Constructing smooth tests without estimating the eigenpairs of the limiting process

Shih-Hsun Hsu [a], Chung-Ming Kuan [b,*]

[a] *Department of Economics, National Chengchi University, Taipei 116, Taiwan*
[b] *Department of Finance and CRETA, National Taiwan University, Taipei 106, Taiwan*

ARTICLE INFO

ABSTRACT

Based on the well known Karhunen–Loève expansion, it can be shown that many omnibus tests lack power against "high frequency" alternatives. The smooth tests of Neyman (1937) may be employed to circumvent this power deficiency problem. Yet, such tests may be difficult to compute in many applications. In this paper, we propose a more operational approach to constructing smooth tests. This approach hinges on a Fourier representation of the postulated empirical process with known Fourier coefficients, and the proposed test is based on the normalized principal components associated with the covariance matrix of finitely many Fourier coefficients. The proposed test thus needs only standard principal component analysis that can be carried out using most econometric packages. We establish the asymptotic properties of the proposed test and consider two data-driven methods for determining the number of Fourier coefficients in the test statistic. Our simulations show that the proposed tests compare favorably with the conventional smooth tests in finite samples.

© 2013 Published by Elsevier B.V.

1. Introduction

Specification tests are indispensable tools in the process of model searching. There are basically two types of specification tests: directional tests and omnibus tests. A directional test focuses on certain alternatives of interest. While this test is powerful against the postulated alternatives, it is not a consistent test in general because it may not have power against some other alternatives. On the other hand, when researchers do not have any particular alternative in mind, they may prefer an omnibus test that is capable of detecting any potential deviations from the null hypothesis. There are numerous omnibus tests in the literature, such as the tests of martingale difference (e.g., Durlauf, 1991; Deo, 2000; Domínguez and Lobato, 2003) and general specification tests (e.g., Bierens, 1982, 1990; Bierens and Ploberger, 1997).

It can be verified that the limits of many omnibus tests are a functional of some stochastic (possibly Gaussian) process. By the well known Karhunen–Loève (KL) expansion (Karhunen, 1946; Loève, 1955), the limiting process in an omnibus test can be represented as a weighted sum of the products of the normalized principal components and eigenfunctions associated with the covariance operator, with the weights being the corresponding eigenvalues that diminish to zero. This suggests that such omnibus tests mainly have local power against a few orthogonal directions determined by the eigenfunctions with larger eigenvalues, but lack power against "high frequency" alternatives, i.e., the directions related to very small eigenvalues. See, e.g., Eubank and LaRiccia (1992), Bierens and Ploberger (1997), Janssen (2000), and Escanciano (2009) for more discussions.

The aforementioned power deficiency may be circumvented by employing "smooth" tests, in the sense of Neyman (1937); see also Eubank and LaRiccia (1992), Fan (1996), Ghosh and Bera (2001), and Escanciano and Mayoral (2010).[1] By construction, such smooth tests avoid diminishing weights in the limit and hence have more even power against a collection of directions. There have been many smooth tests in the literature, such as the tests of goodness of fit (Eubank and LaRiccia, 1992; Delgado and Stute, 2008), tests of martingale difference (Delgado et al., 2005; Escanciano and Mayoral, 2010), and general specification tests (Stute, 1997; Escanciano, 2009). A major difficulty of smooth tests is that they may not be easy to implement, because the statistics rely on the eigenpairs (eigenfunctions and corresponding eigenvalues) of the limiting process, which are usually unknown. It is, however, technically involved to estimate these eigenpairs; see, e.g., William and Seeger

---

* Corresponding author.
  E-mail addresses: shhsu@nccu.edu.tw (S.-H. Hsu), ckuan@ntu.edu.tw (C.-M. Kuan).

[1] Another class of tests based on "kernel smoothing" also has better power against "high frequency" alternatives, e.g., Fan and Li (1996, 2000) and Fan (1998). In this paper, by smooth test we mean Neyman's smooth test.

(2000, 2001), Carrasco et al. (2007), and Escanciano (2009). Thus, smooth tests are not readily available in many applications.

In this paper, we propose a more operational approach to constructing smooth tests. This approach hinges on a Fourier representation of the postulated empirical process with known Fourier coefficients. The proposed smooth test is based on the normalized principal components associated with the covariance matrix of finitely many Fourier coefficients. We thus need only a standard principal component analysis that can be carried out using most econometric and statistics packages. This is much simpler than estimating the eigenpairs of the limiting process. We establish the asymptotic properties of the proposed test and consider two data-driven methods for determining the number of Fourier coefficients in the test statistic. The first method, proposed by Inglot and Ledwina (2006), employs a model-selection criterion; the second method, studied in Inglot et al. (1994) and Fan (1996), is designed to maximize the asymptotic power. Monte Carlo simulations show that the proposed smooth tests compare favorably with the conventional smooth tests in finite samples.

This paper proceeds as follows. We review the conventional smooth test and propose a new smooth test in Section 2. The asymptotic properties of the proposed test and two data-driven tests are discussed in Section 3. Section 4 reports simulation results. Section 5 concludes the paper. All technical conditions and proofs are deferred to Appendix.

2. The proposed smooth test

2.1. The CvM and smooth tests

Many omnibus tests for model specification can be expressed in terms of a functional of an empirical process. In these tests, the behavior of the empirical process is essentially governed by its limiting process under the null hypothesis but tends to deviate from the limiting process otherwise; the chosen functional is then used to measure these deviations. The well known functionals include the Kolmogorov–Smirnov (KS) functional, i.e., the supremum functional, and Cramér–von Mises (CvM) functional, i.e., $\int f^2(s)ds$ for a square integrable function f.

Let X_n denote a square integrable empirical process on $[a, b]$ such that $X_n \Rightarrow X$ on $[a, b]$, where $\Rightarrow$ denotes weak convergence (of the associated probability measure) and X is also a square integrable process with zero mean. An omnibus test based on the CvM functional (hereafter the CvM test) is such that

$$\int_a^b X_n^2(\tau)d\tau \xrightarrow{d} \int_a^b X^2(\tau)d\tau,$$

where $\xrightarrow{d}$ stands for convergence in distribution. The covariance operator of X, $\mathbb{K}_X$, with the kernel $K_X(s, \tau) = \mathbb{E}[X(s)X(\tau)]$, is such that

$$\mathbb{K}_X f(\tau) := \int_a^b K_X(s, \tau)f(s)ds.$$

Corresponding to $\mathbb{K}_X$, there exist orthonormal eigenfunctions $\{\varepsilon_m(\cdot)\}$ and the associated eigenvalues $\{\alpha_m\}$ that satisfy

$$\int_a^b K_X(s, \tau)\varepsilon_m(s)ds = \alpha_m \varepsilon_m(\tau),$$

where $\alpha_1 \geq \alpha_2 \geq \cdots$.[2]

When X is quadratic mean continuous on $[a, b]$, its KL expansion is, in the quadratic mean sense,

$$\begin{aligned} X(\tau) &= \lim_{M \to \infty} \sum_{m=1}^M z_m \varepsilon_m(\tau) \\ &= \lim_{M \to \infty} \sum_{m=1}^M \sqrt{\alpha_m} z_m^* \varepsilon_m(\tau), \quad \tau \in [a, b], \end{aligned} \quad (1)$$

where $z_m = \int_b^a X(s)\varepsilon_m(s)ds$ are the principal components, which are mutually uncorrelated with variance α_m, and $z_m^* = z_m/\sqrt{\alpha_m}$ are the normalized principal components with variance one. It is readily seen that (1) is also a Fourier representation in the eigenfunctions $\{\varepsilon_m(\cdot)\}$, with z_m the Fourier coefficients. It follows from (1) and the Parseval Theorem that, in the quadratic mean sense, the limit of the CvM test is:

$$\int_a^b X^2(\tau)d\tau = \lim_{M \to \infty} \sum_{m=1}^M z_m^2 = \lim_{M \to \infty} \sum_{m=1}^M \alpha_m(z_m^*)^2. \quad (2)$$

When $\mathbb{K}_X$ is square integrable, $\alpha_m \to 0$ as m tends to infinity.[3] Therefore, the CvM test based on X_n virtually has no local power against "high frequency" alternatives, i.e., the directions corresponding to very small eigenvalues (i.e., α_m with large m).

