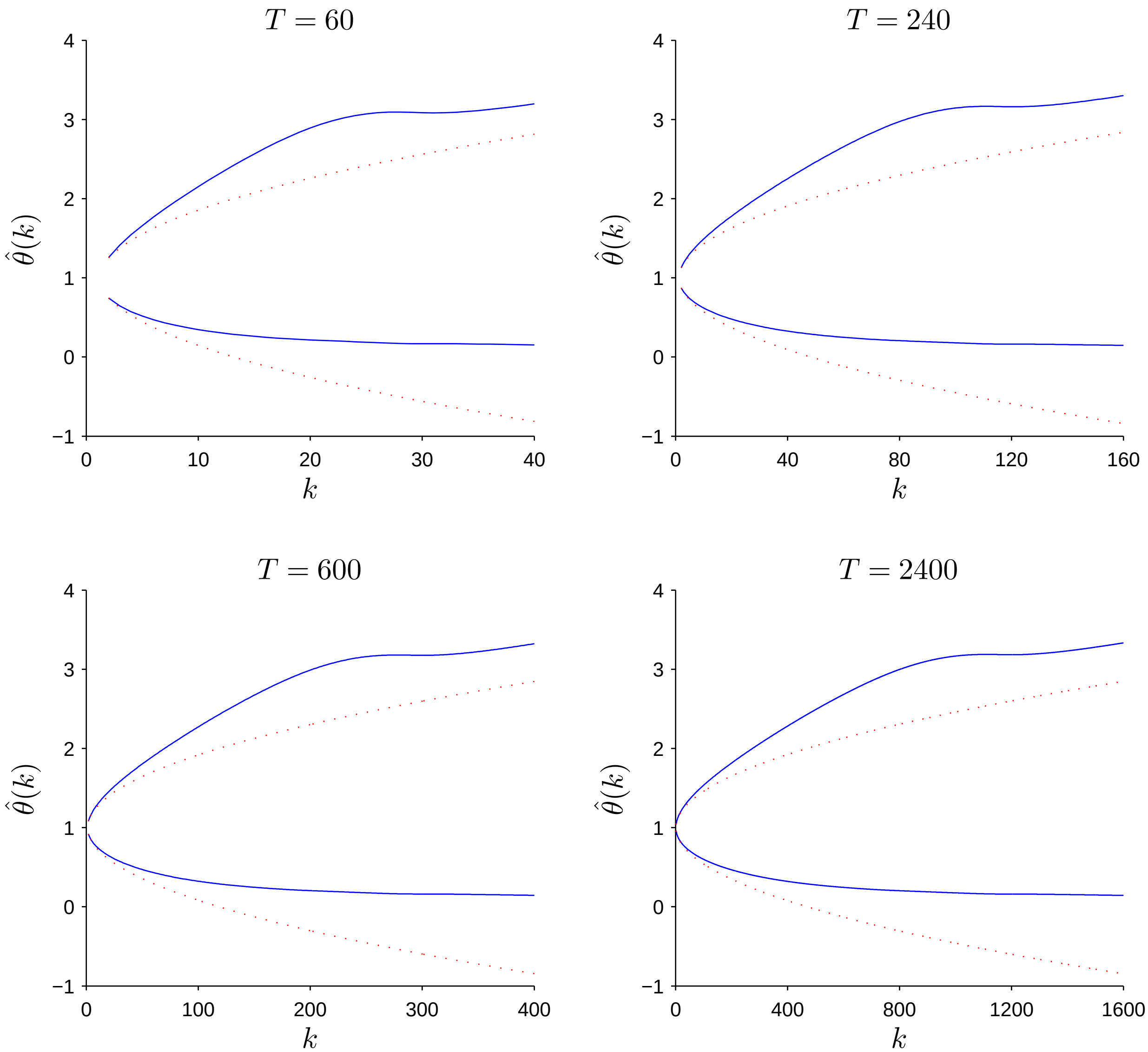

Figure 3: Plots of Lower and Upper 2.5 Percentiles of Sample Variance Ratio as a Function of the Length of the Multi-period Return

The figure presents the plots of lower and upper 2.5 percentiles of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are uncorrelated and jointly elliptically distributed with constant mean and variance. The solid lines represent the percentiles based on the finite sample distribution. The dotted lines represent the percentiles based on the fixed-k asymptotic theory.

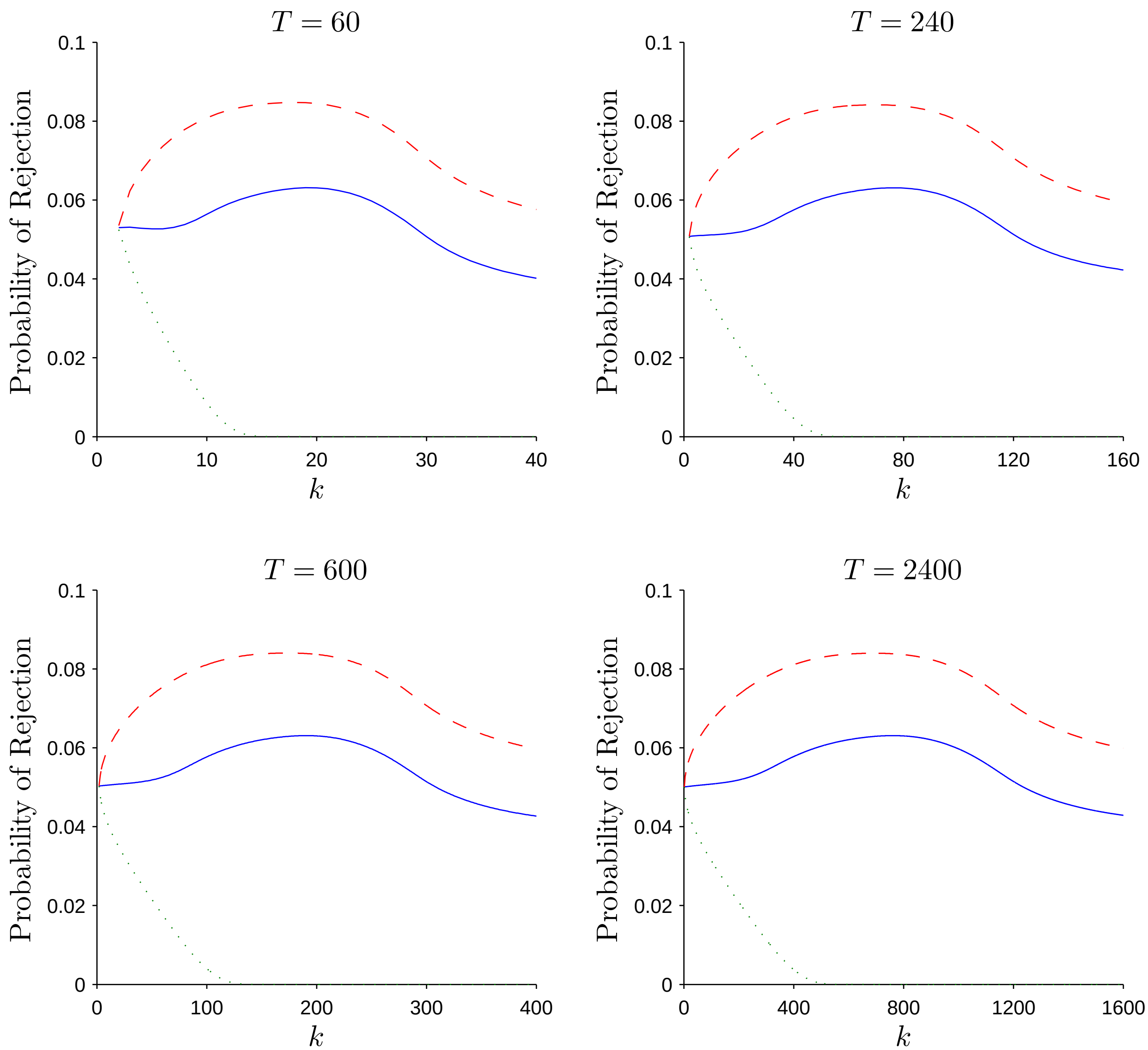

Figure 4: Actual Probabilities of Rejection when Using Asymptotic Tests of Variance Ratio as a Function of the Length of the Multi-period Return

The figure presents the plots of actual probabilities of rejection of fixed-k asymptotic tests of $H_0 : \theta(k) = 1$ as a function of k for different lengths of the sample period (T), where $\theta(k)$ is the variance ratio and k is the length of the multi-period return. The returns are assumed to be uncorrelated and jointly elliptically distributed with constant mean and variance, and the nominal size of the tests is 5%. The solid line represents the actual size of a two-tailed test, the dotted line represents the actual size of a left-tailed test, and the dashed line represents the actual size of a right-tailed test.

