ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Specification Testing Driven by Orthogonal Series In Nonstationary Time Series Models

Chaohua Dong and Jiti Gao

November 2012

Working Paper 20/12

Specification Testing Driven by Orthogonal Series in Nonstationary Time Series Models*

Chaohua Dong† and Jiti Gao†

Monash University†

November 12, 2012

Abstract

This paper establishes two simple and new specification tests based on the use of an orthogonal series. The paper then establishes an asymptotic theory for each of the proposed tests. The first test is initially proposed for the case where the regression function involved is integrable and the second test is an extended version of the first test for covering a class of non–integrable functions. The finite sample performance of the proposed tests is examined through using several simulated examples. Meanwhile, an application of the second test shows that a second–order polynomial model is more appropriate than a commonly used linear model for modelling the relationship between the United States consumers' consumption expenditure and disposable income over the time period of 1960–2009. Our experience generally shows that the proposed tests are easily implementable and also have stable sizes and good power properties even when the 'distance' between the null hypothesis and a sequence of local alternatives is asymptotically negligible.

JEL classification: C12; C14

Keywords: Consumption–income model; Integrated time series; Local–time process; Orthogonal series estimation; Parametric specification

*The authors first acknowledge constructive comments and suggestions from Professor Peter Phillips and Professor Peter Robinson during their visits at Monash University in September 2012. Thanks from the authors go to participants at a seminar at Monash University in October 2012 for their comments and suggestions. Thanks also go to the Australian Research Council Discovery Grants Program for its support under Grant number: DP1096374. Corresponding author: Chaohua Dong, Department of Econometrics and Business Statistics, Monash University, Caulfield East, Victoria 3145, Australia. Email: chaohua.dong@monash.edu

1 Introduction

Econometric model estimation and specification for nonlinear and nonstationary time series data is an increasingly active area in recent years. To avoid possible model mis–specification issues, nonparametric estimation is normally used as the first step to suggest a close parametric approximation when there is no prior information about which particular model should be used for a given set of data. Then, a rigorous specification procedure should be used to test whether the suggested parametric model may be accepted statistically. Moreover, there is another reason to explain why it is necessary to consider such model specification problems in the integrated time series case. As shown in existing literature (see, e.g., Wang and Phillips, 2009a,b, 2011; Karlsen and Tjøstheim, 2001; Karlsen et al., 2007), nonparametric kernel estimation of integrated time series results in a slow rate of convergence. As a consequence, there are some limited applications of such nonparametric kernel estimated models in practice. The same reason is applicable to other nonparametric estimation methods, such as nonparametric series estimation as discussed in Dong and Gao (2011). By contrast, parametric estimation can provide a fast rate of convergence as shown in Park and Phillips (2001), and thus parametric models may be practically more applicable than their nonparametric counterparts. We thus suggest using this type of estimation and specification procedures particularly in some empirical modelling problems, as we will discuss in an empirical example in Section 6 below, by which we are partially motivated to propose a simple model specification test for checking whether a commonly used linear model is appropriate for modeling the relationship between the United States consumers' consumption expenditure and disposable income over the time period of 1960–2009.

With respect to the literature of model specification, various nonparametric tests have already been proposed and studied extensively for the stationary time series case (see, for example, Chapter 3 of Gao, 2007). In the nonstationary time series case, by contrast, there are only a few nonparametric tests available for parametrically specifying univariate integrated time series models. The first set of these tests are constructed by employing the nonparametric kernel method (see, e.g., Gao et al., 2009a,b; Wang and Phillips, 2012). The second set of such tests are based on the nonparametric series method, such as, Hong and Phillips (2010), Kasparis (2008), and Kasparis (2010). The proposed test of this paper has some similar spirits to those considered in the second set. By comparison, moreover, the proposed test takes into account of the full characteristics of both the orthogonality of the series involved and the nonstationarity of the time series under consideration. Meanwhile, the proposed test has a closed–form expression. As a result, a simple asymptotic distribution is being established for the proposed test and its implementation becomes feasible and computationally less–expensive. This paper, along with the majority of the literature of nonparametric estimation and specification testing, only discusses the univariate integrated time series. The main reason is because of the null–recurrent structure of integrated time series and possible "curse of dimensionality" when using nonparametric estimation and specification testing. As a consequence, nonparametric estimation and specification testing methods are not applicable to deal with the case where the nonstationary time series involved is multivariate. One only exception is a recent paper by Gao et al. (2012), in which a vector of univariate nonstationary and multivariate stationary time series regressors are accommodated in a nonparametric kernel–based specification test.

Consider an econometric model of the form

$$y_t = m(x_t) + e_t, \qquad t = 1, \cdots, n, \tag{1.1}$$

where n is sample size, $m(\cdot)$ is an unknown function, x_t is a univariate nonstationary time series and e_t is an error process.

Before we propose our test, we introduce our hypotheses of the form

$$H_0: \quad P(m(x_t) = g(x_t; \theta_0)) = 1 \quad \text{versus} \quad H_1: \quad P(m(x_t) = g(x_t; \theta_0) + \triangle_n(x_t)) = 1 \tag{1.2}$$

for all $t = 1, \cdots, n$, where $g(\cdot, \theta)$ is a real function on $\mathbb{R}$ with $\theta_0 \in \Theta$ where $\Theta \subset \mathbb{R}^d$ is a parameter space for $d \geq 1$, and $\{\Delta_n(x)\}$ is a sequence of unknown functions satisfying $\lim_{n\to\infty} \triangle_n(x) = 0$ for every $x \in \mathbb{R}$.

The main contribution of this paper is the proposal of two consistent tests driven by an orthonormal basis to test the null hypothesis, H_0, against a sequence of local alternatives under H_1 for the case where $m(x) \in L^2(\mathbb{R})$ and the case where $m(x) \in L^2(\mathbb{R}, \exp(-x^2))$, respectively. One advantage of the proposed tests is that they are computationally simple and their formulations are explicit and concise, since, unlike such tests available in the literature, these tests do not involve any random denominator. The initial test proposed for the case where $m(\cdot)$ is integrable basically measures the total difference of $y_t - g(x_t, \widehat{\theta})$ at each observation point with an orthonormal basis function in $L^2(\mathbb{R})$ being a weight, where $\widehat{\theta}$ is a consistent estimator of θ_0 under H_0. It is then extendable to the case where $m(\cdot)$ is non-integrable. The main idea is to employ a simple transformation of the form $M(x) = m(x)\exp(-x^2/2)$ such that $M(x)$ is integrable in $\mathbb{R}$ even though $m(x)$ itself is not integrable in $\mathbb{R}$. As a consequence, a large class of functions, such as high–order polynomials, can be allowed for $m(x)$.

Our experience shows that while each of the tests is of a simple quadratic form, it is not necessarily easy to establish and then prove an asymptotic distribution. This is mainly because existing central limit theorems available for standardised versions of quadratic forms (see, e.g., Theorem A.1 of Gao, 2007) may only be applicable to the stationary regressor case (see, for example, Gao et al., 2002). Another advantage of the proposed tests is that there is no need to centralise and then standardise our quadratic form, as our experience once again shows that a standardised version of such a quadratic form in this kind of integrated situation is not a good and powerful test. A recent paper by Gao et al. (2012) reveals similar findings in the nonparametric kernel case. By comparison with this natural competitor proposed in Gao et al. (2012), the proposed tests do not need to involve any other arbitrary weight function, because of the orthogonality and integrability of the series involved. In equations (3.6)–(3.8) in Section 3 below, we discuss this key feature and the relationship between the stationary case and the integrated time series case considered in this paper.

The organisation in the rest of the paper is as follows. Section 2 gives some preliminaries about Hermite orthogonal polynomial system. Meanwhile, we state assumptions for model (1.1). Our specification test for the case where $m(x)$ is integrable in $\mathbb{R}$ is proposed and then studied in Section 3.1. Section 3.2 discusses the non–integrability case. Section 4 employs a bootstrap simulation procedure, which will be used to investigate the finite–sample performance of the proposed tests in Section 5. Section 6 analyses a set of data for the US consumption expenditure and disposable income and then examines possible empirical models. Section 7 concludes the main parts of this paper with some concluding remarks. Appendices A through F give all the necessary mathematical details for the proofs of the main theorems.

2 Preliminaries and assumptions

Let $\{H_i(x)\}$ be the Hermite orthogonal polynomial system with respect to the density $\exp(-x^2)$. As a classical orthogonal system, it is known that $\{H_i(x)\}$ is a complete orthogonal system in Hilbert space $L^2(\mathbb{R}, \exp(-x^2))$, in which the conventional inner product is used, i.e. $(f, g) = \int f(x)g(x)\exp(-x^2)dx$. In addition, the orthogonality of the system is expressed as the equation $(H_i(x), H_j(x)) = \sqrt{\pi}2^i i!\delta_{ij}$, where δ_{ij} is the Kronecker delta. Put $\varphi(x) = \exp(-\frac{1}{2}x^2)$. Define

$$h_i(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{2^i i!}} H_i(x) \quad \text{and} \quad \mathscr{H}_i(x) = h_i(x)\varphi(x), \quad i \geq 0. \tag{2.1}$$

Note that $\mathscr{H}_i(x)$ are the so-called Hermite functions in the related literature. It follows that $\{\mathscr{H}_i(x)\}$ is complete orthonormal basis in $L^2(\mathbb{R})$ in which the inner product is defined as $(f, g) = \int f(x)g(x)dx$. Therefore, any continuous function $f(x) \in L^2(\mathbb{R})$ has orthogonal expansion $f(x) = \sum_{i=0}^{\infty} b_i \mathscr{H}_i(x)$ with $b_i = \int f(x)\mathscr{H}_i(x)dx$ and the convergence of the series is guaranteed by the continuity of $f(x)$.

Let us recall some facts about the aforementioned orthogonal systems. Firstly, we have $H_i(x) = (-1)^i e^{x^2} D^i e^{-x^2}$, $i = 0, 1, 2, \cdots$, where D stands for differentiation operation. Also, $H_i'(x) = 2iH_{i-1}(x)$. According to Nevai (1986, p.86), we have $\max_{x\in\mathbb{R}} \mathscr{H}_i^2(x) \sim i^{-1/6}$ which implies the uniform boundedness of the system $\{\mathscr{H}_i(x)\}$. Moreover, the Christoffel-Darboux formula for Hermite polynomials spells out

$$\sum_{i=0}^{k-1} h_i^2(x) = \sqrt{k/2}(h_k'(x)h_{k-1}(x) - h_k(x)h_{k-1}'(x)). \tag{2.2}$$

The following assumptions are preparation for the null and the alternative hypotheses of the test in the section below.

Assumption A

(a) Let $x_t = u_t + \cdots + u_1 + x_0$ with $\{u_s\}$ being a linear process defined by $u_s = \sum_{j=0}^{\infty} \psi_j \varepsilon_{s-j}$ satisfying that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and $\psi = \sum_{j=0}^{\infty} \psi_j \neq 0$, in which $\{\varepsilon_j, -\infty < j < \infty\}$ is a sequence of i.i.d. random variables with $E\varepsilon_1 = 0$ and $E\varepsilon_1^2 = 1$. Meanwhile, the characteristic function $\psi(t)$ of ε_1 is integrable, $\int |\psi(t)|dt < \infty$.

(b) There is a sequence of σ–fields $\mathcal{F}_{n,t} = \sigma(e_1, \cdots, e_t, x_1, \cdots, x_n)$ generated by $\{(e_i, x_j) : 1 \le i \le t, 1 \le j \le n\}$ such that for all $1 \le t \le n$, $E(e_t|\mathcal{F}_{n,t-1}) = 0$ almost surely (a.s.), $E(e_t^2|\mathcal{F}_{n,t-1}) = \sigma_e^2$ a.s. and $\mu_4 := \sup_{1\le t\le n} E(e_t^4|\mathcal{F}_{n,t-1}) < \infty$ a.s..

We simply set $x_0 = O_P(1)$ throughout the paper, although $x_0 = o_{a.s.}(\sqrt{n})$ is sufficient for the following development. Other choices which does not affect the result materially are possible (See Phillips and Magdalinos, 2009). Note that since comprising many special cases, Assumption A is engaged in considerable studies in recent years. See, Dong and Gao (2011), Wang and Phillips (2011) and Shi and Phillips (2012). We have the following remarks on the information sequence $\mathcal{F}_{n,t}$.

Remark 2.1. The design of the information sequence $\mathcal{F}_{n,t}$ in the literature aims at allowing some mixture structure between x_s and e_t, where $1 \le s, t \le n$. Currently, there are several versions of $\mathcal{F}_{n,t}$ in the literature which enable us to do so. In Gao et al. (2009b) the authors required independence between two classes of x_t and e_t; while in Wang and Phillips (2012) the authors imposed independence for $\mathcal{F}_{n,t}$ and ε_j whenever $j \ge t + 1$ in Assumption 2 and

$\mathcal{F}_{n,t} = \sigma(e_1, \cdots, e_t, x_1, \cdots, x_n)$ in Assumption 2*. For the sake of convenience, in this paper let $\mathcal{F}_{n,t} = \sigma(e_1, \cdots, e_t, x_1, \cdots, x_n)$.

Our simulation results in Section 5 show that the test L_n and Π_n proposed and then studied in Section 3 below are applicable to the case where there is some type of endogeneity between e_t and ε_s in the sense that $E[e_t \varepsilon_t] \neq 0$ but $E\left[\frac{1}{\sqrt{t}} x_t e_t\right] \to 0$ as $t \to \infty$. Meanwhile, our experience with the proofs of Theorems 3.1–3.4 listed in Section 3 below also suggests that the conclusions of Theorems 3.1–3.4 may remain true when Assumption A(b)(c) is replaced by

Assumption A* (b). There is a sequence of σ–fields $\mathcal{G}_{n,t}$ such that, for $i = 1, 2$ and 4, $\nu_{i,t} = E\left[e_t^i | \mathcal{G}_{n,t-1}\right] \to \nu_i$ a.s. when $t \to \infty$, where $\nu_1 = 0$, $\nu_2 > 0$ and $0 < \nu_4 < \infty$.

However, our experience suggests that it would need to involve many more technicalities to complete the proofs of Theorems 3.1–3.4 for the case where Assumption A(b)(c) is replaced by Assumption A*(b). We therefore decide to focus on the current structure and wish to leave any of such relaxations for future research.

Write $d_n^2 = E x_n^2$. According to Wang et al. (2003), under Assumption A(a) we have

$$d_n^2 = \psi^2 n (1 + o(1)). \tag{2.3}$$

In addition, it follows from Wang and Phillips (2011) that, with $x_{t,n} := \frac{1}{d_n} x_t$ by definition,

$$x_{[nr],n} \to_D B(r), \quad \text{on } D[0,1], \tag{2.4}$$

where $B(r)$ is a Brownian motion on [0,1]. Note that $\frac{1}{d_s} x_s$ has a density $g_s(x)$ that is uniformly bounded in both s and x, and as $s \to \infty$, $\max_x |g_s(x) - \phi(x)| = o(1)$, where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. It is also true that for $\frac{1}{d_{ts}}(x_t - x_s)$ with $d_{ts} = \psi\sqrt{t-s}$ for $t > s$, $\frac{1}{d_{ts}}(x_t - x_s)$ have density functions $g_{ts}(x)$ that are uniformly bounded over all t, s and $x \in \mathbb{R}$, and $g_{ts}(x)$ converges uniformly to $\phi(x)$ as $|t - s| \to \infty$. Such results follow from the proof of Corollary 2.2 in Wang and Phillips (2009a).

In the rest of the paper, denote by k the truncation parameter for function orthogonal expansion and we have the following assumption for k.

Assumption B

(a) Under H_0, there exists a consistent estimator $\widehat{\theta}$ of θ_0 such that $\|\widehat{\theta} - \theta_0\| = o_P(\zeta_n)$ for some sequence ζ_n satisfying $\zeta_n \to 0$ as $n \to \infty$, where $\|\cdot\|$ is the conventional Euclidean norm.

(b) $k = [n^\kappa]$ for some $\kappa : 0 < \kappa < \frac{1}{2}$ satisfying $\sqrt{n/k}\,\zeta_n^2 \to 0$ as $n \to \infty$.

Remark 2.2. Both (a) and (b) are achievable since there already exist such results in Park and Phillips (1999, 2001). Indeed, if $g(x, \theta_0)$ is an I–regular regression function (see Park and Phillips (2001) for more detail about I–regular functions), we then have $\sqrt[4]{n}(\widehat{\theta} - \theta) = O_P(1)$. In this case, we may choose $\zeta_n = n^{-1/4+\epsilon}$ with small $\epsilon > 0$ and $\kappa \in (4\epsilon, 1)$, which implies that κ has a variety of choices.

Assumption C

(a) Suppose that $g(x;\theta)$ is twicely differentiable with respect to θ and $g(x;\theta) \in L^2(\mathbb{R})$ for every fixed $\theta \in \Theta$. Denote that $l_1(x,\theta) := \frac{\partial}{\partial\theta} g(x;\theta)$ and $l_2(x,\theta) := \frac{\partial^2}{\partial\theta\partial\theta'} g(x;\theta)$. Suppose further that $\|l_1(x,\cdot)\|, \|l_2(x,\cdot)\| \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and that there exists a positive function $l(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\|l_2(x,\cdot)\| \le l(x)$.

Remark 2.3. Here, for vector $\|\cdot\|$ is still Euclidean norm, while for matrix $\|\cdot\|$ stands for entrywise norm, i.e. $\|A\| = (\sum_{i=1}^n \sum_{j=1}^m a_{ij}^2)^{1/2}$ if $A = (a_{ij})_{nm}$.

The conditions imposed here are a set of standard conditions used in the literature (see, for example, Assumption 2.4 of Gao et al. (2011) and Assumption 4 of Wang and Phillips (2012)). It is also easily seen that the following type of functions for $g(x;\theta)$ satisfy these conditions: (1) $g(x;\theta) = \theta g_1(x)$, where $g_1(x)$ is integrable and $\theta \in \mathbb{R}$ (so called leaner-in-parameter regression function in Park and Phillips (2001)); (2) $g(x;\theta) = \frac{1}{1+\theta x^2}$ where $\theta > c > 0$; (3) $g(x;\alpha,\beta) = \alpha \exp(-\beta x^2)$ where $\theta = (\alpha,\beta) \in \Theta \subset \mathbb{R} \times \mathbb{R}^+$.

Assumption D Suppose that $\triangle_n(x) = \delta_n \triangle(x)$ satisfying the following conditions.

(a) $\triangle(x)$ is a continuous function such that $\int \triangle^2(x) e^{-x^2/2} dx < \infty$.

(b) $\delta_n \to 0$ as $n \to \infty$; $\delta_n^2 \sqrt{n}/k \to \infty$ as $n \to \infty$.

Remark 2.4. Note that Condition (a) is so weak that $\triangle(x)$ is allowed to belong to a large class of real functions of x. For instance, $\triangle(x)$ may be an integrable function or a bounded function; $\triangle(x)$ may be not integrable function, such as, any power function and polynomials of x with fixed order. Exclusion of zero function for $\triangle(x)$ is trivial. Note also that Condition (a) implies the integrability of $|\triangle(x)|e^{-x^2/2}$ and $[\triangle(x)\mathscr{H}_i(x)]^2$ for every $i \ge 0$, which is important for the proofs of Theorem 3.2 and 3.4 below.