To alleviate the power deficiency in the CvM test, it is natural to construct a test whose limit does not involve the diminishing weights α_m. To this end, consider the process $\Xi_M(\tau) = \sum_{m=1}^M z_m^* \varepsilon_m(\tau)$ and note that

$$\int_a^b \Xi_M^2(\tau)d\tau = \sum_{m=1}^M (z_m^*)^2, \quad (3)$$

cf. (2). Letting $\hat{z}_{m,n}^*$ be consistent estimates of z_m^* based on the sample of size n, we may construct the following test: for a given M,

$$T_{n,M} = \sum_{m=1}^M (\hat{z}_{m,n}^*)^2. \quad (4)$$

It is clear that the limit of $T_{n,M}$ is (3). This is a smooth test in the sense of Neyman (1937); see also Ghosh and Bera (2001) for a review of Neyman's smooth test. When X is a Gaussian process, it is well known that z_m are independent Gaussian random variables so that z_m^* are i.i.d. $\mathcal{N}(0, 1)$. In this case, (3) has a $\chi^2(M)$ distribution.

Compared with the CvM test with the limit (2), the smooth test $T_{n,M}$ ought to have more even power against the directions corresponding to the first M principal components. For the remaining directions corresponding to other components (z_m^* with $m > M$), $T_{n,M}$ would have no power. Yet, the power loss may be minimal because the CvM test itself has little power against these directions, due to the presence of the diminishing weights α_m. On the other hand, $T_{n,M}$ cannot be easily implemented unless the eigenpairs of the covariance operator $\mathbb{K}_X$, hence the normalized principal components z_m^*, are known. Unfortunately, except for some special X processes, such as the standard Wiener process and Brownian bridge, the eigenpairs are unknown and need to be estimated. Estimating the eigenpairs of the covariance operator is, however, technically involved; see, e.g., William and Seeger (2000, 2001), Carrasco et al. (2007), and Escanciano (2009). Therefore, constructing smooth tests may not be as straightforward as one would think.

---

[2] The orthonormal eigenfunctions $\{\varepsilon_m(\cdot)\}$ satisfy $\int_a^b e_m(s)e_n(s)ds = 0$ and $\int_a^b e_m^2(s)ds = 1$.

---

[3] By square integrability of $\mathbb{K}_X$,

$$\int_a^b \int_a^b K_X^2(s, \tau)d\tau ds = \int_a^b \left(\sum_{m=1}^\infty \alpha_m^2 \varepsilon_m^2(s) \right) ds = \sum_{m=1}^\infty \alpha_m^2 < \infty,$$

where the first equality follows from the Parseval's theorem. It follows that $\alpha_m \to 0$ as $m \to \infty$. It can also be shown that $\alpha_m \to 0$ when $\mathbb{K}_X$ is integrable.

2.2. The proposed smooth test

As discussed in the preceding section, it may be difficult to compute the conventional smooth test when the normalized principal components, which are also the Fourier coefficients associated with the eigenfunctions $\{\varepsilon(\cdot)\}$ of X, are unknown. To circumvent this problem, we consider a different Fourier representation of X such that the sample counterparts of its Fourier coefficients have an analytic form. It is then straightforward to compute the normalized principal components of the first M sample Fourier coefficients and construct a smooth test based on these principal components.

The Fourier representation of X in the orthonormal basis functions $\{e_m(\cdot)\}$ is, in the quadratic mean sense,

$$X(\tau) = \lim_{M \to \infty} \sum_{m=1}^{M} \zeta_m e_m(\tau), \tag{5}$$

where $\zeta_m = \int_a^b X(\tau) e_m(\tau) d\tau$ are the associated Fourier coefficients, cf. the KL expansion (1). Given the first M Fourier coefficients $\boldsymbol{\zeta}_M = [\zeta_1 \zeta_2 \ldots \zeta_M]'$, let $\text{var}(\boldsymbol{\zeta}_M)$ denote its variance–covariance matrix, with $\lambda_{i,M}$ and $\boldsymbol{u}_{i,M}$, $i = 1, \ldots, M$, its associated eigenvalues and eigenvectors, respectively, such that $\lambda_{1,M} \geq \lambda_{2,M} \geq \cdots \geq \lambda_{M,M} > 0$. Then,

$$\text{var}(\boldsymbol{\zeta}_M) = \boldsymbol{U}_M \boldsymbol{\Lambda}_M \boldsymbol{U}_M',$$

where $\boldsymbol{U}_M = [\boldsymbol{u}_{1,M} \boldsymbol{u}_{2,M} \ldots \boldsymbol{u}_{M,M}]$ satisfies $\boldsymbol{U}_M' \boldsymbol{U}_M = \boldsymbol{U}_M \boldsymbol{U}_M' = \boldsymbol{I}_M$, and $\boldsymbol{\Lambda}_M$ is the diagonal matrix with $\lambda_{i,M}$ on the principal diagonal. The normalized principal components are

$$\boldsymbol{\zeta}_M^* := [\zeta_{1,M}^* \zeta_{2,M}^* \ldots \zeta_{M,M}^*]' = \boldsymbol{\Lambda}_M^{-1/2} \boldsymbol{U}_M' \boldsymbol{\zeta}_M. \tag{6}$$

Similar to z_m^* in the KL expansion, $\zeta_{m,M}^*$ are uncorrelated random variables with mean zero and variance 1, and they are i.i.d. $\mathcal{N}(0, 1)$ when X is Gaussian.

Note that $\sum_{m=1}^{M} \zeta_{m,M}^2 = \sum_{m=1}^{M} \lambda_{m,M}(\zeta_{m,M}^*)^2$ which also involves diminishing weights $\lambda_{m,M}$, cf. (2). In the light of the conventional smooth test, we may construct a test with the limit $\sum_{m=1}^{M} (\zeta_{m,M}^*)^2$, which does not involve the diminishing weights $\lambda_{m,M}$. To this end, we consider the Fourier representation of the empirical process X_n in $\{e_m(\cdot)\}$:

$$X_n(\tau) = \lim_{M \to \infty} \sum_{m=1}^{M} \zeta_{m,n} e_m(\tau), \tag{7}$$

with $\zeta_{m,n}$ the Fourier coefficient of X_n associated with the basis function $e_m(\cdot)$. In particular, when X_n is such that

$$X_n(\tau) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \phi(\boldsymbol{\eta}_t, \tau) + o_{\mathbb{P}}(1),$$

where $\mathbb{E}[\phi(\boldsymbol{\eta}_t, \cdot)|\mathcal{F}_{t-1}] = 0$, the Fourier coefficients are:

$$\zeta_{m,n} = \int_a^b X_n(\tau) e_m(\tau) d\tau$$
$$= \int_a^b \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \phi(\boldsymbol{\eta}_t, \tau) e_m(\tau) d\tau := \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Phi_m(\boldsymbol{\eta}_t),$$

where $\Phi_m(\boldsymbol{\eta}_t) = \int_a^b \phi(\boldsymbol{\eta}_t, \tau) e_m(\tau) d\tau$. With proper choice of basis functions $\{e_m(\cdot)\}$, Φ_m may have analytic forms; see next subsection for examples.

