ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

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

Model Specification between Parametric and Nonparametric Cointegration

Jiti Gao, Dag Tjøstheim and Jiying Yin

August 2012

Working Paper 18/12

Model Specification between Parametric and Nonparametric Cointegration

Jiti Gao**and Dag Tjøstheim† and Jiying Yin*

*Monash University** *and The University of Bergen*†

August 24, 2012

Abstract

This paper considers a general model specification between a parametric co–integrating model and a nonparametric co–integrating model in a multivariate regression model, which involves a univariate integrated time series regressor and a vector of stationary time series regressors. A new and simple test is proposed and the resulting asymptotic theory is established. The test statistic is constructed based on a natural distance function between a nonparametric estimate and a smoothed parametric counterpart. The asymptotic distribution of the test statistic under the parametric specification is proportional to that of a local–time random variable with a known distribution. In addition, the finite sample performance of the proposed test is evaluated through using both simulated and real data examples.

Key words: Cointegration, nonparametric kernel estimation, parametric model specification, time series.

JEL Classification: C12, C14, C22.

Abbreviated Title: Model Specification in Nonstationary Cointegration.

* *Correspondence*: Jiti Gao, Department of Econometrics and Business Statistics, Monash University, Caulfield East, Victoria 3145, Australia. Email: jiti.gao@monash.edu.

1 Introduction

In recent years, there has been an increasing interest in discussing model estimation and specification testing problems involving nonparametric regression models associated with integrated time series. Recent literature includes Park and Phillips (1998), Karlsen and Tjøstheim (2001), Karlsen *et al* (2007), Wang and Phillips (2009a, 2009b) and Wang and Phillips (2011) in the area of nonparametric estimation. Such existing studies are all limited to the case where each of the integrated time series regressors is univariate, mainly because the null recurrent structure of integrated time series typically reduces the amount of time that such time series spend in the vicinity of any one point, thereby exacerbating the sparse data problem or the "curse of dimensionality" in nonparametric modelling of multivariate integrated time series. As indicated by equation (2.4) below, meanwhile, nonparametric kernel estimation may not be working in the multivariate integrated time series case. Therefore, some semiparametric regression models are being proposed to deal with modelling multivariate integrated time series. Existing studies include Gao and Phillips (2011), and Chen, Gao and Li (2012) in the field of semiparametric regression modelling of multivariate integrated time series. Meanwhile, Cai *et al* (2009) and Xiao (2009) consider using varying–coefficient models as an alternative.

In these latter studies, time series regressors involved in the nonparametric part of the model are only univariate nonstationary without including any stationary regressors. In the parametric integrating case, however, both stationary and nonstationary time series regressors can be involved in the same regression model (see, for example, Chang, Park and Phillips 2001), and there are in fact good reasons for studying such models in addressing empirical problems. Examples include modelling the relationship between the consumption time series and the income time series, in which a short–term interest rate variable can be naturally involved as a stationary time series regressor, while both the consumption and income time series regressors are known to be nonstationary.

A main objective of using a parametric model specification is to find a best available parametric function to approximate an unknown nonparmetric function. As shown in the literature (such as, Karlsen and Tjøstheim 2001; Wang and Phillips 2009a), nonparametric kernel estimation for the integrated time series case often results in a rate of convergence at the order of $\sqrt{\sqrt{n}\ h}$, slower than the rate of $\sqrt{n\ h}$ for the stationary time series case, where h is a bandwidth parameter. By contrast, parametric estimation in the integrated time series case can achieve the conventional rate–$\sqrt{n}$ and even faster than it. As a consequence, one would prefer a parametric co–integrating model to a nonparametric cointegrating model when possible. This thus means that using parametric specification in the integrated time series case may be more relevant and necessary than that for the stationary time series case.

In this paper, we are interested in a multivariate time series model of the form

$$y_t = m(x_t, z_t) + e_t, \tag{1.1}$$

where x_t is a univariate nonstationary time series, $z_t = (z_{t1}, \cdots, z_{td})^\tau$ is a d–dimensional vector of stationary time series regressors, and e_t is a time series error process, and $m(\cdot, \cdot)$ is an unknown function over R^{d+1}. To emphasise the main ideas and avoid the so–called "curse of dimensionality", this paper focuses on the case of $1 \leq d \leq 3$ for our discussion of model specification.

We are then interested in testing the null hypothesis:

$$H_0: \ P\left(m(x_t, z_t) = g(x_t, z_t; \theta_0)\right) = 1, \tag{1.2}$$

versus a sequence of local alternatives of the form:

$$H_1: \ P\left(m(x_t, z_t) = g(x_t, z_t; \theta_0) + \Delta_n(x_t, z_t)\right) = 1, \tag{1.3}$$

where $g(\cdot, \cdot; \theta)$ is a known parametric function indexed by θ, a vector of unknown parameters, $\theta_0 \in \Theta_0$ with Θ_0 being a compact subset of R^c, and $\Delta_n(x, z)$ is a sequence of unknown departure functions.

Recent studies in the field of nonparametric model specification of integrated time series models include Gao *et al* (2009a), and Wang and Phillips (2012). Other related studies include Kasparis (2008), Gao *et al* (2009b), and Hong and Phillips (2010). To the best of our knowledge and experience, the proposed tests by Gao *et al* (2009a), and Wang and Phillips (2012) use exactly the same type of tests as those originally developed for the stationary time series case (see, for example, Chapter 3 of Gao 2007). In other words, the full nature of nonstationarity of $\{x_t\}$ has not been taken into account in the construction of the proposed tests. As a consequence, it now looks that both the establishment and the proof of the asymptotic theory in Gao *et al* (2009a), and Wang and Phillips (2012) are unnecessarily complicated and technical. This paper therefore takes the full feature of the integrated structure of $\{x_t\}$ into account and proposes a new and simple test for testing H_0 versus H_1. Through both theoretical and empirical comparisons, moreover, we show that the proposed test of this paper is preferred to the test proposed in Gao *et al* (2009a) and Wang and Phillips (2012) when $\{\Delta_n(\cdot, \cdot)\}$ is "asymptotically small".

The contributions and organisation of this paper are given as follows. Section 2 constructs our new test and then establishes a general asymptotic theory for the case where the probabilistic structure of (x_t, z_t, e_t) is quite general such that x_t and z_t can be correlated and $\{e_t\}$ is a sequence of martingale differences. Section 3 discusses some power properties of the proposed test and then compares such properties with those of an existing test. A set of simulated examples are given in Section 4. Section 5 considers an empirical application. The paper concludes in Section 6. All mathematical proofs are given in an appendix.

2 Nonparametric specification test

Before we construct our test, we have a look at how to estimate θ_0 and $m(\cdot,\cdot)$, respectively. It follows from model (1.1) that

$$y_t = m(x_t, z_t) + e_t = g(x_t, z_t; \theta_0) + e_t \quad \text{under } H_0. \tag{2.1}$$

Under H_0, model (2.1) suggests estimating θ_0 by $\widehat{\theta}$ that minimises

$$\frac{1}{n}\sum_{t=1}^{n}\left[y_t - g(x_t, z_t; \theta)\right]^2 \quad \text{over all possible } \theta. \tag{2.2}$$

Meanwhile, model (2.1) suggests estimating $m(\cdot,\cdot)$ by

$$\widehat{m}(x,z) = \frac{\sum_{t=1}^{n} K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) y_t}{\sum_{t=1}^{n} K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right)}, \tag{2.3}$$

where K_1 and K_2 are probability kernel functions, and h_1 and h_2 are bandwidth parameters.

Note that the conventional estimation method used in (2.3) may not be extendable to the case where both x_t and z_t are integrated time series. In fact, consider the case where both x_t and z_t are univariate integrated time series. For fixed (x, z)

$$\sum_{t=1}^{n} E\left[K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right)\right] \sim h_1 h_2 \sum_{t=1}^{n}\left(\frac{1}{\sqrt{2\pi t}}\, e^{-\frac{x^2}{2t}}\right)\cdot\left(\frac{1}{\sqrt{2\pi t}}\, e^{-\frac{z^2}{2t}}\right)$$

$$\sim\ C h_1 h_2 \sum_{t=1}^{n} \frac{1}{\sqrt{t}}\cdot\frac{1}{\sqrt{t}} \sim C(1+o(1))\log(n)\ h_1\ h_2 \to 0 \tag{2.4}$$

as $n \to \infty$, using $\frac{1}{\sqrt{t}} x_t \sim N(0,1)$ and $\frac{1}{\sqrt{t}} z_t \sim N(0,1)$, when $x_t = x_{t-1} + \eta_t$ and $z_t = z_{t-1} + \zeta_t$, with $x_0 = z_0 = 0$ and $\eta_t \sim N(0,1)$ and $\zeta_t \sim N(0,1)$. This implies that multivariate nonparametric kernel estimation may not be working in the multivariate $I(1)$ case. A recent paper by Myklebust, Karlsen and Tjøstheim (2012) discusses a similar issue.

To test H_0, model (2.1) suggests constructing a test based on a kind of distance between $\widehat{m}(x,z)$ and $g(x,z;\widehat{\theta})$. In order to construct our test, we introduce a smoothed version of

$g(\cdot,\cdot;\theta_0)$ of the form

$$\widetilde{g}(x,z;\theta_0) = \frac{\sum_{t=1}^n K_1\left(\frac{x_t-x}{h_1}\right) K_2\left(\frac{z_t-z}{h_2}\right) g(x_t,z_t;\theta_0)}{\sum_{t=1}^n K_1\left(\frac{x_t-x}{h_1}\right) K_2\left(\frac{z_t-z}{h_2}\right)}. \tag{2.5}$$

We may then introduce a distance function between $\widehat{m}(x,z)$ and $\widetilde{g}(x,z;\widehat{\theta})$. To avoid introducing some random denominator problem, we propose using a modified distance function by comparing the following quantities:

$$\begin{aligned}
\widehat{q}(x,z) &= \frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^n K_1\left(\frac{x_t-x}{h_1}\right) K_2\left(\frac{z_t-z}{h_2}\right) y_t \ \text{ and} \\
\widetilde{q}(x,\theta_0) &= \frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^n K_1\left(\frac{x_t-x}{h_1}\right) K_2\left(\frac{z_t-z}{h_2}\right) g(x_t,z_t;\theta_0).
\end{aligned} \tag{2.6}$$

This paper now proposes using a test statistic of the form

$$\begin{aligned}
\widehat{L}_n(h_1,h_2) &= \sqrt{nh_1^2h_2^{2d}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\widehat{q}(x,z)-\widetilde{q}\left(x,z;\widehat{\theta}\right)\right)^2\pi_1(x)\pi_2(z)dzdx \\
&= \sqrt{nh_1^2h_2^{2d}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\widehat{m}(x,z)-\widetilde{g}\left(x,z;\widehat{\theta}\right)\right)^2\widehat{p}^2(x,z)\ \pi_1(x)\pi_2(z)dzdx,
\end{aligned} \tag{2.7}$$

which is similar to the original proposal discussed in Härdle and Mammen (1993) for the independent sample case, where $\pi_i(u)$ are both known probability weight functions satisfying $0<\int_{-\infty}^{\infty}\pi_i^2(u)du<\infty$ for $i=1,2$, and $\widehat{p}(x,z)=\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^n K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$.