Condition (b) guarantees the discrepancy for the test proposed in the following section, although the null and the alternative hypotheses are getting closer and closer when the sample size is getting larger and larger. Observe that δ_n has a variety of possibilities, regardless of whether $\triangle(x)$ is integrable on $\mathbb{R}$ or not, as long as $\kappa < \frac{1}{2}$. Whence, the proposed statistic can even detect bounded functions and integrable functions in the alternative from the null in spite that their boundary blurs as the sample size increases. As can be seen from the proofs in Appendices A and B, δ_n can be allowed to satisfy $\liminf_{n\to\infty} |\delta_n| > 0$. In other words, Theorems 3.2 and 3.4 are remaining valid for other cases where $\delta_n = O(1)$, let alone where δ_n diverges.

3 Specification testing

As mentioned in last section, $\{\mathscr{H}_i(x)\}_{i\geq 0}$ is a complete orthonormal basis in $L^2(\mathbb{R})$. Thus, any continuous function $f(x) \in L^2(\mathbb{R})$ has orthogonal expansion $f(x) = \sum_{i=0}^{\infty} b_i \mathscr{H}_i(x)$ with $b_i = \int f(x)\mathscr{H}_i(x)dx$. Clearly, such function $f(x)$ can be approximated by the truncation (or projection) series $f_k(x) = \sum_{i=0}^{k-1} b_i \mathscr{H}_i(x)$ and as $k \to \infty$, $\|f(x) - f_k(x)\| \to 0$, where $\|\cdot\|$ stands for the norm in $L^2(\mathbb{R})$. In addition, the convergence rate depends on the smoothness order of $f(x)$.

In the rest of this section, we consider the case where $m(x)$ is integrable in $\mathbb{R}$ and then the case where $M(x) = m(x)\exp\left(-\frac{x^2}{2}\right)$ is square integrable in $\mathbb{R}$.

3.1 Integrable cases

Suppose that $m(\cdot) \in L^2(\mathbb{R})$ in this subsection. Let $m(x) = \sum_{i=0}^{\infty} b_i \mathscr{H}_i(x)$ and the truncated series of $m(x)$ be defined as $m_k(x) = \sum_{i=0}^{k-1} b_i \mathscr{H}_i(x)$ with the truncation parameters $k = [n^\kappa]$ where κ satisfies Assumption B. Denote $Z(x) = (\mathscr{H}_0(x), \cdots, \mathscr{H}_{k-1}(x))'$ and $\beta = (b_0, \cdots, b_{k-1})'$. Then, $m_k(x) = Z(x)'\beta$ and model (1.1) under H_0 can be written as

$$y_t = Z(x_t)'\beta + \gamma_k(x_t) + e_t, \qquad t = 1, \cdots, n, \tag{3.1}$$

where $\gamma_k(x_t) = \sum_{i=k}^{\infty} b_i \mathscr{H}_i(x_t)$ is the residue in the expansion of m function after truncation.

Denote $Y = (y_1, \cdots, y_n)'$, $Z = (Z(x_1), \cdots, Z(x_n))'$ an $n \times k$ matrix, $e = (e_1, \cdots, e_n)'$ and $\gamma = (\gamma_k(x_1), \cdots, \gamma_k(x_n))'$. The equations in (3.1) are formulated into the following matrix form:

$$Y = Z\beta + \gamma + e. \tag{3.2}$$

We then have the estimator of β by ordinary least squares (OLS)

$$\widehat{\beta} = (Z'Z)^{-1}Z'Y. \tag{3.3}$$

Thereby, $\widehat{m}(x) = Z(x)'\widehat{\beta}$ is an estimator for $m(x)$ at $\forall\, x \in \mathbb{R}$.

For the purpose of constructing out test, we may avoid involving the inverse of $Z'Z$ in $\widehat{\beta}$ by defining $\widetilde{\beta} = Z'Z\widehat{\beta}$. Correspondingly, the estimator $\widehat{m}(x) = Z(x)'\widehat{\beta}$ can be replaced by

$\widetilde{m}(x) = Z(x)'\widetilde{\beta}$ for testing purposes and, invoking the expression of $\widehat{\beta}$,

$$\widetilde{m}(x) = \sum_{t=1}^{n} [Z(x_t)'Z(x)] y_t.$$

Consequently, we also may have a similar version $\widetilde{g}(x;\theta)$ for $g(x;\theta)$:

$$\widetilde{g}(x;\theta) = \sum_{t=1}^{n} [Z(x_t)'Z(x)] g(x_t;\theta).$$

Let $\widehat{\theta}$ be a consistent estimator of θ_0 under H_0. Instead of comparing $\widehat{m}(x)$ with $g(x,\widehat{\theta})$, we measure the distance between $\widetilde{m}(x)$ and $\widetilde{g}(x;\widehat{\theta})$ and then propose using a test statistic of the form:

$$L_n = \int_{-\infty}^{\infty} \left(\widetilde{m}(x) - \widetilde{g}(x;\widehat{\theta})\right)^2 dx. \tag{3.4}$$

Note that

$$\begin{aligned}\left[\widetilde{m}(x) - \widetilde{g}(x;\widehat{\theta})\right]^2 &= \left[\sum_{t=1}^{n} [Z(x_t)'Z(x)](y_t - g(x_t;\theta))\right]^2 \\ &= \sum_{t=1}^{n}\sum_{s=1}^{n} (y_t - g(x_t;\widehat{\theta}))Z(x_t)'Z(x)Z(x)'Z(x_s)(y_s - g(x_s;\widehat{\theta})).\end{aligned}$$

By invoking the orthogonality $\int \mathscr{H}_i(x)\mathscr{H}_j(x)dx = \delta_{ij}$, $\int Z(x)Z(x)'dx$ becomes an identity matrix, and L_n reduces to an explicit and concise expression of the form:

$$L_n = \sum_{t=1}^{n}\sum_{s=1}^{n} (y_t - g(x_t;\widehat{\theta}))Z(x_t)'Z(x_s)(y_s - g(x_s;\widehat{\theta})). \tag{3.5}$$

Note that

$$\begin{aligned}L_n &= \sum_{t=1}^{n} (y_t - g(x_t;\widehat{\theta}))^2 Z(x_t)'Z(x_t) + 2\sum_{t=2}^{n}\sum_{s=1}^{t-1} (y_t - g(x_t;\widehat{\theta}))Z(x_t)'Z(x_s)(y_s - g(x_s;\widehat{\theta})) \\ &\equiv L_{an} + L_{bn}.\end{aligned} \tag{3.6}$$

In the case where x_t is stationary, as shown in Gao et al. (2002), a suitably standardised version of L_{bn} of the form

$$\widehat{L}_{bn} = \frac{\sum_{t=2}^{n}\sum_{s=1}^{t-1} \widehat{e}_s Z(x_t)'Z(x_s)\widehat{e}_t}{\sqrt{\sum_{t=1}^{n}\sum_{s=1}^{n} \widehat{e}_s^2 \left(Z(x_t)'Z(x_s)\right)^2 \widehat{e}_t^2}} \tag{3.7}$$

converges in distribution to $U \sim N(0,1)$, where $\widehat{e}_t = y_t - g(x_t;\widehat{\theta})$. Note that the series 'weight function' of the form $w_k(x_s, x_t) = Z(x_t)'Z(x_s) = \sum_{j=0}^{k-1} \mathscr{H}_j(x_s)\mathscr{H}_j(x_t)$ plays a similar role to $K\left(\frac{x_t - x_s}{h}\right)$ involved in a nonparametric kernel test of the form

$$\widehat{L}_{\text{kernel}} = \frac{\sum_{t=2}^{n}\sum_{s=1}^{t-1} \widehat{e}_s K\left(\frac{x_t - x_s}{h}\right)\widehat{e}_t}{\sqrt{\sum_{t=1}^{n}\sum_{s=1}^{n} \widehat{e}_s^2 K^2\left(\frac{x_t - x_s}{h}\right)\widehat{e}_t^2}} \tag{3.8}$$

which has been discussed in Gao et al. (2009a,b) and Wang and Phillips (2012) for the integrated time series case, in which $K(\cdot)$ is a probability kernel function and h is a bandwidth parameter.

In a recent paper by Gao et al. (2012), the authors show that $\widehat{L}_{\text{kernel}}$ is less powerful than the kernel corresponding version of L_{an}. This is confirmed by our finite–sample comparison in Tables 1 and 5 as well as Tables 2 and 6 in Section 5 below. In the case where x_t is integrated, therefore, this paper proposes using a suitably normalised version of L_{an} and then shows that such a suitably normalised version converges in distribution to a local–time random variable while L_{bn} becomes asymptotically negligible compared with L_{an}. These are some of the key features of the proposed test and its large and small–sample properties as discussed in Sections 3–5 below. Additionally, in comparison with the natural competitor proposed in Gao et al. (2012) for the kernel case, there is no need for the proposed test to involve any weight function. As a consequence, both the size and power properties of L_n become stable and robust. These, along with the large and small–sample properties discussed in Sections 3–5 below, are the key features of L_n.

Notice that under the null, $L_n = \sum_{t=1}^n \sum_{s=1}^n Z(x_t)'Z(x_s)\widehat{e}_t\widehat{e}_s$ where $\widehat{e}_t = g(x_t;\theta_0) - g(x_t;\widehat{\theta}) + e_t$, while $L_n = \sum_{t=1}^n \sum_{s=1}^n Z(x_t)'Z(x_s)(\widehat{e}_t + \triangle_n(x_t))(\widehat{e}_s + \triangle_n(x_s))$ under the hypothesis H_1.

Also, observe that L_n is the squared Euclidean norm of a vector, i.e. , $L_n = \|Z'(Y-\widehat{G})\|^2$ where $\widehat{G} = (g(x_1;\widehat{\theta}), \cdots, g(x_n;\widehat{\theta}))$ and $\|\cdot\|$ stands for Euclidean norm of vectors, which implies that L_n actually measures the squared length of the vectors $Z'(Y-\widehat{G})$.

Let $G = (g(x_1,\theta_0), \cdots, g(x_n,\theta_0))$ and $\triangle = (\triangle(x_1), \cdots, \triangle(x_n))$. We then have

$$Y - \widehat{G} = \begin{cases} (G-\widehat{G}) + e, & \text{under } H_0, \\ (G-\widehat{G}) + \triangle + e, & \text{under } H_1. \end{cases}$$

The key point is that the transformation $Z'(Y-\widehat{G})$ of $(Y-\widehat{G})$ will reveal which hypothesis is true, since $Z'(G-\widehat{G})$ is negligible while $Z'e$ and $Z'(\triangle + e)$ are not, after normalization the first one converges whereas the second diverges. These different behaviors give us the possibility to detect the two hypotheses by investigating the transformation of the deviation vector $Z'(Y-\widehat{G})$. This is the rationale how statistic L_n works.

Theorem 3.1. *Suppose that Assumptions A–C hold. Under Hypothesis H_0, as $n \to \infty$,*

$$\frac{d_n}{nk} L_n \to_D \sigma_e^2 L_B(1,0) \tag{3.9}$$

where $L_B(1,0)$ is the local time process of Brownian motion $B(r)$.

Remark 3.1. It follows from Theorem 3.1 that, at the significance level α, we may reject H_0 if $\frac{d_n}{nk\sigma_e^2} L_n > l_{1-\alpha}$ with $l_{1-\alpha}$ satisfying $P(L_B(1,0) > l_{1-\alpha}) = \alpha$, which means that the proposed test statistic has size α. Note that σ_e^2 can be estimated by $\widehat{\sigma}_e^2 = \frac{1}{n}\sum_{t=1}^n (y_t - g(x_t, \widehat{\theta}))^2$ under H_0 when it is unknown. For the definition and properties of the local–time processes, Revuz and Yor (1999) is a standard reference.

Theorem 3.2. *Under Hypothesis H_1 as well as Assumptions A–D, we have*

$$\frac{d_n}{nk} L_n \to_P \infty \tag{3.10}$$

as $n \to \infty$.

Remark 3.2. Theorem 3.2 shows that the test statistic has nontrivial power against a sequence of local alternatives that may have a distance from the null approaching to zero at a rate slower than $(\sqrt{n}/k)^{1/2}$, as required in Assumption D for δ_n. This is similar to the stationary case (see, for example, Chapter 3 of Gao (2007)). With the results in Theorem 3.1 and 3.2, we can differentiate the models underpinning by H_0 and H_1, respectively. We shall reject H_0 in favour of H_1 if $\frac{d_n}{nk} L_n$ is large enough. Moreover, $\lim_{n\to\infty} P(\frac{d_n}{nk} L_n > C_n) = 1$ provided that C_n is a positive non–stochastic sequence with $C_n = o(\delta_n^2 \sqrt{n}/k)$.

3.2 Beyond integrability

We are about in this subsection to relax the restriction of integrability for regression function $m(x)$ in model (1.1). As a matter of fact, we no longer need the square integrability of $m(x)$ on $\mathbb{R}$ of $m(x)$, as long as $m(x)$ satisfies $\int m^2(x) e^{-x^2} dx < \infty$, that is, $m(x) \in L^2(\mathbb{R}, \exp(-x^2))$. This considerably extends the applicability of this study since $m(x)$ now could belong to a large class of real functions, including all polynomial and trigonometric functions.

Given observations $\{(x_t, y_t), t = 1, 2, \cdots, n\}$ from model (1.1), we are interested to test in model (1.1) that

$$H_{10}: \quad P(m(x_t) = g(x_t; \theta_0)) = 1 \quad \text{versus} \quad H_{11}: \quad P(m(x_t) = g(x_t; \theta_0) + \triangle_n(x_t)) = 1 \tag{3.11}$$

for all $t = 1, \cdots, n$, where for any fixed $\theta \in \Theta \subset \mathbb{R}^d$ $(d \geq 1)$, $g(x, \theta) \in L^2(\mathbb{R}, \exp(-x^2))$, and $\triangle_n(x)$ is the same as in Assumption D.

We propose the following test statistic for (3.11):

$$\Pi_n = \sum_{t=1}^n \sum_{s=1}^n (Y_t - G(x_t; \widehat{\theta})) Z(x_t)' Z(x_s) (Y_s - G(x_s; \widehat{\theta})), \tag{3.12}$$

where $\widehat{\theta}$ is a consistent estimation of θ_0 under the null in (3.11), $Y_t = \varphi(x_t)y_t$, $G(x, \widehat{\theta}) = g(x, \widehat{\theta})\varphi(x)$ with $\varphi(x) := \exp(-\frac{1}{2}x^2)$, and $Z(x)$ is the vector defined in the last subsection.

The motivation of proposing Π_n is as follows. Because the test statistic for the case where $m(x)$ is integrable is already available, it would make sense if we could employ it for the case where $m(\cdot)$ is not integrable. To this end, we first need to transform the dataset of observations from 'non-integrable' to 'integrable'. Specifically, multiplying both side of model (1.1) by $\varphi(x_t)$, denoting that, $M(x_t) = \varphi(x_t)m(x_t)$ and $\varepsilon_t = \varphi(x_t)e_t$, we have

$$Y_t = M(x_t) + \varepsilon_t, \qquad t = 1, \cdots, n. \tag{3.13}$$

Equations in (3.13) are completely equivalent to the equations in (1.1), but in terms of regression function, $M(x) = \varphi(x)m(x)$ is now integrable with respect to x. In other words, we have transformed the dataset from 'non–integrable' (y_t, x_t) to 'integrable' (Y_t, x_t), and this transformation does not affect our specification test at all since the hypotheses in (3.11) can be equivalently written as

$$\mathcal{H}_0: \ P(M(x_t) = G(x_t; \theta_0)) = 1 \ \text{ versus } \ \mathcal{H}_1: \ P(M(x_t) = G(x_t; \theta_1) + \varphi(x_t)\triangle_n(x_t)) = 1 \tag{3.14}$$

for all $t = 1, \cdots, n$.

In view of this transformation, it is natural to propose Π_n following the idea of proposing L_n in last subsection.

Note that $\varphi(x)$ could be substituted by other probability density functions as long as $M(x)$ is integrable and $\varphi(x)$ has support which has nonempty intersection with [-1,1]. If the intersection is empty, we would suffer from the so-called zero energy issue (Wang and Phillips, 2011).

It is clear that $\Pi_n = \sum_{t=1}^n \sum_{s=1}^n \varphi(x_t)\widehat{e}_t Z(x_t)' Z(x_s)\widehat{e}_s \varphi(x_s)$ under H_{10}, while under H_{11}, $\Pi_n = \sum_{t=1}^n \sum_{s=1}^n \varphi(x_t)(\widehat{e}_t + \triangle_n(x_t)) Z(x_t)' Z(x_s)(\widehat{e}_s + \triangle_n(x_s))\varphi(x_s)$, where $\widehat{e}_t = y_t - g(x_t; \widehat{\theta})$.

The following two assumptions are similar to Assumptions B and C but stipulated for the cases where m belongs to $L^2(\mathbb{R}, \exp(-x^2))$, rather than $L^2(\mathbb{R})$.

Assumption B*

(a) Under the null in (3.11), there exists a consistent estimator, $\widehat{\theta}$, of θ_0 such that $\|\widehat{\theta} - \theta_0\| = o_P(\zeta_n)$ for some sequence ζ_n satisfying $\zeta_n \to 0$ as $n \to \infty$.

(b) $k = [n^\kappa]$ for some $\kappa : 0 < \kappa < 1$; and $\sqrt{n/k}\zeta_n^2 \to 0$ as $n \to \infty$.

Remark 3.3. The only difference between Assumptions B and B* is that the estimation of θ_0 is under H_{10} in (3.11). However, it does not require a faster convergence rate. This is because non-integrable $m(x)$ has been transformed as an integrable function $M(x)$, so that the Π_n works in the same environment as the L_n does. Thus, all regression functions studied in the literature easily satisfy this assumption, since for $m(x)$ functions like polynomials, power functions, $\theta \log |x|$, $\theta e^x/(1+e^x)$ and $x(1+\theta x)^{-1}1(x \geq 0)$, the rates of convergence are at least equal to even faster than $1/\sqrt{n}$. A detailed discussion can be found in Theorems 5.2 and 5.3 of Park and Phillips (2001, p. 135).

Assumption C*

(a) $g(x;\theta)$ is twice differentiable with respect to θ and $\int g^2(x,\theta_0)e^{-x^2/k}dx < \infty$ for any positive integer k. Denote that $l_1(x,\theta) := \frac{\partial}{\partial\theta}g(x;\theta)$ and $l_2(x,\theta) := \frac{\partial^2}{\partial\theta\partial\theta'}g(x;\theta)$. Suppose further that $\|l_1(x,\cdot)\|, \|l_2(x,\cdot)\| \in L^2(\mathbb{R}, e^{-x^2/2})$, while $\|l_2(x,\theta)\| \leq l(x)$ uniformly over θ with positive $l(x) \in L^2(\mathbb{R}, e^{-x^2/2})$.