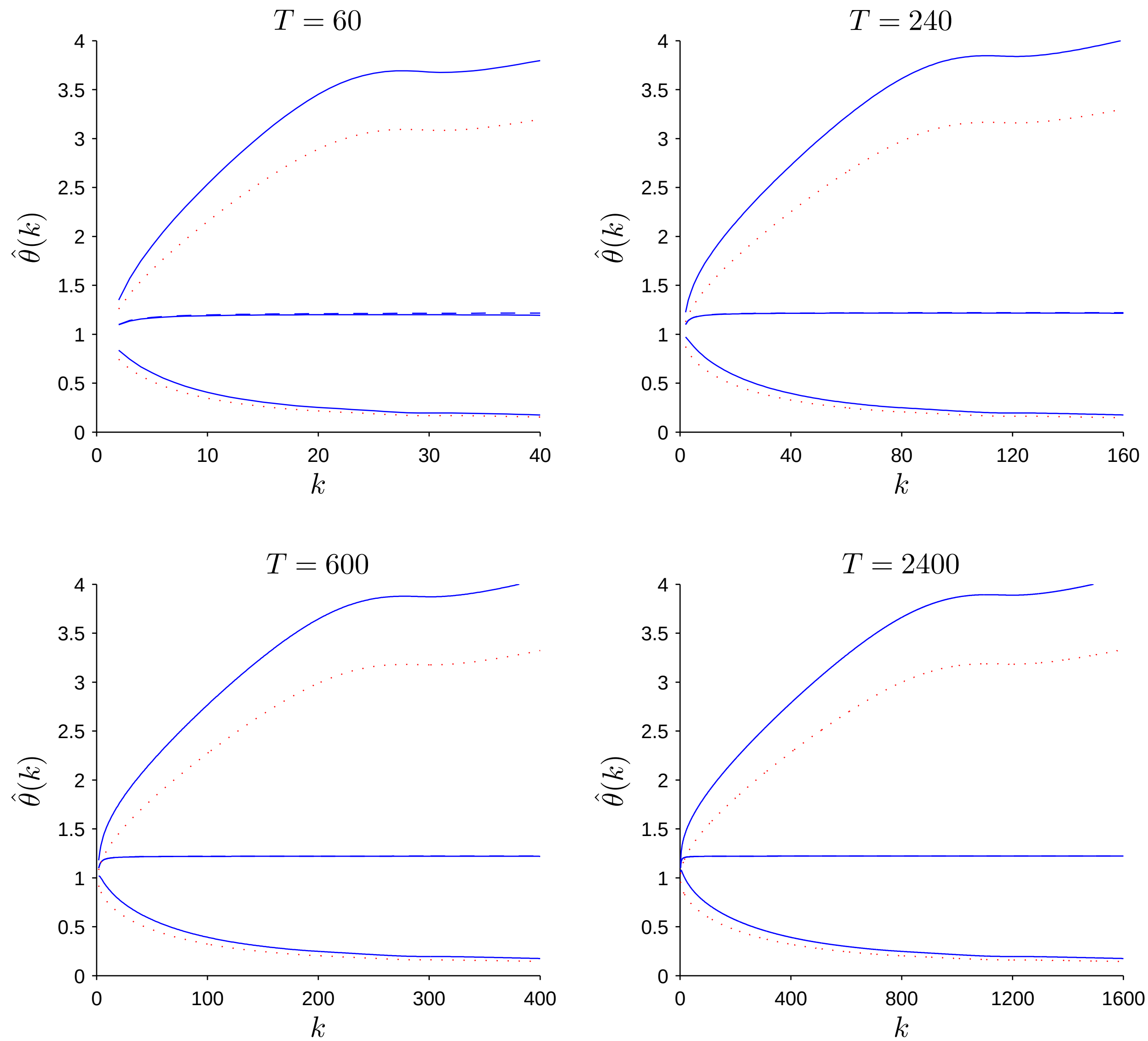

Figure 5: Exact Distribution of Sample Variance Ratio as a Function of the Length of the Multi-period Return when Return Follows an AR(1) Process with $\phi = 0.1$
The figure presents the distribution of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and follow an AR(1) process with parameter $\phi = 0.1$. The three solid lines from bottom to top are the lower 2.5 percentile, the mean, and the upper 2.5 percentile of the sample variance ratio. The dashed line is the population variance ratio, and the dotted lines are the lower and upper 2.5 percentiles of the sample variance ratio under the null hypothesis of random walk.

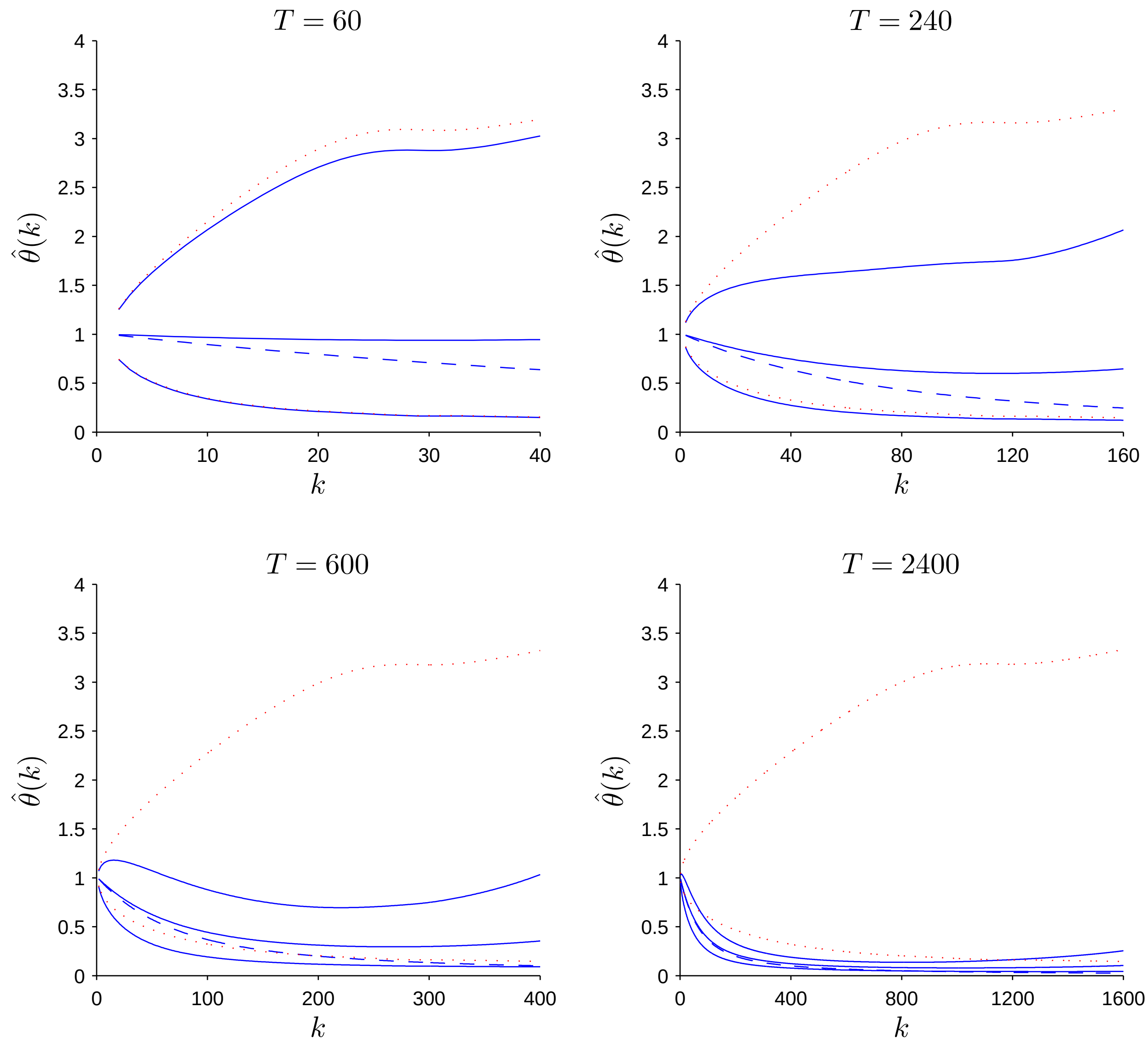

Figure 6: Exact Distribution of Sample Variance Ratio as a Function of the Length of the Multi-period Return when Stock Price Follows an AR(1) Process with $\phi = 0.975$
The figure presents the distribution of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and stock price follows an AR(1) process with parameter $\phi = 0.975$. The three solid lines from bottom to top are the lower 2.5 percentile, the mean, and the upper 2.5 percentile of the sample variance ratio. The dashed line is the population variance ratio, and the dotted lines are the lower and upper 2.5 percentiles of the sample variance ratio under the null hypothesis of random walk.