Given $X_n \Rightarrow X$, $\boldsymbol{\zeta}_{n,M} = [\zeta_{1,n} \zeta_{2,n} \ldots \zeta_{M,n}]'$ converges in distribution to $\boldsymbol{\zeta}_M$, and hence,

$$\text{var}(\boldsymbol{\zeta}_M) = \lim_{n \to \infty} \text{var}(\boldsymbol{\zeta}_{n,M}) = \lim_{n \to \infty} \text{var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{\Phi}_M(\boldsymbol{\eta}_t)\right),$$

where $\boldsymbol{\Phi}_M(\boldsymbol{\eta}_t) = [\Phi_1(\boldsymbol{\eta}_t) \Phi_2(\boldsymbol{\eta}_t) \ldots \Phi_M(\boldsymbol{\eta}_t)]'$. When $\zeta_{m,n}$, $m = 1, \ldots, M$, have analytic forms, a consistent estimator for $\text{var}(\boldsymbol{\zeta}_{n,M})$, denoted as $\widehat{\text{var}(\boldsymbol{\zeta}_{n,M})}$, can be conveniently computed using an Eicker–White estimator or a Newey–West estimator. Such an estimator provides a good approximation to $\text{var}(\boldsymbol{\zeta}_M)$ without knowing $\boldsymbol{\zeta}_M$ or $\text{var}(\boldsymbol{\zeta}_M)$. The eigenpairs of $\widehat{\text{var}(\boldsymbol{\zeta}_{n,M})}$, the matrix of eigenvectors ($\widehat{\boldsymbol{U}}_{n,M}$) and the diagonal matrix with eigenvalues on the principal diagonal ($\widehat{\boldsymbol{\Lambda}}_{n,M}$), are then readily computed. The resulting normalized principal components are:

$$\widehat{\boldsymbol{\zeta}}_{n,M}^* := [\widehat{\zeta}_{1,n,M}^* \widehat{\zeta}_{2,n,M}^* \ldots \widehat{\zeta}_{M,n,M}^*]' = \widehat{\boldsymbol{\Lambda}}_{n,M}^{-1/2} \widehat{\boldsymbol{U}}_{n,M}' \boldsymbol{\zeta}_{n,M}, \tag{8}$$

cf. (6). In analogy with the conventional smooth test (4), the proposed test is:

$$J_{n,M} = \sum_{m=1}^{M} (\widehat{\zeta}_{m,n,M}^*)^2. \tag{9}$$

Our approach relies on the first M Fourier coefficients of a Fourier representation of the empirical process X_n and their principal components; the knowledge of the covariance structure of the limiting process X is not needed.[4] Note that a key element of our approach is to find suitable basis functions such that the Fourier coefficients have analytic forms; see the examples in the subsection below. Then, consistent estimation of $\text{var}(\boldsymbol{\zeta}_{n,M})$ and estimation of its eigenpairs can all be done using the standard econometric packages. By contrast, the conventional smooth test, in general, requires estimating the eigenpairs of the covariance operator of the limiting process X, which is practically more cumbersome. The proposed test is thus more operational and hence a useful complement when the conventional smooth test is not readily available.

2.3. Examples

In this section, we illustrate the proposed test using the following examples. In particular, we show how basis functions may be chosen so as to deliver Fourier coefficients with analytic forms. The first example is concerned with testing model specification, where the proposed test works but the conventional smooth test does not. The second example focuses on the hypothesis of martingale difference, where both the conventional and proposed smooth tests are applicable.

Example 1: testing model specification

Consider the null hypothesis of correct model specification:

$$H_0 : \mathbb{E}[y_t | \mathcal{F}_{t-1}] = f(x_t, \theta_o), \quad \text{for some } \theta_o \in \Theta,$$

where $\mathcal{F}_{t-1}$ is the information set up to time $t-1$ containing the variable x_t. For simplicity, we assume y_t and x_t to be one-dimensional. A model specification test may be based on the following empirical process:

$$X_n(\tau) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} [y_t - f(\boldsymbol{x}_t, \theta_o)] w(x_t, \tau) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \phi(\boldsymbol{\eta}_t, \tau),$$

where $w(x_t, \tau)$ is a misspecification indicator indexed by τ ($0 \leq \tau \leq 1$), and $\phi(\boldsymbol{\eta}_t, \tau) = [y_t - f(x_t, \theta_o)] w(x_t, \tau)$ with $\boldsymbol{\eta}_t = (y_t, x_t)'$.

Two common choices of $w(x_t, \tau)$ in the literature are: the indicator function $\mathbf{1}(x_t \leq \tau)$, where $\mathbf{1}(S) = 1$ if S occurs and 0 otherwise, and the exponential function $\exp(x_t \tau)$. Taking the

---

[4] Escanciano and Mayoral (2010) also suggest to construct smooth tests based on an expansion in general orthogonal basis functions. Their approach is quite different from ours, as it still depends on the covariance kernel of the limiting process.

classical Fourier series as basis functions, we have

$$w(x_t, \tau) = \varphi_0(x_t)$$
$$+ \sum_{k=1}^{\infty} \left[\varphi_k^c(x_t) \cos(2\pi k\tau) + \varphi_k^s(x_t) \sin(2\pi k\tau) \right],$$

where $\varphi_0(x_t)$, $\varphi_k^c(x_t)$ and $\varphi_k^s(x_t)$ are the Fourier coefficients. It can be shown that, when $w(x_t, \tau) = \exp(x_t\tau)$, the Fourier coefficients are

$$\varphi_0(x_t) = \frac{1}{x_t}[\exp(x_t) - 1],$$

$$\varphi_k^c(x_t) = \frac{x_t}{x_t^2 + (2\pi k)^2}[\exp(x_t) - 1], \quad k = 1, 2, \ldots,$$

$$\varphi_k^s(x_t) = \frac{-2\pi k}{x_t^2 + (2\pi k)^2}[\exp(x_t) - 1], \quad k = 1, 2, \ldots.$$

Then, the empirical process X_n is such that

$$X_n(\tau) = \zeta_{0,n} + \lim_{K \to \infty} \sum_{k=1}^{K} \left[\zeta_{k,n}^c \cos(2\pi k\tau) + \zeta_{k,n}^s \sin(2\pi k\tau) \right],$$

where

$$\zeta_{0,n} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} [y_t - f(x_t, \theta_o)]\varphi_0(x_t),$$

$$\zeta_{k,n}^c = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} [y_t - f(x_t, \theta_o)]\varphi_k^c(x_t), \quad k = 1, 2, \ldots,$$

$$\zeta_{k,n}^s = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} [y_t - f(x_t, \theta_o)]\varphi_k^s(x_t), \quad k = 1, 2, \ldots.$$

In the notations of preceding subsections, we have:

$$\zeta_{n,M} = \begin{cases} \left[\zeta_{0,n} \zeta_{1,n}^c \zeta_{1,n}^s \cdots \zeta_{M/2,n}^c \right]', & M \text{ is even,} \\ \left[\zeta_{0,n} \zeta_{1,n}^c \zeta_{1,n}^s \cdots \zeta_{(M-1)/2,n}^c \zeta_{(M-1)/2,n}^s \right]', & \\ \quad \text{otherwise;} \end{cases} \quad (10)$$

$$\varphi_M = \begin{cases} \left[\varphi_0 \varphi_1^c \varphi_1^s \cdots \varphi_{M/2}^c \right]', & M \text{ is even,} \\ \left[\varphi_0 \varphi_1^c \varphi_1^s \cdots \varphi_{(M-1)/2}^c \varphi_{(M-1)/2}^s \right]', & \text{otherwise.} \end{cases} \quad (11)$$

Under the null hypothesis,

$$\text{var}(\zeta_{n,M}) = \text{var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} [y_t - f(x_t, \theta_o)]\varphi_M(x_t) \right)$$
$$= \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}\left([y_t - f(x_t, \theta_o)]^2 \varphi_M(x_t)\varphi_M'(x_t) \right).$$

We can consistently estimate $\text{var}(\zeta_{n,M})$ using an Eicker–White estimator:

$$\widehat{\text{var}(\zeta_{n,M})} = \frac{1}{n} \sum_{t=1}^{n} [\hat{u}_t^2 \varphi_M(x_t)\varphi_M'(x_t)], \quad (12)$$

with $\hat{u}_t = y_t - f(x_t, \hat{\theta}_n)$ and $\hat{\theta}_n$ a consistent estimator for θ_o. The eigenpairs and principal components of (12) are then readily computed. We can also derive the analytic-forms of $\varphi_M(x_t)$ when $\mathbf{1}(x_t \le \tau)$ is employed as the misspecification indicator and compute $\widehat{\text{var}(\zeta_{n,M})}$ accordingly; see Example 2 below. Note that the conventional smooth test is not available in this case, because X_n converges to a Gaussian process whose covariance kernel is unknown.