Remark 3.4. This condition is, like Assumption C, a common requirement in the related studies in the literature. Since the regression functions are in the class of $L^2(\mathbb{R}, e^{-x^2})$, we basically stipulate that their derivatives with respect to the parameter are in $L^2(\mathbb{R}, e^{-x^2/2})$, which is a subclass of $L^2(\mathbb{R}, e^{-x^2})$. Clearly, the examples given above satisfy this condition.

Theorem 3.3. *Suppose Assumptions A, B* and C* hold. Under H_{10} of (3.11) we have*

$$\frac{d_n}{nk}\Pi_n \to_D \sigma_e^2 \int f(x)\varphi(x)^2 dx L_B(1,0), \tag{3.15}$$

as $n \to \infty$, where $f(x) = \frac{2}{\pi}\sqrt{1-x^2}$ on [-1,1] and 0 elsewhere, and $L_B(1,0)$ is the local time process of Brownian motion $B(r)$.

Theorem 3.4. *Under Hypothesis H_{11} as well as Assumptions A, B*, C* and D , we have*

$$\frac{d_n}{nk}\Pi_n \to_P \infty \tag{3.16}$$

as $n \to \infty$.

4 Bootstrap theory

In order to investigate the finite–sample performance of the proposed tests, a bootstrap schedule is proposed to find simulated critical values l_α^* $(0 < \alpha < 1)$ for L_n and Π_n to approximate the exact $1-\alpha$ quantiles of the exact sample distributions of L_n and Π_n, respectively. The schedule for L_n is proposed as follows.

Step 1 Generate $e_t^* = \widehat{e}_t\eta_t$ with η_t being iid sequence possessing two-point distribution of $P(\eta_1 = \mp(\sqrt{5}\mp 1)/2) = (\sqrt{5}\pm 1)/2\sqrt{5}$ and $\widehat{e}_t = y_t - m(x_t, \widehat{\theta})$ where $\widehat{\theta}$ is the consistent estimation of θ_0 based on the original sample (x_t, y_t). Then, generate $y_t^* = m(x_t, \widehat{\theta}) + e_t^*$ for $1 \le t \le n$.

Step 2 Use the dateset (x_t, y_t^*) to re-estimate θ_0 to obtain $\widehat{\theta}^*$. Compute L_n^* that is a corresponding version of L_n with $\widehat{\theta}$ and $\{(x_t, y_t), 1 \le t \le n\}$ being replaced by $\widehat{\theta}^*$ and $\{(x_t, y_t^*), 1 \le t \le n\}$ in the expression of L_n.

Step 3 Repeat the above steps M times and thereby produce M versions of L_n^*, signified by $L_{n\ell}^*$ for $\ell = 1, \cdots, M$. Use the M values of $L_{n\ell}^*$ to construct their empirical bootstrap distribution function. The bootstrap distribution of $\frac{d_n}{nk}L_n^*$ given $\mathcal{W}_n = \{(x_t, y_t), 1 \le t \le n\}$ is defined by $P^*(\frac{d_n}{nk}L_n^* < x) = P(\frac{d_n}{nk}L_n^* < x|\mathcal{W}_n)$. Let l_α^* be the quantile such that $P^*(\frac{d_n}{nk}L_n^* \ge l_\alpha^*) = \alpha$ and then l_α^* is the estimator of l_α.

Step 4 Define the size and power functions by

$$\alpha = P\left(\frac{d_n}{nk}L_n \ge l_\alpha^*|H_0\right) \quad \text{and} \quad \beta = P\left(\frac{d_n}{nk}L_n \ge l_\alpha^*|H_1\right).$$

In practice, both α and β may be approximated by using Edgeworth expansions similarly to (3.23) and (3.24) of Gao (2007).

Notice that, in Step 1 the choice of the iid sequence $\{\eta_t\}$ is according to the principle that $E(\eta_t) = 0$, and $E(\eta_t^i) = 1$ for $i = 2, 3$. It is noteworthy that the distribution of η_1 is not essential theoretically. The two-point distributional structure has been used in the literature (Gao et al., 2011).

Bootstrap schedule for statistic Π_n can be conducted similarly, so it is omitted for brevity.

Assumption E

(a) Under H_0, there exists a consistent estimator, $\widehat{\theta}^*$, of $\widehat{\theta}$ such that $\|\widehat{\theta}^* - \widehat{\theta}\| = o_P(\zeta_n)$, where ζ_n is the same as in Assumption B.

(b) Under H_{10}, there exists a consistent estimator, $\widehat{\theta}^*$, of $\widehat{\theta}$ such that $\|\widehat{\theta}^* - \widehat{\theta}\| = o_P(\zeta_n)$, where ζ_n is the same as in Assumption B*.

Remark 4.1. Assumption E is the counterpart of Assumptions B and B* in the bootstrap schedule, so it is reasonable and necessary.

Theorem 4.1. *(1) Given the conditions in Theorem 3.1 and Assumption E(a), we have under H_0,*

$$\lim_{n\to\infty} P\left(\frac{d_n}{nk}L_n > l_\alpha^*\right) = \alpha$$

for all $\alpha \in [0,1]$.

(2) Given the conditions in Theorem 3.2 and Assumption E(a), we have under H_1,

$$\lim_{n\to\infty} P\left(\frac{d_n}{nk}L_n > l_\alpha^*\right) = 1$$

for all $\alpha \in [0,1]$.

(3) Given the conditions in Theorem 3.3 and Assumption E(b), we have under H_{10},

$$\lim_{n\to\infty} P\left(\frac{d_n}{nk}\Pi_n > l_\alpha^*\right) = \alpha$$

for all $\alpha \in [0,1]$.

(4) Given the conditions in Theorem 3.4 and Assumption E(b), we have under H_{11},

$$\lim_{n\to\infty} P\left(\frac{d_n}{nk}L_n > l_\alpha^*\right) = 1$$

for all $\alpha \in [0,1]$.

5 Simulation examples

Monte Carlo simulations are conducted in this section to evaluate the empirical sizes and powers of the proposed statistics L_n and Π_n in the finite–sample situations. A bootstrap schedule is used to generate critical values l_α^* for them where α are 1% and 5%, although it is known that the distribution of the local time process $L_B(1,0)$ is $P(L_B(1,0) \le x) = 2\Phi(x) - 1$, where $\Phi(x)$ is the distribution function of a standard normal variable $U \sim N(0,1)$. The number of replications for Monte Carlo simulation is $M = 1000$ and that for the bootstrap procedure is $M_b = 200$.

Let $x_t = x_{t-1} + u_t$ with $x_0 = O_P(1)$, and u_t be a linear process generated by $u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$ with $\psi = \psi_1 + \psi_2 + \psi_3 \neq 0$, where $(\varepsilon_t, e_t) \sim iidN(\mathbf{0}, \Sigma)$ with $\Sigma = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ and $\rho \in \{\pm 0.75, \pm 0.5, \pm 0.25, 0\}$. In addition, the truncation parameter is selected by $k = [n^\kappa]$ with $\kappa \in \{1/5, 1/4, 1/3\}$. We choose $\triangle_n(x) = \delta_n \frac{1}{1+x^2}$, in which $\delta_n = \frac{1}{2}\sqrt{k\log(n)/\sqrt{n}}$. We shall examine the finite–sample performance of L_n and Π_n for these sample sizes $n = 300, 500, 800$.

Example 5.1. This example is about to examine the integrable regression function case. The model for simulation is $y_t = m(x_t) + e_t$, and the null hypothesis H_0 is $m(x) = g(x; \theta_0) = \exp(-\theta_0 x^2)$, $\theta_0 = 1$; the alternative is H_1: $m(x) = \exp(-\theta_1 x^2) + \triangle_n(x)$ with $\theta_1 = 1$.

Tables 1 and 3 show the sizes of the statistic $\frac{d_n}{nk} L_n$ at 1% and 5% levels for various sample sizes n, truncation parameters k by choosing different powers of κ and a verity of regressor cases generated by taking a series of covariances ρ, $\psi_1 = 1$, $\psi_2 = \psi_3 = 0$ in Tables 1, while $\psi_1 = -0.01$, $\psi_2 = 0.3, \psi_3 = 0.5$ in Tables 3, respectively. It can be seen that all sizes are reasonably close to the given significance levels. Overall, Tables 1 and 3 show that the sizes become stable when the sample size increases.

Tables 2 and 4 give all corresponding powers of the statistic in the same situations as for their sizes. Although the deviation function $\triangle_n(x) = \delta_n \triangle(x)$ is extremely small in the sense that δ_n approaches zero but $\delta_n^2 \sqrt{n}/k = \frac{1}{4}\log(n)$ and we select $\triangle(x)$ integrable, the powers are relatively strong. It is clear that in most cases the power increases as either the sample size or κ in truncation parameter increases. This is because the increase of sample size or κ would bring more terms of orthogonal basis into statistic L_n and, as can be seen in the proof of Theorem 2, every term may make contribution to the power. In addition, there is no clear difference in terms of power when the correlation level varies, but the choice of the coefficients in the linear process affect the powers which can be seen clearly from Tables 2 and 4. This phenomenon however is not evident in sizes.

We are then interested in comparing the proposed test L_n with a normalised version given in equation (3.7) by $\widehat{L}_{bn} = \frac{L_{bn}}{\sqrt{2\widehat{\sigma}_n^2}}$, where $L_{bn} = \sum_{t=}^{n} \sum_{s=1,\neq t}^{n} (y_t - g(x_t; \widehat{\theta})) Z(x_t)' Z(x_s)(y_s - g(x_s; \widehat{\theta}))$ and $\widehat{\sigma}_n^2 = \sum_{t=}^{n} \sum_{s=1,\neq t}^{n} [(y_t - g(x_t; \widehat{\theta})) Z(x_t)' Z(x_s)(y_s - g(x_s; \widehat{\theta}))]^2$. This is because $\widehat{L}_{bn}$ has been frequently used as a model specification test for the case where the regressors may be stationary processes (see, i.e., Gao et al. (2002)) or nonstationary processes (see, i.e. Gao et al. (2009a,b) and Wang and Phillips (2012)). We compare $\widehat{L}_{bn}$ in the same settings as for L_n that $\psi_1 = 1$ and $\psi_2 = \psi_3 = 0$. The results for the size and the power are reported in Tables 5 and 6, respectively.

In terms of size, $\widehat{L}_{bn}$ performs similarly to L_n, although $\widehat{L}_{bn}$ is slightly more oversized than L_n, especially at the 1% level. While on the power performance, L_n is slightly more powerful than $\widehat{L}_{bn}$. Our experience in Gao et al. (2012) suggests that the fact that L_n is more powerful than $\widehat{L}_{bn}$ may be justified theoretically, but we wish to leave this for future research.

Example 5.2. This example is the case of non-integrable regression function. The model for simulation is $y_t = m(x_t) + e_t$, and the null hypothesis H_{10} is $m(x) = g(x;\theta_0) = \theta_0 x^2$, here $\theta_0 = 1$; the alternative is H_{11}: $m(x) = g(x;\theta_1) + \triangle_n(x) = \theta_1 x^2 + \delta_n/(1+x^2)$ in which $\theta_1 = 1$, $\delta_n = \frac{1}{2}\sqrt{k\log(n)/\sqrt{n}}$.

Simulations for Π_n in this framework corresponding to different regressors with various generating mechanisms (by taking different covariance ρ and coefficients ψ_j, $j = 1,2,3$, in the linear process) are conducted for different sample sizes n, a series of powers κ in truncation parameter k, and their sizes and powers are reported in Tables 7-15, which additionally include the results at the level of 10% in every situation.

It can be seen that all sizes of the tests for $\frac{d_n}{nk}\Pi_n$ are fluctuated reasonably around respective nominal levels, although there are more oversized tests happen in Tables 13 and 15. In majority cases, the sizes of all tests seem approach the critical levels respectively with the increase of sample size. By large, the performance of $\frac{d_n}{nk}\Pi_n$ is satisfactory.

The deviation function in the alternative hypothesis in this example is the same as in Example 5.1. However, the powers of Π_n are slightly stronger than those of L_n in the last example. In terms of power Π_n still performs stable which are increasing with the increase of either sample size n or the truncation parameter k. This once again is mainly because the increase of either n or k results in more terms of the basis employed in the statistic.

6 Empirical analysis

This section provides an empirical example about the application of the proposed testing statistic.

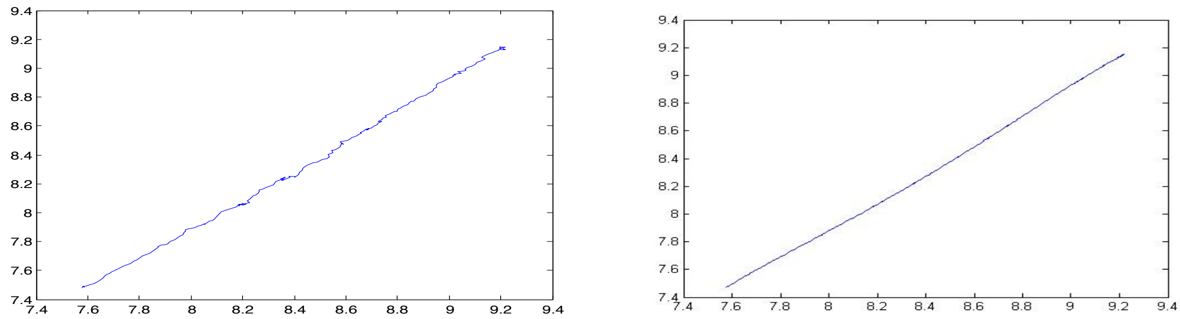
Example 6.1. Let us investigate the relationship between the United States customers' consumption expenditure and disposable income over time span 1960-2009. The data set is quarterly data from the Bureau of Economic Analysis at the website (http://www.bea.gov). Let y_t be the logarithm of the consumption expenditure and x_t be the logarithm of the disposable income, $t = 1, \cdots, 200$.

Consider a nonlinear time series regression model of the form

$$y_t = m(x_t) + e_t, \tag{6.1}$$

where $m(\cdot)$ is an unknown function, and it is to be approximated by a truncated orthogonal series.

Looking at the plot for (x_t, y_t) in Figure 1a, it is unlikely that $m(x)$ is integrable on $\mathbb{R}$. Thus, we will estimate $m(x)$ in the space $L^2(\mathbb{R}, \exp(-x^2))$ by $m_k(x) = \sum_{i=0}^{k-1} b_i h_i(x)$ where



(a) plot (x_t, y_t)

(b) plot $(x_t, \widehat{m}_5(x_t))$

Figure 1: Comparison of the real data and the estimated data

denote $\beta = (b_0, \cdots, b_{k-1})'$ and $h_i(x)$, $i \geq 0$, are defined as

$$h_i(x) = \frac{1}{\sqrt{\sqrt{\pi}2^i i!}} H_i(x)$$

with $H_i(x)$ being Hermite polynomial of order i.

In order to choose a suitable order for the polynomial function, we propose using the Generalised Cross–Validation (GCV) method (see Gao et al., 2002) to select an optimal value for k. Let $\widehat{k}$ denote the optimal value,

$$\widehat{k} = \arg\min_{k \in K_n} \left(1 - \frac{k}{n}\right)^{-2} \widehat{\sigma}^2(k), \tag{6.2}$$

where $K_n = \{[cn^{1/(2(q+1)+1)-\epsilon}], \cdots, [dn^{1/(2(q+1)+1)+\epsilon}]\}$, in which $0 < c < d < \infty$, q is the smoothness order of $m(x)$, $0 < \epsilon < 1/[2(q+1)(2(q+1)+1)]$, and $\widehat{\sigma}^2(k) = \frac{1}{n}\sum_{t=1}^{n}(y_t - m_k(x_t))^2$, in which $\widehat{m}_k(x) = \sum_{i=0}^{k-1} \widehat{b}_i h_i(x)$ with $\widehat{b}_i$ being the i–th component of $\widehat{\beta}$.

In this example, $c = 0.8$, $d = 2$, $q = 1$, so that $K_n = [2, \cdots, 7]$ and we have $\widehat{k} = 5$ as well as $\widehat{\beta} = (-925.9612, 310.1278, -55.0042, 5.35, -0.2263)$ by OLS. The points $(x_t, \widehat{m}_{\widehat{k}}(x_t))$, $t = 1, 2, \cdots, 200$, are plotted in Figure 1b as a comparison with the real dataset. These graphs motivate us to test the following hypothesis,

$$H_0: \ P\left(m(x_t) = g(x_t, \theta_0)\right) = 1. \tag{6.3}$$

where $g(x, \theta_0) = \theta_{00} + \theta_{01}x + \theta_{02}x^2 + \theta_{03}x^3 + \theta_{04}x^4$ with $\theta_0 = (\theta_{00}, \theta_{01}, \cdots, \theta_{04})'$. Under H_{10}, we have the following parametric model

$$y_t = g(x_t, \theta_0) + e_t. \tag{6.4}$$

Thus, θ_0 is estimated as $\widehat{\theta} = (-666.9834, 322.7505, -58.0615, 4.6441, -0.1389)$.

The proposed test Π_n is then applicable to deal with H_0. We invoke bootstrap scheme to generate the empirical distribution for Π_n. A commonly used bootstrap method is described as follows.

Step 1 Generate the bootstrap residuals $\{e_t^*\}$ by $e_t^* = \widehat{e}_t \eta_t^*$, where $\{\eta_t^*, 1 \le t \le n\}$ is a sequence of i.i.d. random variables possessing two-point distribution $P(\eta_t^* = \mp(\sqrt{5} \mp 1)/2) = (\sqrt{5} \pm 1)/2\sqrt{5}$

Step 2 Obtain $y_t^* = g(x_t; \widehat{\theta}) + e_t^*$. Use the data set $\{(x_t, y_t^*), 1 \le t \le n\}$ to re-estimate θ_0 and denote their estimators by $\widehat{\theta}^*$. Then calculate the test statistic Π_n^*, which is the corresponding version of Π_n by replacing $\{(x_t, y_t)\}$ and $\widehat{\theta}$ with $\{(x_t, y_t^*)\}$ and $\widehat{\theta}^*$, respectively.

Step 3 Repeat Steps 1-2 $M = 250$ times. We then have $\Pi_{n\ell}^*$, $\ell = 1, \cdots, 250$ and compute the proportion of $\Pi_n < \Pi_{n\ell}^*$ for model (6.4). This proportion is an approximate P-value of Π_n.

According to our simulated P–values for model (6.4), the corresponding P–values for the first–order, third–order and fourth–order polynomials are all under 3%. When we try a second–order polynomial of the form: $g(x, \theta_0) = \theta_{00} + \theta_{01}x + \theta_{02}x^2$, we have $\widehat{\theta} = (4.6673, -0.1700, 0.0714)$, and a P–value of 74.2% for Π_n. This implies that a second–order polynomial function of the form $g(x, \theta_0) = 4.6673 - 0.17x + 0.0714x^2$ may be suitable for model (6.4). Figure 2 gives the plots of both the original data and estimated data in one picture.

