# An Asymptotic F Test for Uncorrelatedness in the Presence of Time Series Dependence*

Xuexin Wang
School of Economics and WISE
Xiamen University, China

Yixiao Sun
Department of Economics
UC San Diego, USA

**Abstract**

We propose a simple asymptotically F-distributed Portmanteau test for zero autocorrelations in an otherwise dependent time series. By employing the orthonormal series variance estimator of the variance matrix of sample autocovariances, our test statistic follows an F distribution asymptotically under fixed-smoothing asymptotics. The asymptotic F theory accounts for the estimation error in the underlying variance estimator, which the asymptotic chi-squared theory ignores. Monte Carlo simulations reveal that the F approximation is much more accurate than the corresponding chi-squared approximation in finite samples. Compared with the nonstandard test proposed by Lobato (2001), the asymptotic F test is as easy to use as the chi-squared test: There is no need to obtain critical values by simulations. Further, Monte Carlo simulations indicate that Lobato's (2001) nonstandard test tends to be heavily undersized under the null and suffers from substantial power loss under the alternatives.

*Keywords:* Lack of autocorrelations; Portmanteau test; Fixed-smoothing asymptotics; F distribution; Orthonormal series variance estimator

---

*Email: xuexinwang@xmu.edu.cn; yisun@ucsd.edu. Wang gratefully acknowledges the Fundamental Research Funds for the Central Universities [Grant No. 20720171043]. Sun gratefully acknowledges partial research support from the NSF [Grant No. SES-1530592]. We thank the anonymous Co-Editor and reviewers for constructive comments.

# 1 INTRODUCTION

In this study, we test the null hypothesis that a time series process is uncorrelated up to a certain order. This is a basic problem in time series analysis. There is a large body of literature on testing zero autocorrelations in a time series. These tests can be roughly categorized into *time-domain autocorrelation-based* and *frequency-domain periodogram-based* tests. The latter tests consider an infinite number of autocorrelations (see, e.g., Hong (1996), Shao (2011), and the references therein). They are consistent in the sense that all nonzero autocorrelations can be detected. On the other hand, the former tests conventionally focus on a finite number of autocorrelations. They *can* obtain optimal power in some directions that practitioners are interested in.

Among the time-domain tests, as early as the 1970s, Box and Pierce (1970) and Ljung and Box (1978) proposed the so-called $Q$ test for a covariance stationary time series under the independent and identically distributed (i.i.d.) assumption. However, both economic theories and empirical studies have revealed that the i.i.d. assumption may be too restrictive. For example, the market efficiency hypothesis, rational expectation models, and optimal consumption smoothing theory all imply that the relevant process is a martingale difference (MD) sequence instead of an i.i.d. sequence. Empirical studies have further found that conditional heteroskedasticity is prevalent in financial time series.

In general, when a time series process is only uncorrelated, its sample autocorrelations are not necessarily asymptotically standard normal. Instead, the asymptotic covariance matrix of the sample autocorrelations depends on the data generation process (see, e.g., Romano and Thombs (1996)). In this case, Romano and Thombs (1996) and Horowitz et al. (2006) used the $Q$ test statistic with critical values obtained by block bootstrapping procedures, such as the single moving block and the blocks-of-blocks bootstraps. However, the difficulty in choosing the block size and high computational costs limit the scope of bootstrap-based tests.

Lobato, Nankervis, and Savin (2002) later proposed a modified $Q$ test based on a nonparametric estimator of the asymptotic variance matrix of the sample autocovariances. They considered the conventional increasing-smoothing asymptotics, whereas the amount of the nonparametric smoothing is assumed to grow with the sample size, but at a slower rate. Under this type of asymptotics, the proposed test statistic has a convenient asymptotic chi-squared distribution under the null hypothesis. By invoking the consistency argument, the asymptotic chi-squared theory approximates the distribution of the nonparametric variance estimator by a degenerate distribution concentrated at the true variance matrix. Effectively, the theory completely ignores the estimation error in the variance estimator. Even with a delicate choice of the underlying smoothing parameter, the estimation error in the nonparametric variance estimator can still be substantial in finite samples. This explains why the chi-square-based test often exhibits a large size distortion in finite samples.

Recently, the literature has introduced alternative asymptotics to combat the aforementioned problem. Unlike the conventional increasing-smoothing asymptotics, the alternative asymptotics hold the amount of nonparametric smoothing to be fixed. Hence, they are also called the fixed-smoothing asymptotics. There is ample numerical evidence, along with theoretical results, on the higher accuracy of fixed-smoothing asymptotic approximations relative to conventional asymptotic approximations (see, e.g., Sun, Phillips, and Jin (2008) and Zhang and Shao (2013) for location models, and Sun (2014a, 2014b) for the generalized method of moments framework). In testing uncorrelatedness of time series, Lobato (2001) employed a different approach to studentization, but his approach can be regarded as employing the fixed-smoothing asymptotics implicitly. Although the asymptotic distribution of Lobato's (2001) test statistic is pivotal, it is not standard, and critical values have to be tabulated by Monte Carlo simulations.

In this study, we employ the orthonormal series approach to variance estimation in the construction of a new Portmanteau test for zero autocorrelations, while allowing the time series to be otherwise serially dependent. This approach involves projecting the time series onto a sequence of orthonormal basis functions, and then taking the simple average of the outer products of the projection coefficients as the variance estimator. The number of basis functions is the smoothing parameter that underlies this orthonormal series variance estimator.

By employing the orthonormal series variance estimator of the asymptotic variance matrix of the sample autocovariances, the new Portmanteau test statistic follows an F distribution asymptotically under the fixed-smoothing asymptotics. The asymptotic F theory accounts for the estimation error in the underlying variance estimator, which the asymptotic chi-squared theory ignores. Monte Carlo simulations reveal that the F approximation is much more accurate than the corresponding chi-squared approximation in finite samples. Compared with the nonstandard test proposed by Lobato (2001), the asymptotic F test has the same ease of use as the chi-squared test, since the critical values of the F distributions are readily available in standard programming environments and software packages. Further, Monte Carlo simulations indicate that Lobato's (2001) test tends to be heavily undersized under the null and suffers from substantial power loss under the alternatives.

The remaining article is organized as follows. In section 2, we lay out the preliminaries. In section 3, we propose our new Portmanteau test and establish its asymptotic properties. In section 4, we conduct comprehensive Monte Carlo simulations, followed by an empirical application in section 5. In section 6, we conclude our study.

## 2 PRELIMINARIES

Let $\{y_t\}_{t\in\mathbb{Z}}$ be a real-valued covariance stationary time series with mean $\mu$. Define the autocovariance and autocorrelation functions:

$$\gamma(j) = \mathbb{E}\left[(y_t-\mu)(y_{t-j}-\mu)\right], \, j \in \mathbb{Z},$$
$$\rho(j) = \frac{\gamma(j)}{\gamma(0)}, \, j \in \mathbb{Z}\backslash\{0\}.$$

Given the observations $\{y_t\}_{t=1}^{T}$, we can estimate $\gamma(j)$ and $\rho(j)$ by their sample analogues:

$$\hat{\gamma}_T(j) = \frac{1}{T}\sum_{t=j+1}^{T}(y_t-\bar{\mu})(y_{t-j}-\bar{\mu}), j \in 0,\ldots,T-1,$$

$$\hat{\rho}_T(j) = \frac{\hat{\gamma}_T(j)}{\hat{\gamma}_T(0)},\ j \in 1, \dots, T-1,$$

where $\bar{\mu} = T^{-1} \sum_{t=1}^{T} y_t$ is the sample mean.