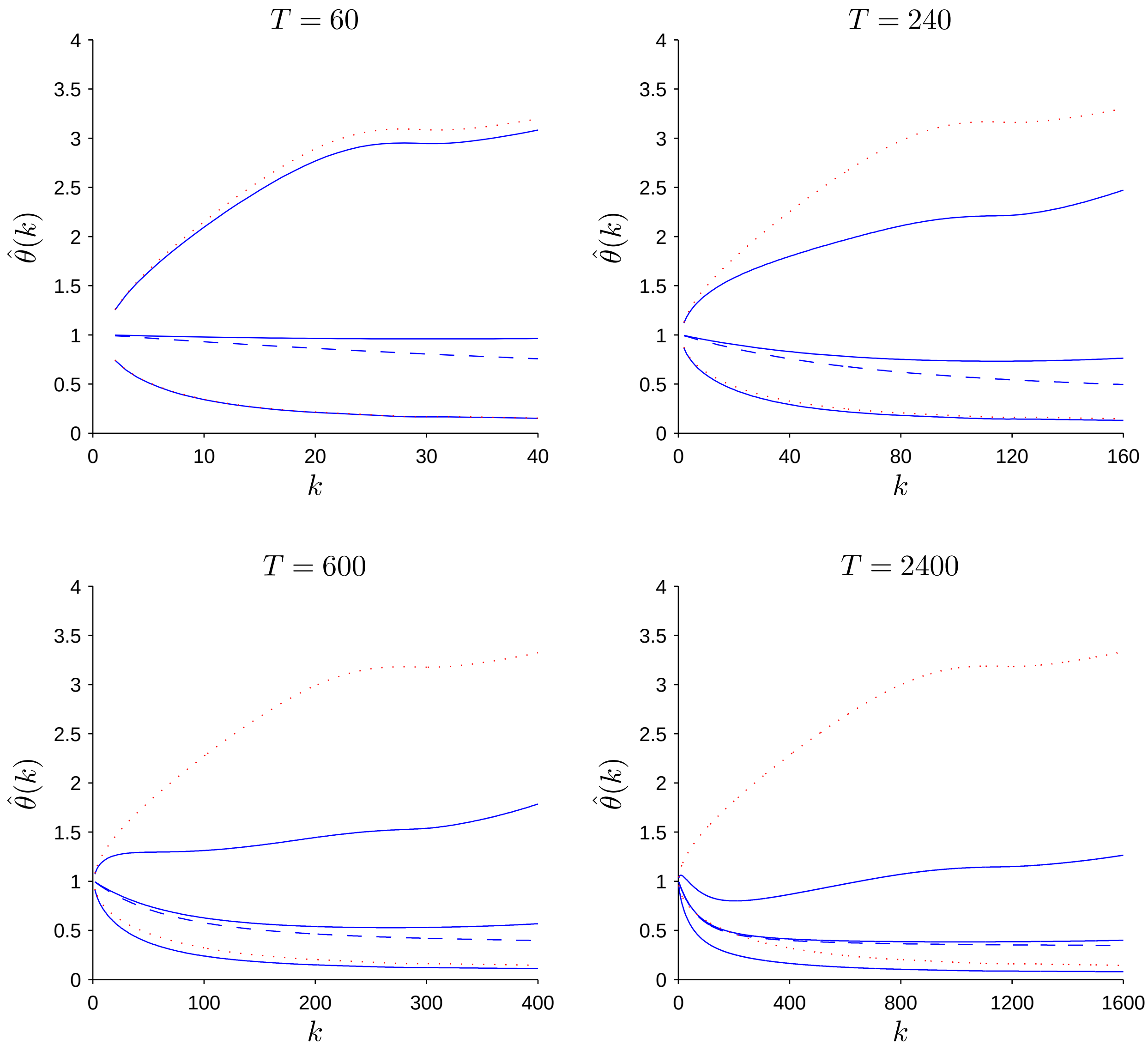

Figure 7: Exact Distribution of Sample Variance Ratio as a Function of the Length of the Multi-period Return when Stock Price is the Sum of a Random Walk Component and an AR(1) Component with $\phi = 0.975$

The figure presents the distribution of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and stock price is the sum of a random walk component and an AR(1) component with parameter $\phi = 0.975$. The variance of the innovation of the random walk component is assumed to be half of the variance of the innovation of the AR(1) component. The three solid lines from bottom to top are the lower 2.5 percentile, the mean, and the upper 2.5 percentile of the sample variance ratio. The dashed line is the population variance ratio, and the dotted lines are the lower and upper 2.5 percentiles of the sample variance ratio under the null hypothesis of random walk.

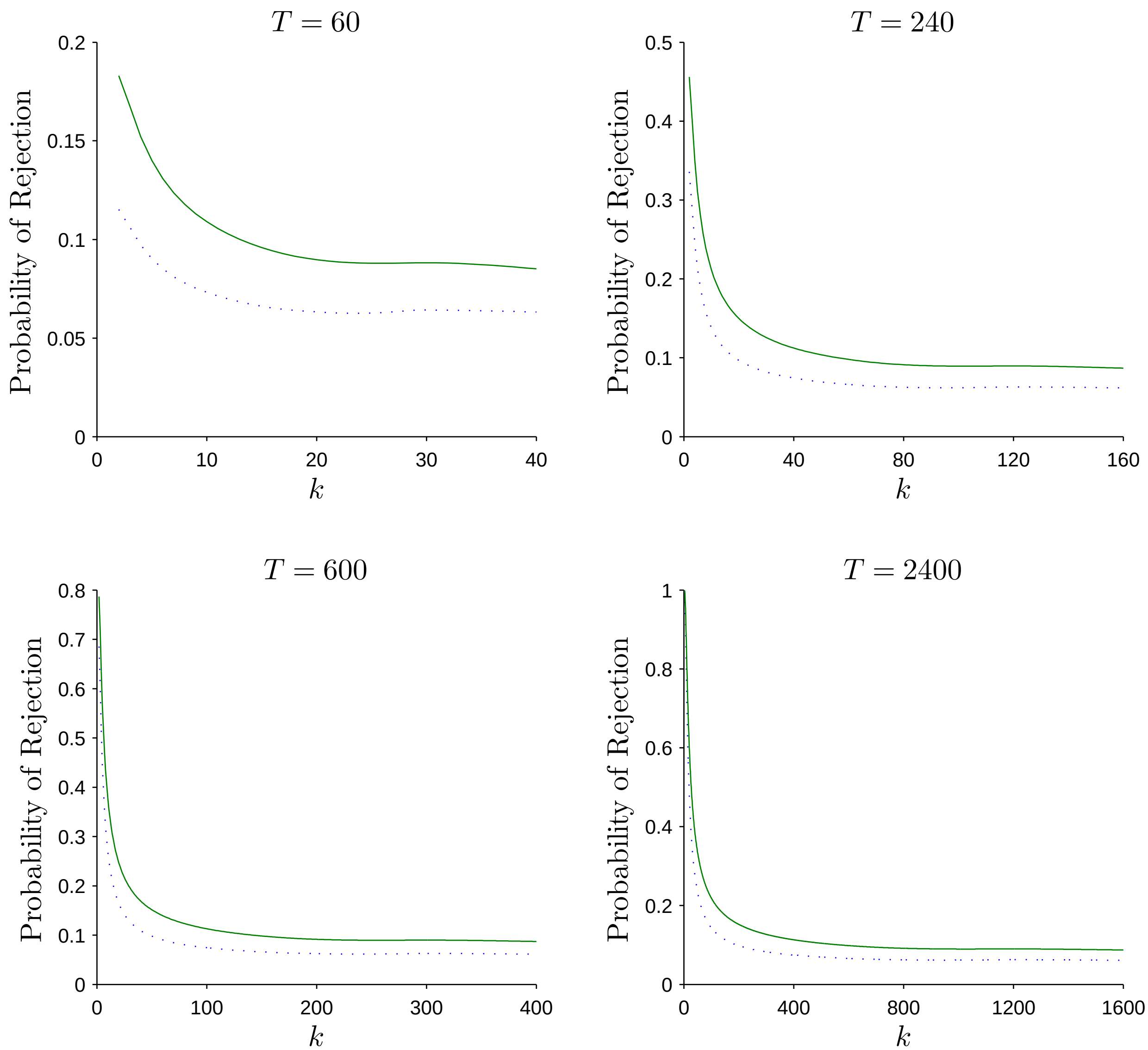

Figure 8: Probabilities of Rejection Using Exact Tests on Variance Ratio as a Function of the Length of the Multi-period Return when Return Follows an AR(1) Process with $\phi = 0.1$

The figure presents the actual probabilities of rejection of using exact sample variance ratio test as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and follow an AR(1) process with parameter $\phi = 0.1$. The solid lines represent the probabilities of rejection for the one-tailed test and the dotted lines represent the probabilities of rejection for the two-tailed test. The nominal size of the tests is 5%.

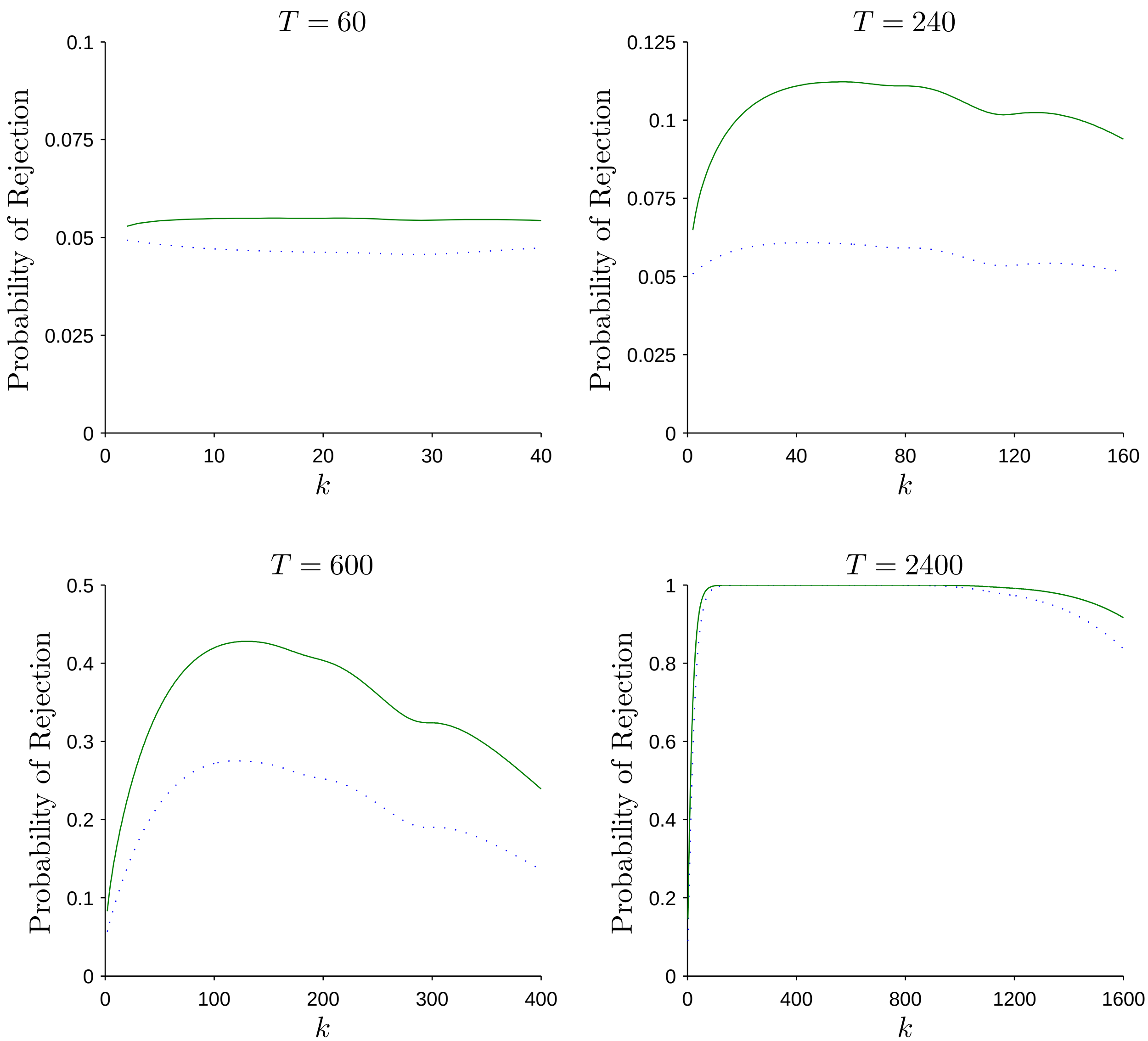

Figure 9: Probabilities of Rejection Using Exact Tests on Variance Ratio as a Function of the Length of the Multi-period Return when Stock Price Follows an AR(1) Process with $\phi = 0.975$

The figure presents the actual probabilities of rejection of using exact sample variance ratio test as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and stock price follows an AR(1) process with parameter $\phi = 0.975$. The solid lines represent the probabilities of rejection for the one-tailed test and the dotted lines represent the probabilities of rejection for the two-tailed test. The nominal size of the tests is 5%.