Since existing literature assumes that y_t and x_t are I(1) processes, their differences $\triangle y_t$ and $\triangle x_t$ are thus stationary. Thus, we would like to test if there exists a linear relationship between them using $\widehat{L}_{bn}$ as given in equation (3.7). The simulated P–value of the test statistic $\widehat{L}_{bn}$ is 0.5%, implying that there is no evidence to support the linearity between $\triangle y_t$ and $\triangle x_t$. This outcome, on the one hand, coincides with the intuition of the plot ($\triangle y_t$, $\triangle x_t$) in Figure 3, on the other hand, partially confirms our finding that y_t and x_t may have a quadratic relationship.

7 Concluding remarks

We have proposed two concise and computationally simple tests for parametric specification of time series models with nonlinearity and nonstationarity in both cases of integrable and non-integrable regression functions. An asymptotic theory for each of the proposed test statistics has been established. A bootstrap scheme has been proposed to find a simulated critical value as an approximation to that of the exact sample distribution of the proposed test statistic under consideration. Several Monte Carlo simulation examples have been used

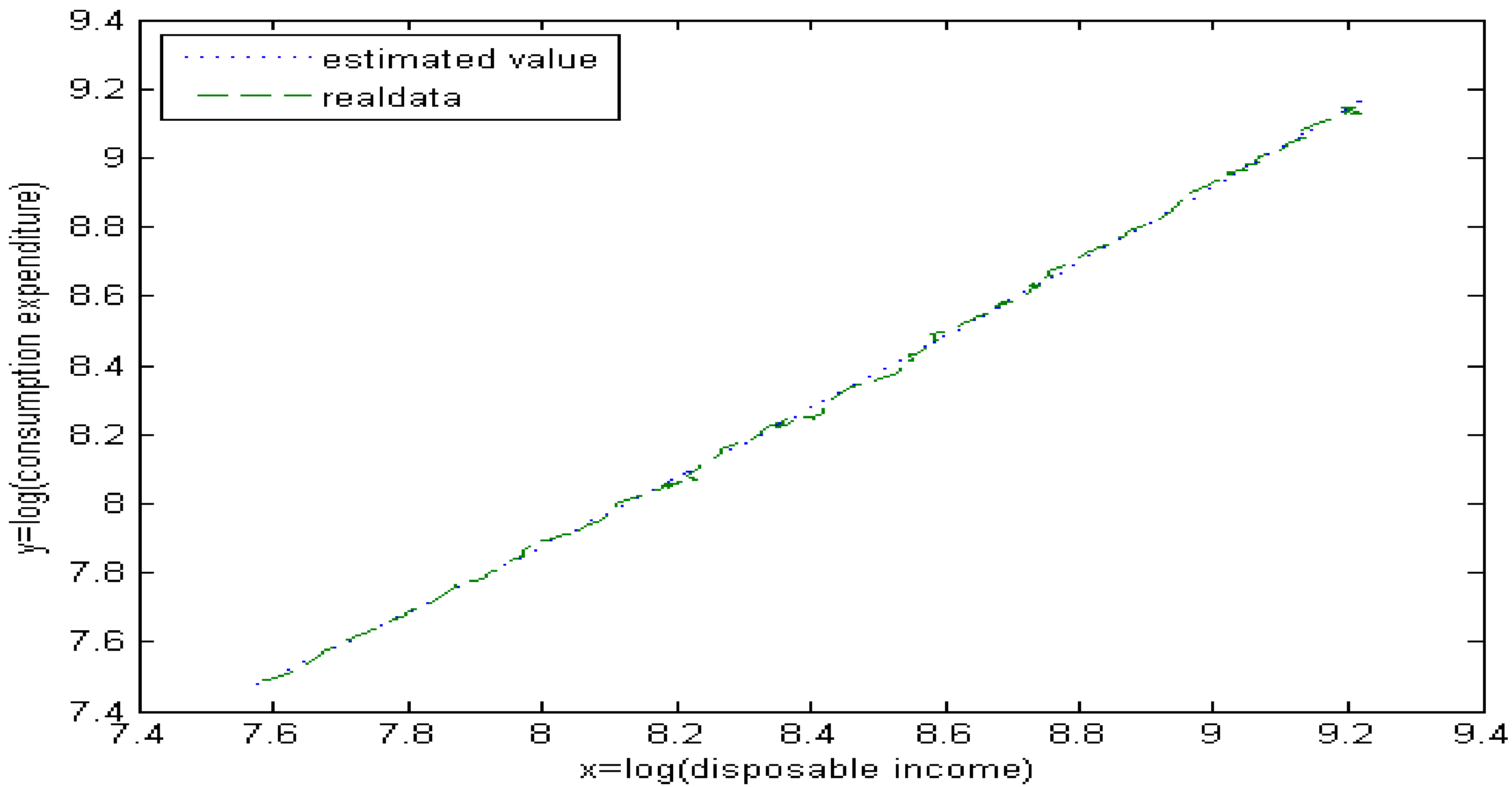


Figure 2: Plots of the real data (x_t, y_t) and the estimated data $(x_t, \widehat{m}_3(x_t))$

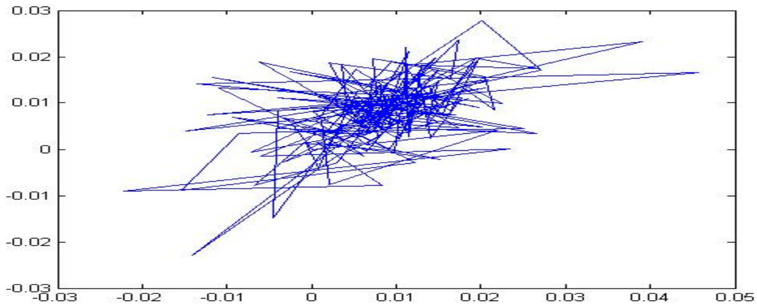


Figure 3: plot $\triangle y_t$ against $\triangle x_t$

to evaluate the finite–sample performance of the proposed tests. Overall, the empirical sizes and power values are found to be satisfactory. In addition, an empirical analysis has been provided to conclude that a simple linear relationship between the United States consumers' consumption expenditure and disposable income may not be justifiable.

As pointed out in the introductory section, the proposed tests have been developed for the univariate time series regressor case. In the case where one will need to model the relationship between y_t and a vector of integrated time series regressors, one may consider using a varying–coefficient model of the form (see, e.g., Cai et al., 2009)

$$y_t = u_t' \beta(v_t) + e_t, \tag{7.1}$$

where $x_t = (u_t', v_t)'$ with u_t being a vector of integrated time series regressors and v_t being a univariate integrated time series regressor, e_t is an error time series and $\beta(\cdot)$ is a vector of unknown functions defined on $\mathbb{R}$.

In this case, there is need in empirical applications to check whether the functional form of $\beta(\cdot)$ may be parametrically specified through testing $H_{20}: \ P\left(\beta(v_t) = \beta(v_t;\theta_0)\right) = 1$. If H_{20} is true, then model (7.1) will become a parametric time series model of the form

$$y_t = u_t' \beta(v_t;\theta_0) + e_t. \tag{7.2}$$

While the construction of a test for H_{20} follows similarly from that of L_n, the establishment of an asymptotic theory requires some additional conditions and proofs. Letting y_t and $u_t = x_t$ be defined in the same way as in Example 6.1 above and v_t be the real interest rate, an important empirical application is to examine whether model (7.2) is more appropriate for the data than what has been discussed in the literature. A recent paper by Gao et al. (2012) considers such an empirical problem, but it treats the real interest rate, which probably should be considered as nonstationary, as a stationary time series. We wish to leave such further discussion for future research.

There are also some other extensions. One of them is whether the proposed test is extendable to accommodate the case where x_t is a vector of multivariate regressors in a nonparametric multivariate time series case. Our experience suggests that it may be possible to construct a multivariate series expansion and then a multivariate version of the proposed test before a resulting theory may be established. Another issue is how to develop a data–driven method for the choice of the truncated parameter k for the integrated time series case, although the literature (e.g.,Gao et al., 2012;Wang and Phillips, 2012) indicates that fixed–bandwidths work well numerically in the integrated time series case. In the stationary time series, Chapter 3 of Gao (2007) shows that one may choose a suitable truncated parameter such that the power function of the test under consideration is maximised while the size function is under control by a significance level. Since extensions require developing new techniques, we also wish to leave them for future research.

A Technical lemmas

The first section in Appendix is to present two technical lemmas which are crucial for the proofs of the theorems. Let us first introduce function sequence $\{f_k(x)\}$ and function $f(x)$:

$$f_k(x) := \frac{1}{k}\|Z(x)\|^2 = \frac{1}{k}\sum_{i=0}^{k-1}\mathscr{H}_i^2(x), \quad \text{and} \quad f(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2}, & \text{if } |x| \le 2, \\ 0, & \text{if } |x| > 2. \end{cases}$$

Lemma A.1. *The sequence $f_k(x)$ converges to $f(x)$ for any $x \in \mathbb{R}$ as $k \to \infty$. Moreover, as $k \to \infty$*

$$\int_{-\infty}^{\infty} |f_k(x) - f(x)|\, dx \to 0. \tag{A.1}$$

Proof. The convergence of $f_k(x) \to f(x)$ is the consequence of the celebrated Wigner's semicircle law (Wigner, 1958), which also can be found in Edelman and Rao (2005, p. 29).

Since $f_k(x)$ for each $k \ge 1$ can be viewed as a probability density and $f(x)$ is actually the so-called Wigner semicircle law with radius 2, namely, a probability density as well, it follows from Scheffe's Theorem of Billingsley (1968, p. 224) that (A.1) holds. $\square$

Lemma A.2. *(1) For integrable function $F(x)$, $\int |F(x)|dx < \infty$, we have $\int \|Z(x)\|^2|F(x)|dx = O(1)\sqrt{k}$.*

(2) For continuous function $F(x)$ satisfying $\int |F(x)|e^{-x^2}dx < \infty$, we have $\int \|Z(x)\|^2F(x)dx = O(1)k$.

Proof. (1) In order to prove the assertion, we need to introduce the following function sequence, $\ell_k(x) := \frac{c_k}{k}\|Z(c_kx)\|^2 = \frac{c_k}{k}\sum_{i=0}^{k-1}\mathscr{H}_i^2(c_kx)$ for $k \geq 1$ with $c_k = \sqrt{2k}$.

In view of the positivity and integration to 1 for $\ell_k(x)$ on $\mathbb{R}$, we define a sequence of absolutely continuous probability measures $\{\nu_k, k = 1, 2, \cdots\}$ by $\nu_k(A) = \int_A \ell_k(x)dx$, where A is any Borel set on $\mathbb{R}$.

According to Theorem 5.3 of Assche (1987, p. 152), ν_k converge weekly to a measure ν with support $[-1, 1]$ as $k \to \infty$. Here, measure ν is the so–called Ullman measure ν^α with parameter $\alpha = 2$. In this particular case, the Ullman measure takes the form of $\nu(A) = \int_A \ell(x)dx$, where $\ell(x) = \frac{2}{\pi}\sqrt{1-x^2}$ on $[-1, 1]$ and 0 elsewhere, and A is any Borel set in $[-1, 1]$.

It follows that $\ell_k(x)$ and $\ell(x)$ are actually the densities (Radon-Nikodým derivatives) of the measures ν_k and ν, respectively. Let ϵ be a positive real number. By virtue of the continuity of $\ell_k(x)$ and $\ell(x)$ and the weak convergence of ν_k to ν, for any x,

$$\begin{aligned}\lim_{k\to\infty}\ell_k(x) &= \lim_{k\to\infty}\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_x^{x+\epsilon}\ell_k(t)dt = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\lim_{k\to\infty}\int_x^{x+\epsilon}\ell_k(t)dt\\ &=\lim_{\epsilon\to 0}\frac{1}{\epsilon}\lim_{k\to\infty}\nu_k(x, x+\epsilon) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\nu(x, x+\epsilon) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_x^{x+\epsilon}\ell(t)dt = \ell(x),\end{aligned}$$

which implies $\|Z(c_kx)\|^2 = O(1)\sqrt{k}$. Thus, making variable change $x = c_ky$ twicely, $\int\|Z(x)\|^2|F(x)|dx = \int\|Z(c_ky)\|^2|F(c_ky)|dc_ky = O(1)\sqrt{k}\int|F(c_ky)|dc_ky = O(1)\sqrt{k}\int|F(x)|dx = O(1)\sqrt{k}$.

(2) Noting from (2.1) that $h_i(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{2^i i!}}H_i(x)$ and $H_i'(x) = 2iH_{i-1}(x)$, we have $h_i'(x) = \sqrt{2i}h_{i-1}(x)$ for $i \geq 1$. It follows from Christoffel-Darboux formula (2.2) for Hermite polynomials that

$$\begin{aligned}\|Z(x)\|^2 &= \sum_{i=0}^{k-1}\mathscr{H}_i^2(x) = e^{-x^2}\sum_{i=0}^{k-1}h_i^2(x)\\ &=e^{-x^2}\sqrt{k/2}(h_k'(x)h_{k-1}(x) - h_k(x)h_{k-1}'(x))\\ &=e^{-x^2}\sqrt{k/2}(\sqrt{2k}h_{k-1}^2(x) - \sqrt{2(k-1)}h_k(x)h_{k-2}(x))\\ &=k(\mathscr{H}_{k-1}^2(x) - \sqrt{1-1/k}\mathscr{H}_k(x)\mathscr{H}_{k-2}(x)).\end{aligned}$$

By virtue of Theorem 5.7 in Assche (1987, p.160), we have

$$\lim_{k\to\infty}\int_{-\infty}^{\infty}\mathscr{H}_k^2(x)F(x)dx = \frac{1}{\pi}\int_{-1}^{1}\frac{F(x)}{\sqrt{1-x^2}}dx,$$

which, along with Cauchy-Schwarz inequality, implies $\int\|Z(x)\|^2F(x)dx = O(1)k$. □

B Proof of Theorem 3.1

Proof. Under the hypothesis H_0, $y_t = g(x_t;\theta_0) + e_t$ for all $t = 1, \cdots, n$. Hence, it is easy to rewrite $L_n = L_{1n} + L_{2n} + L_{3n}$, where

$$L_{1n} = \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_t e_s, \tag{B.1}$$

$$L_{2n} = 2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_t \widehat{g}(x_s), \tag{B.2}$$

$$L_{3n} = \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s) \tag{B.3}$$

where $\widehat{g}(x) := g(x;\theta_0) - g(x;\widehat{\theta})$ for notational convenience.

We first deal with L_{1n}.

$$\begin{aligned} L_{1n} =& \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_t e_s = \sum_{t=1}^{n} Z(x_t)'Z(x_t)e_t^2 + \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} Z(x_t)'Z(x_s)e_t e_s \\ =& \sigma_e^2 \sum_{t=1}^{n} Z(x_t)'Z(x_t) + \sum_{t=1}^{n} Z(x_t)'Z(x_t)(e_t^2 - \sigma_e^2) + 2\sum_{t=2}^{n}\sum_{s=1}^{t-1} Z(x_t)'Z(x_s)e_s e_t \\ =& \sigma_e^2 I_1 + I_2 + 2I_3, \qquad \text{say.} \end{aligned}$$

We then have

$$\begin{aligned} \frac{d_n}{nk} I_1 =& \frac{d_n}{nk}\sum_{t=1}^{n} Z(x_t)'Z(x_t) = \frac{d_n}{n}\sum_{t=1}^{n}\frac{1}{k} Z(x_t)'Z(x_t) = \frac{d_n}{n}\sum_{t=1}^{n} f_k(x_t) \\ =& \frac{d_n}{n}\sum_{t=1}^{n} f(x_t) + \frac{d_n}{n}\sum_{t=1}^{n}\left[f_k(x_t) - f(x_t)\right]. \end{aligned}$$

It follows from Wang and Phillips (2009a) that, since $\int f(x)dx = 1$,

$$\frac{d_n}{n}\sum_{t=1}^{n} f(x_t) = \frac{d_n}{n}\sum_{t=1}^{n} f(d_n x_{tn}) \to_D L_B(1,0).$$

Next, we shall prove that $\frac{d_n}{n}\sum_{t=1}^{n}[f_k(x_t) - f(x_t)] \to_P 0$, and $\frac{d_n}{nk} I_2 \to_P 0$, $\frac{d_n}{nk} I_3 \to_P 0$.

In fact, noting that x_t/d_t have density functions $g_t(x)$ which have a uniform bound, say C, over all x and t,

$$\begin{aligned} & E\left|\frac{d_n}{n}\sum_{t=1}^{n}[f_k(x_t) - f(x_t)]\right| \leq \frac{d_n}{n}\sum_{t=1}^{n} E\left|f_k(x_t) - f(x_t)\right| \\ =& \frac{d_n}{n}\sum_{t=1}^{n}\int \left|f_k(d_t x) - f(d_t x)\right|\; g_t(x)dx = \frac{d_n}{n}\sum_{t=1}^{n}\frac{1}{d_t}\int \left|f_k(x) - f(x)\right|\; g_t(x/d_t)dx \end{aligned}$$

$$\leq C\frac{d_n}{n}\sum_{t=1}^{n}\frac{1}{d_t}\int|f_k(x)-f(x)|\,dx = O(1)\int|f_k(x)-f(x)|\,dx \to 0,$$

by Lemma A.1.

Moreover, by virtue of Assumption A,

$$E\left(\frac{d_n}{nk}I_2\right)^2 = E\left(\frac{d_n}{nk}\sum_{t=1}^{n}Z(x_t)'Z(x_t)(e_t^2-\sigma_e^2)\right)^2 = (\mu_4-\sigma_e^4)\frac{d_n^2}{n^2k^2}\sum_{t=1}^{n}E(Z(x_t)'Z(x_t))^2$$
$$=O(1)\frac{d_n^2}{n^2k^2}\sum_{t=1}^{n}\int[Z(d_tx)'Z(d_tx)]^2g_t(x)dx = O(1)\frac{d_n^2}{n^2k^2}\sum_{t=1}^{n}\frac{1}{d_t}\int[Z(x)'Z(x)]^2g_s(x/d_t)dx$$
$$\leq O(1)\frac{1}{\sqrt{n}k^2}\int[Z(x)'Z(x)]^2dx \leq O(1)\frac{\sqrt{k}}{\sqrt{n}k^2}\int Z(x)'Z(x)dx = O(1)\frac{1}{\sqrt{nk}} \to 0,$$

using Lemma A.2 and the orthogonality $\int Z(x)'Z(x)dx = k$.