As in Box and Pierce (1970) and Ljung and Box (1978), we are interested in testing whether the first $s$ autocorrelations or autocovariances are zero. That is, we are interested in the null

$$H_0^{(s)} : \boldsymbol{\gamma}^{(s)} = \mathbf{0},$$

where $\boldsymbol{\gamma}^{(s)} = (\gamma(1), \dots, \gamma(s))'$.

The $Q$ test statistic in Ljung and Box (1978) takes the following form:

$$\hat{Q}_T(s) = T(T+2) \sum_{j=1}^{s} (T-j)^{-1} \hat{\rho}_T^2(j).$$

Denote $\hat{\boldsymbol{\gamma}}_T^{(s)} = (\hat{\gamma}_T(1), \dots, \hat{\gamma}_T(s))'$ and $\hat{\boldsymbol{\rho}}_T^{(s)} = (\hat{\boldsymbol{\rho}}_T(1), \dots, \hat{\boldsymbol{\rho}}_T(s))'$. Under the assumption that $y_t \sim iid$ as well as other mild regularity conditions, it can be shown that, under $H_0^{(s)}$,

$$\sqrt{T}\hat{\boldsymbol{\rho}}_T^{(s)} \Rightarrow N(\mathbf{0}, \boldsymbol{I}_s) \text{ and } \sqrt{T}\hat{\boldsymbol{\gamma}}_T^{(s)} \Rightarrow N\left(\mathbf{0}, \boldsymbol{I}_s \gamma(0)^2\right),$$

where $\boldsymbol{I}_s$ is the identity matrix of dimension $s$. It then follows that $\hat{Q}_T(s) \Rightarrow \chi_s^2$, the chi-squared distribution with $s$ degrees of freedom.

The chi-squared approximation is convenient. However, in the absence of further restrictions on the dependence structure of the time series apart from $H_0^{(s)}$, we can only expect that $\sqrt{T}\hat{\boldsymbol{\gamma}}_T^{(s)} \Rightarrow N(\mathbf{0}, \boldsymbol{\Omega})$ where $\boldsymbol{\Omega}$ is a general nondiagonal matrix $\boldsymbol{\Omega} = \left[\omega^{(i,j)}\right]_{i,j=1}^{s}$ for

$$\omega^{(i,j)} = \sum_{l=-\infty}^{\infty} \mathbb{E}\left[(y_t - \mu)(y_{t-i} - \mu)(y_{t+l} - \mu)(y_{t+l-j} - \mu)\right].$$

In this case, Romano and Thombs (1996) pointed out that the $Q$ test based on the chi-squared critical values can deliver misleading results. They proposed using computer-intensive methods, such as bootstrapping and subsampling, to obtain more reliable critical values. See also Horowitz et al. (2006).

On the other hand, Lobato et al. (2002) proposed a modified $Q$ test, which is based on the test statistic:

$$\tilde{Q}_T(s) = T\hat{\boldsymbol{\gamma}}_T^{(s)\prime}\hat{\boldsymbol{\Omega}}^{-1}\hat{\boldsymbol{\gamma}}_T^{(s)},$$

where $\hat{\boldsymbol{\Omega}}$ is a nonparametric estimator of $\boldsymbol{\Omega}$. To be precise, denote $f_t = (f_{1t}, \ldots, f_{st})'$ with the $j$th element given by $f_{jt} = (y_t - \bar{\mu})(y_{t-j} - \bar{\mu})$ and $\tilde{f}_t = (\tilde{f}_{1t}, \ldots, \tilde{f}_{st})'$ with the $j$th element given by

$$\tilde{f}_{jt} = (y_t - \bar{\mu})(y_{t-j} - \bar{\mu}) - \hat{\gamma}_T(j).$$

$\hat{\boldsymbol{\Omega}}$ takes the quadratic form of

$$\hat{\boldsymbol{\Omega}} = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{T} w_h\left(\frac{t}{T}, \frac{j}{T}\right)\tilde{f}_t\tilde{f}_j',$$

where $w_h(\cdot,\cdot)$ is a weighting function and $h$ is the smoothing parameter indicating the amount of nonparametric smoothing. For example, we can take $w_h(t/T, j/T) = k((t-j)/(hT))$ for a kernel function $k(\cdot)$, leading to the usual kernel variance estimator. The Newey and West (1987) estimator, for instance, has this form with $k(\cdot)$ equal to the Bartlett kernel $k(x) = \max\{1 - |x|, 0\}$.

By allowing the amount of nonparametric smoothing in $\hat{\boldsymbol{\Omega}}$ to grow with the sample size, but at a slower rate, Lobato et al. (2002) established that $\tilde{Q}_T(s)$ is asymptotically chi-squared under the null hypothesis. However, such an increasing-smoothing asymptotic approximation ignores the estimation error in the variance estimator. As a result, it can be highly inaccurate in finite samples.

Lobato (2001) proposed another test statistic:

$$\hat{L}_T(s) = T\hat{\boldsymbol{\gamma}}_T^{(s)\prime}\left(\frac{1}{T^2}\sum_{t=1}^{T}\Upsilon_t\Upsilon_t'\right)^{-1}\hat{\boldsymbol{\gamma}}_T^{(s)},$$

where

$$\Upsilon_t = \sum_{j=1}^{t}\left(\begin{array}{c} f_{1j}-\hat{\gamma}_T(1),\ldots,f_{sj}-\hat{\gamma}_T(s)\end{array}\right)'.$$

We can show that $T^{-2}\sum_{t=1}^{T}\Upsilon_t\Upsilon_t'$ is asymptotically equivalent to $\hat{\boldsymbol{\Omega}}$ with $w_h(t,\tau) = 1-\max(t/h,\tau/h)$ and $h=1$. Effectively, Lobato (2001) employed a quadratic long run variance estimator with the truncation lag parameter set equal to the sample size. This is in the same spirit of Kiefer and Vogelsang (2002). Lobato (2001) can, therefore, be regarded as among the first to use the fixed-smoothing asymptotics in testing serial uncorrelatedness.

Shao (2010) proposed a new test statistic $\tilde{L}_T(s)$, which is a variation of $\hat{L}_T(s)$. Both $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ take a self-normalization form (see, Shao (2015)). The difference between $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ lies in the different forms of their self-normalization matrices. It has been shown that $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ share the same asymptotic distribution under fixed-smoothing asymptotics. Indeed, the alternative asymptotics here belong to the fixed-b asymptotics developed by Kiefer and Vogelsang (2005). Although the fixed-smoothing asymptotic approximation for $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ is more accurate than the chi-squared approximation, the asymptotic distribution is not standard, and the associated critical values have to be simulated.

# 3 MAIN RESULTS

In this study, we employ an alternative variance estimator—the orthonormal series variance estimator—such that

$$w_h\left(\frac{t}{T},\frac{j}{T}\right) = \frac{1}{K}\sum_{\ell=1}^{K}\Phi_\ell\left(\frac{t}{T}\right)\Phi_\ell\left(\frac{j}{T}\right),$$

where $K$ is the smoothing parameter for this estimator and $\Phi_\ell(\cdot)$ satisfies the following assumption:

**Assumption 1.** *For $\ell = 1, 2, \ldots, K$, the basis functions $\Phi_\ell(\cdot)$ are continuously differentiable and orthonormal in $L^2[0,1]$ and satisfy $\int_0^1 \Phi_\ell(x)\,dx = 0$.*

Denote

$$\Lambda_\ell = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Phi_\ell\left(\frac{t}{T}\right) f_t.$$

Then, we have

$$\hat{\boldsymbol{\Omega}}_{\mathrm{OS}} = \frac{1}{K} \sum_{\ell=1}^{K} \Lambda_\ell \Lambda_\ell'$$

after a straightforward rearrangement. Notably, demeaning is not necessary here, since $\int_0^1 \Phi_\ell(x)\,dx = 0$. To ensure that $\hat{\boldsymbol{\Omega}}$ is positive semidefinite, we assume that $K \geq s$.