Example 2: testing the martingale difference hypothesis

Consider now the hypothesis of martingale difference:

$$H_0 : \mathbb{E}[y_t|\mathcal{F}_{t-1}] = 0. \quad (13)$$

Following Escanciano and Mayoral (2010), the marked empirical process is

$$X_n(\xi) = \frac{1}{\sigma_o\sqrt{n}} \sum_{t=1}^{n} y_t \mathbf{1}(y_{t-1} \le \xi),$$

where $\sigma_o^2 = \mathbb{E}[y_t^2|\mathcal{F}_{t-1}]$. Let $F(\cdot)$ be a non-decreasing transformation such that $0 \le F(\xi) \le 1$, $\forall \xi \in \mathbb{R}$. Then, $X_n(\xi)$ can be expressed as

$$X_n(\tau) = \frac{1}{\sigma_o\sqrt{n}} \sum_{t=1}^{n} y_t \mathbf{1}(F_{t-1} \le \tau),$$

with $F_{t-1} := F(Y_{t-1})$ and $F(\xi) := \tau$.

In this case, the Fourier representation for X_n is

$$X_n(\tau) = \zeta_{0,n} + \lim_{K \to \infty} \sum_{k=1}^{K} \left[\zeta_{k,n}^c \cos(2\pi k\tau) + \zeta_{k,n}^s \sin(2\pi k\tau) \right],$$

where

$$\zeta_{0,n} = \frac{1}{\sigma_o\sqrt{n}} \sum_{t=1}^{n} y_t \varphi_0(F_{t-1}),$$

$$\zeta_{k,n}^c = \frac{1}{\sigma_o\sqrt{n}} \sum_{t=1}^{n} y_t \varphi_k^c(F_{t-1}), \quad k = 1, 2, \ldots,$$

$$\zeta_{k,n}^s = \frac{1}{\sigma_o\sqrt{n}} \sum_{t=1}^{n} y_t \varphi_k^s(F_{t-1}), \quad k = 1, 2, \ldots,$$

and

$$\varphi_0(F_{t-1}) = 1 - F_{t-1},$$

$$\varphi_k^c(F_{t-1}) = \frac{-1}{2\pi k} \sin(2\pi k F_{t-1}), \quad k = 1, 2, \ldots,$$

$$\varphi_k^s(F_{t-1}) = \frac{1}{2\pi k}[\cos(2\pi k F_{t-1}) - 1], \quad k = 1, 2, \ldots.$$

Using the same notations in (10) and (11), we have under the null hypothesis that

$$\text{var}(\zeta_{n,M}) = \text{var}\left(\frac{1}{\sigma_o\sqrt{n}} \sum_{t=1}^{n} y_t \varphi_M(F_{t-1}) \right)$$
$$= \frac{1}{\sigma_o^2 n} \sum_{t=1}^{n} \mathbb{E}\left[y_t^2 \varphi_M(F_{t-1})\varphi_M'(F_{t-1}) \right].$$

A consistent estimator of $\text{var}(\zeta_{n,M})$ is, with $\hat{\sigma}_n^2 = \sum_{t=1}^{n} y_t^2/n$,

$$\widehat{\text{var}(\zeta_{n,M})} = \frac{1}{\hat{\sigma}_n^2 n} \sum_{t=1}^{n} y_t^2 \varphi_M(F_{t-1})\varphi_M'(F_{t-1}), \quad (14)$$

from which the eigenpairs can be easily computed. Note that the conventional smooth test works in this example because X_n converges to the standard Wiener process whose covariance kernel and eigenpairs are well known; see Section 4.1 below and also Escanciano and Mayoral (2010) for details.

3. Asymptotic properties

In this section, we will establish the asymptotic properties of the proposed smooth test. Consider the general empirical process:

$$X_n(\tau) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \phi(\eta_t, \tau) + o_{\mathbb{P}}(1), \quad \tau \in [a, b],$$

where $\mathbb{E}[\phi(\eta_t, \cdot)|\mathcal{F}_{t-1}] = 0$, and $\mathcal{F}_t$ is the sigma-algebra generated by $\{\eta_i : 1 \le i \le t\}$. We will maintain the assumption that $X_n \Rightarrow X$ on $[a, b]$, where X is a quadratic mean continuous, Gaussian

process with zero mean and the covariance kernel $K_X(\cdot, \cdot)$. This convergence holds under mild conditions on the data generating process. We summarize the required regularity conditions in the Appendix; see also Stute (1997) and Escanciano and Mayoral (2010) for such conditions in different contexts.

Given the basis functions $\{e_m(\cdot)\}$, ζ_m and $\zeta_{m,n}$ are the Fourier coefficients of the Fourier representations (5) and (7), respectively. The result below follows easily from the maintained assumption that X_n converges weakly to a Gaussian process X with mean zero.

Lemma 3.1. *For a given M, $\boldsymbol{\zeta}_{n,M} \xrightarrow{d} \boldsymbol{\zeta}_M$, as $n \to \infty$, where $\boldsymbol{\zeta}_M$ is a vector of M normal random variables with mean zero.*

It is then not difficult to see that the normalized principal components of $\widehat{\mathrm{var}(\boldsymbol{\zeta}_{n,M})}$ also converge in distribution to those of $\mathrm{var}(\boldsymbol{\zeta}_M)$.

Lemma 3.2. *For a given M,*

$$\widehat{\boldsymbol{\zeta}}^*_{n,M} \xrightarrow{d} \boldsymbol{\zeta}^*_M \sim \mathcal{N}(0, \boldsymbol{I}_M),$$

as $n \to \infty$, where $\widehat{\boldsymbol{\zeta}}^*_{n,M}$ is given by (8).

The limiting distribution of $J_{n,M}$ is an immediate consequence of Lemma 3.2 and is the same as that of a conventional smooth test.

Theorem 3.3. *For a given M,*

$$J_{n,M} = (\widehat{\boldsymbol{\zeta}}^*_{n,M})'(\widehat{\boldsymbol{\zeta}}^*_{n,M}) \xrightarrow{d} (\boldsymbol{\zeta}^*_M)'(\boldsymbol{\zeta}^*_M) \sim \chi^2(M),$$

as $n \to \infty$.

It remains to determine the number of components, M, in the proposed smooth test. Two data-driven methods are considered in this paper. The first one is based on a model-selection type criterion:

$$J^{ms}_{n,\widetilde{M}} = \sum_{m=1}^{\widetilde{M}} (\widehat{\zeta}^*_{m,n,M})^2, \tag{15}$$

where "ms" in the notation stands for "model selection", $\widetilde{M}$ is a number between the lower bound $\underline{m}$ and the upper bound $\overline{M}$, such that

$$\sum_{m=1}^{\widetilde{M}} (\widehat{\zeta}^*_{m,n,M})^2$$
$$= \max \left(\sum_{m=1}^{M} (\widehat{\zeta}^*_{m,n,M})^2 - \mathcal{P}(n, M, q), \underline{m} \leq M \leq \overline{M} \right),$$

with $\mathcal{P}(n, M, q)$ a penalty term combining Schwarz's and Akaike's selection rules:

$$\mathcal{P}(n, M, q) = \begin{cases} 2M \ln(n), & \text{if } \max_{1 \leq m \leq \overline{M}} |\widehat{\zeta}^*_{m,n,M}| \leq \sqrt{q \ln(m)}, \\ 2M, & \text{otherwise,} \end{cases}$$

and q a fixed number. This data-driven method was also considered by Inglot and Ledwina (2006), Escanciano and Lobato (2009), and Escanciano and Mayoral (2010), among others. When the proposed $\widehat{\zeta}^*_{m,n,M}$ in (15) are replaced with the normalized principal components $\widehat{z}^*_{m,n}$ in (4), we have exactly the data-driven smooth test of Escanciano and Mayoral (2010), which will be referred to as $T^{ms}_{n,\widetilde{M}}$ in our subsequent simulations. It has also been shown that, under the null, $\widetilde{M} \to \underline{m}$ in probability. The result below is analogous to that of Escanciano and Mayoral (2010).