Before we impose certain conditions to establish an asymptotic distribution for $\widehat{L}_n(h_1,h_2)$, we have a look at a closed–form approximation to $\widehat{L}_n(h_1,h_2)$ under H_0. As shown in the proof of Theorem 2.1 given in Appendix A below, we have under H_0

$$\begin{aligned}
\widehat{L}_n(h_1,h_2) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n \widehat{e}_t^2\ \pi_1(x_t)\pi_2(z_t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)dvdu \\
&+ \frac{2}{\sqrt{n}}\sum_{t=2}^n\sum_{s=1}^{t-1}\widehat{e}_t\ \widehat{e}_s\ \pi_1(x_s)\pi_2(z_s)\ L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)+o_P(1) \\
&\equiv \widetilde{L}_n(h_1,h_2)+o_P(1),
\end{aligned} \tag{2.8}$$

where $L_i(v)=\int_{-\infty}^{\infty}K_i(u)K_i(u+v)du$ for $i=1,2$, and

$$\begin{aligned}
\widetilde{L}_n(h_1,h_2) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n \widehat{e}_t^2\ \pi_1(x_t)\pi_2(z_t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)dvdu \\
&+ \frac{2}{\sqrt{n}}\sum_{t=2}^n\sum_{s=1}^{t-1}\widehat{e}_t\ \widehat{e}_s\ \pi_1(x_s)\pi_2(z_s)\ L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right) \\
&\equiv \widetilde{S}_{1n}+\widetilde{S}_{2n}
\end{aligned} \tag{2.9}$$

is a closed form approximation to $\widehat{L}_n(h_1, h_2)$. Our experience in Sections 4 and 5 below shows that it is computationally easier to use $\widetilde{L}_n(h_1, h_2)$ than $\widehat{L}_n(h_1, h_2)$, which involves an integral in $R^2 = (-\infty, \infty) \times (-\infty, \infty)$.

Mainly because of the fact that $\widetilde{S}_{1n}$ converges in distribution to a random variable, there is no need to standardise $\widetilde{L}_n(h_1, h_2)$ to establish asymptotic normality as the limiting distribution of the standardised version of $\widetilde{L}_n(h_1, h_2)$. Moreover, existing literature (Gao *et al* 2009a; Wang and Phillips 2012) shows that it is much harder and complicated to show that a standardised version of $\widetilde{S}_{2n}$ converges in distribution to a standard normal random variable than to prove that $\widetilde{S}_{2n}$ converges in probability to zero as will be done in this paper. In the stationary case where $\{x_t\}$ is also stationary, however, we will need to use a standardised version of the form

$$\overline{M}_n(h_1, h_2) = \frac{\widetilde{L}_n(h_1, h_2) - \widetilde{S}_{1n}}{\widetilde{\sigma}_n} \tag{2.10}$$

as a test statistic (see, for example, Härdle and Mammen 1993; Fan and Yao 2003; Gao 2007; Li and Racine 2007), where $\widetilde{\sigma}_n$ is a normalised quantity. This is mainly due to the fact that $\frac{\widetilde{S}_{1n}}{\sqrt{n}} \to_P C(K, \pi; \sigma_e^2)$, where $C(K, \pi; \sigma_e^2)$ is a non–random constant. In other words, $\widetilde{S}_{1n}$ itself cannot be normalised to be a test statistic.

In order to precisely establish and show an asymptotic distribution for $\widehat{L}_n(h_1, h_2)$, we need to introduce the following assumptions; their justifications are given below. For notational simplicity, an integral of a d–dimensional function $L(\cdot)$ is denoted by the form $\int_{-\infty}^{\infty} L(u)du$, which is the same as in the one–dimensional case.

Assumption 2.1. (i) *Consider $x_t = x_{t-1} + u_t$, where $\{u_t\}$ is a sequence of time series errors with $E[u_t] = 0$. Let $\{(z_t, u_t)\}$ be a vector of stationary time series with $E\left[|u_1|^{2+\delta}\right] < \infty$ and $E\left[||z_1||^{2+\delta}\right] < \infty$ for some $\delta > 0$, where $||\cdot||$ denotes the conventional Euclidean norm. In addition, $\{(z_t, u_t)\}$ is assumed to be α–mixing with mixing coefficient $\alpha_{zu}(k)$ satisfying $\sum_{k=0}^{\infty} \alpha_{zu}^{\frac{\delta}{2+\delta}}(k) < \infty$.*

(ii) *Let $p(z)$ be the marginal density function of z_1 and $p(v_2, v_1)$ be the joint density of (z_t, z_s) for $t > s$. Suppose that $p(z)$ is continuous in z and $p(v_2, v_1)$ is continuous in (v_2, v_1). For $t > s$, let $q\left(v_2, v_1 | \frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}\right)$ and $q\left(v_2 | \frac{u_2}{\sqrt{t}}\right)$ be the conditional density functions of (z_t, z_s) given $\left(\frac{x_t}{\sqrt{t}}, \frac{x_s}{\sqrt{s}}\right) = \left(\frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}\right)$, and z_t given $\frac{x_t}{\sqrt{t}} = \frac{u_2}{\sqrt{t}}$, respectively. Suppose that $q\left(v_2, v_1 | \frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}\right) \leq 2p(v_2, v_1)$ uniformly in (s, t) and $q\left(v_2 | \frac{u_2}{\sqrt{t}}\right) \leq 2p(v_2)$ uniformly in t. Let $g\left(\frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}\right)$ be the joint density of $\left(\frac{x_t}{\sqrt{t}}, \frac{x_s}{\sqrt{s}}\right)$, and $g(\cdot)$ be the density of $\frac{x_s}{\sqrt{s}}$. Suppose that there are two constants $0 < c_1, c_2 < \infty$ such that $\sup_{(u_1,u_2)} g\left(\frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}\right) \leq c_1$ uniformly in (s, t) and $\sup_{u_1} g\left(\frac{u_1}{\sqrt{s}}\right) \leq c_2$ uniformly in s.*

Assumption 2.2. *Let $\{\mathcal{F}_{n,t} : t \geq 1\}$ be a sequence of increasing σ–fields generated by $\{(e_i, u_j, z_j) : 1 \leq i \leq t, 1 \leq j \leq n\}$. Suppose that $E\left[e_t|\mathcal{F}_{n,t-1}\right] = 0$ and $E\left[e_t^2|\mathcal{F}_{n,t-1}\right] = \sigma_e^2 < \infty$ almost surely (a.s.). In addition, there is some $\lambda > 0$ such that $E\left[|e_t|^{4+\lambda}|\mathcal{F}_{n,t-1}\right] < \infty$ almost surely.*

Assumption 2.3. (i) *Suppose that $g(x, z; \theta)$ is differentiable with respect to θ and that there are some function $G_i(x, z; \theta_0)$ for $i = 1, 2$ and small $\varepsilon > 0$ such that for $i, j = 1, 2$*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ||G_i(x, z; \theta_0)||^2 \pi_1^j(x)\pi_2^j(z)p^{2j}(z)dzdx < \infty \quad \text{and}$$
$$|g(x, z; \theta) - g(x, z; \theta_0) - (G_1(x, z; \theta_0))^\tau (\theta - \theta_0)| \leq G_2(x, z; \theta_0)\, ||\theta - \theta_0||^2$$

for all $\theta \in \Theta(\varepsilon) = \{\theta : ||\theta - \theta_0|| \leq \varepsilon\}$.

(ii) *Let $\widehat{\theta}$ be a consistent estimator of θ_0 such that for any small $\epsilon > 0$, we have $P\left(\sqrt{n}h_1h_2^d\left|\left|\widehat{\theta} - \theta_0\right|\right|^2 > \epsilon\right) \to 0$ as $n \to \infty$.*

Assumption 2.4. (i) *Let $K_i(\cdot)$ be symmetric probability kernel functions with*

$$\int_{-\infty}^{\infty} ||u||^j K_i^2(u)du < \infty \quad \text{and} \quad \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} ||v||^j K_i(u + v)K_i(v)dv\right)^2 du < \infty$$

for $i = 1, 2$ and $j = 0, 1, 2$.

(ii) *The bandwidths h_i satisfies $h_i \to 0$ and $nh_1^2h_2^{2d} \to \infty$ for $i = 1, 2$.*

(iii) *Let $\pi_1(\cdot)$ and $\pi_2(\cdot)$ be known probability weight functions such that $\int_{-\infty}^{\infty} \pi_1^{2+\delta}(u)du < \infty$ and $\int_{-\infty}^{\infty} \pi_2^{2+\delta}(z)p(z) < \infty$ for the same $\delta > 0$ as in Assumption 2.1(i).*

(iv) *Let also $\int_{-\infty}^{\infty} \pi_2^i(z)p^{2i}(z)dz < \infty$ for $i = 1, 2$. In addition, there are functions $D_i(x)$ for $i = 1, 2$ with $\int_{-\infty}^{\infty} D_1(x)dx < \infty$ and $\int_{-\infty}^{\infty} D_2(z)p^i(z)dz < \infty$ for $i = 1, 2$ such that*

$$|\pi_1(y) - \pi_1(x)| \leq D_1(x) \cdot |y - x| \text{ for any } (x, y) \in \Omega_1(\epsilon) = \{(x, y) : |y - x| \leq \epsilon, x, y \in R^1\},$$
$$|\pi_2(y) - \pi_2(x)| \leq D_2(x) \cdot ||y - x|| \text{ for any } (x, y) \in \Omega_2(\epsilon) = \{(x, y) : ||y - x|| \leq \epsilon, x, y \in R^d\}.$$

Justifications about the suitability and verifiability of Assumptions 2.1–2.4 are given below.

Assumption 2.1(i) imposes a stationarity structure on $\{u_t\}$ and then an α–mixing condition on (z_t, u_t). Such conditions are quite commonly used in the stationary time series case. If both u_t and z_t are allowed to be linear processes, some extra conditions are needed to ensure that each of the linear processes is α–mixing (see, for example, Theorem 2.1 of Chanda 1974; Corollary 4 of Withers 1981; Fan and Yao 2003). Assumption 2.1(ii) imposes some necessary and mild smoothness conditions on the marginal and joint density functions. Such conditions are very mild. We define, for $s < t$, the joint density functions of $(x_t, x_s; z_t, z_s)$ and (z_t, z_s) by $p(u_2, u_1; v_2, v_1)$ and $p(v_2, v_1)$, respectively, and the conditional density functions of x_t given (x_s, z_t, z_s), x_s given (z_t, z_s), and x_s given z_s by $p(u_2|u_1, v_2, v_1)$,

$p(u_1|v_2, v_1)$ and $p(u_1|v_1)$, respectively. We also define, for $s < t$, the joint density function of $\left(\frac{x_t}{\sqrt{t}}, \frac{x_s}{\sqrt{s}}; z_t, z_t\right)$ by $q(u_2, u_1; v_2, v_1)$, and the conditional density functions of $\frac{x_t}{\sqrt{t}}$ given $\left(\frac{x_s}{\sqrt{s}}, z_t, z_s\right)$, $\frac{x_s}{\sqrt{s}}$ given (z_t, z_s), and $\frac{x_s}{\sqrt{s}}$ given z_s by $q(u_2|u_1, v_2, v_1)$, $q(u_1|v_2, v_1)$ and $q(u_1|v_1)$, respectively. For notational simplicity, the notation involves no t and s as indices.