For some recent use of the orthonormal series variance estimator in the econometric literature, see Sun (2011, 2013, 2014a, 2014b) and references therein. The simplest and most familiar example of this estimator is the average periodogram estimator, which involves taking a simple average of a few periodograms. In this case,

$$\Phi_\ell(r) = \begin{cases} \sqrt{2}\cos(\pi \ell r), & \text{if } \ell \text{ is even}, \\ \sqrt{2}\sin(\pi(\ell+1) r), & \text{if } \ell \text{ is odd}. \end{cases}$$

We will use the above basis functions for an even $K$ in our simulation study.

With the orthonormal series variance estimator $\hat{\boldsymbol{\Omega}}_{\mathrm{OS}}$, we construct the test statistic:

$$\tilde{Q}_T^*(s) = \frac{K-s+1}{Ks} T \hat{\boldsymbol{\gamma}}_T^{(s)\prime} \hat{\boldsymbol{\Omega}}_{\mathrm{OS}}^{-1} \hat{\boldsymbol{\gamma}}_T^{(s)} = \frac{K-s+1}{Ks} T \hat{\boldsymbol{\gamma}}_T^{(s)\prime} \left( \frac{1}{K} \sum_{\ell=1}^{K} \Lambda_\ell \Lambda_\ell' \right)^{-1} \hat{\boldsymbol{\gamma}}_T^{(s)}.$$

Note that there is a multiplicative correction term $(K-s+1)/(Ks)$. When $s = 1$, this term is simply equal to 1. It becomes more important when $s$ is larger.

Herein, we consider the fixed-smoothing asymptotics, wherein the number of basis functions

$K$ is held fixed, as the sample size increases.

### 3.1 Asymptotics under the null

To establish the fixed-smoothing asymptotics of $\tilde{Q}_T^*(s)$ under the null, we introduce the following assumption:

**Assumption 2.** *For $r \in (0,1]$, $T^{-1/2}\sum_{t=1}^{[Tr]} f_t \Rightarrow \mathbf{\Omega}^{1/2} W_f(r)$, where $W_f(r)$ is a standard vector Brownian motion process.*

Assumption 2 requires that a functional central limit theorem holds for the partial sums $T^{-1/2}\sum_{t=1}^{[Tr]} f_t$. Primitive sufficient conditions can be found in Lobato (2001) (Assumption 2, p. 1070) and the references therein.

Under Assumptions 1 and 2, we have

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Phi_\ell\left(\frac{t}{T}\right) f_t \Rightarrow \mathbf{\Omega}^{1/2}\int_0^1 \Phi_\ell(r)\, dW_f(r) := \mathbf{\Omega}^{1/2}\eta_{f,\ell}. \tag{1}$$

Note that $\eta_{f,\ell} = \int_0^1 \Phi_\ell(r)\, dW_f(r)$ is i.i.d. $N(0, \boldsymbol{I}_s)$ over $\ell = 1, 2, \ldots, K$. This holds because $\eta_{f,\ell}$ is normal and

$$\operatorname{cov}\left[\eta_{f,\ell_1}, \eta_{f,\ell_2}\right] = \int_0^1 \Phi_{\ell_1}(r)\,\Phi_{\ell_2}(r)\, dr = 1\left\{\ell_1 = \ell_2\right\}.$$

using the orthonormality of $\{\Phi_\ell(\cdot), \ell = 1, \ldots, K\}$ on $L_2[0,1]$. Also, $W_f(1)$ is independent of $\{\eta_{f,\ell}\}$ because $W_f(1)$ and $\eta_{f,\ell}$ are standard normals and for $\ell = 1, \ldots, K$,

$$\operatorname{cov}(\eta_{f,\ell}, W_f(1)) = \mathbb{E}\left[\int_0^1 \Phi_\ell(r)\, dW_f(r)\, W_f(1)'\right] = \boldsymbol{I}_s \int_0^1 \Phi_\ell(r)\, dr = 0.$$

Using the continuous mapping theorem, we now have

$$\tilde{Q}_T^*(s) \Rightarrow \frac{K-s+1}{s} W_f(1)' \left[\sum_{\ell=1}^{K} \eta_{f,\ell}\eta_{f,\ell}'\right]^{-1} W_f(1).$$

Note that

$$\sum_{\ell=1}^{K} \eta_{f,\ell}\eta_{f,\ell}' \overset{d}{=} \mathbb{W}\left(\boldsymbol{I}_s, K\right),$$

where $\mathbb{W}\left(\boldsymbol{I}_s, K\right)$ is a Wishart random variable that is independent of $W_f\left(1\right)$. Then, by Proposition 8.2 in Bilodeau and Brenner (1999), we have the theorem below.

**Theorem 1.** *Let Assumptions 1 and 2 hold. Under $H_0^{(s)}$, we have*

$$\tilde{Q}_T^*\left(s\right) \Rightarrow F\left(s, K-s+1\right)$$

*where $F(s, K-s+1)$ is the standard $F$ distribution with degrees of freedom $s$ and $K-s+1$.*

This result is intriguing. First, the asymptotic F test has the same ease of use as the chi-squared test, since the critical values of F distributions are readily available in standard programming environments. Second, the asymptotic F distribution is directly related to $K$. This is different from the conventional chi-squared theory, wherein the smoothing parameter must grow with the sample size, but does not influence the asymptotic distribution.

Let $F^{\alpha}\left(s, K-s+1\right)$ be the $(1-\alpha)$-quantile of the F distribution $F\left(s, K-s+1\right)$. Our asymptotic F test is based on the statistic $\tilde{Q}_T^*\left(s\right)$ with $F^{\alpha}\left(s, K-s+1\right)$ as the critical value.

If we follow Lobato (2001) and Lobato et al. (2002) and use the test statistic of the standard form, namely, $T\hat{\boldsymbol{\gamma}}_T^{(s)\prime}\hat{\boldsymbol{\Omega}}_{\mathrm{OS}}^{-1}\hat{\boldsymbol{\gamma}}_T^{(s)}$, we will obtain:

$$T\hat{\boldsymbol{\gamma}}_T^{(s)\prime}\hat{\boldsymbol{\Omega}}_{\mathrm{OS}}^{-1}\hat{\boldsymbol{\gamma}}_T^{(s)} \Rightarrow \frac{(Ks)}{(K-s+1)} \cdot F\left(s, K-s+1\right).$$

Our asymptotic F test is equivalent to the test using $T\hat{\boldsymbol{\gamma}}_T^{(s)\prime}\hat{\boldsymbol{\Omega}}_{\mathrm{OS}}^{-1}\hat{\boldsymbol{\gamma}}_T^{(s)}$ as the test statistic and

$$\frac{K}{(K-s+1)} \cdot sF^{\alpha}\left(s, K-s+1\right)$$

as the critical value. The above modified F critical value is larger than the corresponding chi-squared critical value for two reasons. First, $sF^{\alpha}\left(s, K-s+1\right)$ is larger than the corresponding chi-squared critical value, as the F distribution has a random denominator. Second, the extra multiplicative factor $K/(K-s+1)$ is greater than 1. This factor is larger for a larger $s$ or a smaller $K$. In finite samples, the difference between the chi-squared critical value and the modified F critical value can be substantial, especially when $s$ is large and $K$ is small. On the other hand, when $K$ is large, the chi-squared critical value and the modified F critical value will become close to each other.