Theorem 3.4. *The asymptotic distribution of $J^{ms}_{n,\widetilde{M}}$ under the null is*

$$J^{ms}_{n,\widetilde{M}} \xrightarrow{d} \chi^2(\underline{m}).$$

In addition, we consider the data-driven method of Inglot et al. (1994) and Fan (1996); the resulting test is also known as an "adaptive Neyman's test". Specifically, Fan (1996) suggests maximizing the asymptotic power of $J_{n,M}$ tests by choosing the maximum of the standardized $J_{n,M}$ tests, i.e.,

$$J_{n,M^*} = \max_{1 \leq M \leq \overline{M}} \left\{ \frac{J_{n,M} - M}{\sqrt{2M}} \right\},$$

given the upper bound $\overline{M}$. Following Fan (1996), the second data-driven test is based on an adjustment of J_{n,M^*}:

$$J^{an}_{n,M^*} = \sqrt{2 \ln \ln(\overline{M})} J_{n,M^*}$$
$$- \left\{ 2 \ln \ln(\overline{M}) + 0.5 \ln \ln \ln(\overline{M}) - 0.5 \ln(4\pi) \right\}, \tag{16}$$

where "an" in the notation stands for "adaptive Neyman". This adjustment leads to the following analytic result for J^{an}_{n,M^*}; see Fan (1996, Theorem 2.1).

Theorem 3.5. *Under the null, the asymptotic distribution of J^{an}_{n,M^*} is*

$$\mathbb{P}(J^{an}_{n,M^*} < \varepsilon) \to \exp(-\exp(-\varepsilon)), \quad \text{as } \overline{M}, n \to \infty.$$

Given the significance level α, the corresponding asymptotic critical value, c^{an}_α, is such that $\exp(-\exp(-c^{an}_\alpha)) = 1 - \alpha$. That is, the asymptotic critical region of this data-driven test with significance level α is:

$$J^{an}_{n,M^*} > c^{an}_\alpha = -\ln(-\ln(1 - \alpha)).$$

Although this is an analytic result, Fan and Lin (1998) and Fan and Huang (2001) find in their simulations that this approximation may be poor when the sample size is small or when $\overline{M}$ is small. Instead, one may implement this test based on the simulated distribution of J^{an}_{n,M^*} for each $\overline{M}$.

4. Simulations

In our simulations, we consider testing the martingale difference hypothesis and linearity of model specifications. As the conventional smooth test is available only for the former, the proposed test is compared with that of Escanciano and Mayoral (2010) in this case but not otherwise. As benchmarks, we also compute the CvM and KS tests based on a wild bootstrap procedure in all experiments. In our simulations, we consider three sample sizes $n = 100$, 200, 300. For the proposed smooth tests, we simulate $J_{n,M}$ (9) for $M = 1, 2, 3, 4, 5$, and two corresponding data-driven tests $J^{ms}_{n,\widetilde{M}}$ (15) and J^{an}_{n,M^*} (16), with $q = 2.4$, $\underline{m} = 3$, and $\overline{M} = 5, 8, 11$ for $n = 100$, 200 and 300, respectively.[5] Note that we implement J^{an}_{n,M^*} based on the critical values of the simulated distribution. All nominal sizes are 5%. The number of Monte Carlo replications is 3000; the number of bootstraps is 500.

4.1. Testing the martingale difference hypothesis

For testing the martingale difference hypothesis (13), we follow the simulations in Escanciano and Mayoral (2010). Letting u_t be i.i.d. $\mathcal{N}(0, 1)$, we consider three different data generating processes (DGPs) for size simulations.

(1) IID: $y_t = u_t$.
(2) GARCH: $y_t = \sigma_t u_t$, with $\sigma_t^2 = 0.001 + 0.01 y_{t-1}^2 + 0.90 \sigma_{t-1}^2$.

---

[5] To ease comparison, we set the parameters of our simulations as those in Escanciano and Mayoral (2010).

(3) SV (Stochastic Volatility): $y_t = \exp(\sigma_t)u_t$, where $\sigma_t = 0.936\sigma_{t-1} + 0.32v_t$, and v_t are also i.i.d. $\mathcal{N}(0,1)$ and $\{u_t\}$ and $\{v_t\}$ are mutually independent.

The proposed smooth test (9) and two data-driven tests (15) and (16) are computed using the eigenpairs of (14). Let $\sigma^2 = \mathbb{E}[y_t^2]$ and $\hat{\sigma}_n^2 = \sum_{t=1}^n y_t^2/n$ be a consistent estimate. The conventional CvM and KS tests are:

$$\text{CvM}_n = \frac{1}{n}\sum_{i=2}^n \left[\frac{1}{\hat{\sigma}_n\sqrt{n}}\sum_{t=1}^n y_t \mathbf{1}(y_{t-1} \leq y_{i-1})\right]^2,$$

$$\text{KS}_n = \max_{i=2,\dots,n}\left|\frac{1}{\hat{\sigma}_n\sqrt{n}}\sum_{t=1}^n y_t \mathbf{1}(y_{t-1} \leq y_{i-1})\right|.$$

Under suitable regularity conditions,

$$X_n(\xi) \Rightarrow \mathbf{W}(\tau^2(\xi)),$$

where $\mathbf{W}$ is the standard Wiener process, and $\tau^2(\xi) := \sigma^{-2}\mathbb{E}[y_t^2\mathbf{1}(y_{t-1} \leq \xi)]$. It is well known that the eigenpairs associated with the covariance kernel of $\mathbf{W}$ are:

$$\alpha_m = \frac{1}{(m-1/2)^2\pi^2},$$

$$\varepsilon_m(t) = \sqrt{2}\sin((m-1/2)\pi t), \quad t \in [0,1], \; m = 1, 2, \dots.$$

The conventional smooth test is computed as (4): $T_{n,M} = \sum_{m=1}^M (\hat{z}_{m,n}^*)^2$, with

$$\hat{z}_{m,n}^* = \frac{1}{\sqrt{\alpha_m}}\int_{\mathbb{R}} \varepsilon_m(\tau_n^2(\xi))\left[\frac{1}{\hat{\sigma}_n\sqrt{n}}\sum_{t=1}^n y_t\mathbf{1}(y_{t-1} \leq \xi)\right]\tau_n^2(d\xi)$$

$$= \frac{\sqrt{2}}{\hat{\sigma}_n\sqrt{n}}\sum_{i=2}^n y_i\cos\left((m-1/2)\pi\tau_n^2(y_{i-1})\right),$$

and $\tau_n^2(\xi) = \hat{\sigma}_n^{-2}\sum_{t=2}^n y_t^2\mathbf{1}(y_{t-1} \leq \xi)/n$.[6] The data-driven test, $T_{n,\widetilde{M}}^{ms}$, of Escanciano and Mayoral (2010) is computed according to (15), where $\widehat{\zeta}_{m,n,M}^*$ are replaced with $\hat{z}_{m,n}^*$.