We have $p(u_1|v_1) = \frac{1}{\sqrt{s}} q\left(\frac{u_1}{\sqrt{s}}|v_1\right)$, $p(u_2|u_1, v_2, v_1) = \frac{1}{\sqrt{t}} q\left(\frac{u_2}{\sqrt{t}}|u_1, v_2, v_1\right)$ and $p(u_1|v_2, v_1) = \frac{1}{\sqrt{s}} q\left(\frac{u_1}{\sqrt{s}}|v_2, v_1\right)$. By Lemma B.1 in Appendix B below, we then have that as $t \to \infty$ and $s \to \infty$

$$q\left(\frac{u_2}{\sqrt{t}}|u_1, v_2, v_1\right) \to \phi(0) \text{ and } q\left(\frac{u_1}{\sqrt{s}}|v_2, v_1\right) \to \phi(0) \tag{2.11}$$

when (u_1, u_2, v_1, v_2) are all given, where $\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$ is the density function of the standard normal random variable $U \sim N(0, 1)$.

This implies as $t \to \infty$ and $s \to \infty$

$$\begin{aligned} p(v_2, v_1|u_2, u_1) &= \frac{p(u_2|u_1, v_2, v_1)p(u_1|v_2, v_1)p(v_2, v_1)}{p(u_2|u_1)p(u_1)} \\ &= \frac{q\left(\frac{u_2}{\sqrt{t}}|\frac{u_1}{\sqrt{s}}, v_2, v_1\right) q\left(\frac{u_1}{\sqrt{s}}|v_2, v_1\right) p(v_2, v_1)}{q\left(\frac{u_2}{\sqrt{t}}|\frac{u_1}{\sqrt{s}}\right) q\left(\frac{u_1}{\sqrt{s}}\right)} \to \frac{\phi^2(0) \cdot p(v_2, v_1)}{\phi^2(0)} = p(v_2, v_1), \end{aligned} \tag{2.12}$$

which implies that $p(v_2, v_1|u_2, u_1) \to p(v_2, v_1)$ as $s, t \to \infty$ and $p(v_2|u_2) \to p(v_2)$ as $t \to \infty$. Thus, we assume without loss of generality that $q\left(v_2, v_1|\frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}\right) \leq 2p(v_2, v_1)$ uniformly in (t, s) and $q\left(v_2|\frac{u_2}{\sqrt{t}}\right) \leq 2p(v_2)$ uniformly in t.

The justification of the second part of Assumption 2.1(ii) follows trivially. For example, when $s \to \infty$, we have $g\left(\frac{u_1}{\sqrt{s}}\right) \to \phi(0)$ by Lemma B.1. When s is fixed and the support of x_s is compact, the boundedness of $g\left(\frac{u_1}{\sqrt{s}}\right)$ follows from the common assumption of the continuity of $g(\cdot)$. When s is fixed and $g\left(\frac{u_1}{\sqrt{s}}\right) \to 0$ as $u_1 \to \infty$, the boundedness of $g\left(\frac{u_1}{\sqrt{s}}\right)$ also follows trivially. As shown in the proof of Corollary 2.2 of Wang and Phillips (2009a), moreover, the boundedness assumption is satisfied automatically when both u_t and z_t are additionally assumed to follow linear processes. In summary, it is not unreasonable to assume the boundedness in Assumption 2.1(ii).

Assumption 2.2 imposes a martingale structure on $(e_t, \mathcal{F}_{n,t})$. If $\{u_t\}$ is further assumed to follow a linear process of the form $u_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ with $\{\varepsilon_i : -\infty < i < \infty\}$ be a sequence of independent and identically distributed continuous random variables and $\{\psi_j : j \geq 0\}$ satisfying certain conditions, $\mathcal{F}_{n,t}$ can be replaced by $\mathcal{F}_t = \sigma(e_1, \cdots, e_t; \varepsilon_{-\infty}, \cdots, \varepsilon_{t+1})$ generated by $\{(e_i, \varepsilon_j) : 1 \leq i \leq t, -\infty < j \leq t+1\}$. In this case, one will need to further assume that $\{\varepsilon_s\}$ and $\{e_t\}$ are independent for all $s \geq t+1$. Similar conditions have been assumed in Li *et al* (2011), and Wang and Phillips (2012).

Assumption 2.3(i) imposes some mild conditions to ensure the integrability of the first partial derivative of $g(x, z; \theta)$ with respect to θ. Due to the involvement of $\pi_1(x)$ in particular, various functional forms of $g(x, z; \theta)$, including the conventional integrable functions and non–integrable polynomial functions, can be covered in Assumption 2.3(i) when $\pi_1(x)$ is suitably chosen. Specifically, one may choose $\pi_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ when the partial derivatives of $g(x, z; \theta)$ with respect to θ are of polynomial forms. As a consequence, there may be no need to individually consider the case where $g(x, z; \theta)$ is either integrable with respect to (x, z) or asymptotically homogeneous with respect to (x, z) as has been done in the literature (see, for example, Park and Phillips 2001; Li *et al* 2011). In summary, the differentiability condition on $g(x, z; \theta)$ with respect to θ, along with the integrability of $G_i(x, z; \theta)$, is quite flexible and easily verifiable.

Since weak convergence with a rate is needed to establish the asymptotic distribution in Theorem 2.1, Assumption 2.3(ii) imposes a condition on the rate of convergence directly rather than imposing certain conditions to imply the asymptotic consistency. It is however that this may be easily satisfied when $\widehat{\theta} - \theta_0$ achieves either a slow rate of $n^{-\frac{1}{4}}$ or a fast rate of $n^{-\frac{1}{2}}$ as has been established in the literature (see, for example, Chang, Park and Phillips 2001; Park and Phillips 2001).

Assumption 2.4 imposes smoothness conditions on the weight functions $\pi_1(x)$ and $\pi_2(z)$. Note that all such conditions may not be the weakest possible, but are all quite mild and verifiable.

We now state the main theorem of this paper; its proof is given in Appendix A.

THEOREM 2.1. *Consider model (1.1). Let Assumptions 2.1–2.4 hold. Then under* H_0

$$\begin{aligned} \widehat{L}_n(h_1, h_2) &= \sqrt{n h_1^2 h_2^{2d}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\widehat{q}(x, z) - \widetilde{q}\left(x, z; \widehat{\theta}\right) \right)^2 \pi_1(x) \pi_2(z) dz dx \\ &\to_D \ C_1(K, \pi; \sigma_e^2) \cdot L_{B_u}(1, 0) \quad \text{as } n \to \infty, \end{aligned} \tag{2.13}$$

where $C_1(K, \pi; \sigma_e^2) = \prod_{i=1}^2 \int_{-\infty}^{\infty} K_i^2(v) dv \cdot \int_{-\infty}^{\infty} \pi_2(z)\ p(z) dz \cdot \sigma_e^2$ *with* $p(z)$ *being the marginal density of* $\{z_t\}$, *and* $L_{B_u}(r, 0) = \lim_{\delta \to 0} \frac{1}{2\delta} \int_0^r I\left[|B_u(s)| < \delta\right] ds$ *is the local–time process associated with the Brownian motion* B_u, *which is the weak limit of* $U_n(r)$ *such that* $U_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t \Longrightarrow_D B_u(r)$ *on* $D[0,1]$ *as* $n \to \infty$, *and* $L_{B_u}(1, 0)$ *is a local–time random variable with its cumulative distribution function being given by*

$$F_L(x) = P\left(L_{B_u}(1, 0) \le x\right) = \begin{cases} 2\Phi(x) - 1, & x \ge 0, \\ 0, & x < 0, \end{cases} \tag{2.14}$$

in which $\Phi(x)$ *is the cdf of* $N(0, 1)$.

When σ_e^2 is unknown, it can be estimated by $\widehat{\sigma}_e^2 = \frac{1}{n}\sum_{t=1}^n \left(y_t - g(x_t, z_t; \widehat{\theta})\right)^2$ under H_0. Note that Theorem 2.1 shows that the asymptotic distribution is proportional to $L_{B_u}(1,0)$ that has a known distribution with the known distribution function given in (2.14). Note also that it is quite common in the parametric case to have a functional of Brownian motion as a limiting distribution of a unit–root test statistic. It is therefore not unnatural to the local–time process as the limiting distribution of the proposed test statistic of this paper.

Meanwhile, Section 3 below discusses asymptotic power properties of the proposed test and its natural competitor, and shows that the proposed test is more powerful than the natural competitor. The finite–sample study in Section 4 further confirms this.

3 Asymptotic power properties

Since the methodologies and techniques required for us to rigorously study the power function of the proposed test are not readily available, this section briefly discusses some theoretical properties of the proposed test and a natural competitor under a sequence of asymptotically localised alternatives.

To the best of our knowledge, the only test for the univariate case available in the literature is the test proposed in Gao *et al* (2009a) and then used in Wang and Phillips (2012). An extended form of the existing test to our case can be written as

$$\widehat{M}_n(h_1, h_2) = \frac{\sum_{t=1}^n \sum_{s=1,\neq t}^n K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right)\ \widehat{e}_s\ \widehat{e}_t}{\sqrt{2\sum_{t=1}^n \sum_{s=1}^n K_1^2\left(\frac{x_t - x}{h_1}\right) K_2^2\left(\frac{z_t - z}{h_2}\right)\ \widehat{e}_s^2\ \widehat{e}_t^2}}. \tag{3.1}$$

Let $\widehat{M}_{2n}^2 = 2\sum_{t=1}^n \sum_{s=1}^n K_1^2\left(\frac{x_t - x}{h_1}\right) K_2^2\left(\frac{z_t - z}{h_2}\right)\ \widehat{e}_s^2\ \widehat{e}_t^2$. Similarly to the derivations of equation (3.5) given in Appendix B below, we have under H_0:

$$\begin{aligned}\widehat{M}_{2n}^2 &= 2\sum_{t=2}^n \sum_{s=1}^n K_1^2\left(\frac{x_t - x}{h_1}\right) K_2^2\left(\frac{z_t - z}{h_2}\right)\ e_s^2\ e_t^2 + o_P(1) \equiv M_{2n} + o_P(1),\\ \sigma_{2n}^2 &= E\left[M_{2n}^2\right] = C(1 + o(1)) \cdot n^{\frac{3}{2}} h_1 h_2^d.\end{aligned} \tag{3.2}$$

We then show that $\widehat{L}_n(h_1, h_2)$ is asymptotically more powerful than $\widehat{M}_n(h_1, h_2)$ under a sequence of local alternatives of (1.3) of the form:

$$\Delta_n(x, z) = \delta_n \cdot \Delta(x, z), \tag{3.3}$$

where $\delta_n \to 0$ and $\delta_n^2 \sqrt{n} h_1 h_2^d \to \infty$ as $n \to \infty$, and $\Delta(x, z)$ is chosen such that for $j = 1, 2$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Delta^2(x, z) p(z) dz dx < \infty \quad \text{and} \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Delta^2(x, z) p^2(z) \pi_1^j(x) \pi_2^j(z) dz dx < \infty. \tag{3.4}$$