In practice, we can choose $K$ to minimize the mean square error of $\hat{\mathbf{\Omega}}_{\mathrm{OS}}$. Phillips (2005) proposed such a data-driven procedure. Although it is not necessarily best suited for hypothesis testing, Monte Carlo simulations in the next section reveal that this choice of $K$ delivers good finite sample performances for the asymptotic F test.

### 3.2 Asymptotics under local alternative

We consider the following local alternative:

$$H_{1T}^{(s)} : \boldsymbol{\gamma}^{(s)} = \frac{\boldsymbol{\delta}}{\sqrt{T}}.$$

Under this local alternative, we assume that the following functional central limit theorem (FCLT) holds:

**Assumption 3.** $T^{-1/2}\sum_{t=1}^{[Tr]} f_t \Rightarrow r\boldsymbol{\delta} + \mathbf{\Omega}^{1/2}W_f\left(r\right).$

Under $H_{1T}^{(s)}$, we now have

$$\tilde{Q}_T^*\left(s\right) \Rightarrow \frac{K-s+1}{s}\left[\mathbf{\Omega}^{-1/2}\,\boldsymbol{\delta}+W_f\left(1\right)\right]'\left[\sum_{\ell=1}^{K}\eta_{f,\ell}\eta_{f,\ell}'\right]^{-1}\left[\mathbf{\Omega}^{-1/2}\,\boldsymbol{\delta}+W_f\left(1\right)\right].$$

Then, by Proposition 8.2 in Bilodeau and Brenner (1999), we obtain the following theorem:

**Theorem 2.** *Under $H_{1T}^{(s)}$, and Assumptions 1 and 3, we have*

$$\tilde{Q}_T^*\left(s\right) \Rightarrow F_{\lambda^2}\left(s, K-s+1\right),$$

*where* $\lambda = \left\|\boldsymbol{\Omega}^{-1/2}\boldsymbol{\delta}\right\|$ *and* $F_{\lambda^2}(s, K-s+1)$ *is the noncentral F distribution with noncentrality parameter* $\lambda^2$ *and degrees of freedom* $s$ *and* $K-s+1$.

The theorem establishes that the asymptotic F test based on $\tilde{Q}_T^*(s)$ has nontrivial testing power for the Pitman local alternative. In particular, under $H_{1T}^{(s)}$, $\Pr(\tilde{Q}_T^*(s) > F_{\lambda^2}^{\alpha}(s, K-s+1)) \to 1$ as $\lambda \to \infty$ and so the test is consistent.

For Lobato's (2001) test, we can show that under $H_{1T}^{(s)}$,

$$\hat{L}_T(s) \Rightarrow \left[\boldsymbol{\Omega}^{-1/2}\,\boldsymbol{\delta} + W_f(1)\right]' \left[\int_0^1 \left[W_f(r) - rW_f(1)\right]^{\otimes 2} dr\right]^{-1} \left[\boldsymbol{\Omega}^{-1/2}\,\boldsymbol{\delta} + W_f(1)\right],$$

where for a column vector $a$, $a^{\otimes 2} = aa'$. Let $H = \left(\frac{\boldsymbol{\Omega}^{-1/2}\boldsymbol{\delta}}{\|\boldsymbol{\Omega}^{-1/2}\boldsymbol{\delta}\|}, \tilde{H}\right)$ be an orthogonal matrix. Then

$$\begin{aligned}
\hat{L}_T(s) &\Rightarrow \left[H'(\boldsymbol{\Omega}^{-1/2}\,\boldsymbol{\delta} + W_f(1))\right]' \left[\int_0^1 \left\{H'\left[W_f(r) - rW_f(1)\right]\right\}^{\otimes 2}\right]^{-1} H'\left[\boldsymbol{\Omega}^{-1/2}\,\boldsymbol{\delta} + W_f(1)\right] \\
&= \left[\lambda e_1 + W_f(1))\right]' \left[\int_0^1 \left[W_f(r) - rW_f(1)\right]^{\otimes 2} dr\right]^{-1} \left[\lambda e_1 + W_f(1)\right],
\end{aligned}$$

where $e_1 = (1, 0, \ldots, 0)$ is the first basis vector in $\mathbb{R}^s$. So the asymptotic distributions of $\hat{L}_T(s)$ and $\tilde{Q}_T^*(s)$ under the local alternative depend on the same noncentrality parameter $\lambda^2$.

Given that the asymptotic distribution of $\hat{L}_T(s)$ is nonstandard, it is not easy to analytically compare its local asymptotic power with that of the asymptotic F test. However, it is easy to simulate the local asymptotic powers. Figure 1 reports the local asymptotic powers of the two tests. For the asymptotic F test, we consider a few fixed $K$ values starting with $K = s+4$ so that the approximating F distribution has a finite variance. The figure shows that the F test has a higher local asymptotic power when $K$ is large enough. When $K$ is relatively small or $s$ is relatively large, Lobato's (2001) test may have a higher local asymptotic power. Compared to Lobato's (2001) test, the asymptotic F test is more flexible in that the value of $K$ can be data-driven. In the next section, we show that with a data-driven choice of $K$, the F test is more powerful than Lobato's (2001) test in finite samples for almost all data generating processes under consideration.

## 4 MONTE CARLO EVIDENCE

We now examine the performance of the proposed F test in a set of Monte Carlo experiments. The data generating processes (DGPs) used in the experiments include three MD processes and three non-MD processes. To be more precise, let $\eta_t \sim iid\ N(0,1)$. The six DGPs are

M1: i.i.d. normal process: $\{y_t\}$ is a sequence of i.i.d $N(0,1)$ random variables.

M2: generalized autoregressive conditional heteroskedasticity or GARCH (1,1) process: $y_t = h_t \eta_t$, where $h_t^2 = 0.1 + 0.09 y_{t-1}^2 + 0.9 h_{t-1}^2$.

M3: 1-dependent process: $y_t = \eta_t \eta_{t-1}$.

M4: non-MD-1 (non-martingale difference) process: $y_t = \eta_t^2 \eta_{t-1}$.

M5: nonlinear moving average (NLMA) process: $y_t = \eta_{t-2}\eta_{t-1}(\eta_{t-2} + \eta_t + 1)$.

M6: bilinear process: $y_t = \eta_t + 0.5\eta_{t-1} y_{t-2}$.

M1–M3 are MD sequences. More specifically, the i.i.d process in M1 is a basic benchmark. The GARCH process in M2 is empirically relevant in the financial literature. The 1-dependent process in M3 is considered by Romano and Thombs (1996), Lobato (2001), and Horowitz et al. (2006); for this process, $\mathbf{\Omega}$ is an identity matrix, except that $\omega^{(1,1)} = 3$. M4–M6 are non-MD processes, but with zero autocorrelations. M5 and M6 are considered by Lobato (2001) and Horowitz et al. (2006).