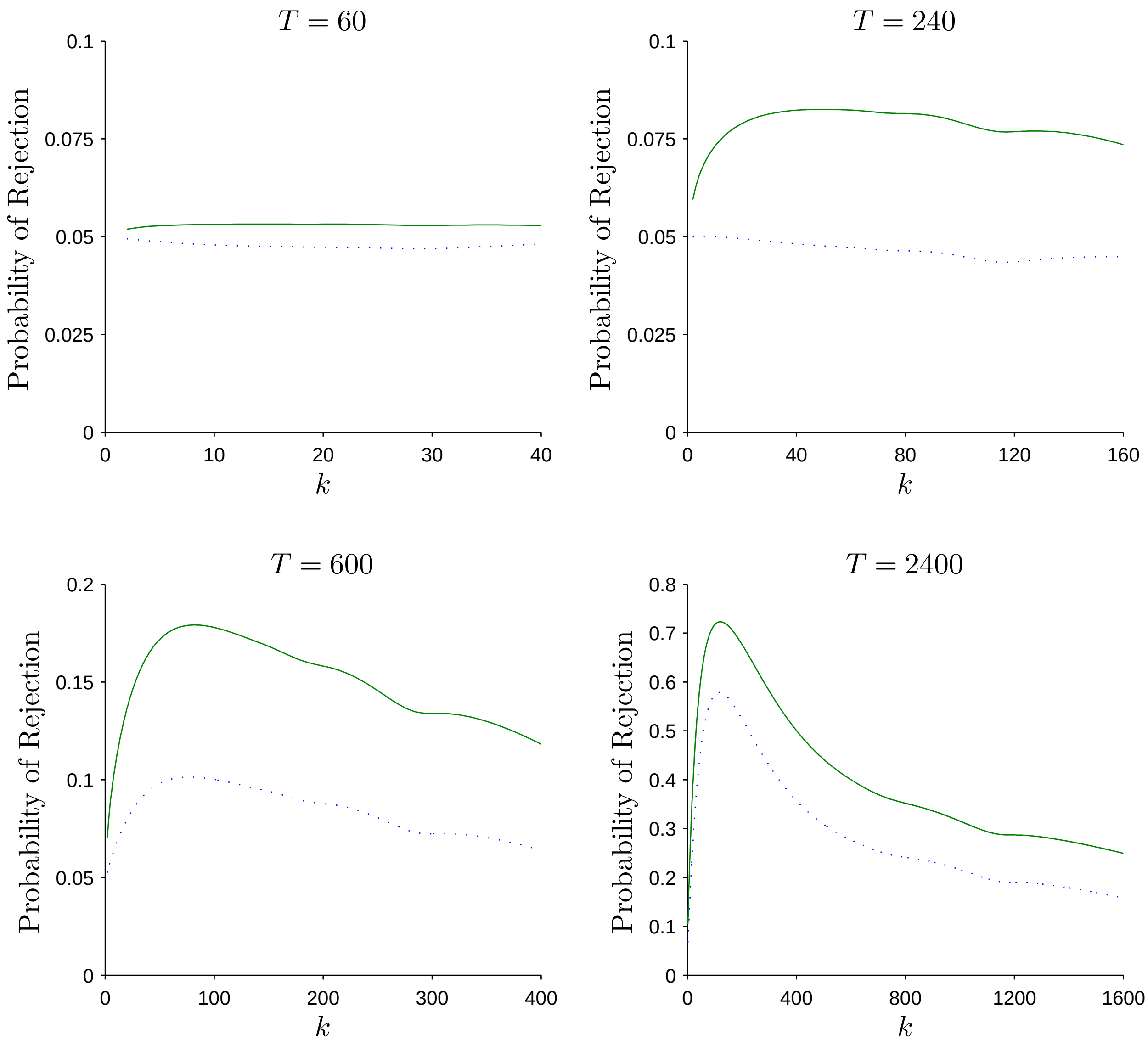

Figure 10: Probabilities of Rejection Using Exact Tests on Variance Ratio as a Function of the Length of the Multi-period Return when Stock Price is the Sum of a Random Walk Component and an AR(1) Component with $\phi = 0.975$

The figure presents the actual probabilities of rejection of using exact sample variance ratio test as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and stock price is the sum of a random walk component and an AR(1) component with parameter $\phi = 0.975$. The variance of the innovation of the random walk component is assumed to be half of the variance of the innovation of the AR(1) component. The solid lines represent the probabilities of rejection for the one-tailed test and the dotted lines represent the probabilities of rejection for the two-tailed test. The nominal size of the tests is 5%.

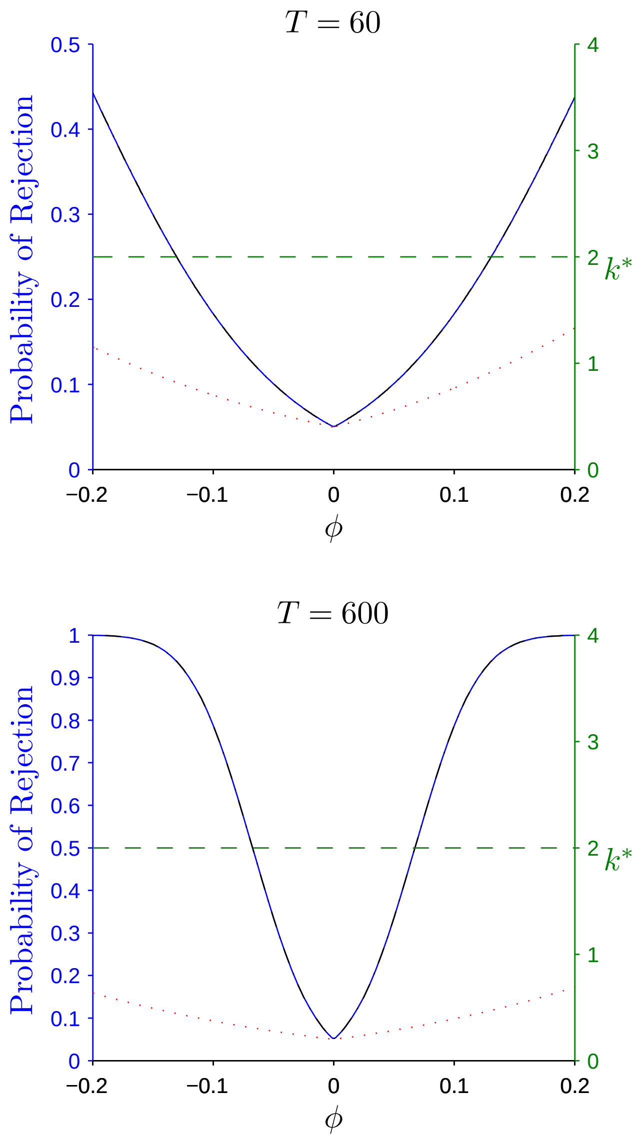

Figure 11: Optimal Length of Multi-period Return when Return Follows an AR(1) Process

The figure presents the optimal length of multi-period return (k^*) for exact variance ratio test as a function of the AR(1) parameter ϕ for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and follow an AR(1) process with parameter ϕ. The dashed line (with the scale on the right hand side of the graph) represents the optimal k^* as a function of ϕ. The solid line represents the probability of rejection under the optimal choice of k. The dashed dotted line (not visible because it overlaps with the dashed line) represents the probability of rejection under the choice of $k = 2$. The dotted line represents the probability of rejection under the choice of $k = T/4$. The k^* and the probabilities of rejection are for the one-tailed test with a nominal size of 5%.