The empirical sizes are summarized in Table 1. As expected, the empirical sizes of the bootstrapped CvM and KS tests are very close to the nominal size 5% in all cases. The smooth tests of Escanciano and Mayoral (2010), $T_{n,M}$, and the proposed smooth tests, $J_{n,M}$, perform reasonably well in most cases but are under-sized when the DGP is SV. It can be seen that the data-driven tests, $T_{n,\widetilde{M}}^{ms}$ of Escanciano and Mayoral (2010) and the proposed $J_{n,\widetilde{M}}^{ms}$ and J_{n,M^*}^{an}, also have quite accurate sizes, except when the DGP is SV.

For power simulations we consider the following DGPs: let u_t be i.i.d. $\mathcal{N}(0,1)$.

(4) NLMA (Nonlinear Moving Average): $y_t = u_{t-1}u_{t-2}(u_{t-2} + u_t + 1)$.
(5) BIL (Bilinear): $y_t = u_t + 0.15u_{t-1}y_{t-1} + 0.05u_{t-1}y_{t-2}$.
(6) TAR-1 (Threshold AR):

$$y_t = \begin{cases} -0.5y_{t-1} + u_t, & \text{if } y_{t-1} \geq 1, \\ 0.4y_{t-1} + u_t, & \text{otherwise.} \end{cases}$$

(7) Exp-AR (Exponential AR): $y_t = 0.6y_{t-1}\exp(-0.5y_{t-1}^2) + u_t$.

The empirical powers are summarized in Table 2. Clearly, KS_n and CvM_n have quite different power performance. While KS_n has no power in all cases but BIL, CvM_n has no power against NLMA but high power against TAR-1. Compared with CvM_n, the conventional

smooth tests $T_{n,M}$ have better empirical powers under NLMA and Exp-AR (except for $T_{n,1}$) and have comparable powers under BIL and TAR-1. It is interesting to observe that the proposed smooth tests, $J_{n,M}$, dominate the conventional smooth tests in most cases. These results are encouraging, as they indicate that deviations from the null may be detected without knowledge of the covariance kernel of the limiting process. Similar results are also obtained when comparing the proposed data-driven tests $J_{n,\widetilde{M}}^{ms}$ and J_{n,M^*}^{an} with $T_{n,\widetilde{M}}^{ms}$ of Escanciano and Mayoral (2010). Yet, the powers of J_{n,M^*}^{an} are, in general, slightly lower than those of $J_{n,\widetilde{M}}^{ms}$, except for the case of TAR-1. Compared with those $J_{n,M}$ tests, we find that the powers of the two proposed data-driven tests are neither the highest nor the lowest. This indicates that the proposed data-driven tests are more robust than the smooth test with a given M, since the directions of deviations from the null may vary from case to case.

4.2. Testing linear model specification

We now consider testing the hypothesis of a correct linear model specification:

$$H_0 : \mathbb{P}(\mathbb{E}[y_t|x_t] = x_t\theta_o) = 1 \quad \text{for some } \theta_o \in \Theta \subset \mathbb{R}.$$

Following Lee et al. (1993), we generate four DGPs for power simulations with u_t i.i.d. $\mathcal{N}(0,1)$ and $y_0 = 0$.

(1) NLAR (Nonlinear AutoRegressive): $y_t = 0.7|y_{t-1}|/[|y_{t-1}| + 2] + u_t$.
(2) STAR (Smooth Transition AutoRegressive):
   $y_t = 0.6\Phi(y_{t-1})y_{t-1} + u_t$, where $\Phi(\cdot)$ denotes the standard normal distribution function.
(3) Threshold AutoRegressive (TAR-2):

$$y_t = \begin{cases} 0.9y_{t-1} + u_t, & \text{if } |y_{t-1}| < 1, \\ -0.3y_{t-1} + u_t, & \text{otherwise.} \end{cases}$$

(4) Sign autoregressive (SGN): $y_t = \text{sgn}(y_{t-1}) + u_t$, where

$$\text{sgn}(x) = \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ -1, & \text{if } x < 0. \end{cases}$$

Constructing misspecification indicators using the indicator function, the proposed smooth test (9) and two data-driven tests (15) and (16) are computed using the eigenpairs of (12). The CvM and KS test statistics are computed as:

$$\text{CvM}_n = \frac{1}{n}\sum_{i=1}^n\left|\frac{1}{\sqrt{n}}\sum_{t=1}^n\left(y_t - x_t\hat{\theta}_n\right)\mathbf{1}(x_t \leq x_i)\right|^2,$$

$$\text{KS}_n = \max_{i=1,\dots,n}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^n\left(y_t - x_t\hat{\theta}_n\right)\mathbf{1}(x_t \leq x_i)\right|,$$

where $\hat{\theta}_n$ is the OLS estimator. As the KL expansion is not available in this case, we do not consider the conventional smooth tests in these simulations.

The empirical powers are summarized in Table 3. Compared with CvM_n and KS_n, $J_{n,M}$ with $M = 2, 3, 4, 5$ perform significantly better under TAR-2 and SGN but are less powerful under NLAR; none of these tests have clear power advantage under STAR. Again, compared with those $J_{n,M}$ tests, the two proposed data-driven tests, $J_{n,\widetilde{M}}^{ms}$ and J_{n,M^*}^{an}, are more robust because their powers are neither the highest nor the lowest. It is also interesting to note that, while $J_{n,\widetilde{M}}^{ms}$ outperforms J_{n,M^*}^{an} in the previous simulations, J_{n,M^*}^{an} performs better in most cases here. This suggests that each data-driven method has its own merit and that these two data-driven tests can complement each other.

---

[6] Note that $\hat{z}_{m,n}^*$ are different from those in the simulations of Escanciano and Mayoral (2010), which are based on a "sample" version of z_m^*; see Eq. (4) in their paper.

Table 1
Size simulations: testing the martingale difference hypothesis.

Test	IID			GARCH			SV		
	$n = 100$	200	300	100	200	300	100	200	300
CvM_n	5.37	4.50	5.47	5.00	4.80	5.40	3.77	4.53	4.43
KS_n	5.20	5.03	5.53	4.83	5.40	5.03	3.13	3.67	4.53
$T_{n,1}$	4.93	4.57	5.50	4.73	4.70	5.40	1.70	2.70	3.53
$T_{n,2}$	4.47	4.63	4.90	4.40	4.97	4.47	1.60	2.47	2.53
$T_{n,3}$	4.00	4.27	4.70	4.03	4.67	4.53	1.40	1.90	2.30
$T_{n,4}$	3.93	4.17	4.90	3.60	4.43	4.20	1.37	1.93	2.73
$T_{n,5}$	3.63	3.97	5.10	3.37	4.50	4.13	1.03	1.67	2.73
$T_{n,\tilde{M}}^{ms}$	4.23	4.37	4.80	4.40	4.87	4.63	1.43	1.97	2.30
$J_{n,1}$	5.13	4.53	5.73	5.07	5.47	4.97	3.30	4.10	4.30
$J_{n,2}$	5.23	4.70	5.77	5.10	5.03	4.87	2.77	3.20	4.40
$J_{n,3}$	4.33	4.53	5.10	3.77	4.93	4.33	1.93	2.63	3.93
$J_{n,4}$	4.00	4.57	4.93	3.53	3.87	4.00	1.60	2.77	4.00
$J_{n,5}$	3.90	4.30	4.70	2.80	3.53	3.70	1.67	3.00	4.00
$J_{n,\tilde{M}}^{ms}$	4.50	4.57	5.27	3.87	5.10	5.73	2.00	2.70	4.07
J_{n,M^*}^{an}	4.60	4.40	4.73	3.70	4.47	6.63	1.93	2.47	3.43

Notes:
1. the entries are rejection frequencies in percentage; the nominal size is 5%;
2. for $T_{n,\tilde{M}}^{ms}$ and $J_{n,\tilde{M}}^{ms}$, we set $q = 2.4$, $\underline{m} = 3$, and $\overline{M} = 5, 8, 11$ for $n = 100, 200$ and 300, respectively;
3. for J_{n,M^*}^{an}, $\overline{M} = 5, 8, 11$ for $n = 100, 200$ and 300, respectively.