Meanwhile, engaging density functions $g_s(x)$ for x_s/d_s and conditional densities $g_{ts}(x)$ for $\frac{1}{d_{ts}}(x_t-x_s)$ on x_s with $d_{ts}=|\psi|\sqrt{t-s}$ and using Assumption A,

$$E\left(\frac{d_n}{nk}I_3\right)^2 = 4\sigma_e^4\frac{d_n^2}{n^2k^2}\sum_{t=2}^{n}\sum_{s=1}^{t-1}E(Z(x_s)'Z(x_t))^2$$
$$=4\sigma_e^4\frac{d_n^2}{n^2k^2}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{d_s}\frac{1}{d_{ts}}\iint(Z(x)'Z(y))^2g_s(y/d_s)g_{ts}((x-y)/d_{ts})dxdy$$
$$\leq O(1)\frac{d_n^2}{n^2k^2}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{d_s}\frac{1}{d_{ts}}\iint(Z(x)'Z(y))^2dxdy$$
$$=O(1)\frac{1}{k^2}\iint(Z(x)'Z(y))^2dxdy$$
$$=O(1)\frac{1}{k^2}k = O(1)\frac{1}{k} \to 0,$$

by orthogonality.

We are now in a position to tackle the term L_{3n}. For any $\delta,\epsilon>0$,

$$P\left(\frac{d_n}{nk}|L_{3n}|>\delta\right)$$
$$\leq P(\|\widehat{\theta}-\theta_0\|>\zeta_n\epsilon)+P\left(\frac{d_n}{nk}|L_{3n}|>\delta, \|\widehat{\theta}-\theta_0\|\leq\zeta_n\epsilon\right)$$
$$=P(\|\widehat{\theta}-\theta_0\|>\zeta_n\epsilon)+P\left(\frac{d_n}{nk}|L_{3n}|>\delta, I(\|\widehat{\theta}-\theta_0\|\leq\zeta_n\epsilon)=1\right)$$
$$=P(\|\widehat{\theta}-\theta_0\|>\zeta_n\epsilon)+P\left(\frac{d_n}{nk}|L_{3n}|I(\|\widehat{\theta}-\theta_0\|\leq\zeta_n\epsilon)>\delta\right)$$
$$\leq P(\|\widehat{\theta}-\theta_0\|>\zeta_n\epsilon)+\frac{d_n}{\delta nk}E[|L_{3n}|I(\|\widehat{\theta}-\theta_0\|\leq\zeta_n\epsilon)],$$

by virtue of Markov's inequality, where $I(\cdot)$ stands for indicator function.

As $\|\widehat{\theta}-\theta_0\| = o_P(\zeta_n)$ assumed in Assumption B, $P(\|\widehat{\theta}-\theta_0\| > \zeta_n\epsilon) \to 0$. Using Taylor expansion for $g(x;\theta)$ with respect to θ in a neighborhood of θ_0, we have

$$
\begin{aligned}
\widehat{g}(x_t) =& g(x_t;\theta_0) - g(x_t;\widehat{\theta}) \\
=& l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta}) - \frac{1}{2}(\theta_0-\widehat{\theta})' l_2(x_t;\bar{\theta})(\theta_0-\widehat{\theta})
\end{aligned}
$$

where $\bar{\theta}$ is on the line segment connecting θ_0 and $\widehat{\theta}$. In view of this, it follows that

$$
\begin{aligned}
& \frac{d_n}{nk} E(|L_{3n}| I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon)) \\
=& \frac{d_n}{nk} E \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s) I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon) \\
=& \frac{d_n}{nk} E \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s) l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta}) \\
& \times l_1(x_s;\theta_0)'(\theta_0-\widehat{\theta}) I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon) \\
& - \frac{d_n}{nk} E \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s) l_1(x_t;\theta_0)'(\theta_0-\widehat{\theta}) \\
& \times (\theta_0-\widehat{\theta})' l_2(x_s;\bar{\theta})(\theta_0-\widehat{\theta}) I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon) \\
& + \frac{d_n}{4nk} E \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)(\theta_0-\widehat{\theta})' l_2(x_t;\bar{\theta})(\theta_0-\widehat{\theta}) \\
& \times (\theta_0-\widehat{\theta})' l_2(x_s;\bar{\theta})(\theta_0-\widehat{\theta}) I(\|\widehat{\theta}-\theta_0\| \le \zeta_n\epsilon) \\
=& T_1 + T_2 + T_3, \qquad \text{say.}
\end{aligned}
$$

Using Cauchy-Schwarz inequality,

$$
\begin{aligned}
T_1 \le& \epsilon^2\zeta_n^2 \frac{d_n}{nk} E \sum_{t=1}^{n} Z(x_t)'Z(x_t)\|l_1(x_t;\theta_0)\|^2 \\
& + 2\epsilon^2\zeta_n^2 \frac{d_n}{nk} E \sum_{t=2}^{n}\sum_{s=1}^{t-1} |Z(x_t)'Z(x_s)| \|l_1(x_s;\theta_0)\| \|l_1(x_t;\theta_0)\| \\
\le& C\epsilon^2\zeta_n^2 \frac{d_n}{nk} \sum_{t=1}^{n} \frac{1}{d_t} \int Z(x)'Z(x)\|l_1(x;\theta_0)\|^2 dx \\
& + 2C\epsilon^2\zeta_n^2 \frac{d_n}{nk} \sum_{t=2}^{n}\sum_{s=1}^{t-1} \frac{1}{d_s}\frac{1}{d_{ts}} \iint |Z(x)'Z(y)| \|l_1(x;\theta_0)\| \|l_1(y;\theta_0)\| dxdy \\
\le& O(1)\epsilon^2\zeta_n^2 \frac{1}{\sqrt{k}} + O(1)\epsilon^2\zeta_n^2 \frac{\sqrt{n}}{k} \left(\int \|Z(x)\| \|l_1(x;\theta_0)\| dx \right)^2 \\
\le& O(1)\zeta_n^2 \frac{1}{\sqrt{k}} + O(1)\zeta_n^2 \sqrt{n}/k \int \|Z(x)\|^2 \|l_1(x;\theta_0)\| dx \int \|l_1(x;\theta_0)\| dx \\
\le& O(1)\zeta_n^2 \frac{1}{\sqrt{k}} + O(1)\zeta_n^2 \sqrt{n/k} \to 0
\end{aligned}
$$

by Lemma A.2, Assumption B, and Cauchy-Schwarz inequality.

Similarly, using Assumptions B and C we have

$$
\begin{aligned}
T_3 \leq & \epsilon^4 \zeta_n^4 \frac{d_n}{nk} E \sum_{t=1}^{n} Z(x_t)' Z(x_t) \|l_2(x_t;\bar{\theta})\|^2 + \epsilon^4 \zeta_n^4 \frac{d_n}{nk} E \sum_{t=2}^{n} \sum_{s=1}^{t-1} |Z(x_t)' Z(x_s)| \|l_2(x_t;\bar{\theta})\| \|l_2(x_s;\bar{\theta})\| \\
\leq & O(1) \zeta_n^4 \frac{d_n}{nk} \sum_{t=1}^{n} \frac{1}{d_t} \int l^2(x) Z(x)' Z(x) dx + O(1) \zeta_n^4 \frac{d_n}{nk} \sum_{t=2}^{n} \sum_{s=1}^{t-1} \frac{1}{d_s} \frac{1}{d_{ts}} \iint l(x) l(y) |Z(x)' Z(y)| dx dy \\
\leq & O(1) \zeta_n^4 \frac{1}{\sqrt{k}} + O(1) \zeta_n^4 \frac{\sqrt{n}}{\sqrt{k}} \to 0.
\end{aligned}
$$

Notice that $T_1 = o_P(1)$ and $T_3 = o_P(1)$ imply $T_2 = o_P(1)$. Hence, it follows that $L_{3n} \to_P 0$ and then from Cauchy-Schwarz inequality that $L_{2n} \to_P 0$ as well. The proof of Theorem 1 is finished. □

C Proof of Theorem 3.2

Proof. Under H_1 and Assumption D, we have $m(x_t) = g(x_t;\theta_1) + \triangle_n(x_t) = g(x_t;\theta_1) + \delta_n \triangle(x_t)$. Similar to (B.1)-(B.3), write $L_n = L_{1n} + L_{2n} + L_{3n}$, where

$$
\begin{aligned}
L_{1n} &= \sum_{t=1}^{n} \sum_{s=1}^{n} Z(x_t)' Z(x_s) (e_t + \delta_n \triangle(x_t))(e_s + \delta_n \triangle(x_s)) \\
L_{2n} &= 2 \sum_{t=1}^{n} \sum_{s=1}^{n} Z(x_t)' Z(x_s) (e_t + \delta_n \triangle(x_t)) \widehat{g}(x_s) \\
L_{3n} &= \sum_{t=1}^{n} \sum_{s=1}^{n} Z(x_t)' Z(x_s) \widehat{g}(x_t) \widehat{g}(x_s),
\end{aligned}
$$

in which $\widehat{g}(x) = g(x;\theta_1) - g(x;\widehat{\theta})$.

Observe that using the same argument as in the proof of Theorem 3.1 we can show that $\frac{d_n}{nk} L_{3n} = o_P(1)$, and Cauchy-Schwarz inequality implies that $|L_{2n}| \leq 2\sqrt{L_{1n} L_{3n}}$. Thus, to fulfill the proof, it suffices to show that $\frac{d_n}{nk} L_{1n} \to_P \infty$. To begin with,

$$
\begin{aligned}
L_{1n} = & \sum_{t=1}^{n} \sum_{s=1}^{n} Z(x_t)' Z(x_s) (e_t + \delta_n \triangle(x_t))(e_s + \delta_n \triangle(x_s)) \\
= & \delta_n^2 \sum_{t=1}^{n} \sum_{s=1}^{n} Z(x_t)' Z(x_s) \triangle(x_t) \triangle(x_s) \\
& + 2\delta_n \sum_{t=1}^{n} \sum_{s=1}^{n} Z(x_t)' Z(x_s) \triangle(x_t) e_s \\
& + \sum_{t=1}^{n} \sum_{s=1}^{n} Z(x_t)' Z(x_s) e_t e_s
\end{aligned}
$$

$$:= L'_{1n} + 2L''_{1n} + L'''_{1n}, \quad \text{say.}$$

Looking at the proof of Theorem 3.1, we have $\frac{d_n}{nk} L'''_{1n} \to_D \sigma_e^2 L_B(1,0)$. Rewrite $L''_{1n} = L''_{1n1} + L''_{1n2} + L''_{1n3}$, where

$$\begin{aligned}
L''_{1n1} &= \delta_n \sum_{t=1}^{n} Z(x_t)' Z(x_t) \triangle(x_t) e_t \\
L''_{1n2} &= \delta_n \sum_{t=2}^{n} e_t \sum_{s=1}^{t-1} Z(x_t)' Z(x_s) \triangle(x_s) \\
L''_{1n3} &= \delta_n \sum_{t=1}^{n-1} e_t \sum_{s=t+1}^{n} Z(x_t)' Z(x_s) \triangle(x_s).
\end{aligned}$$

Notice that $\frac{d_n}{nk} L''_{1n1} = o_P(1)$. In fact,

$$\begin{aligned}
& E\left(\frac{d_n}{nk} L''_{1n1}\right)^2 = \frac{d_n^2}{n^2k^2} \delta_n^2 E\left(\sum_{t=1}^{n} Z(x_t)' Z(x_t) \triangle(x_t) e_t\right)^2 \\
=& \sigma_e^2 \frac{d_n^2}{n^2k^2} \delta_n^2 \sum_{t=1}^{n} E[Z(x_t)' Z(x_t) \triangle(x_t)]^2 \\
\le& \sigma_e^2 C_g \frac{d_n^2}{n^2k^2} \delta_n^2 \sum_{t=1}^{n} \frac{1}{d_t} \int [Z(x)' Z(x) \triangle(x)]^2 dx \\
\le& O(1) \frac{1}{\sqrt{n}k^2} \delta_n^2 \sqrt{k} \int Z(x)' Z(x) \triangle^2(x) dx = O(1) \frac{1}{\sqrt{nk}} \delta_n^2 \to 0
\end{aligned}$$

using Lemma A.2.

Notice also that

$$\begin{aligned}
& E\left(\frac{d_n}{nk} L''_{1n2}\right)^2 = \frac{d_n^2}{n^2k^2} \delta_n^2 E\left(\sum_{t=2}^{n} e_t \sum_{s=1}^{t-1} Z(x_t)' Z(x_s) \triangle(x_s)\right)^2 \\
=& \sigma_e^2 \frac{d_n^2}{n^2k^2} \delta_n^2 \sum_{t=2}^{n} E\left(\sum_{s=1}^{t-1} Z(x_t)' Z(x_s) \triangle(x_s)\right)^2 \\
=& \sigma_e^2 \frac{d_n^2}{n^2k^2} \delta_n^2 \sum_{t=2}^{n} \sum_{s=1}^{t-1} E(Z(x_t)' Z(x_s) \triangle(x_s))^2 \\
&+ 2\sigma_e^2 \frac{d_n^2}{n^2k^2} \delta_n^2 \sum_{t=3}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} E(Z(x_t)' Z(x_{s_1}) \triangle(x_{s_1}) \\
&\qquad \times Z(x_t)' Z(x_{s_2}) \triangle(x_{s_2})) \\
\le& \sigma_e^2 C_g^2 \frac{d_n^2}{n^2k^2} \delta_n^2 \sum_{t=2}^{n} \sum_{s=1}^{t-1} \frac{1}{d_s d_{ts}} \iint (Z(x)' Z(y) \triangle(y))^2 dx dy \\
&+ 2C_g^3 \sigma_e^2 \frac{d_n^2}{n^2k^2} \delta_n^2 \sum_{t=3}^{n} \sum_{s_1=2}^{t-1} \sum_{s_2=1}^{s_1-1} \frac{1}{d_t d_{ts_1} d_{ts_2}}
\end{aligned}$$

$$
\begin{aligned}
&\times \iiint |Z(x)'Z(y)Z(x)'Z(z)\triangle(y)\triangle(z)|dxdydz \\
=&O(1)\frac{1}{k^2}\delta_n^2 \int Z(y)'Z(y)\triangle^2(y)dy \\
&+ O(1)\frac{\sqrt{n}}{k^2}\delta_n^2 \int dx \left(\int |Z(x)'Z(y)\triangle(y)|dy\right)^2 \\
=&O(1)\frac{1}{k}\delta_n^2 + O(1)\frac{\sqrt{n}}{k^2}\delta_n^2 \int dx \left(\int \left|\sum_{i=0}^{k-1} h_i(x)h_i(y)\right| e^{-(x^2+y^2)/2}|\triangle(y)|dy\right)^2 \\
\leq&O(1)\frac{1}{k}\delta_n^2 + O(1)\frac{\sqrt{n}}{k^2}\delta_n^2 \\
&\times \int dx \int \left|\sum_{i=0}^{k-1} h_i(x)h_i(y)\right|^2 e^{-x^2}e^{-y^2/2}dy \int |\triangle(y)|^2 e^{-y^2/2}dy \\
=&O(1)\frac{1}{k}\delta_n^2 + O(1)\frac{\sqrt{n}}{k^2}\delta_n^2 \int dx \int (Z(x)'Z(y))^2 e^{y^2/2}dy \\
=&O(1)\frac{1}{k}\delta_n^2 + O(1)\frac{\sqrt{n}}{k^2}\delta_n^2 \int \|Z(y)\|^2 e^{y^2/2}dy \\
=&O(1)\frac{1}{k}\delta_n^2 + O(1)\frac{\sqrt{n}}{k}\delta_n^2,
\end{aligned}
$$

by orthogonality of the components in $Z(x)$ and Lemma A.3.

Thus, $\frac{d_n}{nk}L''_{1n2} = O_P(n^{1/4}k^{-1/2}\delta_n)$. Similar calculation yields the same result for $\frac{d_n}{nk}L''_{1n3}$. Therefore, $\frac{d_n}{nk}L''_{1n} = O_P(n^{1/4}k^{-1/2}\delta_n)$.

Lastly,

$$
\begin{aligned}
\frac{d_n}{nk}L'_{1n} =& \delta_n^2\frac{d_n}{nk}\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\triangle(x_t)\triangle(x_s) \\
=&\delta_n^2\frac{d_n}{nk}\sum_{t=1}^{n}\sum_{s=1}^{n}\sum_{i=0}^{k-1} \mathscr{H}_i(x_t)\mathscr{H}_i(x_s)\triangle(x_t)\triangle(x_s) \\
=&\delta_n^2\frac{d_n}{nk}\sum_{i=0}^{k-1}\left(\sum_{t=1}^{n}\mathscr{H}_i(x_t)\triangle(x_t)\right)^2 \\
=&\delta_n^2\frac{n}{d_nk}\sum_{i=0}^{k-1}\left(\frac{d_n}{n}\sum_{t=1}^{n}\mathscr{H}_i(d_nx_{tn})\triangle(d_nx_{tn})\right)^2 \\
\geq&\delta_n^2\frac{\sqrt{n}}{k}\left(\frac{d_n}{n}\sum_{t=1}^{n}\mathscr{H}_{i_0}(d_nx_{tn})\triangle(d_nx_{tn})\right)^2 \to_P \infty
\end{aligned}
$$

because by Assumption D, $\delta_n^2\sqrt{n}/k \to \infty$ and from Wang and Phillips (2009a) and continuous mapping theorem $\left(\frac{d_n}{n}\sum_{t=1}^{n}\mathscr{H}_{i_0}(x_t)\triangle(x_t)\right)^2 \to_D \left(\int \mathscr{H}_{i_0}(x)\triangle(x)dxL_B(1,0)\right)^2$, where $i_0 \geq 0$ is the first integer such that $\int \mathscr{H}_{i_0}(x)\triangle(x)dx \neq 0$. Obviously, such i_0 does exist, otherwise for any i, $\int \mathscr{H}_i(x)\triangle(x)dx = \int h_i(x)\triangle(x)e^{-x^2/2}dx = 0$ so that $\triangle(x)e^{-x^2/4}$ is orthogonal with every $h_i(x)e^{-x^2/4}$ in Hilbert space $L^2(\mathbb{R})$. However, $\{h_i(x)e^{-x^2/4}\}$ is a maximal independent set in the space, because if $h(x) \in L^2(\mathbb{R})$, then $h(x)e^{x^2/4} \in L^2(\mathbb{R}, e^{-x^2})$ and hence $h(x)e^{x^2/4} = \sum_i c_i h_i(x)$, which implies that $h(x) = \sum_i c_i h_i(x)e^{-x^2/4}$. Therefore, $\{h_i(x)e^{-x^2/4}\}$ can be orthogonalised as a basis by Gram Schmidt process. This amounts to saying that $\triangle(x)e^{-x^2/4}$ is orthogonal with every element in a basis and hence $\triangle(x)e^{-x^2/4}$ must be a zero function, implying $\triangle(x)$ is a zero function. □