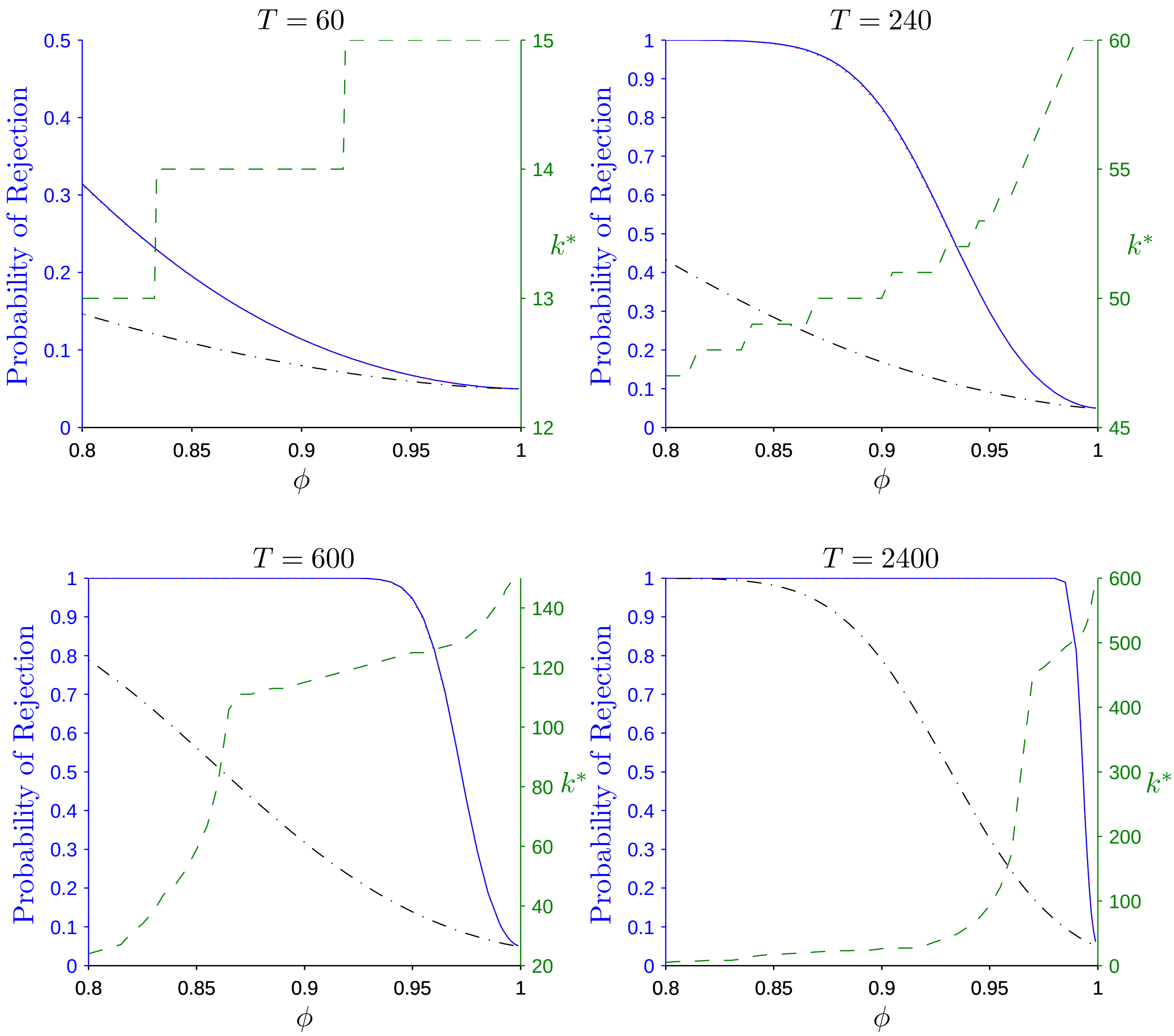

Figure 12: Optimal Length of Multi-period Return when Stock Price Follows an AR(1) Process

The figure presents the optimal choice of length of multi-period return (k^*) for exact variance ratio test as a function of the AR(1) parameter ϕ for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and stock price follows an AR(1) process. The dashed line (with the scale on the right hand side of the graph) represents the optimal k^* as a function of ϕ. The solid line represents the probability of rejection under the optimal choice of k. The dashed dotted line represents the probability of rejection under the choice of $k = 2$. The dotted line (not visible because it overlaps with the solid line) represents the probability of rejection under the choice of $k = T/4$. The k^* and the probabilities of rejection are for the one-tailed test with a nominal size of 5%.

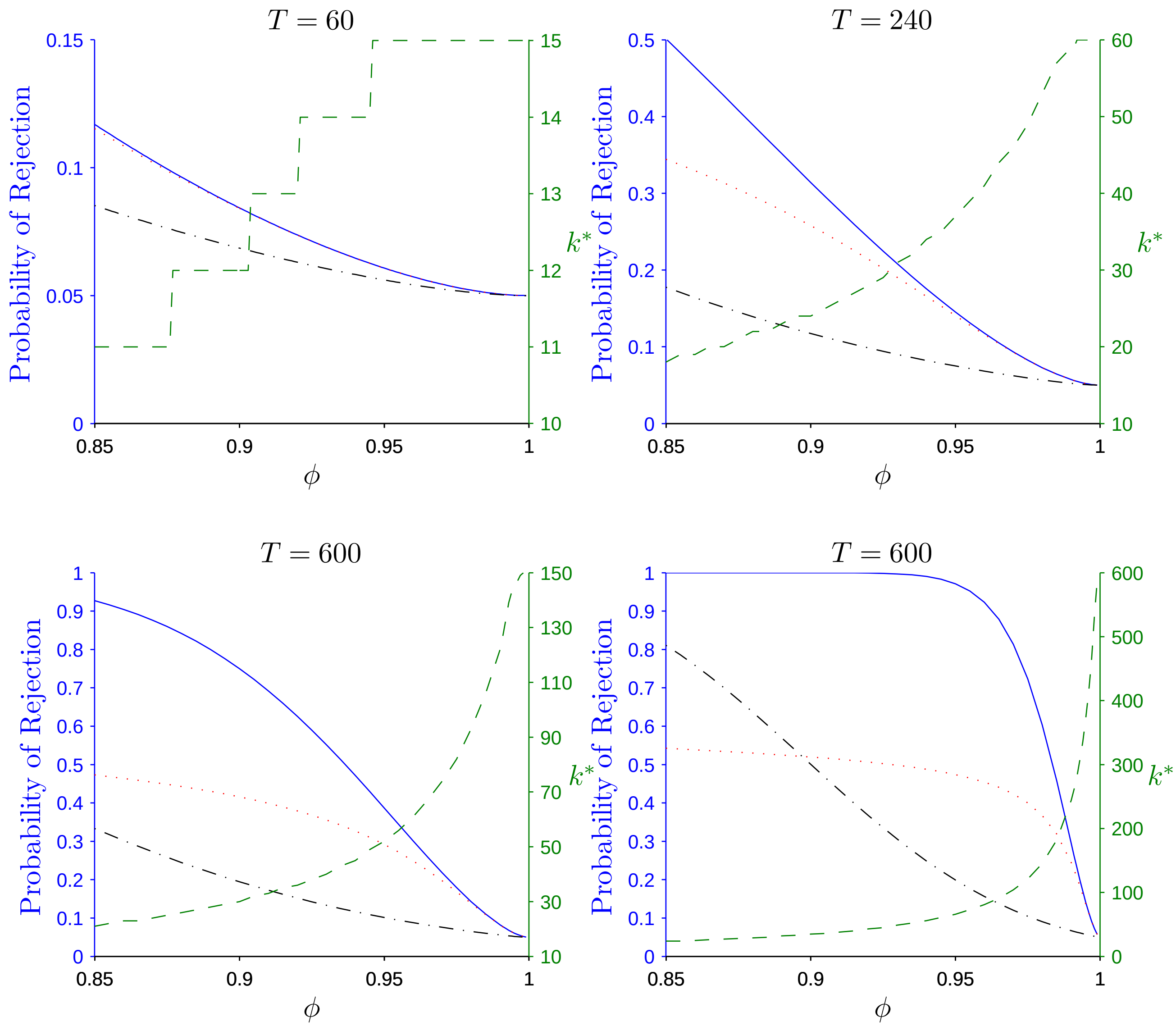

Figure 13: Optimal Length of Multi-period Return when Stock Price is the Sum of a Random Walk Component and an AR(1) Component

The figure presents the optimal length of multi-period return (k^*) for exact variance ratio test as a function of the AR(1) parameter ϕ for different lengths of the sample period (T) under the assumption that returns are jointly elliptically distributed and stock price is the sum of a random walk component and an AR(1) component with parameter ϕ. The variance of the innovation of the random walk component is assumed to be half of the variance of the innovation of the AR(1) component. The dashed line (with the scale on the right hand side of the graph) represents the optimal k^* as a function of ϕ. The solid line represents the probability of rejection under the optimal choice of k. The dashed dotted line represents the probability of rejection under the choice of $k = 2$. The dotted line represents the probability of rejection under the choice of $k = T/4$. The k^* and the probabilities of rejection are for the one-tailed test with a nominal size of 5%.