Let $\widehat{M}_{1n} = \sum_{t=1}^{n}\sum_{s=1}^{n} K_1\left(\frac{x_t-x}{h_1}\right) K_2\left(\frac{z_t-z}{h_2}\right)\ \widehat{e}_s\ \widehat{e}_t$. Note that $\widehat{e}_t = y_t - g(x_t, z_t; \widehat{\theta}) - \Delta_n(x_t, z_t)$ under H_1. As can be deduced from the proof of Lemma A.2 in Appendix B below, we have under H_1

$$
\begin{aligned}
\widehat{M}_{1n} &= \sum_{t=1}^{n}\sum_{s=1}^{n} K_1\left(\frac{x_t-x_s}{h_1}\right) K_2\left(\frac{z_t-z_s}{h_2}\right)\ \widehat{e}_s\ \widehat{e}_t \\
&= \sum_{t=1}^{n}\sum_{s=1}^{n} K_1\left(\frac{x_t-x_s}{h_1}\right) K_2\left(\frac{z_t-z_s}{h_2}\right)\ e_s\ e_t \\
&+ \sum_{t=1}^{n}\sum_{s=1}^{n} K_1\left(\frac{x_t-x_s}{h_1}\right) K_2\left(\frac{z_t-z_s}{h_2}\right)\ \Delta_n(x_s, z_s)\ \Delta_n(x_t, z_t) + o_P(1). \\
&\geq \sum_{t=1}^{n}\sum_{s=1}^{n} K_1\left(\frac{x_t-x_s}{h_1}\right) K_2\left(\frac{z_t-z_s}{h_2}\right)\ \Delta_n(x_s, z_s)\ \Delta_n(x_t, z_t) + o_P(1) \\
&\equiv M_{1n} + o_P(1),
\end{aligned}
$$

where we have as $n \to \infty$

$$
\begin{aligned}
R_{1n} \equiv \frac{E\left[M_{1n}\right]}{\sigma_{2n}} &= C\delta_n^2 \cdot \frac{nh_1h_2^d}{\sqrt{n^{\frac{3}{2}}h_1h_2^d}} \cdot (1 + o(1)) \\
&= C\delta_n^2 \sqrt{\sqrt{n}h_1h_2^d} \cdot (1 + o(1)) \qquad (3.5)
\end{aligned}
$$

when $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Delta^2(x, z)p(z)dzdx < \infty$. An outline of the derivation of (3.5) is given in Appendix B.

As shown in Lemma A.2 in Appendix A below, under H_1, we have as $n \to \infty$

$$
\begin{aligned}
\widehat{L}_n(h_1, h_2) &= \frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} \widehat{e}_s\ \widehat{e}_t \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1\left(\frac{x_t-x}{h_1}\right) K_2\left(\frac{z_t-z}{h_2}\right) \\
&\times K_1\left(\frac{x_s-x}{h_1}\right) K_2\left(\frac{z_s-z}{h_2}\right) \pi_1(x)\pi_2(z)dzdx \\
&\geq \frac{\delta_n^2}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} \Delta(x_s, z_s)\Delta(x_t, z_t) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1\left(\frac{x_t-x}{h_1}\right) K_2\left(\frac{z_t-z}{h_2}\right) \\
&\times K_1\left(\frac{x_s-x}{h_1}\right) K_2\left(\frac{z_s-z}{h_2}\right) \pi_1(x)\pi_2(z)dzdx + o_P(1) \\
&\equiv L_{1n} + o_P(1),
\end{aligned}
$$

where we have as $n \to \infty$

$$
E[L_{1n}] = C(1 + o(1)) \cdot \delta_n^2 \cdot \sqrt{n}h_1h_2^d \qquad (3.6)
$$

when $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Delta^2(x, z)p^2(z)\pi_1(x)\pi_2(z)dzdx < \infty$. An outline of the derivation of (3.6) is given in Appendix B.

In view of Assumption 2.4(ii), equations (3.5) and (3.6) therefore imply that there is some $C_0 > 0$ such that

$$\frac{E\left[L_{1n}\right]}{R_{1n}} = \frac{E\left[L_{1n}\right]\ \sigma_{2n}}{E\left[M_{1n}\right]} = C_0 \sqrt{\sqrt{n} h_1 h_2^d} \to \infty, \tag{3.7}$$

which implies that $\widehat{L}_n(h_1, h_2)$ is more powerful than $\widehat{M}_n(h_1, h_2)$ under a sequence of alternatives of the forms (3.3) and (3.4).

It should be pointed out that there is a kind of trade–off between ensuring that $\widehat{L}_n(h_1, h_2)$ is more powerful than $\widehat{M}_n(h_1, h_2)$ and involving the weight functions $\pi_1(\cdot)$ and $\pi_2(\cdot)$ as well as requiring both parts of (3.4), in addition to requiring Assumption 2.4(ii). This is mainly because $\widehat{M}_n(h_1, h_2)$ can be more powerful than $\widehat{L}_n(h_1, h_2)$ when the second part of (3.4) is satisfied but the first part of (3.4) is not necessarily satisfied. Examples include the case where $\Delta(x, z) = \alpha x^2 + \beta z^2$. In this case, $\Delta(x, z)$ is not integrable with respect to x, but it can be asymptotically homogeneous with respect to x (see, for example, Definition 2.2 of Li *et al* 2011). However, this paper is not interested in such a case for power comparison. The main reason is that the departure function $\Delta_n(x, z)$ can be asymptotically 'large' even though $\delta_n \to 0$ with a rate. Let us just consider the univariate case where $g(x, \theta) = \alpha + \beta x$ and $\Delta_n(x_t) = \delta_n\ \Delta(x_t)$ with

$$\delta_n = \frac{1}{n^{\frac{11}{12}} h^{\frac{1}{4}}} \quad \text{and} \quad \Delta(x_t) = x_t^2, \tag{3.8}$$

where $h = n^{-\frac{1}{3}}$, $x_t = x_{t-1} + u_t$ with $x_0 = 0$ and $u_t \sim N(0, 1)$ (an example of this type has been considered in the simulation section of Wang and Phillips 2012).

Since $E[x_n^2] = n$, we have

$$E\left[\Delta_n(x_n)\right] = \delta_n\ E[x_n^2] = \frac{n}{n^{\frac{11}{12}} n^{-\frac{1}{12}}} = \frac{n}{n^{\frac{10}{12}}} = n^{\frac{1}{6}} \to \infty \tag{3.9}$$

even though $\delta_n = n^{-\frac{5}{6}} \to 0$.

This shows that the choice of a polynomial form for the departure function in the integrated time series case may not be so interesting because of the explosive nature of polynomial functions of such integrated time series. We are therefore only interested in the case where $\Delta(x, z)$ is a 'small' integrable function as required in equation (3.4). As a consequence, the departure function $\Delta_n(x, z)$ can be asymptotically negligible because $\delta_n \to 0$ as $n \to \infty$. As shown in Section 4 below, the proposed test $\widehat{L}_n(h_1, h_2)$ has power to pick up such 'small' departure and is more powerful than $\widehat{M}_n(h_1, h_2)$ when $\Delta_n(x, z)$ is asymptotically negligible. In summary, the theoretical discussion in Sections 2 and 3, along with the finite–sample study in Section 4 below, shows that $\widehat{L}_n(h_1, h_2)$ is a more powerful test than $\widehat{M}_n(h_1, h_2)$.

4 Simulation evaluation

This section uses several simulated examples to show how to implement the proposed test in practice and then examine whether the proposed test works numerically. Example 4.1 considers the case where the model under the null hypothesis is a simple linear model. Some nonlinear models are used in Example 4.2. In both Examples 4.1 and 4.2, the dimensionality of z_t is $d = 1$.

Recall that we are interested in the following hypotheses:

$$H_0: \ P\left(m(x_t, z_t) = g(x_t, z_t; \theta_0)\right) = 1 \quad \text{versus}$$
$$H_1: \ P\left(m(x_t, z_t) = g(x_t, z_t; \theta_0) + \Delta_n(x_t, z_t)\right) = 1.$$

Define the following test statistic:

$$\widehat{L}_{1n}(h_1, h_2) \equiv \widehat{M}_n(h_1, h_2) = \frac{\sum_{t=1}^n \sum_{s=1, \neq t}^n \widehat{e}_s K_1\left(\frac{x_s - x_t}{h_1}\right) K_2\left(\frac{z_s - z_t}{h_2}\right) \widehat{e}_t}{\widehat{\sigma}_{1n}}, \tag{4.1}$$

where $\widehat{\sigma}_{1n}^2 = 2\sum_{t=1}^n \sum_{s=1,\neq t}^n \widehat{e}_s^2 K_1^2\left(\frac{x_s - x_t}{h_1}\right) K_2^2\left(\frac{z_s - z_t}{h_2}\right) \widehat{e}_t^2$ with $\widehat{e}_t = \left(y_t - g(x_t, z_t; \widehat{\theta})\right)$, in which $\widehat{\theta}$ is the nonlinear least squares estimators of θ defined by minimising

$$\frac{1}{n} \sum_{t=1}^n \left(y_t - g(x_t, z_t; \theta)\right)^2 \quad \text{over } \theta.$$

In view of equations (2.8) and (2.9), define another test statistic as an approximation to $\widehat{L}_n(h_1, h_2)$:

$$\widehat{L}_{2n}(h_1, h_2) \equiv \widetilde{L}_n(h_1, h_2) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \widehat{e}_t^2 \ \pi_1(x_t)\pi_2(z_t) \ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_1^2(u) K_2^2(v) dv du$$
$$+ \frac{2}{\sqrt{n}} \sum_{t=2}^n \sum_{s=1}^{t-1} \widehat{e}_t \ \widehat{e}_s \ \pi_1(x_s)\pi_2(z_s) \ L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right), \tag{4.2}$$

where $\pi_1(x)$ and $\pi_2(z)$ are probability weight functions.

Our experience shows that the choice of $\pi_1(x)$ and $\pi_2(z)$ has little impact on both the size and power properties of the proposed test, as long as they both satisfy (3.4) above. In the simulated and real data examples below, we choose $\pi_1(x) = \frac{1}{\pi \ (1+x^2)}$ and $\pi_2(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$. In addition, we choose $K_i(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ for $i = 1, 2$. In this case, we then have $\int_{-\infty}^{\infty} K_i^2(u) du = \frac{1}{2\sqrt{\pi}}$ and $L_i(u) = \int_{-\infty}^{\infty} K_i(v) K_i(u+v) dv = \frac{1}{2\sqrt{\pi}} e^{-\frac{u^2}{4}}$.

Mainly due to the fact that Edgeworth expansions for both $\widehat{L}_{1n}(h_1, h_2)$ and $\widehat{L}_{2n}(h_1, h_2)$ are not readily available, we are unable to adopt the power–function approach for the choice of optimal bandwidths (as discussed in Li *et al* 2011). Instead, we propose using an estimation–based optimal bandwidths of the form:

$$\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right) = \arg \min_{(h_1,h_2)\in H_{\rm cv}} \frac{1}{n} \sum_{t=1}^{n} \left(y_t - \widehat{m}_{-t}(x_t, z_t; h_1, h_2)\right)^2, \tag{4.3}$$

where $\widehat{m}_{-t}(x_t, z_t; h_1, h_2) = \frac{\sum_{s=1,\neq t}^{n} K_1\left(\frac{x_t - x_s}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2}\right) y_s}{\sum_{u=1,\neq t}^{n} K_1\left(\frac{x_t - x_u}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2}\right)}$ and

$$H_{\rm cv} = \left[c_1\, n^{-\frac{1}{12}-c_0}, c_2\, n^{-\frac{1}{12}+c_0}\right] \times \left[d_1\, n^{-\frac{1}{6}-d_0}, d_2\, n^{-\frac{1}{6}+d_0}\right]$$

for some $0 < c_1 < c_2 < \infty$, $0 < c_0 < \frac{1}{48}$, $0 < d_1 < d_2 < \infty$ and $0 < d_0 < \frac{1}{24}$. Before selecting $H_{\rm cv}$, we actually calculated equation (4.3) over all possible intervals. Our computation indicates that $H_{\rm cv}$ is the shortest possible interval on which the CV function attains its smallest value.