To examine the size property, we repeat the experiment $10,000$ times for sample sizes $T = 100$ and 200. The nominal level of all tests is 5%. We use values for $s$ up to 10 and 15 for $T = 100$ and 200, respectively. Figures 2 and 3 report the empirical rejection probabilities of the $\tilde{Q}_T^*(s)$ test, that is, the asymptotic F test based on test statistic $\tilde{Q}_T^*(s)$. As a comparison, the empirical rejection probabilities are also reported for the $\hat{Q}_T(s)$, $\tilde{Q}_T(s)$, $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ tests. The $\hat{Q}_T(s)$ and $\tilde{Q}_T(s)$ tests are based on the chi-squared approximations, and the $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ test are based on the nonstandard distribution established in Lobato (2001). The $\tilde{Q}_T(s)$ test

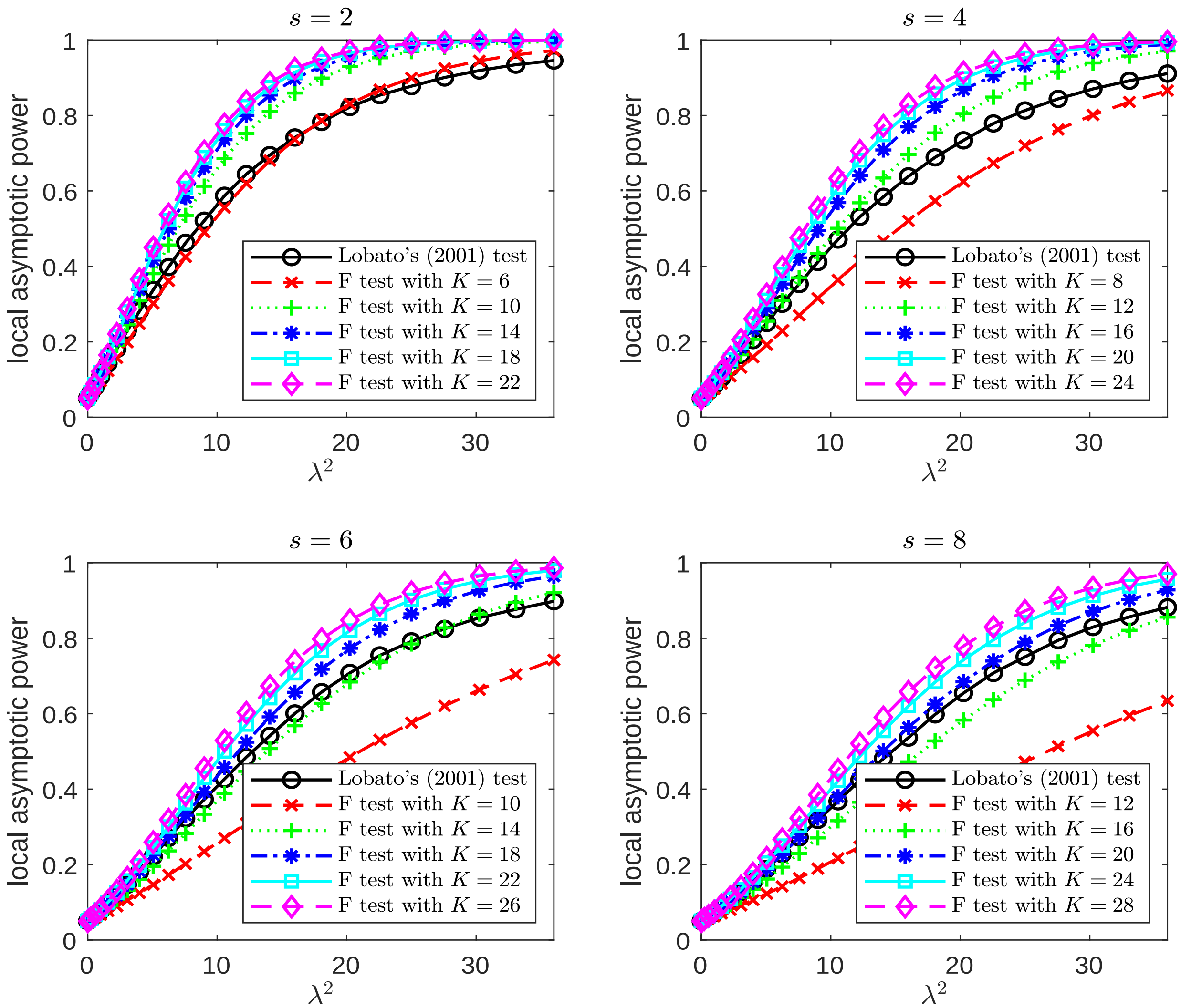


Figure 1: Local asymptotic power of the asymptotic F test based on the modified statistic $\tilde{Q}_T^*(s)$ (this paper) and Lobato's (2001) test based on $\hat{L}_T(s)$ for different values of $s$.

is implemented using the vector autoregression heteroskedasticity and autocorrelation consistent procedure described in Lobato et al. (2002), with the vector autoregression (VAR) order selected by the Bayesian information criterion. The maximum VAR order is set at $1.2T^{1/3}$. For easy references, we identify the tests with their test statistics $\tilde{Q}_T^*(s)$, $\hat{Q}_T(s)$, $\tilde{Q}_T(s)$, $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ with the understanding that different reference distributions are used for different tests.

The main features of the results are as follows:

i. For the i.i.d. process, all tests except for the $\tilde{L}_T(s)$ test tend to work satisfactorily. The $\tilde{L}_T(s)$ test is heavily oversized in this case.

ii. For the GARCH process, the empirical rejection probabilities of the $\tilde{Q}_T^*(s)$ test are close to the nominal size of 5% for all $s$ considered. On the other hand, both the $\hat{Q}_T(s)$ and $\tilde{Q}_T(s)$ tests are heavily oversized, and their empirical rejection probabilities tend to increase for a larger $s$ when $T = 100$. When $T = 200$, the empirical rejection probabilities of $\tilde{Q}_T(s)$ come closer to the nominal size of 5%, while those of the $\hat{Q}_T(s)$ tests are still far away from the nominal size of 5%. In contrast, the $\hat{L}_T(s)$ test is heavily undersized, especially when $s$ is large. The $\tilde{L}_T(s)$ test performs better than the $\hat{L}_T(s)$ test when $s$ is small, but tends to be oversized when $s$ is large.

iii. For the 1-dependent process, the size patterns of these tests are similar to those in the case of the GARCH process. This is expected, since the covariance matrices of the sample autocovariances of both processes are diagonal with heterogeneous diagonal elements.

iv. For the non-MD-1 process, the empirical rejection probabilities of the $\tilde{Q}_T^*(s)$ test are sufficiently controlled. However, those of the $\tilde{Q}_T(s)$ test worsen for a larger $s$, while those of the $\hat{Q}_T(s)$ test improve for a larger $s$, albeit still far from the nominal size of 5%. On the other hand, the $\hat{L}_T(s)$ test is heavily undersized. In some cases, its empirical rejection probabilities are quite close to 0. The empirical rejection probabilities of the $\tilde{L}_T(s)$ test is closer to the nominal size than the $\hat{L}_T(s)$ test, although it is still undersized.

v. For the NLMA process, the size patterns of these tests are reasonably similar to those in the case of the non-MD-1 process.

vi. For the bilinear process, the $\hat{L}_T(s)$ test performs well when $s$ is small ($s \leq 3$ for $T = 100$ and $s \leq 6$ for $T = 200$). The $\tilde{Q}_T^*(s)$ test is slightly oversized for the sample sizes considered here. On the other hand, the $\hat{Q}_T(s)$, $\tilde{Q}_T(s)$ and $\tilde{L}_T(s)$ tests are all heavily oversized.

Overall, the $\tilde{Q}_T^*(s)$ test performs satisfactorily for all the processes considered here. The $\tilde{Q}_T(s)$ test is not reliable for a sample size of 100 or 200. This is in accordance with the simulation findings in Lobato et al. (2002). Finally, the $\hat{L}_T(s)$ test tends to be heavily undersized, especially when $s$ is large. The empirical rejection probabilities of the $\tilde{L}_T(s)$ test are closer to the nominal size in some cases, but it is heavily oversized in many other cases.

For the power comparison, we consider an MA(1) process, $u_t = y_t + 0.25y_{t-1}$, where $\{y_t\}$ are generated according to DGPs M1–M6. Since there exists a substantial difference in the null rejection probabilities for different tests using the asymptotic critical values, in the power comparison we employ the size-adjusted critical values such that the empirical rejection probability of each test under the null is exactly 5%. It should be noted that the critical-value adjustment is not practically feasible.