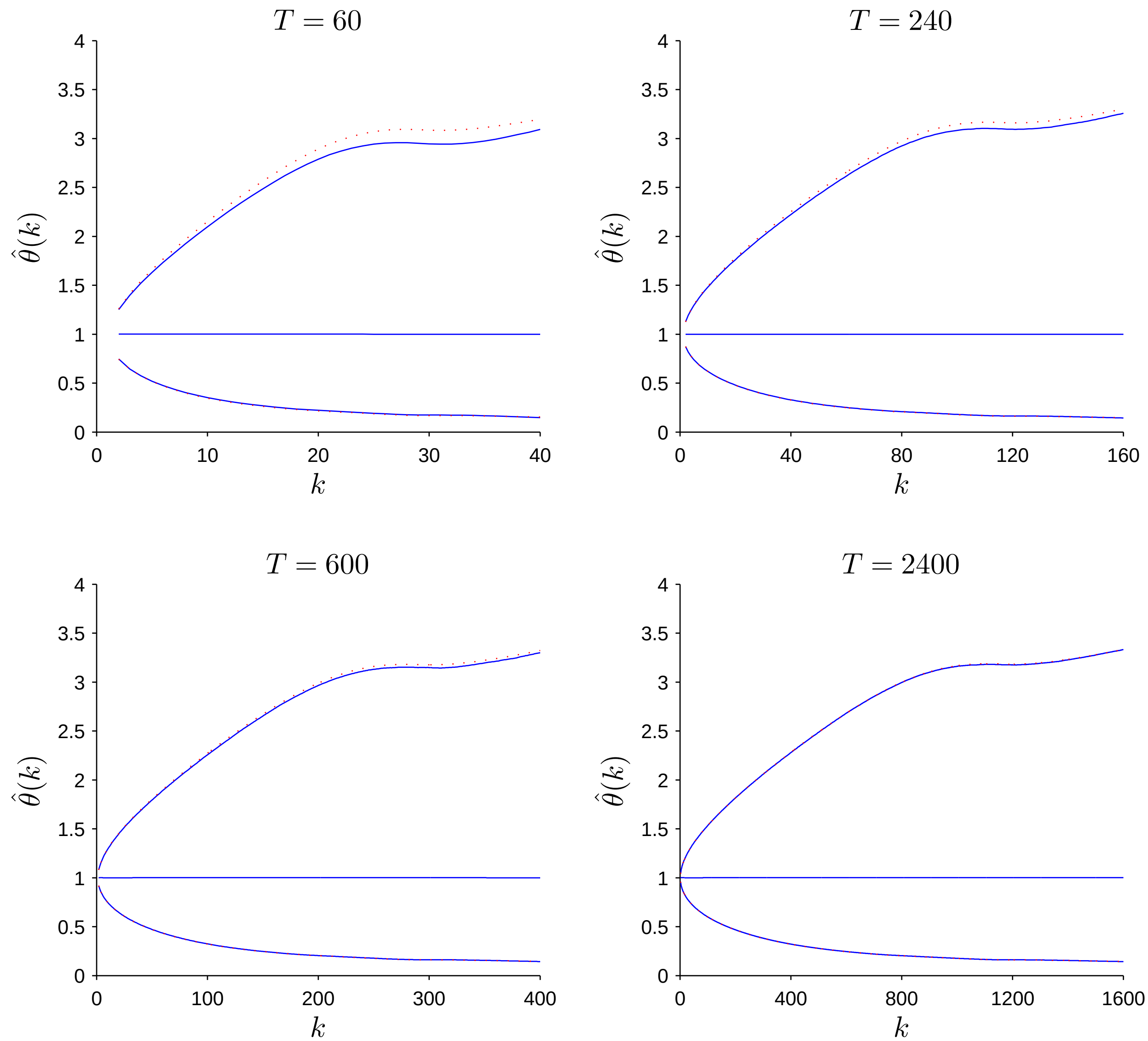

Figure 14: Plots of Lower and Upper 2.5 Percentiles of Sample Variance Ratio as a Function of the Length of the Multi-period Return when Return Follows an Empirical Distribution of the Daily Returns of the NYSE

The figure presents the plots of lower and upper 2.5 percentiles of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that the returns are independently drawn from an empirical distribution based on the daily returns of the NYSE from 1926/1/2–2009/12/31. The three solid lines from bottom to top are the lower 2.5 percentile, the mean and the 97.5 percentile of the sample variance ratio under this distribution assumption. The dotted lines represent the lower and upper 2.5 percentiles of the sample variance ratio when the variance of the return is constant over time.

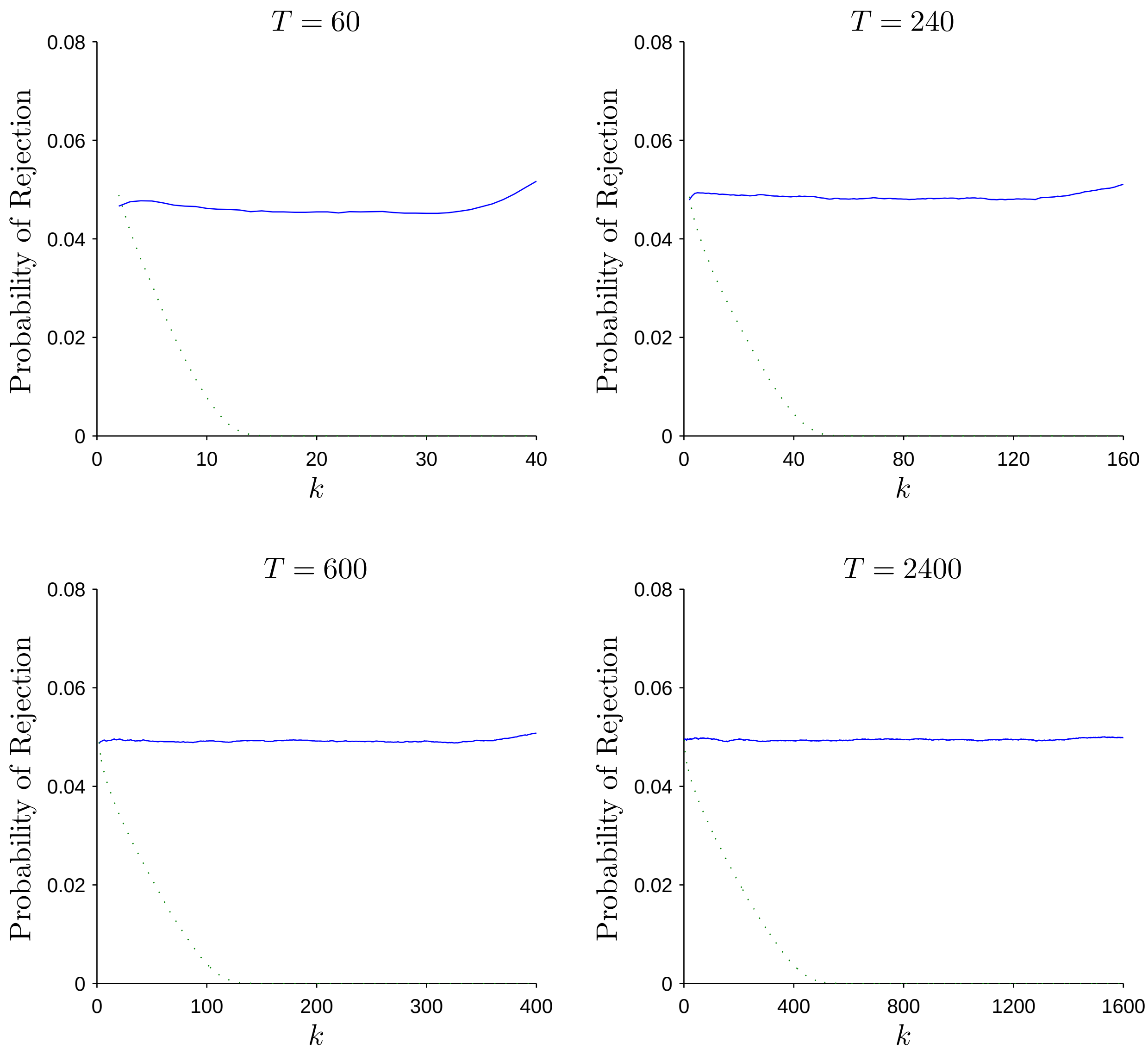

Figure 15: Actual Probabilities of Rejection for Various Tests on Variance Ratio as a Function of the Length of the Multi-period Return when Return Follows an Empirical Distribution of the Daily Returns of the NYSE

The figure presents the plots of actual probabilities of rejection of the finite sample test and the fixed-k asymptotic test of $H_0 : \theta(k) = 1$ as a function of k for different lengths of the sample period (T), where $\theta(k)$ is the variance ratio and k is the length of the multi-period return. The returns are assumed to be independently drawn from an empirical distribution based on the daily returns of the NYSE from 1926/1/2–2009/12/31. The solid line represents the actual size of a left-tailed test using the finite sample distribution under joint elliptical distribution assumption, and the dotted line represents the actual size of a left-tailed test using the fixed k asymptotic test that assumes homoskedasticity. The nominal size of the tests is 5%.