Let $Q_n(h_1, h_2)$ denote either $\widehat{L}_{1n}(h_1, h_2)$ or $\widehat{L}_{2n}(h_1, h_2)$. Our experience with Examples 4.1 and 4.2 shows that $\widehat{L}_{2n}$ already has some stable sizes and good power values under the choice of $\left(\widehat{h}_{1\rm cv}, \widehat{h}_{2\rm cv}\right)$. This may be because this pair of bandwidths may be either exactly identical or very close to such bandwidth values that maximise the power function while controlling the size function. In the stationary time series case, the theory developed in Gao and Gijbels (2008) shows that such estimation–based optimal bandwidth values may also be optimal for testing purposes.

Let q_r be the asymptotic critical value of $Q_n\left(\widehat{h}_{1\rm cv}, \widehat{h}_{2\rm cv}\right)$ at the significance level r. We then propose using the following bootstrap method to find a simulated critical value, q_r^*, to approximate q_r.

Step 1: Generate the bootstrap residuals $\{e_t^*\}$ by $e_t^* = \widehat{\sigma}_e \eta_t^*$, where

$$\widehat{\sigma}_e^2 = \frac{1}{n} \sum_{t=1}^{n} \left(y_t - g(x_t, z_t; \widehat{\theta})\right)^2, \tag{4.4}$$

in which $\{\eta_t^*, 1 \le t \le n\}$ is a sequence of i.i.d. random variables drawn from

$$P\left(\eta_1^* = -\frac{\sqrt{5}-1}{2}\right) = \frac{\sqrt{5}+1}{2\sqrt{5}} \quad \text{and} \quad P\left(\eta_1^* = \frac{\sqrt{5}+1}{2}\right) = \frac{\sqrt{5}-1}{2\sqrt{5}}.$$

Step 2: Obtain $y_t^* = g(x_t; z_t; \widehat{\theta}) + e_t^*$. The resulting sample $\{(y_t^*, x_t, z_t), 1 \le t \le n\}$ is called a bootstrap sample.

Step 3: Use the data set $\{(y_t^*, x_t, z_t), 1 \le t \le n\}$ to re–estimate θ and denote its estimator by $\widehat{\theta}^*$. Then calculate the test statistic $\widehat{Q}_n^*\left(\widehat{h}_{1\rm cv}, \widehat{h}_{2\rm cv}\right)$, which is the corresponding version of $\widehat{Q}_n\left(\widehat{h}_{1\rm cv}, \widehat{h}_{2\rm cv}\right)$ by replacing $\{(y_t, x_t, z_t)\}$ and $\widehat{\theta}$ with $\{(y_t^*, x_t, z_t)\}$ and $\widehat{\theta}^*$, respectively.

Step 4: Repeat Steps 1–3 $M = 250$ times and produce $M = 250$ versions of $\widehat{Q}_n^*\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right)$. Denote the M versions of $\widehat{Q}_n^*\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right)$ by $\widehat{Q}_{n,m}^*(h_1, h_2)$, $m = 1, 2, \cdots, M$. Then, we construct the empirical distributions of $\widehat{Q}_{n,m}^*\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right)$. That is,

$$\mathsf{P}^*\left(\widehat{Q}_n^*\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right) \leq x\right) = \mathsf{P}\left(\widehat{Q}_n^*\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right) \leq x|\mathcal{W}_n\right),$$

where $\mathcal{W}_n = \{(y_t, x_t, z_t), 1 \leq t \leq n\}$.

For each pair $\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right)$, choose q_r^* such that $P^*\left(\widehat{Q}_n^*\left(\widehat{h}_{1cv}, \widehat{h}_{2cv}\right) > q_r^*\right) = r$ and estimate q_r by q_r^*. For $\widehat{L}_{1n}$ or $\widehat{L}_{2m}$, we approximate the asymptotic critical value z_r or l_r by z_r^* or l_r^*, respectively.

For $M = 250$, let also $f_{j\text{cv}}^*$ denote the frequency of $\widehat{L}_{1n}(\widehat{h}_{1cv}, \widehat{h}_{2cv}; j) > z_r^*$ for $j = 0, 1$ under H_0 or H_1, and $g_{j\text{cv}}^*$ denote the frequency of $\widehat{L}_{2n}(\widehat{h}_{1cv}, \widehat{h}_{2cv}; j) > l_r^*$ for $j = 0, 1$ under H_0 or H_1.

In both Examples 4.1 and 4.2, we use the chosen bandwidths and then the simulated critical values z_r^* for $\widehat{L}_{1n}$ and l_r^* for $\widehat{L}_{2n}$. The corresponding simulation results are reported in Tables 4.1–4.2 below.

Example 4.1. Consider a linear time series model of the form:

$$H_0: \ y_t = \alpha + \beta x_t + \gamma z_t + e_t, \quad t = 1, 2, \cdots, n, \tag{4.5}$$

versus

$$H_1: \ y_t = \alpha + \beta x_t + \gamma z_t + \Delta_n(x_t, z_t) + e_t, \quad t = 1, 2, \cdots, n, \tag{4.6}$$

where $x_t = x_{t-1} + u_t$ with $x_0 = 0$, $\alpha = 0$, $\beta = \gamma = 1$, and $\{(e_t, u_t, z_t) : 1 \leq t \leq n\}$ are independent and identically distributed as

$$\begin{pmatrix} e_t \\ u_t \\ z_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}\right), \tag{4.7}$$

with $\rho_i = 0$ or $\rho_i = 0.9$ for $i = 1, 2, 3$, and

$$\Delta_n(x, z) = \frac{\delta_n \ z^2}{\sqrt{1 + x^2}} \quad \text{with } \delta_n = \tfrac{\log(n)}{2n^{\frac{1}{8}}}. \tag{4.8}$$

Note that there is an endogeneity between e_t and (u_t, z_t) when $\rho_i \neq 0$, such as $\rho_1 = E[e_t u_t] = 0.9$ as chosen in Table 4.1. Note also that the choice of δ_n in theory is to ensure that $\delta_n \to 0$ and $\delta_n^2 \sqrt{n} h_1 h_2 \to \infty$ required in equation (3.6). Since the leading orders of h_1 and h_2 are chosen as $n^{-\frac{1}{12}}$ and $n^{-\frac{1}{6}}$, respectively in the cross–validation method in (4.3), the choice of δ_n in (4.8) satisfies the theoretical requirements. Table 4.1 below gives the simulated sizes and power values at the level of $r = 1\%$ and 5%.

Table 4.1: Bootstrap with $M_b = 250$ and $M = 1000$ for Example 4.1

	$\rho_1 = \rho_2 = \rho_3 = 0$				$\rho_1 = \rho_2 = \rho_3 = 0.9$			
H_0	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
n	1%	5%	1%	5%	1%	5%	1%	5%
100	0.008	0.018	0.008	0.056	0.002	0.012	0.011	0.062
300	0.007	0.027	0.010	0.047	0.005	0.022	0.011	0.066
500	0.007	0.028	0.012	0.049	0.005	0.043	0.009	0.051
H_1	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
100	0.458	0.569	0.544	0.716	0.858	0.906	0.924	0.962
300	0.683	0.759	0.692	0.836	0.929	0.952	0.946	0.967
500	0.760	0.815	0.761	0.863	0.938	0.960	0.977	0.981

Example 4.2. Consider one nonlinear time series model of the form for **Case A**:

$$H_0: \ y_t = \alpha \ e^{-\beta \ x_t^2} + \gamma z_t + e_t, \quad t = 1, 2, \cdots, n, \tag{4.9}$$

versus

$$H_1: \ y_t = \alpha \ e^{-\beta \ x_t^2} + \gamma z_t + \Delta_n(x_t, z_t) + e_t, \quad t = 1, 2, \cdots, n, \tag{4.10}$$

and another nonlinear time series model of the form for **Case B**:

$$H_0: \ y_t = \alpha \ (1 + x_t^2)^\beta + \gamma z_t + e_t, \quad t = 1, 2, \cdots, n, \tag{4.11}$$

versus

$$H_1: \ y_t = \alpha \ (1 + x_t^2)^\beta + \gamma z_t + \Delta_n(x_t, z_t) + e_t, \quad t = 1, 2, \cdots, n, \tag{4.12}$$

where $x_t = x_{t-1} + u_t$ with $x_0 = 0$, $\alpha = \beta = \gamma = \frac{1}{2}$, and $\{(e_t, u_t, z_t) : 1 \leq t \leq n\}$ are independent and identically distributed as

$$\begin{pmatrix} e_t \\ u_t \\ z_t \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix} \right), \tag{4.13}$$

with $\rho_i = 0$ or $\rho_i = 0.9$ for $i = 1, 2, 3$, and

$$\Delta_n(x, z) = \frac{\delta_n\ z^2}{\sqrt{1+x^2}} \quad \text{with } \delta_n = \tfrac{\log(n)}{2n^{\frac{1}{8}}}. \tag{4.14}$$

The choice of δ_n is the same as in (4.8). Tables 4.2 below gives the simulated sizes and power values at the level of $r = 1\%$ and 5% for both **Case A** and **Case B**.

Table 4.2a: Bootstrap with $M_b = 250$ and $M = 1000$ for Case A in Example 4.2

	$\rho_1 = \rho_2 = \rho_3 = 0$				$\rho_1 = \rho_2 = \rho_3 = 0.9$			
H_0	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
n	1%	5%	1%	5%	1%	5%	1%	5%
100	0.009	0.022	0.018	0.055	0.009	0.026	0.011	0.051
300	0.007	0.028	0.015	0.063	0.003	0.022	0.009	0.064
500	0.014	0.032	0.014	0.050	0.014	0.030	0.006	0.046
H_1	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
100	0.491	0.578	0.594	0.706	0.827	0.882	0.871	0.912
300	0.692	0.754	0.853	0.901	0.938	0.962	0.941	0.967
500	0.789	0.838	0.854	0.908	0.945	0.962	0.967	0.980

Table 4.2b: Bootstrap with $M_b = 250$ and $M = 1000$ for Case B in Example 4.2

	$\rho_1 = \rho_2 = \rho_3 = 0$				$\rho_1 = \rho_2 = \rho_3 = 0.9$			
H_0	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
n	1%	5%	1%	5%	1%	5%	1%	5%
100	0.004	0.008	0.021	0.070	0.004	0.017	0.024	0.071
300	0.010	0.026	0.013	0.059	0.005	0.017	0.015	0.065
500	0.007	0.018	0.015	0.049	0.003	0.016	0.012	0.054
H_1	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
100	0.565	0.638	0.619	0.727	0.895	0.930	0.895	0.950
300	0.696	0.767	0.723	0.849	0.931	0.955	0.940	0.974
500	0.782	0.824	0.809	0.908	0.961	0.967	0.972	0.985

REMARK 4.1 (i) Tables 4.1 and 4.2 show that both $\widehat{L}_{1n}(h_1, h_2)$ and $\widehat{L}_{2n}(h_1, h_2)$ work well numerically even though the sample size is as small as $n = 100$. Meanwhile, Tables 4.1–4.2 show that the proposed test $\widehat{L}_{2n}(h_1, h_2)$ is more powerful than $\widehat{L}_{1n}(h_1, h_2)$ when the alternative form is chosen as in (4.8).