The simulations are carried out using $1,000$ replications. Again, $s$ ranges from 1 to 10 and 15 for $T = 100$ and 200, respectively. To save space, we only report the results for the case $T = 100$ in Figure 4. The main features of the results are as follows:

i. The $\tilde{Q}_T^*(s)$ test has substantial testing power, which is comparable to the $\tilde{Q}_T(s)$ and $\hat{Q}_T(s)$ tests in many cases. For example, for the non-MD experiments, the $\tilde{Q}_T^*(s)$ test is even more powerful than the $\tilde{Q}_T(s)$ and $\hat{Q}_T(s)$ tests when $s$ is small.

ii. The $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ tests, however, suffer from substantial power loss, even after the critical-value adjustment.

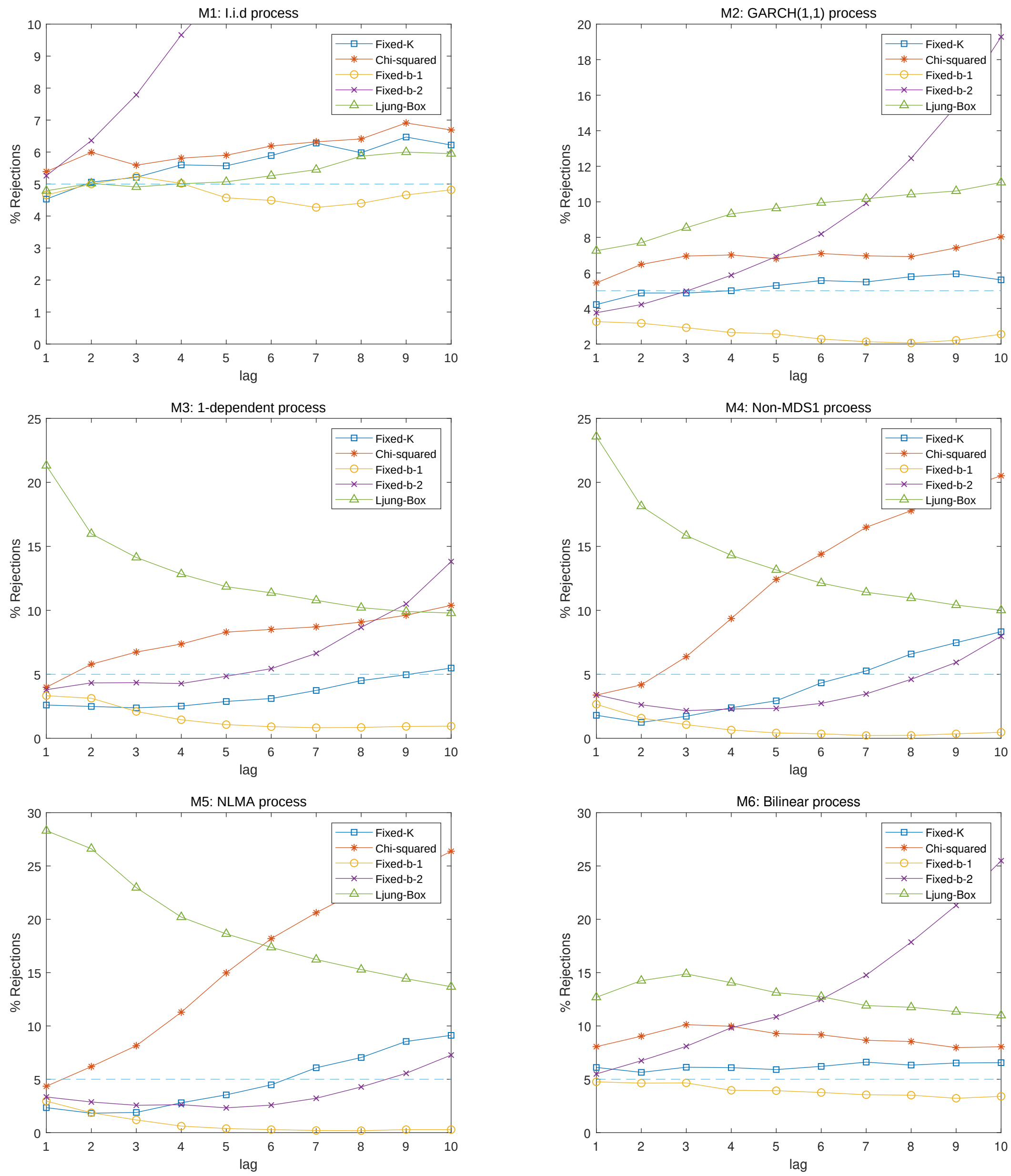


Figure 2: Percentage of rejections of Portmanteau tests in terms of the lag $s$ under the null with $T = 100$. "Fixed-K" represents the asymptotic F test based on the modified statistic $\tilde{Q}_T^*(s)$ (this paper); "Fixed-b-1" represents the nonstandard test based on the modified statistic $\hat{L}_T(s)$ (Lobato, 2001); "Fixed-b-2" represents the nonstandard test based on the modified statistic $\tilde{L}_T(s)$ (Shao, 2010); "Chi-squared" represents the Chi-squared test based on the statistic $\tilde{Q}_T(s)$ (Lobato et al., 2002); and "Ljung-Box" represents the Ljung–Box $Q$ test.