Table 2
Power simulations: testing the martingale difference hypothesis.

Test	NLMA			BIL			TAR-1			Exp-AR		
	$n = 100$	200	300	100	200	300	100	200	300	100	200	300
CvM_n	4.03	6.40	10.77	15.33	30.30	46.30	73.70	94.90	99.23	23.37	38.53	55.73
KS_n	2.87	4.80	8.07	32.07	55.33	73.37	0.07	0.00	0.00	6.77	7.43	6.90
$T_{n,1}$	1.30	1.33	1.77	14.37	29.17	45.00	69.43	92.27	98.13	20.63	30.70	40.70
$T_{n,2}$	6.37	15.53	25.07	12.23	23.63	37.70	73.70	97.53	99.70	44.50	80.67	93.57
$T_{n,3}$	7.93	20.83	35.73	18.37	43.53	64.77	68.80	96.80	99.87	49.00	84.97	96.63
$T_{n,4}$	8.67	24.40	42.80	15.50	39.00	58.93	64.07	96.13	99.73	41.93	80.50	95.00
$T_{n,5}$	8.03	25.37	45.17	14.50	41.17	64.40	62.37	96.90	99.87	38.30	78.57	94.47
$T_{n,\tilde{M}}^{ms}$	8.27	22.07	38.00	18.53	43.93	64.97	69.03	96.90	99.90	49.03	84.97	96.63
$J_{n,1}$	2.47	2.57	3.47	19.37	39.13	58.13	74.53	94.83	99.03	16.27	24.47	33.53
$J_{n,2}$	13.17	23.80	35.77	17.60	34.57	52.40	64.03	91.77	98.33	60.67	91.67	98.57
$J_{n,3}$	12.43	28.23	47.00	26.23	60.47	82.57	77.73	98.63	99.93	53.60	88.03	97.27
$J_{n,4}$	12.17	31.20	48.90	25.10	58.83	82.93	84.83	99.83	100.00	47.33	84.53	96.30
$J_{n,5}$	11.77	34.73	57.17	20.93	54.93	79.37	80.77	99.73	100.00	42.40	80.03	94.67
$J_{n,\tilde{M}}^{ms}$	13.20	29.50	48.23	26.53	60.80	82.70	79.13	99.03	99.97	53.70	88.10	97.27
J_{n,M^*}^{an}	9.27	26.07	46.37	22.03	51.33	74.93	80.23	99.50	100.00	49.57	84.77	96.33

Notes:
1. the entries are rejection frequencies in percentage; the nominal size is 5%;
2. the parameters for the proposed tests are the same as those in Table 1.

Table 3
Power simulations: testing model linearity.

Test	NLAR			STAR			TAR-2			SGN		
	$n = 100$	200	300	100	200	300	100	200	300	100	200	300
CvM_n	34.60	61.33	79.77	28.80	61.67	81.63	10.07	15.63	24.83	19.00	25.13	34.40
KS_n	34.37	62.83	81.70	27.50	60.37	81.73	9.90	18.87	32.23	25.23	53.43	81.50
$J_{n,1}$	36.63	64.87	83.13	34.50	67.83	86.03	3.03	2.97	3.53	4.00	5.13	6.30
$J_{n,2}$	21.10	47.87	70.60	16.57	47.63	72.00	17.63	65.60	91.90	72.23	99.30	100.00
$J_{n,3}$	20.80	52.17	77.23	26.20	73.77	93.43	26.13	71.30	92.83	73.03	99.03	100.00
$J_{n,4}$	17.77	46.73	73.03	22.93	69.97	92.47	86.53	99.70	100.00	95.53	100.00	100.00
$J_{n,5}$	14.03	42.03	68.47	17.23	62.93	89.60	82.23	99.57	100.00	93.77	100.00	100.00
$J_{n,\tilde{M}}^{ms}$	21.10	52.27	77.30	26.43	73.87	93.50	63.87	96.67	99.70	83.23	99.63	100.00
J_{n,M^*}^{an}	26.93	54.37	76.53	27.80	68.83	89.57	73.07	99.00	100.00	88.77	99.93	100.00

Note:
1. the entries are rejection frequencies in percentage; the nominal size is 5%;
2. the parameters for the proposed tests are the same as those in Table 1.

5. Concluding remarks

In this paper, we propose a more operational approach to constructing Neyman's smooth tests without knowing the covariance kernel of the limiting process. This approach greatly expands the scope of smooth tests because it enables researchers to compute a smooth test even when the limiting process is non-standard and the conventional smooth test is not available. It is also found from our simulations that, together with one of the data-driven methods, the proposed test has nice finite-sample performance. Smooth

tests may be further extended. Note that, smooth tests enjoy power advantage in certain directions by sacrificing test consistency. It is therefore important to construct an omnibus test that is consistent and also carries the spirit of smooth tests; see, e.g., Fan (1996). This topic is currently being investigated.

Acknowledgments

Appendix

For the general empirical process:

$$X_n(\tau) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \phi(\eta_t, \tau) + o_\mathbb{P}(1), \quad \tau \in [a, b],$$

where $\mathbb{E}[\phi(\eta_t, \cdot)|\mathcal{F}_{t-1}] = 0$, and $\mathcal{F}_t$ is the sigma-algebra generated by $\{\eta_i : 1 \leq i \leq t\}$. The following conditions ensure $X_n \Rightarrow X$ on $[a, b]$, where X is a quadratic mean continuous, Gaussian process with zero mean and the covariance kernel $K_X(\cdot, \cdot)$. See Stute (1997) and Escanciano and Mayoral (2010) for similar conditions.

(A1) $\{\eta_t\}$ is a strictly stationary and ergodic sequence of random vectors defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}[\eta_{t,j}]^{2+\delta} < \infty$ for some δ and for all j.

(A2) Under $\mathbb{E}[\phi(\eta_t, \cdot)|\mathcal{F}_{t-1}] = 0$, $n^{-1} \sum_{t=1}^{n} \phi(\eta_t, s)\phi(\eta_t, \tau) - K_X(s, \tau) = o_\mathbb{P}(1)$ uniformly in $(s, \tau) \in [a, b] \times [a, b]$.

(A3) For every $\epsilon = 2^{-\nu} \in (0, 1)$, there exists a finite partition $P_\nu = \{A_k; 1 \leq k \leq N_\epsilon\}$ of every compact subset of $[a, b]$, such that

$$\int_0^1 \sqrt{\log(N_\epsilon)} d\epsilon < \infty, \qquad \sup_{\nu \in N} \frac{V_n(P_\nu)}{2^{-2\nu}} = O_\mathbb{P}(1),$$

where

$$V_n(P_\nu) = \max_{1 \leq k \leq N_\epsilon} \frac{1}{n} \sum_{t=1}^{n} \sup_{s,\tau \in A_k} \\ \times \mathbb{E}\left[|\phi(\eta_t, s) - \phi(\eta_t, \tau)|^2 |\mathcal{F}_{t-1}\right].$$

Proof of Lemma 3.1. Given $X_n(\cdot) \Rightarrow X(\cdot)$ on $[a, b]$, the continuous mapping theorem ensures that, as n tends to infinity,

$$\zeta_{m,n} = \int_a^b X_n(\tau) e_m(\tau) d\tau$$
$$\xrightarrow{d} \int_a^b X(\tau) e_m(\tau) d\tau = \zeta_m, \quad m = 1, \ldots, M.$$

As X is a Gaussian process with mean zero, ζ_m is a normal random variable with mean zero. $\square$