(ii) In both examples, we also used an asymptotic critical value and the fixed bandwidths $h_1 = n^{-\frac{1}{12}}$ and $h_2 = n^{-\frac{1}{6}}$ in each case. For $\widehat{L}_{1n}$, we used $z_{0.01} = 2.33$ at the 1% level and $z_{0.05} = 1.645$ at the 5% level. For $\widehat{L}_{2n}$, we used the critical value, l_r, of $\sigma^2(K)\ \widehat{\sigma}_e^2\ L_B(1, 0)$ at the 1% level and at the 5% level. Our simulation results show that both tests have relatively stable sizes and good power values, but both tests are slightly under sized and the power values are uniformly smaller than the corresponding values reported in Tables 4.1 and 4.2.

(iii) Tables 4.1 and 4.2 also show that both $\widehat{L}_{1n}(h_1, h_2)$ and $\widehat{L}_{2n}(h_1, h_2)$ work well when there is endogeneity between (x_t, z_t) and e_t, although this case has not been covered in the theory for either $\widehat{L}_{1n}(h_1, h_2)$ or $\widehat{L}_{2n}(h_1, h_2)$. Similar observations from their simulation study are given in Wang and Phillips (2012) for the univariate version of $\widehat{L}_{1n}(h_1, h_2)$. A positive implication of Tables 4.1 and 4.2 is that the asymptotic theory may remain true for the case where there is an endogeneity between (x_t, z_t) and e_t. Since both test statistics involve quadratic forms of nonstationary time series, it is not clear whether one could extend existing methods (such as, Wang and Phillips 2009b) to rigorously prove Theorem 2.1 in an endogeneity case.

5 An empirical application

Example 5.1. This example considers the data set from the Bureau of Economic Analysis (USA Economic Accounts) at the website: http://www.bea.gov/. Let c_t = log(consumption expenditure), I_t = log(disposable income), z_t =(nominal interest rate) or $z_t = w_t$ =(real interest rate). Note that the data sets used were quarterly data of 199 observations. The period considered here is from the first quarter of 1960 to the last quarter of 2009. Note also that the real interest rate was calculated by deducting the inflation rate over the following quarter from the nominal interest rate. Figures 1 and 2 below give the plots of the relevant data sets.

Let $y_t = c_t$, $x_t = I_t$, and $\{w_t\}$ is the real interest rate. Consider using a simple linear model of the form

$$y_t = \alpha + \beta x_t + \gamma w_t + e_t, \tag{5.1}$$

where (α, β, γ) is a vector of unknown parameters.

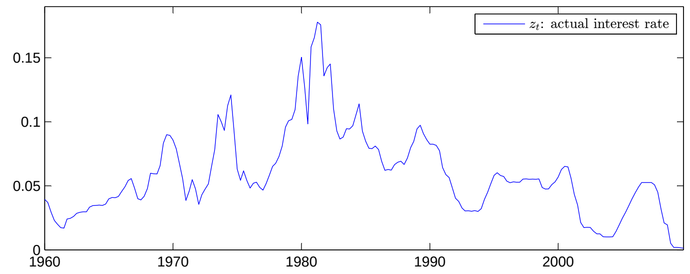

Figure 1a. Nominal interest rate

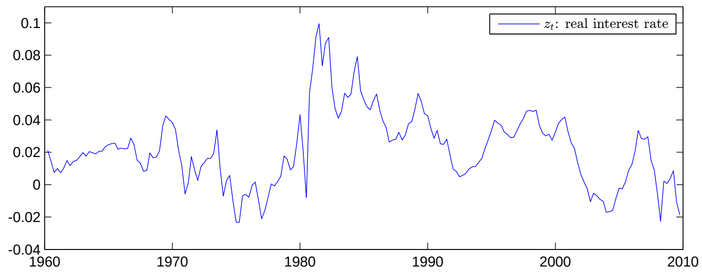

Figure 1b. Real interest rate

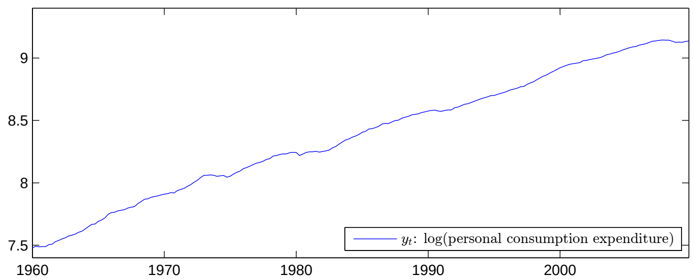

Figure 2a. Plot of c_t

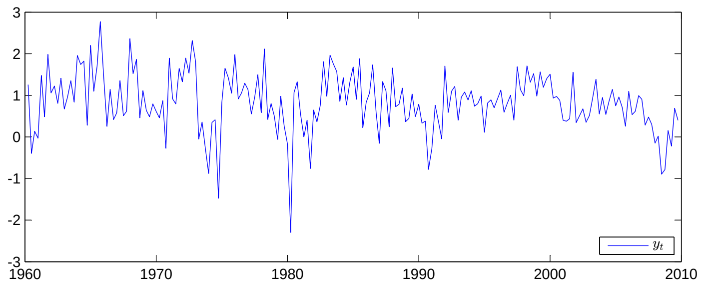

Figure 2b. Plot of $c_t - c_{t-1}$

Meanwhile, there is a growing literature (see, for example, Gylfason 1981; Faff and Brooks 1998; Hahm and Steigerwald 1999; Cai, Li and Park 2009; Xiao 2009) to support that $\beta = \beta(\cdot)$ should be treated as a function of w_t. With regard to the issue of which particular form should be chosen for $\beta(\cdot)$, polynomial functions have been commonly used (see, for example, Faff and Brooks 1998).

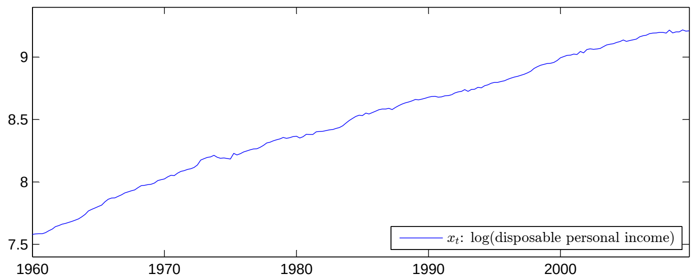

Figure 2c. Plot of I_t

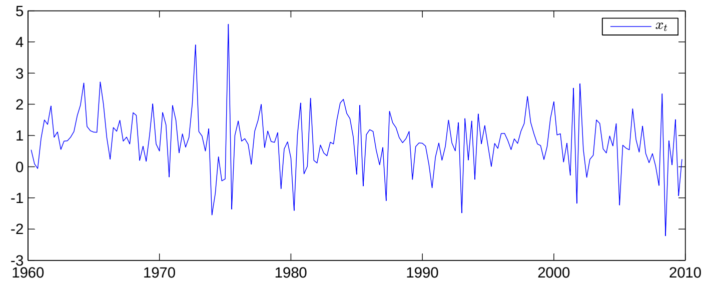

Figure 2d. Plot of $I_t - I_{t-1}$

This section then proposes using a varying–coefficient model of the form

$$y_t = \alpha + \beta(w_t)x_t + \gamma w_t + e_t, \tag{5.2}$$

where $\beta(\cdot)$ is an unknown function, and γ is still an unknown parameter.

Existing estimation methods (see, for example, Chapter 2 of Gao 2007) then produce a semiparametric estimate of the form $\widehat{\beta}(w)$. Its plot is given in Figure 3 below. Meanwhile, a second–order polynomial approximate form, $\widetilde{\beta}(w)$, of $\widehat{\beta}(w)$ is also given in Figure 3.

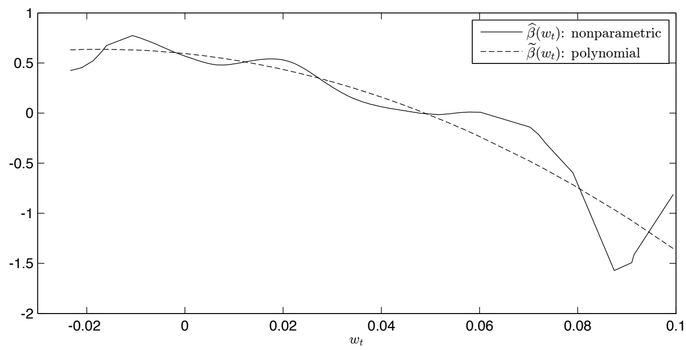

Figure 3. Plots of $\widehat{\beta}(w)$ and $\widetilde{\beta}(w)$

Figure 3, along with model (5.2), motivates us to rigorously support this parametric specification by testing

$$
\begin{aligned}
H_0: \ & y_t = \theta_0 + \theta_1 x_t + \theta_2 x_t w_t + \theta_3 x_t w_t^2 + \theta_4 w_t + e_t \quad \text{versus} \\
H_1: \ & y_t = \theta_0 + \theta_1 x_t + \theta_2 x_t w_t + \theta_3 x_t w_t^2 + \theta_4 w_t + \Delta_n(x_t, z_t) + e_t, \qquad (5.3)
\end{aligned}
$$

where $\Delta_n(x, z)$ is probably unknown and can be estimated under H_1.

An application of the proposed tests $\widehat{L}_{1n}(h_1, h_2)$ and $\widehat{L}_{2n}(h_1, h_2)$ shows that the simulated P–values are 0.1024 and 0.1436, respectively. This indicates that there is some evidence to suggest accepting a second–order polynomial form to approximate $\beta(w)$.

Models (5.2) and (5.3) show that the slope parameter $\beta(w)$ should be treated as a second–order polynomial function of w_t rather than as a constant parameter. In other words, a simple linear model of the form:

$$y_t - y_{t-1} = \alpha_0 + \alpha_1 (x_t - x_{t-1}) + \alpha_2 w_t + \varepsilon_t \qquad (5.4)$$

commonly used in the literature (see, for example, Campbell and Mankiw 1990; Campbell, Lo and MacKinlay 1997), may not be justifiable and suitable for such data sets.

Let $\widetilde{y}_t = y_t - y_{t-1}$ and $\widetilde{x}_t = x_t - x_{t-1}$. In view of models (5.1)–(5.4), we propose using a first–order polynomial function of w_t to replace α_1 in model (5.4) and then compare the following model

$$\widetilde{y}_t = \beta_0 + (\beta_1 + \beta_2 w_t)\, \widetilde{x}_t + \beta_3 w_t + \eta_t \qquad (5.5)$$

with a commonly used model of the form

$$\widetilde{y}_t = \gamma_0 + \gamma_1 \widetilde{x}_t + \gamma_2 w_t + \zeta_t. \qquad (5.6)$$

The estimated versions of models (5.5) and (5.6) become respectively

$$
\begin{aligned}
\widetilde{y}_t &= 0.554 + (0.345 - 0.366 w_t)\, \widetilde{x}_t + 0.116 w_t \quad \text{and} \qquad (5.7) \\
\widetilde{y}_t &= 0.558 + 0.341 \widetilde{x}_t - 0.199 w_t, \qquad (5.8)
\end{aligned}
$$

with the estimated standard deviations of the parameter estimates being between 0.0241 and 0.0632. An application of $\widehat{L}_{1n}(h_1, h_2)$ to check whether either model (5.5) or model (5.6) is appropriate as a parametric model gives the simulated P–values of 0.1209 and 0.04387, respectively. This indicates that there is some evidence to support using model (5.5) in practice for such data sets.