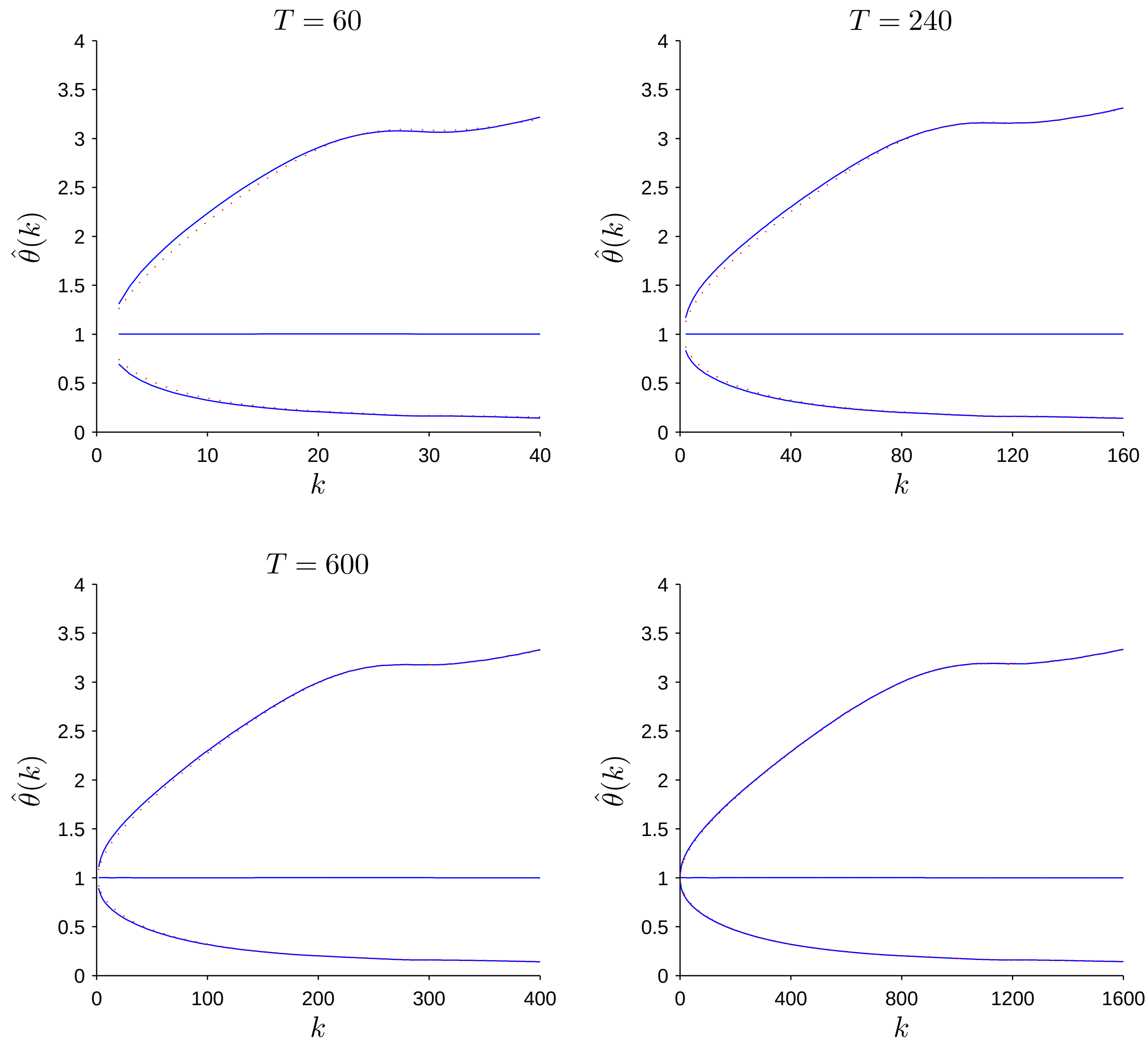

Figure 16: Plots of Lower and Upper 2.5 Percentiles of Sample Variance Ratio as a Function of the Length of the Multi-period Return when Return Variance Follows an AR(1) Process

The figure presents the plots of lower and upper 2.5 percentiles of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that returns are uncorrelated over time but the natural log of its variance follows an AR(1) process with an AR(1) parameter of 0.5. The three solid lines from bottom to top are the lower 2.5 percentile, the mean and the 97.5 percentile of the sample variance ratio under this distribution assumption. The dotted lines represent the lower and upper 2.5 percentiles of the sample variance ratio when the variance of the return is constant over time.

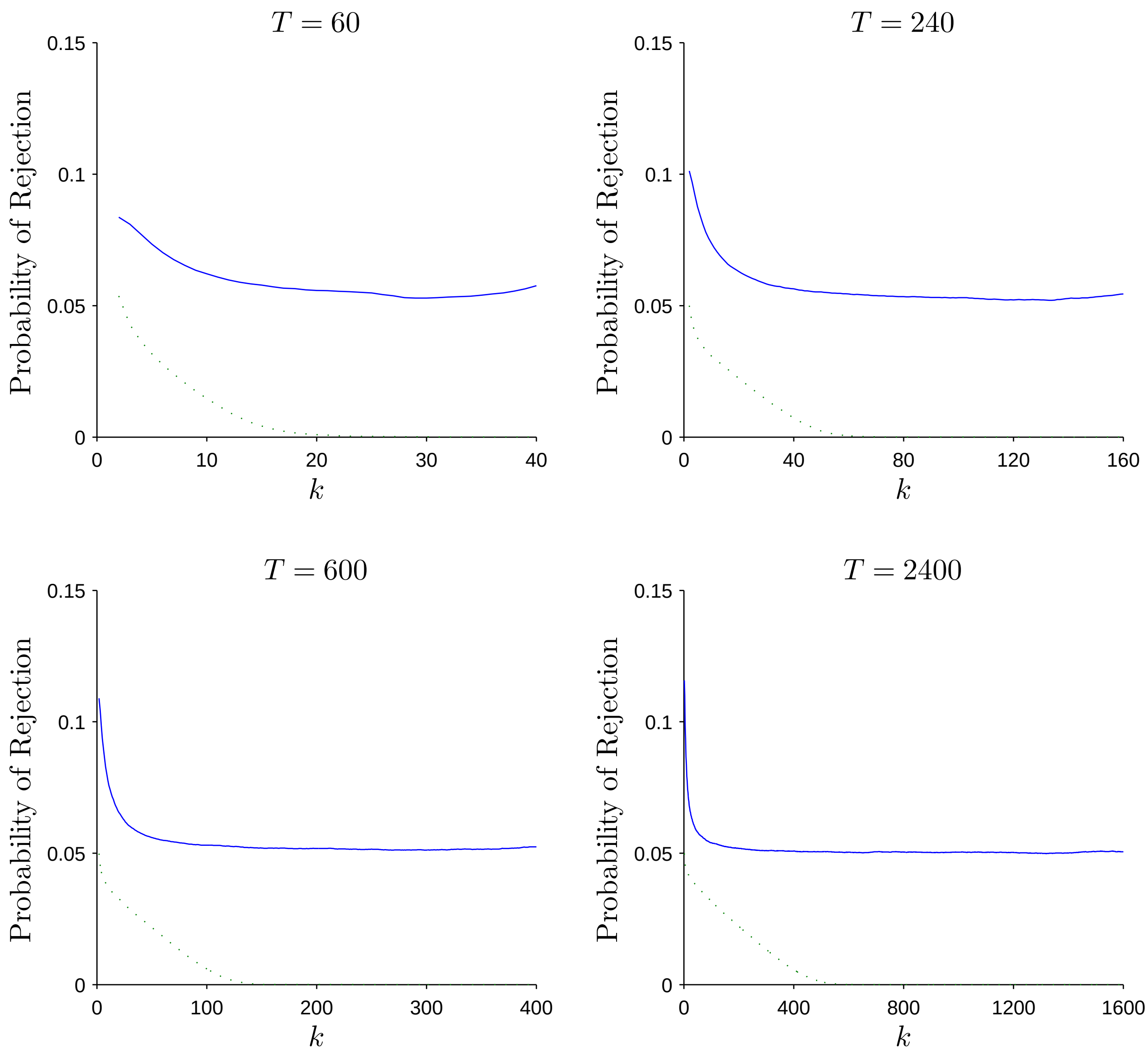

Figure 17: Actual Probabilities of Rejection for Various Tests on Variance Ratio as a Function of the Length of the Multi-period Return when Return Variance Follows an AR(1) Process

The figure presents the plots of actual probabilities of rejection of the finite sample test and Lo and MacKinlay (1988) fixed-k (heteroskedasticity consistent) asymptotic test of $H_0 : \theta(k) = 1$ as a function of k for different lengths of the sample period (T), where $\theta(k)$ is the variance ratio and k is the length of the multi-period return. The returns are assumed to be uncorrelated over time but the natural log of its variance follows an AR(1) process with an AR(1) parameter of 0.5. The solid line represents the actual size of a left-tailed test using the finite sample distribution under joint elliptical distribution assumption, and the dotted line represents the actual size of a left-tailed test using Lo and MacKinlay (1988) test. The nominal size of the tests is 5%.

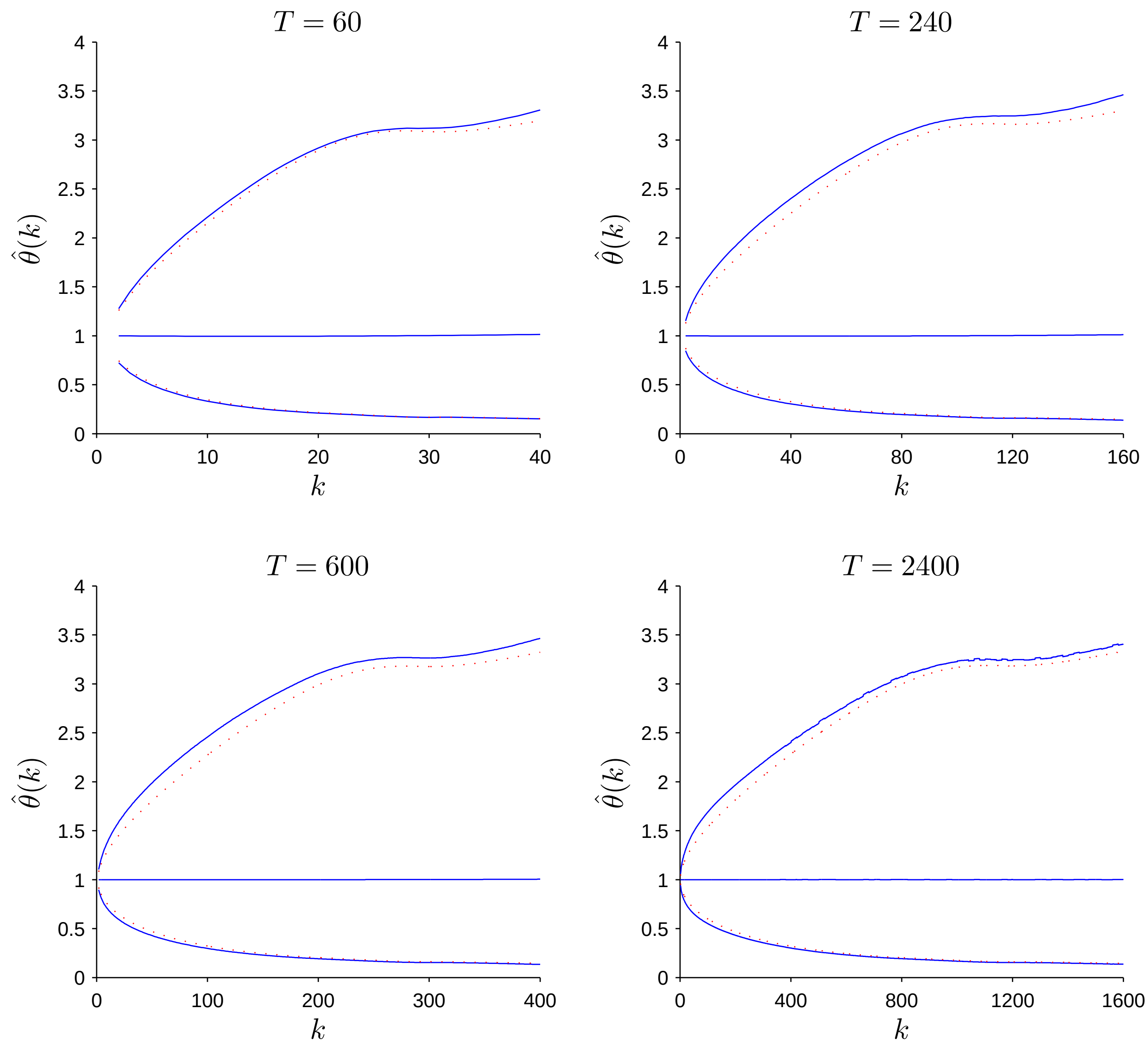

Figure 18: Plots of Lower and Upper 2.5 Percentiles of Sample Variance Ratio as a Function of the Length of the Multi-period Return when Return Follows a GARCH(1,1) Model

The figure presents the plots of lower and upper 2.5 percentiles of sample variance ratio $\hat{\theta}(k)$ as a function of the length of the multi-period return (k) for different lengths of the sample period (T) under the assumption that return follows a GARCH(1,1) model with parameters $\alpha = 0.1171$ and $\beta = 0.8575$. The three solid lines from bottom to top are the lower 2.5 percentile, the mean and the 97.5 percentile of the sample variance ratio under this GARCH(1,1) assumption on returns. The dotted lines represent the lower and upper 2.5 percentiles of the sample variance ratio when the variance of the return is constant over time.