D Proof of Theorem 3.3

Proof. Under H_{10},

$$\begin{aligned}
\Pi_n =& \sum_{t=1}^{n}\sum_{s=1}^{n}(Y_t - G(x_t;\widehat{\theta}))Z(x_t)'Z(x_s)(Y_s - G(x_s;\widehat{\theta})) \\
=& \sum_{t=1}^{n}\sum_{s=1}^{n}(y_t - g(x_t;\widehat{\theta}))Z(x_t)'Z(x_s)(y_s - g(x_s;\widehat{\theta}))\varphi(x_t)\varphi(x_s) \\
=& \sum_{t=1}^{n}\sum_{s=1}^{n}(e_t + \widehat{g}(x_t))Z(x_t)'Z(x_s)(e_s + \widehat{g}(x_s))\varphi(x_t)\varphi(x_s) \\
=& \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)e_te_s + 2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\widehat{g}(x_s)e_t \\
&+ \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\widehat{g}(x_t)\widehat{g}(x_s) := \Pi_{1n} + 2\Pi_{2n} + \Pi_{3n},
\end{aligned}$$

where $\widehat{g}(x) := g(x;\theta_0) - g(x;\widehat{\theta})$ for notational convenience. Moreover,

$$\begin{aligned}
\Pi_{1n} =& \sum_{t=1}^{n} Z(x_t)'Z(x_t)\varphi^2(x_t)e_t^2 + \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)e_te_s \\
=& \sigma_e^2\sum_{t=1}^{n} Z(x_t)'Z(x_t)\varphi^2(x_t) + \sum_{t=1}^{n} Z(x_t)'Z(x_t)\varphi^2(x_t)(e_t^2 - \sigma_e^2) \\
&+ 2\sum_{t=2}^{n}\sum_{s=1}^{t-1} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)e_te_s \\
:=& \sigma_e^2\Pi_{1n}' + \Pi_{1n}'' + 2\Pi_{1n}'''.
\end{aligned}$$

It is readily seen that

$$\begin{aligned}
\frac{d_n}{nk}\Pi_{1n}' =& \frac{d_n}{n}\sum_{t=1}^{n}\frac{1}{k}Z(x_t)'Z(x_t)\varphi^2(x_t) = \frac{d_n}{n}\sum_{t=1}^{n} f_k(x_t)\varphi^2(x_t) \\
=& \frac{d_n}{n}\sum_{t=1}^{n} f(x_t)\varphi^2(x_t) + \frac{d_n}{n}\sum_{t=1}^{n}[f_k(x_t) - f(x_t)]\varphi^2(x_t)
\end{aligned}$$

$$
\begin{aligned}
=&\frac{d_n}{n}\sum_{t=1}^{n} f(d_n x_{tn})\varphi^2(d_n x_{tn}) + \frac{d_n}{n}\sum_{t=1}^{n}[f_k(x_t) - f(x_t)]\varphi^2(x_t) \\
\to_D &\int f(x)\varphi^2(x)dxL_B(1,0),
\end{aligned}
$$

by Wang and Phillips (2009a) for the first term and by showing that the second term converges to zero in probability. In fact, noting that the density functions $g_t(x)$ of x_t/d_t have a uniform bound,

$$
\begin{aligned}
&E\left|\frac{d_n}{n}\sum_{t=1}^{n}[f_k(x_t) - f(x_t)]\varphi^2(x_t)\right| \\
\le&\frac{d_n}{n}\sum_{t=1}^{n} E|f_k(x_t) - f(x_t)| \\
=&\frac{d_n}{n}\sum_{t=1}^{n}\int |f_k(d_t x) - f(d_t x)|g_t(x)dx \\
=&\frac{d_n}{n}\sum_{t=1}^{n}\frac{1}{d_t}\int |f_k(x) - f(x)|g_t(x/d_t)dx \\
\le&O(1)\int |f_k(x) - f(x)|dx \to 0
\end{aligned}
$$

by Lemma 1 as n (so that k) approaches infinity.

We are now going to show that $\frac{d_n}{nk}\Pi''_{1n} \to_P 0$ and $\frac{d_n}{nk}\Pi'''_{1n} \to_P 0$. Invoking Assumption A and the density $g_t(x)$ of x_t/d_t,

$$
\begin{aligned}
E\left(\frac{d_n}{nk}\Pi''_{1n}\right)^2 =&(\mu_4 - \sigma_e^4)\frac{d_n^2}{n^2k^2}\sum_{t=1}^{n} E[Z(x_t)'Z(x_t)\varphi(x_t)^2]^2 \\
=&O(1)\frac{d_n^2}{n^2k^2}\sum_{t=1}^{n}\int [Z(d_t x)'Z(d_t x)\varphi(d_t x)^2]^2 g_t(x)dx \\
\le&O(1)\frac{d_n^2}{n^2k^2}\sum_{t=1}^{n}\frac{1}{d_t}\int [Z(x)'Z(x)]^2\varphi(x)^4dx = O(1)\frac{1}{\sqrt{n}k} \to 0,
\end{aligned}
$$

due to Lemma A.2. Similarly,

$$
\begin{aligned}
E\left(\frac{d_n}{nk}\Pi'''_{1n}\right)^2 =&\sigma_e^4\frac{d_n^2}{n^2k^2}\sum_{t=2}^{n}\sum_{s=1}^{t-1} E\left(Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\right)^2 \\
\le&O(1)\frac{d_n^2}{n^2k^2}\sum_{t=2}^{n}\sum_{s=1}^{t-1} E(Z(x_t)'Z(x_s))^2 \\
\le&O(1)\frac{d_n^2}{n^2k^2}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{d_{ts}}\frac{1}{d_s}\iint (Z(x)'Z(y))^2dxdy = O(1)\frac{1}{k} \to 0.
\end{aligned}
$$

It therefore follows that $\frac{d_n}{nk}\Pi_{1n} \to_D \int f(x)\varphi^2(x)dxL_B(1,0)$.

We are now in a position to tackle the term Π_{3n}. For any $\delta, \epsilon > 0$, similarly to the proof of Theorem 3.1, we deduce

$$P\left(\frac{d_n}{nk}|\Pi_{3n}| > \delta\right) \leq P(\|\widehat{\theta} - \theta_0\| > \zeta_n\epsilon) + P\left(\frac{d_n}{nk}|\Pi_{3n}| > \delta, \|\widehat{\theta} - \theta_0\| \leq \zeta_n\epsilon\right)$$
$$\leq P(\|\widehat{\theta} - \theta_0\| > \zeta_n\epsilon) + \frac{d_n}{\delta nk}E[|\Pi_{3n}|I(\|\widehat{\theta} - \theta_0\| \leq \zeta_n\epsilon)].$$

As $\|\widehat{\theta} - \theta_0\| = o_P(\zeta_n)$ required in Assumption B*, $P(\|\widehat{\theta} - \theta_0\| > \zeta_n\epsilon) \to 0$. Using Taylor expansion for $g(x;\theta)$ with respect to θ in a neighborhood of θ_0, we have

$$\widehat{g}(x_t) = g(x_t;\theta_0) - g(x_t;\widehat{\theta}) = l_1(x_t;\theta_0)'(\theta_0 - \widehat{\theta}) - \frac{1}{2}(\theta_0 - \widehat{\theta})'l_2(x_t;\bar{\theta})(\theta_0 - \widehat{\theta})$$

where $\bar{\theta}$ is on the line segment connecting θ_0 and $\widehat{\theta}$. It follows that

$$\begin{aligned}
&\frac{d_n}{nk}E(|\Pi_{3n}|I(\|\widehat{\theta} - \theta_0\| \leq \zeta_n\epsilon))\\
=&\frac{d_n}{nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{g}(x_t)\widehat{g}(x_s)\varphi(x_t)\varphi(x_s)I(|\widehat{\theta} - \theta_0| \leq \zeta_n\epsilon)\\
=&\frac{d_n}{nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)l_1(x_t;\theta_0)'(\theta_0 - \widehat{\theta})\\
&\times l_1(x_s;\theta_0)'(\theta_0 - \widehat{\theta})I(\|\widehat{\theta} - \theta_0\| \leq \zeta_n\epsilon)\\
&-\frac{d_n}{nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)l_1(x_t;\theta_0)'(\theta_0 - \widehat{\theta})\\
&\times (\theta_0 - \widehat{\theta})'l_2(x_s;\bar{\theta})(\theta_0 - \widehat{\theta})I(\|\widehat{\theta} - \theta_0\| \leq \zeta_n\epsilon)\\
&+\frac{d_n}{4nk}E\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)(\theta_0 - \widehat{\theta})'l_2(x_t;\bar{\theta})(\theta_0 - \widehat{\theta})\\
&\times (\theta_0 - \widehat{\theta})'l_2(x_s;\bar{\theta})(\theta_0 - \widehat{\theta})I(\|\widehat{\theta} - \theta_0\| \leq \zeta_n\epsilon)\\
=&T_1 + T_2 + T_3, \qquad \text{say.}
\end{aligned}$$

Observe that

$$\begin{aligned}
T_1 \leq&\epsilon^2\zeta_n^2\frac{d_n}{nk}E\sum_{t=1}^{n} Z(x_t)'Z(x_t)\|l_1(x_t;\theta_0)\|^2\varphi^2(x_t)\\
&+ 2\epsilon^2\zeta_n^2\frac{d_n}{nk}E\sum_{t=2}^{n}\sum_{s=1}^{t-1}|Z(x_t)'Z(x_s)|\|l_1(x_s;\theta_0)\|\|l_1(x_t;\theta_0)\|\varphi(x_t)\varphi(x_s)\\
\leq&C\epsilon^2\zeta_n^2\frac{d_n}{nk}\sum_{t=1}^{n}\frac{1}{d_t}\int Z(x)'Z(x)\|l_1(x;\theta_0)\|^2\varphi^2(x)dx + 2C^2\epsilon^2\zeta_n^2\frac{d_n}{nk}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{d_s}\frac{1}{d_{ts}}\\
&\times\iint |Z(x)'Z(y)|\|l_1(x;\theta_0)\|\|l_1(y;\theta_0)\|\varphi(x)\varphi(y)dxdy\\
\leq&O(1)\epsilon^2\zeta_n^2\frac{1}{\sqrt{k}} + O(1)\epsilon^2\zeta_n^2\frac{\sqrt{n}}{k}\left(\int \|Z(x)\|\|l_1(x;\theta_0)\|\varphi(x)dx\right)^2
\end{aligned}$$

$$
\begin{aligned}
&\leq O(1)\epsilon^2\zeta_n^2\frac{1}{\sqrt{k}} + O(1)\epsilon^2\zeta_n^2\frac{\sqrt{n}}{k}\int \|Z(x)\|^2\varphi(x)dx \\
&= O(1)\epsilon^2\zeta_n^2\frac{1}{\sqrt{k}} + O(1)\epsilon^2\zeta_n^2\sqrt{n/k} \to 0
\end{aligned}
$$

by Assumptions B* and C*, and Lemma A.2 for $\|l_1(x;\theta_0)\|^2\varphi(x)$ is integrable.

Similarly, using Assumptions B* and C* we have

$$
\begin{aligned}
T_3 \leq & O(1)\zeta_n^4\frac{d_n}{nk}\sum_{t=1}^{n}\frac{1}{d_t}\int l^2(x)Z(x)'Z(x)\varphi^2(x)dx + O(1)\zeta_n^4\frac{d_n}{nk}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{d_t}\frac{1}{d_{ts}} \\
& \times \iint |Z(x)'Z(y)|l(x)l(y)\varphi(x)\varphi(y)dxdy \leq O(1)\zeta_n^4\frac{1}{\sqrt{k}} + O(1)\zeta_n^4\sqrt{n/k} \to 0.
\end{aligned}
$$

Notice that $T_1 = o_P(1)$ and $T_3 = o_P(1)$ imply $T_2 = o_P(1)$. Hence, it follows that $\Pi_{3n} \to_P 0$ and then from Cauchy-Schwarz inequality that $\Pi_{2n} \to_P 0$ as well. The proof is finished. □

E Proof of Theorem 3.4

Proof. We have $M(x_t) = G(x_t;\theta_1) + \varphi(x_t)\triangle_n(x_t) = G(x_t;\theta_1) + \delta_n\varphi(x_t)\triangle(x_t)$ under Hypothesis H_{11} and Assumption D. Write $\Pi_n = \Pi_{1n} + \Pi_{2n} + \Pi_{3n}$, where

$$
\begin{aligned}
\Pi_{1n} &= \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)Z(x_s)(e_t + \delta_n\triangle(x_t))(e_s + \delta_n\triangle(x_s))\varphi(x_t)\varphi(x_s), \\
\Pi_{2n} &= 2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)Z(x_s)(e_t + \delta_n\triangle(x_t))\varphi(x_t)\varphi(x_s)\widehat{g}(x_s), \\
\Pi_{3n} &= \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)Z(x_s)\varphi(x_t)\varphi(x_s)\widehat{g}(x_t)\widehat{g}(x_s),
\end{aligned}
$$

where $\widehat{g}(x) = g(x;\theta_1) - g(x;\widehat{\theta})$.

Observe that using the same argument as in the proof of Theorem 3.3 we can show that $\frac{d_n}{nk}\Pi_{3n} = o_P(1)$, and Cauchy-Schwarz inequality implies that $|\Pi_{2n}| \leq 2\sqrt{\Pi_{1n}\Pi_{3n}}$. Thus, to fulfill the proof, it suffices to show that $\frac{d_n}{nk}\Pi_{1n} \to_P \infty$. However,

$$
\begin{aligned}
\Pi_{1n} =& \delta_n^2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\triangle(x_t)\triangle(x_s)\varphi(x_t)\varphi(x_s) \\
& + 2\delta_n\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\triangle(x_t)e_s \\
& + \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)e_te_s \\
:=& \Pi_{1n}' + 2\Pi_{1n}'' + \Pi_{1n}''', \quad \text{say.}
\end{aligned}
$$

Moreover, we have $\frac{d_n}{nk}\Pi'''_{1n} \to_D \sigma_e^2 \int f(x)\varphi^2(x)dxL_B(1,0)$ looking at the proof in Theorem 3.3.

Rewrite $\Pi''_{1n} = \Pi''_{1n1} + \Pi''_{1n2} + \Pi''_{1n3}$, where

$$\begin{aligned}
\Pi''_{1n1} &= \delta_n \sum_{t=1}^{n} Z(x_t)'Z(x_t)\varphi^2(x_t)\triangle(x_t)e_t, \\
\Pi''_{1n2} &= \delta_n \sum_{t=2}^{n} e_t \sum_{s=1}^{t-1} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\triangle(x_s), \\
\Pi''_{1n3} &= \delta_n \sum_{t=1}^{n-1} e_t \sum_{s=t+1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\triangle(x_s).
\end{aligned}$$

Notice that $\frac{d_n}{nk}\Pi''_{1n1} = o_P(1)$. In fact, using Assumption A and Lemma A.2,

$$\begin{aligned}
&E\left(\frac{d_n}{nk}\Pi''_{1n1}\right)^2 = \sigma_e^2 \frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=1}^{n} E[Z(x_t)'Z(x_t)\varphi^2(x_t)\triangle(x_t)]^2 \\
\leq& O(1)\frac{d_n^2}{n^2k^2}\delta_n^2 \sum_{t=1}^{n}\frac{1}{d_t}\int [Z(x)'Z(x)\triangle(x)]^2 dx \\
\leq& O(1)\frac{1}{\sqrt{n}k^2}\delta_n^2\sqrt{k}\int Z(x)'Z(x)\triangle^2(x)dx = O(1)\frac{1}{\sqrt{nk}}\delta_n^2 \to 0.
\end{aligned}$$

Note also that

$$\begin{aligned}
&E\left(\frac{d_n}{nk}\Pi''_{1n2}\right)^2 = \sigma_e^2\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t=2}^{n} E\left(\sum_{s=1}^{t-1} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\triangle(x_s)\right)^2 \\
=&\sigma_e^2\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t=2}^{n}\sum_{s=1}^{t-1} E(Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\triangle(x_s))^2 \\
&+ 2\sigma_e^2\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t=3}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1} E(Z(x_t)'Z(x_{s_1})\varphi^2(x_t)\varphi(x_{s_1})\triangle(x_{s_1})Z(x_t)'Z(x_{s_2})\varphi(x_{s_2})\triangle(x_{s_2})) \\
\leq&\sigma_e^2 C_g^2\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{d_s d_{ts}}\iint (Z(x)'Z(y)\varphi(y)\triangle(y))^2 dxdy \\
&+ 2\sigma_e^2 C_g^3\frac{d_n^2}{n^2k^2}\delta_n^2\sum_{t=3}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1}\frac{1}{d_t d_{ts_1} d_{ts_2}} \\
&\quad\times\iiint |Z(x)'Z(y)Z(x)'Z(z)\triangle(y)\triangle(z)|\varphi(x)^2\varphi(y)\varphi(z)dxdydz \\
=&O(1)\frac{1}{k^2}\delta_n^2\int \|Z(y)\|^2(\varphi(y)\triangle(y))^2dy + O(1)\frac{\sqrt{n}}{k^2}\delta_n^2\int\varphi(x)^2dx\left(\int |Z(x)'Z(y)\triangle(y)|\varphi(y)dy\right)^2 \\
\leq&O(1)\frac{1}{k^{3/2}}\delta_n^2 + O(1)\frac{\sqrt{n}}{k^2}\delta_n^2\int\varphi(x)^2dx\int (Z(x)'Z(y))^2\varphi(y)dy \\
\leq&O(1)\frac{1}{k^{3/2}}\delta_n^2 + O(1)\frac{\sqrt{n}}{k^2}\delta_n^2\int\|Z(x)\|^2\varphi(x)^2dx\int\|Z(y)\|^2\varphi(y)dy
\end{aligned}$$

$$\leq O(1)\frac{1}{k^{3/2}}\delta_n^2 + O(1)\frac{\sqrt{n}}{k}\delta_n^2.$$

Thus, $\frac{d_n}{nk}\Pi''_{1n2} = O_P(n^{1/4}k^{-1/2}\delta_n)$. Similar calculation yields the same result for $\frac{d_n}{nk}\Pi''_{1n3}$. Therefore, $\frac{d_n}{nk}\Pi''_{1n} = O_P(n^{1/4}k^{-1/2}\delta_n)$.

Lastly,

$$\begin{aligned}
\frac{d_n}{nk}\Pi'_{1n} &= \delta_n^2\frac{d_n}{nk}\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\varphi(x_t)\varphi(x_s)\triangle(x_t)\triangle(x_s)\\
&=\delta_n^2\frac{d_n}{nk}\sum_{i=0}^{k-1}\left(\sum_{t=1}^{n}\mathscr{H}_i(x_t)\triangle(x_t)\varphi(x_t)\right)^2 \geq \delta_n^2\frac{d_n}{nk}\left(\sum_{t=1}^{n}\mathscr{H}_{i_0}(x_t)\triangle(x_t)\varphi(x_t)\right)^2\\
&=\delta_n^2\frac{\sqrt{n}}{|\psi|k}\left(\frac{d_n}{n}\sum_{t=1}^{n}\mathscr{H}_{i_0}(x_t)\triangle(x_t)\varphi(x_t)\right)^2 \to_P \infty,
\end{aligned}$$

due to Assumption D and the convergence of the square term, where i_0 is the first nonnegative integer such that $\int \mathscr{H}_{i_0}(x)\triangle(x)\varphi(x) \neq 0$, which certainly exists since otherwise $\triangle(x)\varphi(x)$ and therefore $\triangle(x)$ become zero function.