In summary, our findings show that the β parameter involved in model (5.1) should be treated as a varying–coefficient function of w and an appropriately chosen polynomial form may be appropriate for this kind of empirical analysis.

6 Conclusions and discussions

We have proposed a new testing method for a general model specification in a nonlinear time series model with multivariate regressors. A new asymptotic theory has been established for the proposed test. Simulated examples have been used to evaluate the finite–sample performance of the proposed test as well as comparison with a natural competitor. Meanwhile, the proposed test has also been applied to test the suitability of a simple linear model commonly used in the consumption–income literature.

Further studies are needed in the following two directions. The first is how to theoretically justify the applicability of the proposed test for the case where there is some endogeneity between the regressors and the errors, as the simulated studies support the employability of the proposed test. The second direction is whether the new approach proposed in this paper is applicable to deal with the autoregressive case where $x_t = y_{t-1}$ in model (1.1). Gao and King (2012) discuss two special forms of $m(x, z)$ in the autoregressive case. Such issues are left for future research.

7 Acknowledgments

The authors acknowledge constructive comments and suggestions from participants at several seminars and conferences, particularly to those who attended the seminar held at Monash University in April 2012, and then two conferences held at Yale University in June 2012 and Xiamen University in July 2012, respectively. The authors also acknowledge constructive discussions with Dr Qiying Wang. Thanks also go to the Australian Research Council Discovery Grants Program for its support under Grant number: DP1096374.

8 Appendix

Appendix A then gives some useful lemmas before the proof of Theorem 2.1 is given. The proofs of these lemmas, along with the derivations of (3.5) and (3.6), are then given in Appendix B. Note that some of the derivations in Appendix B are based on techniques explained in the proofs of Appendix A.

Appendix A

Lemma A.1. *Let the conditions of Theorem 2.1 hold. Under H_0, we then have as $n \to \infty$*

$$\begin{aligned}
\widehat{S}_{1n} = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}K_1^2\left(\frac{x_t-x}{h_1}\right)K_2^2\left(\frac{z_t-z}{h_2}\right)\widehat{e}_t^2\right)\pi_1(x)\pi_2(z)dzdx \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}K_1^2\left(\frac{x_t-x}{h_1}\right)K_2^2\left(\frac{z_t-z}{h_2}\right)\ e_t^2\right)\pi_1(x)\pi_2(z)dzdx + o_P(1) \\
\equiv & \ S_{1n} + o_P(1), \qquad \text{(A.1)} \\
\widehat{S}_{2n} = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n}\widehat{e}_s\widehat{e}_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right) \\
\times & \ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n}e_s\ e_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right) \\
\times & \ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx + o_P(1) \equiv S_{2n} + o_P(1). \qquad \text{(A.2)}
\end{aligned}$$

Lemma A.2. *Let the conditions of Theorem 2.1 hold. If, in addition, equations (3.3) and (3.4) are satisfied, under H_1, we have as $n \to \infty$*

$$\begin{aligned}
\widehat{S}_{1n} = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}K_1^2\left(\frac{x_t-x}{h_1}\right)K_2^2\left(\frac{z_t-z}{h_2}\right)\widehat{e}_t^2\right)\pi_1(x)\pi_2(z)dzdx \\
\geq & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}K_1^2\left(\frac{x_t-x}{h_1}\right)K_2^2\left(\frac{z_t-z}{h_2}\right)\right) \\
\times & \ \Delta_n^2(x_t,z_t)\pi_1(x)\pi_2(z)dzdx + o_P(1), \qquad \text{(A.3)} \\
\widehat{S}_{2n} = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n}\widehat{e}_s\widehat{e}_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right) \\
\times & \ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
\geq & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n}K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right) \\
\times & \ \Delta_n(x_s,z_s)\Delta_n(x_t,z_t)\pi_1(x)\pi_2(z)dzdx + o_P(1). \qquad \text{(A.4)}
\end{aligned}$$

Lemma A.3. *Let Assumptions 2.1, 2.2 and 2.4 hold. Then as $n \to \infty$*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\psi_1(x_t)\psi_2(z_t) \to_D L_{B_u}(1,0)\cdot\sigma_e^2\cdot\int_{-\infty}^{\infty}\psi_1(x)dx\cdot\int_{-\infty}^{\infty}\psi_2(z)p(z)dz, \qquad \text{(A.5)}$$

where $\psi_i(\cdot) = \pi_i(\cdot)$ or $D_i(\cdot)$ for $i = 1, 2$ are all defined in Assumption 2.4.

Before the proofs of Lemmas A.1–A.3 are given in Appendix B below, we give the proof of Theorem 2.1. Note that Lemma A.3 and its proof may be of general interest. Without loss of generality, we let $\sigma_e^2 \equiv 1$ throughout Appendices A and B.

Proof of Theorem 2.1. In view of Lemma A.1, in order to prove Theorem 2.1, it suffices to show that as $n \to \infty$

$$S_{1n} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}K_1^2\left(\frac{x_t-x}{h_1}\right)K_2^2\left(\frac{z_t-z}{h_2}\right)\ e_t^2\right)\pi_1(x)\pi_2(z)dzdx$$
$$\to_D L_{B_u}(1,0)\cdot\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)dvdu\right)\cdot\int_{-\infty}^{\infty}\pi_2(z)p(z)dz, \tag{A.6}$$
$$S_{2n} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n}e_s\ e_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx = o_P(1). \tag{A.7}$$

We start with the proof of (A.6). Under Assumptions 2.1, 2.2 and 2.4, using Lemma A.3, we have as $n \to \infty$

$$S_{1n} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)\pi_1(x_t-uh_1)\pi_2(z_t-vh_2)dudv$$
$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)\pi_1(x_t)\pi_2(z_t)dudv$$
$$+\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)\left(\pi_1(x_t-uh_1)-\pi_1(x_t)\right)\left(\pi_2(z_t-vh_2)-\pi_2(z_t)\right)dudv$$
$$+\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)\left(\pi_1(x_t-uh_1)-\pi_1(x_t)\right)\pi_2(z_t)dvdu$$
$$+\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)\pi_1(x_t)\left(\pi_2(z_t-vh_2)-\pi_2(z_t)\right)dudv$$
$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\pi_1(x_t)\pi_2(z_t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)dudv$$
$$+\ O(h_1h_2^d)\cdot\frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2D_1(x_t)D_2(z_t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u||v|K_1^2(u)K_2^2(v)dvdu$$
$$+\ O(h_1)\cdot\frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2D_1(x_t)\pi_2(z_t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u|K_1^2(u)K_2^2(v)dvdu$$
$$+\ O(h_2^d)\cdot\frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\pi_1(x_t)D_2(z_t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|v|K_1^2(u)K_2^2(v)dvdu$$
$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}e_t^2\pi_1(x_t)\pi_2(z_t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)dudv + o_P(1)$$
$$\to_D\ L_{B_u}(1,0)\cdot\int_{-\infty}^{\infty}\pi_2(z)p(z)dz\cdot\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)dudv, \tag{A.8}$$

which completes the proof of (A.6).

We then prove (A.7). Let

$$B(s,t) \equiv B(x_s, x_t; z_s, z_t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right)$$
$$\times\ \pi_1(x)\pi_2(z)dzdx,$$
$$A_t \equiv A(x_1, \cdots, x_t; z_1, \cdots, z_t; e_1, \cdots, e_{t-1}) = \frac{2}{h_1 h_2^d}\sum_{s=1}^{t-1} B(s,t)\ e_s,$$
$$S_{2n} = \frac{1}{\sqrt{n}}\sum_{t=2}^{n} A_t\ e_t. \tag{A.9}$$

Similarly to the derivations in (A.8), we have

$$B(s,t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1\left(\frac{x_t - x_s}{h_1} + \frac{x_s - x}{h_1}\right) K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2} + \frac{z_s - z}{h_2}\right) K_2\left(\frac{z_s - z}{h_2}\right)$$
$$\times\ \pi_1(x)\pi_2(z)dzdx$$
$$=\ h_1 h_2^d \cdot \int_{-\infty}^{\infty} K_1\left(\frac{x_t - x_s}{h_1} + u\right) K_1(u)\, \pi_1(x_s - uh_1)du$$
$$\times\ \int_{-\infty}^{\infty} K_2\left(\frac{z_t - z_s}{h_2} + v\right) K_2(v)\, \pi_2(z_s - vh_2)dv$$
$$=\ h_1 h_2^d\ \pi_1(x_s)\ \pi_2(z_s) \cdot L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right)$$
$$+\ h_1 h_2^d \cdot \int_{-\infty}^{\infty} K_1\left(\frac{x_t - x_s}{h_1} + u\right) K_1(u)\left(\pi_1(x_s - uh_1) - \pi_1(x_s)\right) du$$
$$\times\ \int_{-\infty}^{\infty} K_2\left(\frac{z_t - z_s}{h_2} + v\right) K_2(v)\left(\pi_2(z_s - vh_2) - \pi_2(z_s)\right) dv$$
$$+\ h_1 h_2^d\ \pi_2(z_s) \cdot L_2\left(\frac{z_t - z_s}{h_2}\right) \cdot \int_{-\infty}^{\infty} K_1\left(\frac{x_t - x_s}{h_1} + u\right) K_1(u)\left(\pi_1(x_s - uh_1) - \pi_1(x_s)\right) du$$
$$+\ h_1 h_2^d\ \pi_1(x_s) \cdot L_1\left(\frac{x_t - x_s}{h_1}\right) \cdot \int_{-\infty}^{\infty} K_2\left(\frac{z_t - z_s}{h_2} + v\right) K_2(v)\left(\pi_2(z_s - vh_2) - \pi_2(z_s)\right) dv$$
$$\equiv\ h_1 h_2^d \left(B_1(s,t) + B_2(s,t) + B_3(s,t) + B_4(s,t)\right), \tag{A.10}$$

where $L_i(v) = \int_{-\infty}^{\infty} K_i(u+v)K_i(u)du$ for $i = 1, 2$.