Proof of Lemma 3.2. Given that $\widehat{\text{var}(\zeta_{n,M})}$ is a consistent estimator of $\text{var}(\zeta_M)$ the corresponding eigenpairs, $\widehat{U}_{n,M}$ and $\widehat{\Lambda}_{n,M}$, are consistent for U_M and Λ_M, respectively. Consequently, the normalized principal components are:

$$\widehat{\zeta}_{n,M}^* = [\widehat{\zeta}_{1,n,M}^* \widehat{\zeta}_{2,n,M}^* \cdots \widehat{\zeta}_{M,n,M}^*]' = \widehat{\Lambda}_{n,M}^{-1/2} \widehat{U}_{n,M}' \zeta_{n,M}.$$

Since $\widehat{\Lambda}_{n,M} \xrightarrow{P} \Lambda_M$ and $\widehat{U}_{n,M} \xrightarrow{P} U_M$, $\widehat{\zeta}_{n,M}^*$ is such that

$$\widehat{\zeta}_{n,M}^* = \Lambda_M^{-1/2} U_M' \zeta_{n,M} + o_\mathbb{P}(1) \xrightarrow{d} \Lambda_M^{-1/2} U_M' \zeta_M,$$

by Lemma 3.1. The limit is nothing but ζ_M^* which contains M uncorrelated random variables with mean zero and variance one. When X is Gaussian, we immediately have $\zeta_M^* \sim \mathcal{N}(0, I_M)$. $\square$

Proof of Theorem 3.3. The result is immediate from Lemma 3.2. $\square$

Proof of Theorem 3.4. Let $\widehat{M}$ is a number between the bounds $\underline{m}$ and $\overline{M}$ such that

$$\sum_{m=1}^{\widehat{M}} (\widehat{\zeta}_{m,n,M}^*)^2 = \max\left(\sum_{m=1}^{M} (\widehat{\zeta}_{m,n,M}^*)^2 - 2M\ln(n), \underline{m} \leq M \leq \overline{M}\right),$$

then by the proof of Theorem 1 of Escanciano and Mayoral (2010),

$$\lim_{n \to \infty} \mathbb{P}(\widetilde{M} = \widehat{M}) = 1 \quad \text{and} \quad \lim_{n \to \infty} \mathbb{P}(\widehat{M} = \underline{m}) = 1.$$

We immediately have $J_{n,\widetilde{M}}^{ms} \xrightarrow{d} \chi^2(\underline{m})$ by Theorem 3.3. $\square$

Proof of Theorem 3.5. From Theorem 1 in Darling and Erdös (1956), we have

$$\mathbb{P}\left(J_{n,M^*} < \sqrt{2\ln\ln(\overline{M})} + \frac{\ln\ln\ln(\overline{M})}{2\sqrt{2\ln\ln(\overline{M})}} + \frac{t}{\sqrt{2\ln\ln(\overline{M})}}\right)$$
$$= \exp\left(-(\sqrt{4\pi})^{-1} \exp(-\nu)\right),$$

as $n, \overline{M} \to \infty$. Letting $\varepsilon = \nu + 0.5\ln(4\pi)$, the desired result is immediate. $\square$

References

Bierens, H.J., 1982. Consistent model specification tests. Journal of Econometrics 20, 105–134.

Bierens, H.J., 1990. A consistent conditional moment test of functional form. Econometrica 58, 1443–1458.

Bierens, H.J., Ploberger, W., 1997. Asymptotic theory of integrated conditional mean test. Econometrica 65, 1127–1151.

Carrasco, M., Flofrens, J.P., Renault, E., 2007. Linear inverse problems in structural econometrics: estimation based on spectral decomposition and regularization. In: Heckman, J.J., Learner, E.E. (Eds.), Handbook of Econometrics, Vol. 6B. North-Holland, Amsterdam, pp. 5633–5751.

Darling, D.A., Erdös, P., 1956. A limit theorem for the maximum of normalized sums of independent random variables. Duke Mathematical Journal 23, 143–155.

Delgado, M.A., Hidalgo, J., Velasco, C., 2005. Distribution free goodness-of-fit tests for linear models. Annals of Statistics 33, 2568–2609.

Delgado, M.A., Stute, W., 2008. Distribution-free specification tests of conditional models. Journal of Econometrics 143, 37–55.

Deo, R.S., 2000. Spectral tests of the martingale hypothesis under conditional heteroscedasticity. Journal of Econometrics 99, 291–315.

Domínguez, M.A., Lobato, I.N., 2003. A consistent test for the martingale difference hypothesis. Econometric Reviews 22, 351–377.

Durlauf, S., 1991. Spectral based testing of the martingale hypothesis. Journal of Econometrics 50, 355–376.

Escanciano, J.C., 2009. On the lack of power of Omnibus specification tests. Econometric Theory 25, 162–194.

Escanciano, J.C., Lobato, I.N., 2009. An automatic data-driven portmanteau test for serial correlation. Journal of Econometrics 151, 140–149.

Escanciano, J.C., Mayoral, S., 2010. Data-driven smooth tests for the martingale difference hypothesis. Computational Statistics & Data Analysis 54, 1983–1998.

Eubank, R.L., LaRiccia, V.M., 1992. Asymptotic comparison of Cramér–von Mises and nonparametric function estimation techniques for testing goodness-of-fit. Annals of Statistics 20, 2071–2086.

Fan, J., 1996. Test of significance based on Wavelet thresholding and Neyman's truncation. Journal of the American Statistical Association 91, 674–688.

Fan, Y., 1998. Goodness of fit tests based on kernel density estimators with fixed smoothing parameters. Econometric Theory 14, 604–621.

Fan, J., Huang, L.S., 2001. Goodness-of-fit tests for parametric regression models. Journal of the American Statistical Association 96, 640–652.

Fan, Y., Li, Q., 1996. Consistent model specification tests: omitted variables and semi-parametric functional forms. Econometrica 64, 865–890.

Fan, Y., Li, Q., 2000. Consistent model specification tests: kernel based tests versus Bierens' ICM tests. Econometric Theory 16, 1016–1041.

Fan, J., Lin, S.K., 1998. Test of significance when data are curves. Journal of the American Statistical Association 93, 1007–1021.

Ghosh, A., Bera, A., 2001. Neyman's smooth test and its applications in econometrics. In: Ullah, A., Wan, A., Chaturvedi, A. (Eds.), Handbook of Applied Econometrics and Statistical Inference. Marcel Dekker, New York, pp. 170–230 (Chapter 10).

Inglot, T., Kallenberg, W.C.M., Ledwina, T., 1994. Power approximations to and power comparison of smooth goodness-of-fit tests. Scandinavian Journal of Statistics 21, 131–145.

Inglot, T., Ledwina, T., 2006. Towards data driven selection of a penalty function for data driven Neyman tests. Linear Algebra and its Applications 417, 124–133.

Janssen, A., 2000. Global power functions of goodness of fit tests. Annals of Statistics 28, 239–253.

Karhunen, K., 1946. Zur Spektraltheorie Stochastischer Prozess. In: Annals of Academic Science Fennicae Series, vol. 34. pp. 1–7.

Lee, T., White, H., Granger, C., 1993. Testing for neglected nonlinearity in times series models: a comparison of neural network methods and alternative tests. Journal of Econometrics 56, 269–290.

Loève, M., 1955. Probability Theory. D. Van Nostrand, Princeton, NJ.

Neyman, J., 1937. Smooth test for goodness of fit. Skandinavisk Aktuarietidskrift 20, 149–199.

Stute, W., 1997. Nonparametric model checks for regression. The Annals of Statistics 25, 613–641.

William, C.K.I., Seeger, M., 2000. The effect of the input density distribution on Kernel-based classifiers. In: Proceedings of the 17th International Conference on Machine Learning. Morgan Kaufmann, pp. 1159–1166.

William, C.K.I., Seeger, M., 2001. Using the Nyström method to speed up kernel machines. In: Leen, T.K., Dietterich, T.G., Tresp, V. (Eds.), Annual Advances in Neural Information Processing Systems 13: Proceedings of the 2000 Conference. MIT Press, pp. 682–688.