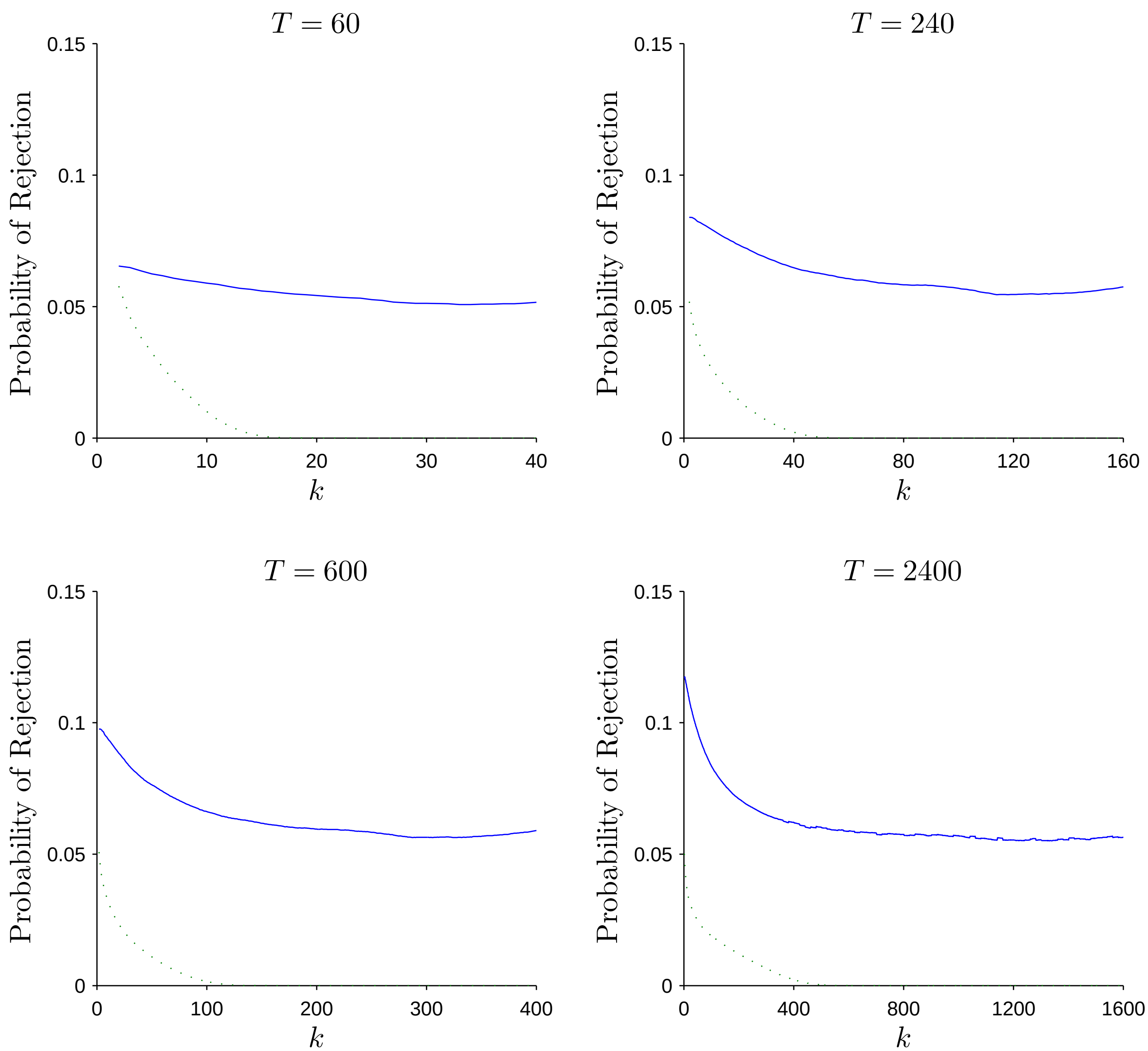

Figure 19: Actual Probabilities of Rejection for Various Tests on Variance Ratio as a Function of the Length of the Multi-period Return when Return Follows a GARCH(1,1) Model

The figure presents the plots of actual probabilities of rejection of the finite sample test and Lo and MacKinlay (1988) fixed-k (heteroskedasticity consistent) asymptotic test of $H_0 : \theta(k) = 1$ as a function of k for different lengths of the sample period (T), where $\theta(k)$ is the variance ratio and k is the length of the multi-period return. The return is assumed to follow a GARCH(1,1) model with parameters $\alpha = 0.1171$ and $\beta = 0.8575$. The solid line represents the actual size of a left-tailed test using the finite sample distribution under joint elliptical distribution assumption, and the dotted line represents the actual size of a left-tailed test using Lo and MacKinlay (1988) test. The nominal size of the tests is 5%.

Table I

Approximation Errors of Various Asymptotic Standard Errors of Variance Ratios

The table presents the approximation errors (in percentage) of three different asymptotic standard errors of sample variance ratio for different lengths of sample period (T) and lengths of the multi-period returns (k) under the assumption that returns are uncorrelated and jointly elliptically distributed with the same mean and variance. The standard errors in Panel A are based on the asymptotic theory that assumes k is fixed but $T \to \infty$. The standard errors in Panel B are based on the asymptotic theory that assumes both $k \to \infty$ and $T \to \infty$ but k/T is fixed. The standard errors in Panel C are based on the asymptotic theory that assumes both $k \to \infty$ and $T \to \infty$ but $k/T \to 0$.

Panel A: k fixed, $T \to \infty$

	k									
T	2	4	12	24	60	120	240	360	480	600
60	−1.59	−3.15	−8.73	−7.90	n/a	n/a	n/a	n/a	n/a	n/a
120	−0.81	−1.65	−5.16	−9.18	0.80	n/a	n/a	n/a	n/a	n/a
240	−0.41	−0.84	−2.75	−5.38	−10.63	0.41	n/a	n/a	n/a	n/a
360	−0.28	−0.57	−1.87	−3.74	−8.36	−10.94	9.53	n/a	n/a	n/a
480	−0.21	−0.43	−1.41	−2.86	−6.67	−10.76	0.21	11.40	n/a	n/a
600	−0.17	−0.34	−1.14	−2.31	−5.52	−9.57	−8.75	6.95	12.32	n/a
1200	−0.08	−0.17	−0.58	−1.18	−2.93	−5.57	−9.62	−11.30	−8.80	0.08
2400	−0.04	−0.09	−0.29	−0.60	−1.50	−2.95	−5.59	−7.85	−9.65	−10.86

Panel B: $k \to \infty$, $T \to \infty$, k/T fixed

T	2	4	12	24	60	120	240	360	480	600
60	64.02	24.40	7.80	4.32	n/a	n/a	n/a	n/a	n/a	n/a
120	63.65	23.90	7.21	3.79	2.08	n/a	n/a	n/a	n/a	n/a
240	63.47	23.67	6.94	3.49	1.56	1.04	n/a	n/a	n/a	n/a
360	63.41	23.59	6.85	3.40	1.45	0.82	0.93	n/a	n/a	n/a
480	63.39	23.55	6.81	3.36	1.40	0.77	0.52	0.77	n/a	n/a
600	63.37	23.53	6.79	3.33	1.37	0.74	0.43	0.51	0.66	n/a
1200	63.33	23.49	6.74	3.28	1.32	0.68	0.37	0.27	0.21	0.21
2400	63.32	23.46	6.72	3.26	1.29	0.65	0.34	0.24	0.18	0.15

Panel C: $k \to \infty$, $T \to \infty$, $k/T \to 0$

T	2	4	12	24	60	120	240	360	480	600
60	60.71	19.56	−2.62	−4.92	n/a	n/a	n/a	n/a	n/a	n/a
120	61.97	21.40	1.19	−6.24	2.08	n/a	n/a	n/a	n/a	n/a
240	62.63	22.40	3.76	−2.32	−9.50	1.04	n/a	n/a	n/a	n/a
360	62.85	22.74	4.70	−0.62	−7.20	−10.38	9.87	n/a	n/a	n/a
480	62.96	22.92	5.18	0.28	−5.49	−10.20	0.52	11.63	n/a	n/a
600	63.03	23.02	5.48	0.85	−4.32	−9.00	−8.46	7.17	12.49	n/a
1200	63.16	23.23	6.08	2.01	−1.70	−4.97	−9.34	−11.11	−8.66	0.21
2400	63.23	23.34	6.38	2.62	−0.25	−2.34	−5.30	−7.66	−9.50	−10.75