□

F Proof of Theorem 4.1

Proof. (1) Denote $G_t^* = g(x_t;\widehat{\theta}) - g(x_t;\widehat{\theta}^*)$ and hence $y_t^* - g(x_t;\widehat{\theta}^*) = G_t^* + e_t^*$.

$$\begin{aligned}
L_n^* &= \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)(y_t^* - g(x_t;\widehat{\theta}^*))(y_s^* - g(x_s;\widehat{\theta}^*))\\
&= \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)(G_t^* + e_t^*)(G_s^* + e_s^*)\\
&= \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_t^*e_s^* + 2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)G_t^*e_s^*\\
&\quad + \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)G_t^*G_s^*\\
&:= L_{1n}^* + 2L_{2n}^* + L_{3n}^*, \qquad \text{say.}
\end{aligned}$$

Notice that

$$\begin{aligned}
L_{1n}^* &= \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_t^*e_s^*\\
&= \sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)\widehat{e}_t\widehat{e}_s\eta_t\eta_s
\end{aligned}$$

$$
\begin{aligned}
=&\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)(e_t + G(x_t;\theta_0,\widehat{\theta}))(e_s + G(x_s;\theta_0,\widehat{\theta}))\eta_t\eta_s \\
=&\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)e_te_s\eta_t\eta_s \\
&+\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)G(x_t;\theta_0,\widehat{\theta})G(x_s;\theta_0,\widehat{\theta})\eta_t\eta_s \\
&+2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)G(x_t;\theta_0,\widehat{\theta})e_s\eta_t\eta_s \\
=&\sum_{t=1}^{n} Z(x_t)'Z(x_t)e_t^2 + \sum_{t=1}^{n} Z(x_t)'Z(x_t)e_t^2(\eta_t^2-1) \\
&+\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} Z(x_t)'Z(x_s)e_te_s\eta_t\eta_s \\
&+\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)G(x_t;\theta_0,\widehat{\theta})G(x_s;\theta_0,\widehat{\theta})\eta_t\eta_s \\
&+2\sum_{t=1}^{n}\sum_{s=1}^{n} Z(x_t)'Z(x_s)G(x_t;\theta_0,\widehat{\theta})e_s\eta_t\eta_s
\end{aligned}
$$

where, as defined before, $G(x_t;\theta_0,\widehat{\theta}) = g(x_t;\theta_0) - g(x_t;\widehat{\theta})$.

Let $F(x)$ be the distribution function of $\sigma_e^2 L_B(1,0)$ and l_α be the $1-\alpha$ quantile of $F(x)$, that is, $F(l_\alpha) = 1-\alpha$. It follows from Theorem 1 and the properties of the sequence η_t that as $n \to \infty$

$$P^*\left(\frac{d_n}{nk}L_n^* < x\right) \to F(x) \tag{F.1}$$

holds for any $x \in \mathbb{R}$ in probability with respect to the distribution of the sample $\mathcal{W}_n$. Hence,

$$P^*\left(\frac{d_n}{nk}L_n^* > l_\alpha\right) \to 1 - F(l_\alpha) = \alpha, \qquad \text{in probability} \tag{F.2}$$

which, together with $P^*\left(\frac{d_n}{nk}L_n^* > l_\alpha^*\right) = \alpha$ by definition, implies that $l_\alpha^* \to_P l_\alpha$.

Note that the result of Theorem 1 and (F.1) are tantamount to that as $n \to \infty$

$$P^*\left(\frac{d_n}{nk}L_n^* < x\right) - P\left(\frac{d_n}{nk}L_n < x\right) \to_P 0, \quad \forall\, x \in \mathbb{R}. \tag{F.3}$$

Recalling the definition of l_α^* again, (F.3) indicates $\lim_{n\to\infty} P\left(\frac{d_n}{nk}L_n > l_\alpha^*\right) = \alpha$, as required.

(2) This part is the implication of Theorem 3.2.

(3) The proof is similar to that of (1), so we omit for brevity.

(4) It is the consequence of Theorem 3.4. □

References

Assche, W. V. (1987). *Asymptotics for Orthogonal Polynomials.* Lecture Notes in Mathematics. Springer-Verlag, Berlin.

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sons, Inc., New York.

Cai, Z., Li, Q., and Park, J. (2009). Functional-coefficient cointegration models for nonstationary time series data. *Journal of Econometrics*, 148:101–113.

Dong, C. and Gao, J. (2011). Orthogonal expansion of functionals of lévy processes and its application in statistical estimation. Working paper available at http://www.buseco.monash.edu.au/ebs/pubs/wpapers/2012/wp2-12.pdf.

Edelman, A. and Rao, N. R. (2005). Random matrix theory. *Acta Numerica*, Cambridge University Press:1–65.

Gao, J. (2007). *Nonlinear Time Series: Semiparametric and Nonparametric Methods.* Monographs on statistics and applied probability. Chapman & Hall, New York.

Gao, J., King, M., Lu, Z., and Tjøstheim, D. (2009a). Nonparametric specification testing for nonlinear time series with nonstationarity. *Econometric Theory*, 25:1869–1892.

Gao, J., King, M., Lu, Z., and Tjøstheim, D. (2009b). Specification testing in nonlinear and nonstationary time series autoregression. *Annals of Statistics*, 37(68):3893–3928.

Gao, J., Tjøstheim, D., and Yin, J. (2012). Parametric model specification testing between parametric and nonparametric co–integrating models. Working paper at http://www.buseco.monash.edu.au/ebs/pubs/wpapers/2012/wp18-12.pdf.

Gao, J., Tong, H., and Wolff, R. (2002). Model specification tests in nonparametric stochastic regression models. *Journal of Multivariate Analysis*, 83:324–359.

Gao, J., Wang, Q., and Yin, J. (2011). Specification testing in nonlinear time series with long-range dependence. *Econometric Theory*, 27:260–284.

Hong, S. H. and Phillips, P. C. B. (2010). Testing linearity of cointegriting relations with an application to purchasing power parity. *Journal of Business and Economic Statistics*, 28(1):96–114.

Karlsen, H. A., Mykelbust, T., and Tjøstheim, D. (2007). Nonparametric estimation in a nonlinear cointegration type model. *Annals of Statistics*, 35:252–299.

Karlsen, H. A. and Tjøstheim, D. (2001). Nonparametric estimation in null recurrent time series. *Annals of Statistics*, 29:372–416.

Kasparis, I. (2008). Detection of functional form misspecification in cointegrating relations. *Econometric Theory*, 24:1373–1403.

Kasparis, I. (2010). The Bierens test for certain nonstationary models. *Journal of Econometrics*, 158(2):221–230.

Nevai, P. (1986). Geza Frued, orthogonal polynomials and Christoffel functions. A case study. *Journal of Approximation Theory*, 48:3–167.

Park, J. Y. and Phillips, P. C. B. (1999). Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory*, 15:269–298.

Park, J. Y. and Phillips, P. C. B. (2001). Nonlinear regression with integreted time series. *Econometrica*, 69(1):117–161.

Phillips, P. C. B. and Magdalinos, T. (2009). Unit root and cointegrating limit theory when initialization is in the infinite past. *Econometric Theory*, 25:1682–1715.

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion.* A series of comprehensive studies in mathematics 293. Springer-Verlag.

Shi, X. and Phillips, P. C. B. (2012). Nonlinear cointegrating regression under weak identification. *Econometric Theory*, 28:509–547.

Wang, Q., Lin, Y. X., and Gulati, C. M. (2003). Asymptotics for general fractionally integrated processes with applications to unit root tests. *Econometric Theory*, 19:143–164.

Wang, Q. and Phillips, P. C. B. (2009a). Asymptotic theory for local time density estimation and nonparametric cointegreting regression. *Econometric Theory*, 25:710–738.

Wang, Q. and Phillips, P. C. B. (2009b). Structure nonparametric cointegrating regression. *Econometrica*, 77:1901–1948.

Wang, Q. and Phillips, P. C. B. (2011). Asymptotic theory for zero energy functionals with nonparametric regression applications. *Econometric Theory*, 27:235–259.

Wang, Q. and Phillips, P. C. B. (2012). A specification test for nonlinear nonstationary models. *Annals of Statistics*, 40(2):727–758.

Wigner, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327.

Table 1: Size: $g(x;\theta_0) = \exp(-\theta_0 x^2)$, $\psi_1 = 1$, $\psi_2 = \psi_3 = 0$

n		Nominal size 1%			Nominal size 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.0100	0.0190	0.0110	0.0550	0.0540	0.0540
500		0.0070	0.0130	0.0120	0.0500	0.0760	0.0480
800		0.0140	0.0090	0.0070	0.0490	0.0500	0.0480
				$\rho = -0.75$			
300		0.0160	0.0130	0.0130	0.0680	0.0550	0.0470
500		0.0200	0.0140	0.0110	0.0610	0.0600	0.0560
800		0.0050	0.0110	0.0070	0.0400	0.0570	0.0460
				$\rho = 0.50$			
300		0.0100	0.0070	0.0150	0.0480	0.0530	0.0540
500		0.0070	0.0120	0.0110	0.0540	0.0590	0.0420
800		0.0150	0.0120	0.0120	0.0650	0.0420	0.0450
				$\rho = -0.50$			
300		0.0140	0.0070	0.0100	0.0420	0.0410	0.0410
500		0.0150	0.0100	0.0140	0.0480	0.0490	0.0400
800		0.0150	0.0140	0.0060	0.0500	0.0600	0.0510
				$\rho = 0.25$			
300		0.0130	0.0090	0.0150	0.0570	0.0570	0.0500
500		0.0070	0.0140	0.0120	0.0390	0.0530	0.0480
800		0.0120	0.0160	0.0130	0.0400	0.0500	0.0430
				$\rho = -0.25$			
300		0.0130	0.0120	0.0090	0.0460	0.0500	0.0500
500		0.0070	0.0090	0.0200	0.0680	0.0530	0.0560
800		0.0090	0.0090	0.0170	0.0450	0.0430	0.0420
				$\rho = 0$			
300		0.0120	0.0130	0.0110	0.0460	0.0560	0.0470
500		0.0060	0.0170	0.0080	0.0400	0.0620	0.0470
800		0.0150	0.0080	0.0090	0.0540	0.0360	0.0510

$u_t = \psi_1 \varepsilon_t + \psi_2 \varepsilon_{t-1} + \psi_3 \varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$

Table 2: Power: $g(x;\theta_0) = \exp(-\theta_0 x^2) + \triangle_n(x)$, $\psi_1 = 1$, $\psi_2 = \psi_3 = 0$

n		Level 1%			Level 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.3880	0.4390	0.5110	0.5760	0.6130	0.6650
500		0.4310	0.4570	0.5640	0.6100	0.6270	0.7090
800		0.4230	0.5340	0.6050	0.6120	0.6710	0.7230
				$\rho = -0.75$			
300		0.3890	0.4420	0.5290	0.5830	0.6100	0.6600
500		0.4030	0.4460	0.5680	0.5840	0.6290	0.7020
800		0.4270	0.5180	0.5960	0.6080	0.6820	0.7150
				$\rho = 0.50$			
300		0.3590	0.4240	0.5060	0.5590	0.6130	0.6640
500		0.3780	0.4370	0.5400	0.5620	0.5910	0.7010
800		0.4200	0.5290	0.6280	0.6100	0.6760	0.7560
				$\rho = -0.5$			
300		0.3830	0.4220	0.4990	0.5590	0.5700	0.6640
500		0.3900	0.4520	0.5410	0.5850	0.6260	0.6850
800		0.4260	0.5250	0.6150	0.5810	0.6800	0.7190
				$\rho = 0.25$			
300		0.4060	0.4310	0.4720	0.5900	0.6020	0.6310
500		0.3980	0.4460	0.5580	0.6020	0.6110	0.7060
800		0.4000	0.4930	0.6400	0.5750	0.6580	0.7490
				$\rho = -0.25$			
300		0.3950	0.4180	0.4990	0.5850	0.6180	0.6450
500		0.4060	0.4370	0.5590	0.5970	0.6170	0.7000
800		0.4200	0.5220	0.6200	0.6080	0.6670	0.7360
				$\rho = 0$			
300		0.3830	0.4260	0.5080	0.5730	0.5890	0.6630
500		0.4110	0.4550	0.5370	0.5870	0.6070	0.7020
800		0.3970	0.4790	0.6010	0.5740	0.6350	0.7170

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$ and $\triangle_n(x) = \delta_n/(1+x^2)$ in which $\delta_n = \frac{1}{2}\sqrt{k\log(n)/\sqrt{n}}$

Table 3: Size: $g(x;\theta_0) = \exp(-\theta_0 x^2)$, $\psi_1 = -0.01$, $\psi_2 = 0.3$, $\psi_3 = 0.5$

n		Nominal size 1%			Nominal size 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.0130	0.0130	0.0160	0.0650	0.0540	0.0540
500		0.0090	0.0190	0.0070	0.0430	0.0540	0.0360
800		0.0140	0.0110	0.0140	0.0450	0.0610	0.0500
				$\rho = -0.75$			
300		0.0110	0.0180	0.0080	0.0520	0.0480	0.0550
500		0.0090	0.0140	0.0120	0.0410	0.0590	0.0440
800		0.0130	0.0130	0.0070	0.0590	0.0630	0.0430
				$\rho = 0.50$			
300		0.0120	0.0090	0.0130	0.0470	0.0500	0.0470
500		0.0090	0.0120	0.0110	0.0490	0.0550	0.0480
800		0.0110	0.0110	0.0130	0.0530	0.0580	0.0510
				$\rho = -0.50$			
300		0.0090	0.0120	0.0100	0.0530	0.0620	0.0560
500		0.0080	0.0150	0.0110	0.0410	0.0560	0.0440
800		0.0130	0.0080	0.0110	0.0580	0.0450	0.0460
				$\rho = 0.25$			
300		0.0060	0.0150	0.0100	0.0480	0.0490	0.0430
500		0.0070	0.0090	0.0080	0.0390	0.0600	0.0360
800		0.0110	0.0140	0.0140	0.0490	0.0520	0.0450
				$\rho = -0.25$			
300		0.0110	0.0140	0.0130	0.0410	0.0560	0.0480
500		0.0170	0.0100	0.0130	0.0620	0.0590	0.0550
800		0.0130	0.0090	0.0100	0.0550	0.0620	0.0540
				$\rho = 0$			
300		0.0110	0.0140	0.0150	0.0420	0.0520	0.0500
500		0.0110	0.0090	0.0120	0.0530	0.0570	0.0500
800		0.0110	0.0110	0.0130	0.0490	0.0600	0.0450

$u_t = \psi_1 \varepsilon_t + \psi_2 \varepsilon_{t-1} + \psi_3 \varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \text{cov}(e_t, \varepsilon_t)$

Table 4: Power: $g(x;\theta_0) = \exp(-\theta_0 x^2) + \triangle_n(x)$, $\psi_1 = -0.01$, $\psi_2 = 0.3, \psi_3 = 0.5$

n		Level 1%			Level 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.4870	0.5480	0.6170	0.6850	0.7110	0.7530
500		0.5060	0.5460	0.6510	0.6690	0.6990	0.7680
800		0.5290	0.6140	0.7030	0.7090	0.7540	0.7960
				$\rho = -0.75$			
300		0.5180	0.5580	0.6250	0.6770	0.7090	0.7640
500		0.5280	0.5610	0.6590	0.7020	0.7110	0.7870
800		0.5310	0.6130	0.7160	0.7060	0.7610	0.8170
				$\rho = 0.50$			
300		0.4720	0.5420	0.6050	0.6850	0.7100	0.7520
500		0.5300	0.5620	0.6840	0.7030	0.7200	0.8120
800		0.5430	0.6380	0.6980	0.7020	0.7740	0.7980
				$\rho = -0.5$			
300		0.4850	0.5590	0.5970	0.6780	0.7080	0.7260
500		0.5270	0.5590	0.6770	0.6820	0.7010	0.7860
800		0.5370	0.6300	0.7040	0.6860	0.7740	0.7910
				$\rho = 0.25$			
300		0.5150	0.5590	0.6270	0.6730	0.7210	0.7440
500		0.5440	0.5740	0.6710	0.7150	0.7190	0.8070
800		0.5550	0.6180	0.7010	0.7090	0.7490	0.7950
				$\rho = -0.25$			
300		0.4850	0.5410	0.5960	0.6650	0.7180	0.7490
500		0.5340	0.5540	0.6860	0.7080	0.7220	0.7990
800		0.5470	0.6170	0.7060	0.7220	0.7530	0.8220
				$\rho = 0$			
300		0.5020	0.5690	0.6250	0.6880	0.7260	0.7500
500		0.5050	0.5700	0.6490	0.6800	0.7000	0.7750
800		0.5670	0.6240	0.7130	0.7170	0.7770	0.8120

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$ and $\triangle_n(x) = \delta_n/(1+x^2)$ in which $\delta_n = \frac{1}{2}\sqrt{k\log(n)/\sqrt{n}}$

Table 5: Size of $\widehat{L}_{bn}$: $g(x;\theta_0) = \exp(-\theta_0 x^2)$, $\psi_1 = 1$, $\psi_2 = \psi_3 = 0$

n		Nominal size 1%			Nominal size 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.0140	0.0180	0.0100	0.0580	0.0590	0.0480
500		0.0130	0.0150	0.0200	0.0520	0.0530	0.0510
800		0.0110	0.0110	0.0120	0.0480	0.0460	0.0450
				$\rho = -0.75$			
300		0.0110	0.0120	0.0080	0.0590	0.0510	0.0380
500		0.0130	0.0180	0.0180	0.0580	0.0550	0.0750
800		0.0140	0.0150	0.0260	0.0620	0.0600	0.0570
				$\rho = 0$			
300		0.0120	0.0100	0.0110	0.0410	0.0470	0.0400
500		0.0160	0.0120	0.0190	0.0580	0.0540	0.0540
800		0.0130	0.0080	0.0150	0.0550	0.0360	0.0560

Table 6: Power of $\widehat{L}_{bn}$: $g(x;\theta_0) = \exp(-\theta_0 x^2) + \triangle_n(x)$, $\psi_1 = 1$, $\psi_2 = \psi_3 = 0$

n		Level 1%			Level 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.3470	0.3810	0.4400	0.5440	0.5490	0.5930
500		0.3670	0.3880	0.4910	0.5610	0.5560	0.6460
800		0.3860	0.4900	0.5580	0.5850	0.6440	0.6880
				$\rho = -0.75$			
300		0.3490	0.3690	0.4250	0.5250	0.5600	0.5870
500		0.3830	0.4130	0.5130	0.5720	0.5840	0.6500
800		0.3960	0.4910	0.5910	0.5640	0.6430	0.7150
				$\rho = 0$			
300		0.3330	0.3860	0.4360	0.5200	0.5480	0.6050
500		0.3590	0.3980	0.4900	0.5390	0.5840	0.6370
800		0.3900	0.5040	0.5770	0.5800	0.6450	0.7020