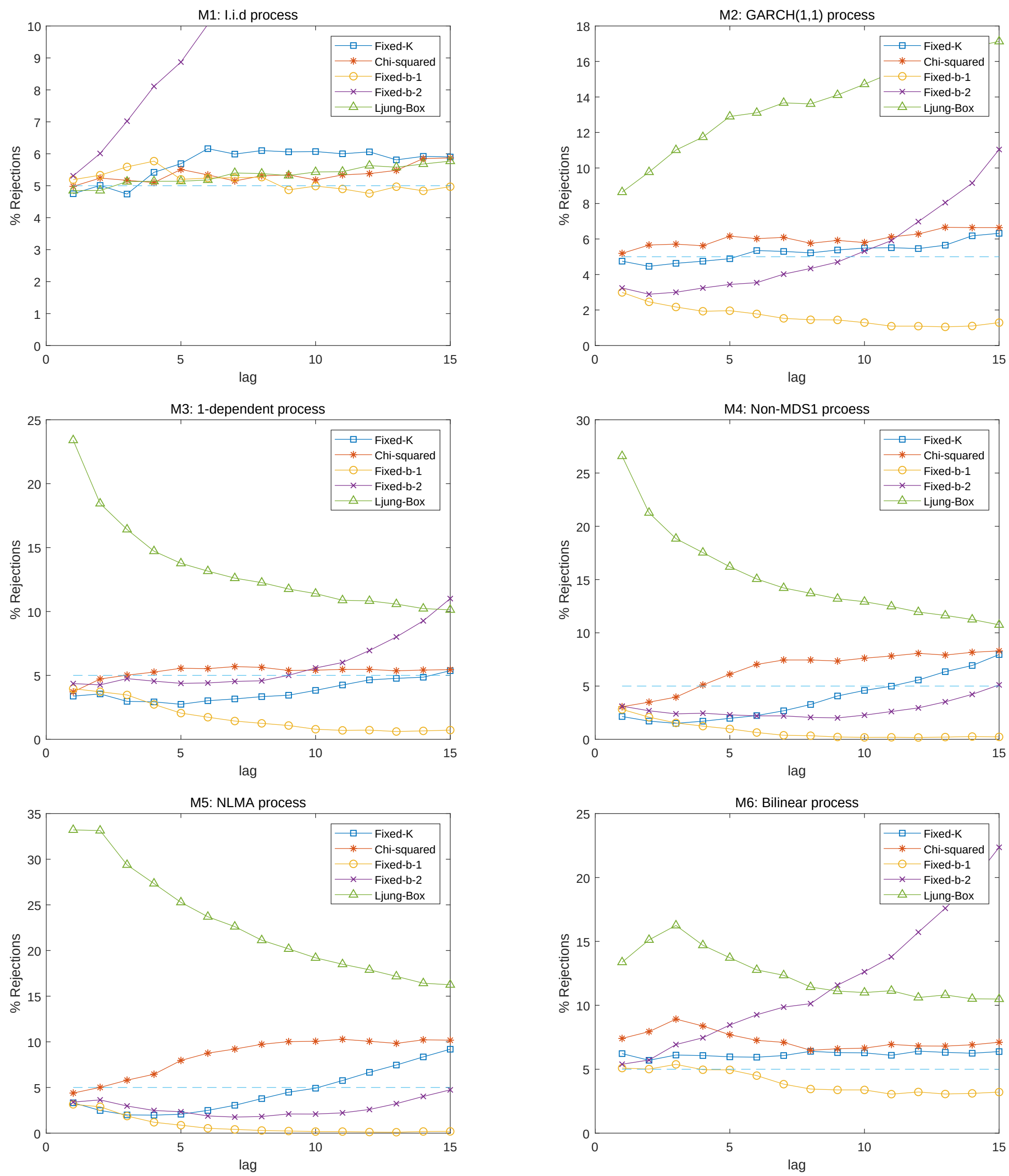


Figure 3: Percentage of rejections of Portmanteau tests in terms of the lag $s$ under the null with $T = 200$. "Fixed-K" represents the asymptotic F test based on the modified statistic $\tilde{Q}_T^*(s)$ (this paper); "Fixed-b-1" represents the nonstandard test based on the modified statistic $\hat{L}_T(s)$ (Lobato, 2001); "Fixed-b-2" represents the nonstandard test based on the modified statistic $\tilde{L}_T(s)$ (Shao, 2010); "Chi-squared" represents the Chi-squared test based on the statistic $\tilde{Q}_T(s)$ (Lobato et al., 2002); and "Ljung-Box" represents the Ljung–Box $Q$ test.

In summary, the $\tilde{Q}_T^*(s)$ test has the most accurate size among all four tests, with only a minor compromise in testing power in some cases.

## 5 APPLICATION

We use the extended Nelson and Plosser (1982) economic data set for our empirical application. The data set contains annual information of 15 relevant economic series for the U.S. economy. Following Lobato (2001), we focus on two time series: the growth of employment and the returns of the S&P 500 stock market index. The employment series covers the period from 1890 to 1988, whereas the stock index covers that from 1871 to 1988. Hence, there are 99 and 118 observations, respectively. Figures 5 plots the employment growth and the stock index returns.

We consider the null hypotheses $H_0^{(s)}$ up to $s = 5$ for both the employment growth series and the S&P500 returns series. The p-values of the $\tilde{Q}_T^*(s)$, $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ tests are reported in Table 1[1]. For completeness, the sample autocorrelation coefficients and their standard errors are also reported. For the employment growth series, the null hypotheses for $s \geq 2$ are rejected at the 10% significance level by our F tests. However, they are not rejected at the 10% significance level by the nonstandard tests of Lobato (2001) and Shao (2010). In macroeconomics, Blanchard et al. (1989) developed a conceptual framework to interpret the dynamic behavior of employment, and unemployment, vacancies and their interactions. So, from a theoretical point of view, the serial dependence of employment growth is plausible. Empirically, the nonzero autocorrelation of the employment growth series is in accordance with the results in Chinn et al. (2014). For the S&P 500 returns, the null hypotheses are rejected at the 10% significance level by our F tests. For Lobato's (2001) and Shao's (2010) tests, the null hypotheses are not rejected at the 10% significance level for some values of $s$.

Our results show that different tests may lead to different conclusions in empirical appli-

[1] We use a sequence of i.i.d. standard normal random vectors with length $10^4$ to approximate one realization of the vector standard Brownian motion path. We employ $10^5$ Monte-Carlo replications to simulate the nonstandard distribution of $\hat{L}_T(s)$.

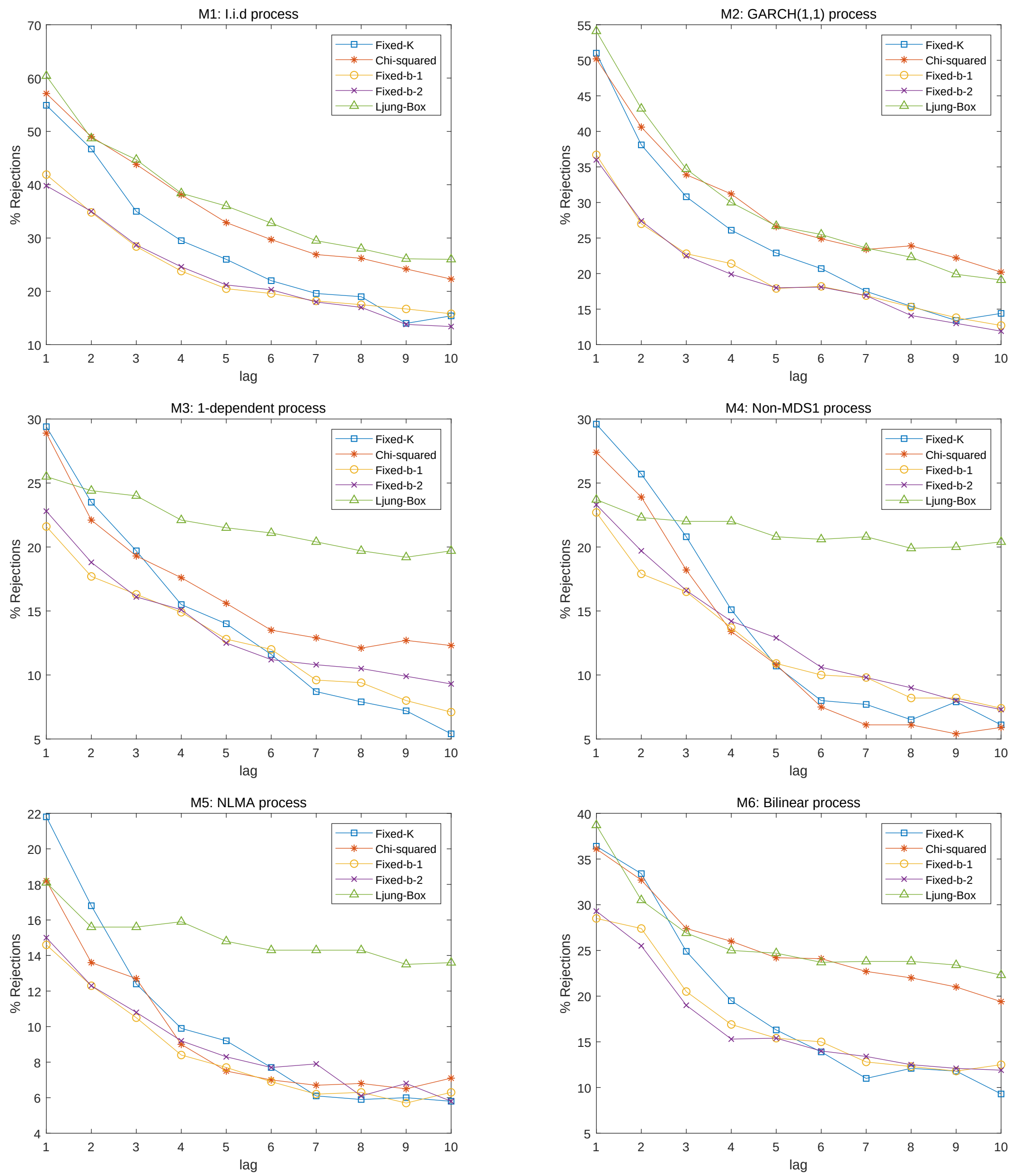


Figure 4: Size-adjusted empirical powers of the 5% tests under an MA(1) process with $T = 100$.