Table 7: Size: $g(x;\theta_0) = \theta_0 x^2$, $\psi_1 = 0.1$, $\psi_2 = -0.01$ and $\psi_3 = 0.8$

n		Level 1%			Level 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.0060	0.0040	0.0040	0.0420	0.0390	0.0330
500		0.0040	0.0070	0.0120	0.0390	0.0440	0.0460
800		0.0090	0.0190	0.0070	0.0440	0.0550	0.0460
				$\rho = -0.75$			
300		0.0090	0.0020	0.0060	0.0440	0.0330	0.0420
500		0.0080	0.0060	0.0110	0.0380	0.0560	0.0420
800		0.0040	0.0070	0.0060	0.0380	0.0430	0.0540
				$\rho = 0.50$			
300		0.0110	0.0090	0.0100	0.0490	0.0420	0.0480
500		0.0110	0.0040	0.0090	0.0460	0.0470	0.0590
800		0.0100	0.0110	0.0090	0.0470	0.0500	0.0440
				$\rho = -0.5$			
300		0.0090	0.0090	0.0030	0.0440	0.0530	0.0460
500		0.0100	0.0100	0.0130	0.0470	0.0530	0.0430
800		0.0080	0.0080	0.0100	0.0540	0.0410	0.0470
				$\rho = 0.25$			
300		0.0040	0.0070	0.0080	0.0470	0.0460	0.0430
500		0.0180	0.0060	0.0140	0.0660	0.0380	0.0500
800		0.0130	0.0120	0.0100	0.0530	0.0450	0.0520
				$\rho = -0.25$			
300		0.0120	0.0120	0.0130	0.0540	0.0550	0.0480
500		0.0120	0.0110	0.0120	0.0630	0.0540	0.0510
800		0.0170	0.0120	0.0120	0.0600	0.0570	0.0500
				$\rho = 0$			
300		0.0060	0.0130	0.0160	0.0410	0.0530	0.0510
500		0.0120	0.0120	0.0090	0.0570	0.0490	0.0500
800		0.0060	0.0130	0.0080	0.0370	0.0470	0.0380

$u_t = \psi_1 \varepsilon_t + \psi_2 \varepsilon_{t-1} + \psi_3 \varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \text{cov}(e_t, \varepsilon_t)$

Table 8: Power: $g(x;\theta_1) = \theta_1 x^2 + \triangle_n(x)$, $\psi_1 = 0.1$, $\psi_2 = -0.01$ and $\psi_3 = 0.8$

n		Nominal size 1%			Nominal size 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.4510	0.5210	0.6710	0.6120	0.6560	0.7850
500		0.4540	0.5210	0.7170	0.6110	0.6560	0.8110
800		0.4700	0.6220	0.7640	0.6130	0.7390	0.8260
				$\rho = -0.75$			
300		0.4270	0.4930	0.6350	0.5800	0.6310	0.7480
500		0.4440	0.5360	0.6610	0.5840	0.6730	0.7710
800		0.4730	0.6240	0.7790	0.6180	0.7430	0.8610
				$\rho = 0.50$			
300		0.4550	0.5250	0.6500	0.6100	0.6550	0.7750
500		0.4650	0.5500	0.7480	0.5970	0.6660	0.8250
800		0.4820	0.6070	0.7980	0.6310	0.7220	0.8620
				$\rho = -0.50$			
300		0.4370	0.5310	0.6460	0.6000	0.6510	0.7710
500		0.4610	0.5420	0.6890	0.6040	0.6770	0.7860
800		0.4590	0.6080	0.7740	0.6070	0.7340	0.8520
				$\rho = 0.25$			
300		0.4090	0.5150	0.6430	0.5670	0.6560	0.7470
500		0.4370	0.5480	0.6850	0.5770	0.6800	0.7790
800		0.4680	0.6330	0.7800	0.6150	0.7520	0.8460
				$\rho = -0.25$			
300		0.4220	0.5510	0.6420	0.5600	0.6820	0.7670
500		0.4320	0.5470	0.6880	0.5950	0.6640	0.7850
800		0.4630	0.5810	0.7710	0.6140	0.7290	0.8610
				$\rho = 0$			
300		0.4320	0.5360	0.6580	0.5750	0.6660	0.7700
500		0.4590	0.5440	0.7000	0.6170	0.6700	0.8050
800		0.4710	0.6450	0.8020	0.6310	0.7600	0.8770

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$, $\triangle_n(x) = \delta_n/(1+x^2)$ in which $\delta_n = \frac{1}{2}\sqrt{k\log(n)/\sqrt{n}}$.

Table 9: Size and Power at 10%: $\psi_1 = 0.1$, $\psi_2 = -0.01$ and $\psi_3 = 0.8$

n		Size at 10%			Power at 10%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.1080	0.0970	0.0980	0.6820	0.7280	0.8280
500		0.0830	0.0930	0.1100	0.6920	0.7440	0.8530
800		0.0890	0.1000	0.1150	0.6910	0.7910	0.8650
				$\rho = -0.75$			
300		0.0970	0.0880	0.0970	0.6630	0.7030	0.8070
500		0.0880	0.1120	0.1070	0.6660	0.7300	0.8070
800		0.1000	0.0930	0.1060	0.7050	0.7970	0.9000
				$\rho = 0.50$			
300		0.1110	0.0970	0.1070	0.6840	0.7360	0.8310
500		0.0960	0.1000	0.1080	0.6820	0.7270	0.8720
800		0.1060	0.1200	0.1000	0.7050	0.7960	0.8860
				$\rho = -0.50$			
300		0.1000	0.1160	0.1130	0.6920	0.7230	0.8320
500		0.1050	0.1110	0.0940	0.6830	0.7390	0.8300
800		0.1200	0.0940	0.1080	0.6980	0.8030	0.8930
				$\rho = 0.25$			
300		0.1000	0.0950	0.1070	0.6600	0.7360	0.8130
500		0.1250	0.0970	0.1030	0.6690	0.7400	0.8410
800		0.1090	0.0970	0.1010	0.6970	0.7950	0.8870
				$\rho = -0.25$			
300		0.1000	0.1030	0.1060	0.6420	0.7420	0.8180
500		0.1210	0.1110	0.0930	0.6830	0.7290	0.8390
800		0.1120	0.1110	0.1050	0.6790	0.7880	0.9040
				$\rho = 0$			
300		0.0960	0.1040	0.1060	0.6730	0.7410	0.8240
500		0.1090	0.1170	0.1180	0.6990	0.7450	0.8590
800		0.0910	0.1090	0.0880	0.7030	0.8070	0.9070

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$.

Table 10: Size: $g(x;\theta_0) = \theta_0 x^2$, $\psi_1 = 0.01$, $\psi_2 = 0.3, \psi_3 = 0.5$

n		Level 1%			Level 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.0080	0.0110	0.0060	0.0580	0.0420	0.0480
500		0.0060	0.0170	0.0100	0.0500	0.0560	0.0580
800		0.0090	0.0040	0.0080	0.0510	0.0390	0.0380
				$\rho = -0.75$			
300		0.0150	0.0090	0.0100	0.0510	0.0490	0.0500
500		0.0150	0.0080	0.0090	0.0620	0.0460	0.0440
800		0.0130	0.0110	0.0110	0.0680	0.0530	0.0530
				$\rho = 0.50$			
300		0.0120	0.0090	0.0070	0.0440	0.0450	0.0350
500		0.0090	0.0120	0.0030	0.0630	0.0510	0.0500
800		0.0080	0.0100	0.0120	0.0520	0.0590	0.0500
				$\rho = -0.5$			
300		0.0180	0.0090	0.0060	0.0570	0.0470	0.0490
500		0.0040	0.0120	0.0100	0.0370	0.0480	0.0510
800		0.0140	0.0180	0.0120	0.0580	0.0530	0.0500
				$\rho = 0.25$			
300		0.0140	0.0140	0.0050	0.0410	0.0580	0.0380
500		0.0110	0.0140	0.0080	0.0450	0.0490	0.0480
800		0.0090	0.0100	0.0120	0.0530	0.0540	0.0440
				$\rho = -0.25$			
300		0.0090	0.0090	0.0090	0.0450	0.0470	0.0570
500		0.0070	0.0070	0.0080	0.0430	0.0360	0.0490
800		0.0150	0.0100	0.0070	0.0400	0.0520	0.0400
				$\rho = 0$			
300		0.0110	0.0170	0.0050	0.0480	0.0550	0.0380
500		0.0160	0.0050	0.0060	0.0520	0.0390	0.0540
800		0.0110	0.0150	0.0060	0.0520	0.0570	0.0470

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$

Table 11: Power: $g(x;\theta_1) = \theta_1 x^2 + \triangle_n(x)$, $\psi_1 = 0.01$, $\psi_2 = 0.3, \psi_3 = 0.5$

n		Nominal size 1%			Nominal size 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.4670	0.5760	0.6640	0.6090	0.7150	0.7620
500		0.4880	0.5750	0.6900	0.6280	0.6990	0.7810
800		0.5150	0.6480	0.7810	0.6510	0.7640	0.8540
				$\rho = -0.75$			
300		0.5060	0.5480	0.6690	0.6410	0.6910	0.7610
500		0.4830	0.5620	0.7060	0.6270	0.6970	0.7920
800		0.4960	0.6460	0.7660	0.6320	0.7460	0.8380
				$\rho = 0.50$			
300		0.4490	0.5250	0.6630	0.5850	0.6510	0.7630
500		0.4730	0.5730	0.7160	0.6100	0.6930	0.8000
800		0.4950	0.6230	0.7690	0.6240	0.7520	0.8440
				$\rho = -0.50$			
300		0.4550	0.5310	0.6770	0.5940	0.6700	0.7710
500		0.4560	0.5590	0.7110	0.6070	0.6860	0.7960
800		0.4580	0.6410	0.7730	0.6150	0.7590	0.8490
				$\rho = 0.25$			
300		0.4320	0.5660	0.6690	0.5910	0.7060	0.7670
500		0.4660	0.5700	0.7390	0.6130	0.6970	0.8270
800		0.4940	0.6270	0.7930	0.6480	0.7450	0.8530
				$\rho = -0.25$			
300		0.4520	0.5390	0.6500	0.5860	0.6580	0.7580
500		0.4600	0.5790	0.7140	0.5950	0.7060	0.8040
800		0.4690	0.6620	0.7780	0.6170	0.7670	0.8540
				$\rho = 0$			
300		0.4640	0.5280	0.6740	0.6190	0.6580	0.7810
500		0.4520	0.5360	0.7020	0.6210	0.6740	0.8100
800		0.4950	0.6330	0.7860	0.6480	0.7360	0.8680

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$, $\triangle_n(x) = \delta_n/(1+x^2)$ in which $\delta_n = \frac{1}{2}\sqrt{k\log(n)/\sqrt{n}}$.

Table 12: Size and Power at 10%: $\psi_1 = 0.01$, $\psi_2 = 0.3, \psi_3 = 0.5$

n		Size at 10%			Power at 10%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.1150	0.0920	0.1010	0.6820	0.7580	0.8150
500		0.1210	0.1180	0.1130	0.6930	0.7640	0.8230
800		0.1120	0.1000	0.0800	0.7110	0.8210	0.8940
				$\rho = -0.75$			
300		0.1140	0.1050	0.1120	0.7220	0.7540	0.8020
500		0.1250	0.1010	0.0990	0.6980	0.7730	0.8390
800		0.1240	0.1120	0.1090	0.7090	0.8000	0.8690
				$\rho = 0.50$			
300		0.1100	0.1090	0.0990	0.6700	0.7250	0.8140
500		0.1240	0.1130	0.1170	0.6890	0.7480	0.8320
800		0.1030	0.1100	0.0960	0.7070	0.7980	0.8770
				$\rho = -0.50$			
300		0.1100	0.1110	0.1050	0.6810	0.7330	0.8200
500		0.0930	0.1030	0.1010	0.6960	0.7490	0.8420
800		0.1100	0.1140	0.0920	0.6980	0.8190	0.8920
				$\rho = 0.25$			
300		0.0960	0.1150	0.0920	0.6790	0.7760	0.8030
500		0.0990	0.1130	0.1120	0.6960	0.7730	0.8700
800		0.1050	0.1180	0.0980	0.7100	0.8020	0.8770
				$\rho = -0.25$			
300		0.1020	0.0970	0.1110	0.6830	0.7340	0.8090
500		0.1040	0.0960	0.1080	0.6800	0.7760	0.8440
800		0.0990	0.1070	0.0960	0.6960	0.8210	0.8830
				$\rho = 0$			
300		0.0930	0.1100	0.1050	0.6950	0.7360	0.8310
500		0.0990	0.1070	0.0960	0.7050	0.7480	0.8530
800		0.0990	0.1140	0.0930	0.7130	0.7930	0.8950

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$.

Table 13: Size: $g(x;\theta_0) = \theta_0 x^2$, $\psi_1 = -0.01$, $\psi_2 = 0.3$ and $\psi_3 = 0.1$

n		Level 1%			Level 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.0160	0.0120	0.0090	0.0710	0.0590	0.0540
500		0.0150	0.0190	0.0130	0.0670	0.0720	0.0570
800		0.0110	0.0130	0.0050	0.0650	0.0600	0.0420
				$\rho = -0.75$			
300		0.0120	0.0080	0.0050	0.0570	0.0510	0.0470
500		0.0100	0.0160	0.0180	0.0580	0.0670	0.0680
800		0.0160	0.0140	0.0060	0.0520	0.0570	0.0510
				$\rho = 0.50$			
300		0.0100	0.0120	0.0120	0.0580	0.0490	0.0640
500		0.0120	0.0150	0.0140	0.0570	0.0710	0.0450
800		0.0150	0.0100	0.0090	0.0660	0.0620	0.0490
				$\rho = -0.5$			
300		0.0150	0.0070	0.0100	0.0590	0.0570	0.0660
500		0.0150	0.0120	0.0150	0.0570	0.0520	0.0470
800		0.0080	0.0110	0.0080	0.0450	0.0610	0.0400
				$\rho = 0.25$			
300		0.0080	0.0030	0.0080	0.0470	0.0340	0.0480
500		0.0080	0.0090	0.0120	0.0450	0.0580	0.0480
800		0.0140	0.0110	0.0150	0.0490	0.0520	0.0410
				$\rho = -0.25$			
300		0.0110	0.0090	0.0120	0.0480	0.0500	0.0550
500		0.0160	0.0160	0.0080	0.0650	0.0500	0.0490
800		0.0120	0.0030	0.0160	0.0430	0.0470	0.0550
				$\rho = 0$			
300		0.0110	0.0070	0.0070	0.0620	0.0490	0.0460
500		0.0080	0.0070	0.0150	0.0520	0.0550	0.0580
800		0.0100	0.0150	0.0050	0.0470	0.0520	0.0400

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$

Table 14: Power: $g(x;\theta_1) = \theta_1 x^2 + \triangle_n(x)$, $\psi_1 = -0.01$, $\psi_2 = 0.3$ and $\psi_3 = 0.1$

n		Nominal size 1%			Nominal size 5%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.4660	0.5650	0.6620	0.5900	0.7000	0.7560
500		0.5060	0.5470	0.7160	0.6420	0.6580	0.7980
800		0.4830	0.6340	0.7820	0.6330	0.7400	0.8440
				$\rho = -0.75$			
300		0.4990	0.5440	0.6490	0.6230	0.6930	0.7710
500		0.4890	0.5940	0.7370	0.6440	0.6900	0.8160
800		0.4630	0.6470	0.7420	0.6140	0.7500	0.8360
				$\rho = 0.50$			
300		0.4810	0.5650	0.6800	0.6320	0.6920	0.7810
500		0.4930	0.5840	0.7090	0.6070	0.7040	0.8110
800		0.5010	0.6440	0.7970	0.6470	0.7500	0.8600
				$\rho = -0.50$			
300		0.5030	0.5380	0.6500	0.6290	0.6730	0.7670
500		0.5100	0.5980	0.7080	0.6300	0.7230	0.8050
800		0.5380	0.6570	0.7930	0.6670	0.7610	0.8400
				$\rho = 0.25$			
300		0.4610	0.5530	0.6740	0.5990	0.6750	0.7830
500		0.4640	0.5570	0.7120	0.6230	0.7000	0.7930
800		0.4880	0.6430	0.8060	0.8200	0.7530	0.8660
				$\rho = -0.25$			
300		0.4740	0.5580	0.6690	0.6250	0.6780	0.7810
500		0.4930	0.5700	0.7400	0.6260	0.6940	0.8250
800		0.5010	0.6420	0.7870	0.6430	0.7590	0.8570
				$\rho = 0$			
300		0.4760	0.5640	0.6630	0.6230	0.6860	0.7590
500		0.4570	0.5590	0.7300	0.6120	0.6980	0.8120
800		0.4740	0.6510	0.7720	0.6350	0.7460	0.8510

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \mathrm{cov}(e_t, \varepsilon_t)$, $\triangle_n(x) = \delta_n/(1+x^2)$ in which $\delta_n = \frac{1}{2}\sqrt{k\log(n)/\sqrt{n}}$.

Table 15: Size and Power at 10%: $\psi_1 = -0.01$, $\psi_2 = 0.3$ and $\psi_3 = 0.1$

n		Size at 10%			Power at 10%		
	$\kappa =$	1/5	1/4	1/3	1/5	1/4	1/3
				$\rho = 0.75$			
300		0.1410	0.1330	0.1190	0.6710	0.7600	0.8060
500		0.1140	0.1290	0.1130	0.7090	0.7250	0.8280
800		0.1310	0.1270	0.0960	0.7080	0.8020	0.8820
				$\rho = -0.75$			
300		0.1350	0.1160	0.1070	0.6920	0.7580	0.8220
500		0.1180	0.1300	0.1110	0.7100	0.7520	0.8620
800		0.1090	0.1080	0.1180	0.7080	0.8040	0.8660
				$\rho = 0.50$			
300		0.1170	0.1130	0.1210	0.7120	0.7580	0.8350
500		0.1270	0.1310	0.0960	0.6940	0.7680	0.8470
800		0.1290	0.1280	0.0970	0.7180	0.7970	0.8970
				$\rho = -0.50$			
300		0.1170	0.1110	0.1090	0.7200	0.7520	0.8190
500		0.1170	0.1060	0.1020	0.7070	0.7710	0.8450
800		0.1070	0.1070	0.0920	0.7260	0.8130	0.8750
				$\rho = 0.25$			
300		0.1060	0.0840	0.1120	0.6870	0.7330	0.8270
500		0.1020	0.1090	0.1140	0.6960	0.7580	0.8380
800		0.1100	0.1100	0.1050	0.7060	0.7970	0.8920
				$\rho = -0.25$			
300		0.1040	0.1120	0.1120	0.7030	0.7360	0.8270
500		0.1220	0.1130	0.1090	0.7030	0.7670	0.8560
800		0.0870	0.1040	0.1110	0.7080	0.8220	0.8960
				$\rho = 0$			
300		0.1180	0.1090	0.1040	0.7060	0.7730	0.8060
500		0.1200	0.1050	0.1200	0.6840	0.7530	0.8520
800		0.1150	0.1010	0.0980	0.6930	0.8050	0.8890

$u_t = \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \psi_3\varepsilon_{t-2}$, $k = [n^\kappa]$, $\rho = \text{cov}(e_t, \varepsilon_t)$.