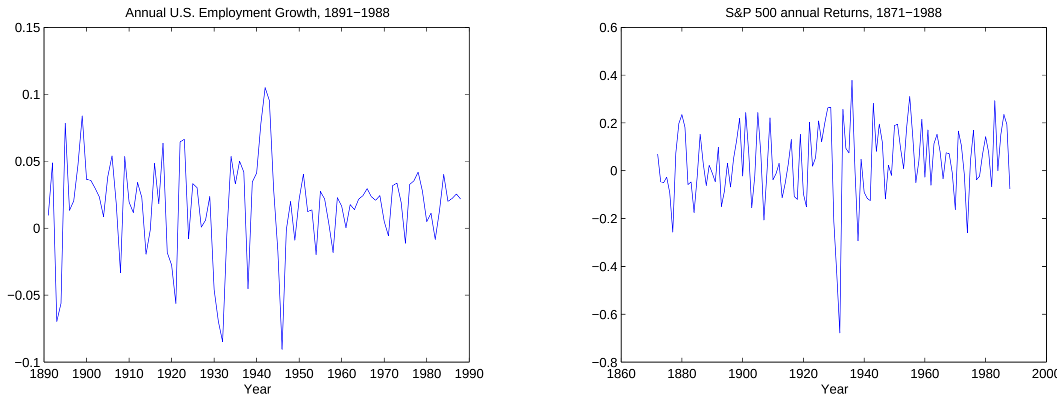


Figure 5: Time series plot of the employment growth and S& P500 returns.

| | Employment growth | | | | | S&P500 returns | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $s$ | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| $\tilde{Q}_T^*(s)$ | 0.11 | 0.02 | 0.08 | 0.01 | 0.02 | 0.10 | 0.05 | 0.06 | 0.04 | 0.09 |
| $\hat{L}_T(s)$ | 0.12 | 0.23 | 0.56 | 0.24 | 0.18 | 0.10 | 0.08 | 0.18 | 0.13 | 0.24 |
| $\tilde{L}_T(s)$ | 0.12 | 0.16 | 0.28 | 0.46 | 0.63 | 0.08 | 0.08 | 0.11 | 0.08 | 0.16 |
| $\widehat{\rho(s)}$ | 0.31 | −0.06 | −0.09 | −0.16 | −0.20 | 0.19 | −0.14 | −0.06 | −0.11 | −0.21 |
| | (0.17) | (0.13) | (0.13) | (0.12) | (0.12) | (0.13) | (0.08) | (0.16) | (0.15) | (0.09) |

Table 1: The p-values of the $\tilde{Q}_T^*(s)$, $\hat{L}_T(s)$ and $\tilde{L}_T(s)$ tests. $\hat{\rho}(s)$ is the $s$th sample autocorrelation coefficient, and the number in parentheses is the standard error of $\hat{\rho}(s)$.

cations. A more accurate test is preferred in practice, as it helps us reach more trustworthy conclusions.

# 6 CONCLUSION

We propose a new test to test zero serial autocorrelations in an otherwise dependent time series. By employing the orthonormal series variance estimator of the covariance matrix of the sample autocovariances, the new test follows an F distribution asymptotically under the fixed-smoothing asymptotics. Monte Carlo simulations show that this convenient F test has accurate size and highly competitive power. It would be interesting to extend this methodology to test the autocorrelations of the residuals from parametric time series models. We will pursue this nontrivial extension in future research.

# References


[1] Bilodeau, M., and Brenner, D. (1999): *Theory of Multivariate Statistics.* Springer.

[2] Blanchard, O., Diamond, P., Hall, R., & Yellen, J. (1989): "The Beveridge curve." *Brookings Papers on Economic Activity*, 1989(1), 1–76.

[3] Box, G. E. P. and Pierce, D. A. (1970): "Distribution of the residual autocorrelations in autoregressive integrated moving average time series models." *Journal of American Statistical Association* 65, 1509–26.

[4] Chinn, M., Ferrara, L., and Mignon, V. (2014): "Explaining US employment growth after the great recession: the role of output–employment non-linearities." *Journal of Macroeconomics* 42, 118–129.

[5] Hong, Y. (1996): "Consistent testing for serial correlation of unknown form." *Econometrica* 64, 837–864.

[6] Horowitz, J. L., Lobato, I. N., Nankervis, J. C. and Savin, N. E. (2006): "Bootstrapping the Box–Pierce Q test: a robust test of uncorrelatedness." *Journal of Econometrics* 133, 841–862.

[7] Kiefer, N. M. and Vogelsang, T. J. (2002): "Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size." *Econometric Theory*, 18:1350–1366.

[8] Kiefer, N. M. and Vogelsang, T. J. (2005): "A new asymptotic theory for heteroskedasticity-autocorrelation robust tests." *Econometric Theory* 21, 113–1164.

[9] Ljung, G. M. and Box, G. E. P. (1978): "On a measure of lack of fit in time series models." *Biometrika* 65, 297–303.

[10] Lobato, I. N. (2001): "Testing that a dependent process is uncorrelated." *Journal of the American Statistical Association* 96, 1066–1076.

[11] Lobato, I. N., Nankervis, J. C., and Savin, N. E. (2002): "Testing for zero autocorrelation in the presence of statistical dependence." *Econometric Theory* 18, 730–743.

[12] Nelson, C. and Plosser, C. (1982): "Trends and Random Walks in Macroeconomics Time Series: Some Evidence and Implications." *Journal of Monetary Economics* 10, 139-162.

[13] Newey, W. K. and West, K. D. (1987): "A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica* 55, 703–708.

[14] Phillips, P. C. B. (2005): "HAC estimation by automated regression." *Econometric Theory* 21, 116–142.

[15] Romano, J. P. and Thombs, L. A. (1996). "Inference for autocorrelations under weak assumptions." *Journal of the American Statistical Association* 91, 590–600.

[16] Shao, X. (2010): "A self-normalized approach to confidence interval construction in time series." *Journal of the Royal Statistical Society B* 72, 343–366.

[17] Shao, X. (2011): "Testing for white noise under unknown dependence and its applications to diagnostic checking for time series models." *Econometric Theory* 27, 312–343.

[18] Shao, X. (2015): "Self-normalization for time series: a review of recent developments." *Journal of the American Statistical Association* 110, 1797–1817.

[19] Sun, Y. (2011): "Robust trend inference with series variance estimator and testing-optimal smoothing parameter." *Journal of Econometrics* 164(2), 345–366.

[20] Sun, Y. (2013): "A heteroskedasticity and autocorrelation robust F test using orthonormal series variance estimator." *Econometrics Journal* 16, 1–26.

[21] Sun, Y. (2014a): "Let's Fix It: fixed-$b$ asymptotics versus small-$b$ asymptotics in heteroscedasticity and autocorrelation robust inference." *Journal of Econometrics* 178(3), 659–677.

[22] Sun, Y. (2014b): "Fixed-smoothing asymptotics in a two-step GMM framework." *Econometrica* 82(6), 2327–2370.

[23] Sun, Y., Phillips, P. C. B., and Jin, S. (2008): "Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing." *Econometrica* 76(1), 175–194.

[24] Zhang, X. and Shao, X. (2013): "Fixed-smoothing asymptotics for time series." *Annals of Statistics* 41, 1329-1349.