Under Assumptions 2.1 and 2.2, we then have

$$E\left[S_{2n}^2\right] = \frac{1}{n}\sum_{t=2}^{n} E[A_t^2] + \frac{2}{n}\sum_{t_1=2}^{n}\sum_{t_2=2}^{t_1-1} E\left(A_{t_1}A_{t_2}e_{t_2}E\left[e_{t_1}|\mathcal{F}_{n,t_1-1}\right]\right]$$
$$= \frac{4}{nh_1^2h_2^{2d}}\sum_{t=2}^{n}\sum_{s=1}^{t-1} E\left[B^2(s,t)e_s^2\right] + \frac{8}{nh_1^2h_2^{2d}}\sum_{t=3}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1} E\left[B(s_1,t)B(s_2,t)e_{s_2}E\left[e_{s_1}|\mathcal{F}_{n,s_1-1}\right]\right]$$
$$= \frac{4}{nh_1^2h_2^{2d}}\sum_{t=2}^{n}\sum_{s=1}^{t-1} E\left[B^2(s,t)E\left[e_s^2|\mathcal{F}_{n,s-1}\right]\right] = \frac{4}{nh_1^2h_2^2}\sum_{t=2}^{n}\sum_{s=1}^{t-1} E\left[B^2(s,t)\right]. \tag{A.11}$$

In view of (A.10) and (A.11), in order to evaluate the order of $E\left[S_{2n}^2\right]$, we need to involve the joint distribution of $(x_t, x_s; z_t, z_s)$. Recall that we define, for $s < t$, the joint density functions of $(x_t, x_s; z_t, z_s)$ and (z_t, z_s) by $p(u_2, u_1; v_2, v_1)$ and $p(v_2, v_1)$, respectively, and the joint density function of $\left(\frac{x_t}{\sqrt{t}}, \frac{x_s}{\sqrt{s}}; z_t, z_t\right)$ by $q(u_2, u_1; v_2, v_1)$. Then, we have

$$p(u_2, u_1; v_2, v_1) = \frac{1}{\sqrt{s}\sqrt{t}} q\left(\frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}; v_2, v_1\right) = \frac{1}{\sqrt{s}\sqrt{t}} q\left(\frac{u_2}{\sqrt{t}}, \frac{u_1}{\sqrt{s}}|v_2, v_1\right) p(v_2, v_1), \tag{A.12}$$

where $q(\cdot,\cdot|v_2,v_1)$ denotes the conditional density of $\left(\frac{x_t}{\sqrt{t}},\frac{x_s}{\sqrt{s}}\right)$ given $(z_t,z_s)=(v_2,v_1)$. For notational simplicity, the notation involves no t and s as indices.

In view of equation (A.12), by Assumptions 2.1(ii) and 2.4 in particular, we then have as $n\to\infty$

$$
\begin{aligned}
&\sum_{t=2}^{n}\sum_{s=1}^{t-1}E\left[B_1^2(s,t)\right]=\sum_{t=2}^{n}\sum_{s=1}^{t-1}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\pi_1^2(u_1)\pi_2^2(v_1)L_1^2\left(\frac{u_2-u_1}{h_1}\right)L_2^2\left(\frac{v_2-v_1}{h_2}\right)\\
&\times p(u_2,u_1;v_2,v_1)du_2du_1dv_2dv_1\\
&= h_1h_2^d\sum_{t=2}^{n}\sum_{s=1}^{t-1}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\pi_1^2(u_1)\pi_2^2(v_1)L_1^2(u_3)L_2^2(v_3)p(u_1+u_3h_1,u_1;v_1+v_3h_2,v_1)\\
&\times du_3du_1dv_3dv_1\\
&= h_1h_2^d\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{t}\sqrt{s}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\pi_1^2(u_1)\pi_2^2(v_1)L_1^2(u_3)L_2^2(v_3)\\
&\times q\left(\frac{u_1+u_3h_1}{\sqrt{t}},\frac{u_1}{\sqrt{s}}|v_1+v_3h_2,v_1\right)p(v_1+v_3h_2,v_1)\ du_3du_1dv_3dv_1\\
&\leq Cnh_1h_2^d\cdot\int_{-\infty}^{\infty}\pi_1^2(x)dx\cdot\int_{-\infty}^{\infty}\pi_2^2(z)p^2(z)dz\cdot\int_{-\infty}^{\infty}L_1^2(u)du\cdot\int_{-\infty}^{\infty}L_2^2(v)dv. \qquad \text{(A.13)}
\end{aligned}
$$

Similarly, we have as $n\to\infty$

$$J_{2n}=O\left(nh_1h_2^d\right). \qquad \text{(A.14)}$$

Equations (A.11)–(A.14) then imply as $n\to\infty$

$$\frac{1}{n}\sum_{t=2}^{n}\sum_{s=1}^{t-1}E\left[B_1^2(s,t)\right]\leq Ch_1h_2^d. \qquad \text{(A.15)}$$

In a similar way to the derivation of (A.15), using Assumption 2.4(iv) in particular, we have for $j=2,3,4$

$$\frac{1}{n}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\left(E\left[B_2^2(s,t)\right]+E\left[B_3^2(s,t)\right]+E\left[B_4^2(s,t)\right]\right)\leq Ch_1h_2^d\cdot(h_1h_2^d+h_1+h_2^d). \qquad \text{(A.16)}$$

Therefore, equations (A.10)–(A.11) and (A.15)–(A.16) imply that as $n\to\infty$

$$E\left[S_{2n}^2\right]\leq Ch_1h_2^d,\quad \text{which deduces } S_{2n}=o_P(1). \qquad \text{(A.17)}$$

The proof of Theorem 2.1 follows from equations (A.6), (A.7) and (A.17).

Appendix B

This appendix gives the proofs of Lemmas A.1–A.3 and then the derivations of equations (3.5) and (3.6) are given in the last part of this appendix.

We first introduce a very useful lemma, which has been used in the proof of Theorem 2.1. The proof of Lemma B.1 below follows from existing results for central limit theorems (see, for example, Rényi 1958; Dam 1998; Denker and Gordin 2003; and Peligrad and Utev 2006). The first part of Lemma B.1 below is a standard central limit theorem. The second part, as pointed out by Rényi

(1958), indicates that a mixing sequence of random variables is asymptotically independent of any random variable.

Let $(\Omega, \mathcal{B}_{zu}, P)$ be a probability space and $(z_i, u_i): \ \Omega \to R^2$ be a vector of stationary time series. Let $\widehat{\phi}_k(x)$ be the probability density function of $L_k = \frac{1}{\sqrt{k}\ \sigma_u}\sum_{j=1}^{k} u_j$ and $\widehat{\phi}_k(x|\mathcal{B}_{zu})$ be the conditional probability density function of L_k given $B \in \mathcal{B}_{zu}$ with $P(B) > 0$, where $\sigma_u^2 = \text{var}(u_1)$.

Lemma B.1. *Under Assumption 2.1, we have for any $B \in \mathcal{B}_{zu}$ with $P(B) > 0$ and as $k \to \infty$*

$$\sup_{x\in R^1}\left|\widehat{\phi}_k(x) - \phi(x)\right| \to 0 \quad \text{and} \quad \sup_{x\in R^1}\left|\widehat{\phi}_k(x|B) - \phi(x)\right| \to 0, \tag{B.1}$$

almost surely, where $\phi(\cdot)$ is the probability density of the standard normal random variable $U \sim N(0,1)$.

Proof of Lemma A.1. We only provide the proof of equation (A.2), as the proof of (A.1) follows similarly. Recall that under H_0:

$$\widehat{e}_t = y_t - g(x_t, z_t; \widehat{\theta}) = e_t + g(x_t, z_t; \theta_0) - g(x_t, z_t; \widehat{\theta}) \equiv e_t + r_n(x_t, z_t; \theta_0), \tag{B.2}$$

where $r_n(x, z; \theta_0) = g(x, z; \theta_0) - g(x, z, \widehat{\theta})$.

Define

$$\begin{aligned}
\widehat{T}_n \equiv \ & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1 h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}\widehat{e}_s\widehat{e}_t \cdot K_1\left(\frac{x_t - x}{h_1}\right)K_2\left(\frac{z_t - z}{h_2}\right) \\
\times\ & K_1\left(\frac{x_s - x}{h_1}\right)K_2\left(\frac{z_s - z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
= \ & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1 h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} e_s e_t \cdot K_1\left(\frac{x_t - x}{h_1}\right)K_2\left(\frac{z_t - z}{h_2}\right) \\
\times\ & K_1\left(\frac{x_s - x}{h_1}\right)K_2\left(\frac{z_s - z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
+ \ & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1 h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} e_s r_n(x_t, z_t; \theta_0) \cdot K_1\left(\frac{x_t - x}{h_1}\right)K_2\left(\frac{z_t - z}{h_2}\right) \\
\times\ & K_1\left(\frac{x_s - x}{h_1}\right)K_2\left(\frac{z_s - z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
+ \ & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1 h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} e_t r_n(x_s, z_s; \theta_0) \cdot K_1\left(\frac{x_t - x}{h_1}\right)K_2\left(\frac{z_t - z}{h_2}\right) \\
\times\ & K_1\left(\frac{x_s - x}{h_1}\right)K_2\left(\frac{z_s - z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
+ \ & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1 h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} r_n(x_t, z_t; \theta_0) r_n(x_s, z_s; \theta_0) \cdot K_1\left(\frac{x_t - x}{h_1}\right)K_2\left(\frac{z_t - z}{h_2}\right) \\
\times\ & K_1\left(\frac{x_s - x}{h_1}\right)K_2\left(\frac{z_s - z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
\equiv \ & \widehat{T}_{1n} + \widehat{T}_{2n} + \widehat{T}_{3n} + \widehat{T}_{4n}.
\end{aligned} \tag{B.3}$$

By Assumption 2.3(ii), we have for some $\epsilon > 0$ and $\delta > 0$, $P\left(||\widehat{\theta} - \theta_0|| > \epsilon\right) < \delta$, as $n \to \infty$. We thus consider the case where $||\widehat{\theta} - \theta_0|| \le \epsilon$ holds in probability in the following derivations. Using

Assumptions 2.3(i) and 2.4(iv) in particular, in view of the second equality of equation (A.10), we then have

$$
\begin{aligned}
\widehat{T}_{4n} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} r_n(x_t,z_t;\theta_0)r_n(x_s,z_s;\theta_0)\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right) \\
&\times K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\
&\leq ||\widehat{\theta}-\theta_0||^2\cdot\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}||G_1(x_t,z_t;\theta_0)||\cdot||G_1(x_s,z_s;\theta_0)||\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \\
&\times K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx+R_n \\
&= \ ||\widehat{\theta}-\theta_0||^2\cdot\frac{(1+o_P(1))}{\sqrt{n}}\sum_{t=1}^{n}\sum_{s=1}^{n}||G_1(x_t,z_t;\theta_0)||\,||G_1(x_s,z_s;\theta_0)||\,\pi_1(x_s)\pi_2(z_s) \\
&\times L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)+R_n\equiv Q_n+R_n, \qquad \text{(B.4)}
\end{aligned}
$$

where R_n is the remainder term that involves $\left(G_2(\cdot,\cdot;\theta_0)\left|\left|\widehat{\theta}-\theta_0\right|\right|^2\right)^2$, which is of an order higher than Q_n.

Similarly to the derivations in equation (A.13), we have as $n\to\infty$

$$
\begin{aligned}
&\sum_{t=1}^{n}\sum_{s=1}^{n}E\left[||G_1(x_t,z_t;\theta_0)||\,||G_1(x_s,z_s;\theta_0)||\,\pi_1(x_s)\pi_2(z_s)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\right] \\
&\leq Cnh_1h_2^d, \qquad \text{(B.5)}
\end{aligned}
$$

which, along with Assumption 2.3(ii) and equation (B.4), implies that $\widehat{T}_{4n}=o_P(1)$.

To show that $\widehat{T}_{jn}=o_P(1)$ for $j=2,3$, we will need to repeatedly use Assumption 2.3(i) and then Assumption 2.3(ii). Without loss of generality, we assume that the dimensionality of Θ is $c=1$ in the following derivations. Similarly to (B.4), we have

$$
\begin{aligned}
\widehat{T}_{2n} &= \frac{(1+o_P(1))}{\sqrt{n}}\sum_{t=1}^{n}\sum_{s=1}^{n}e_t r_n(x_s,z_s;\theta_0)\cdot L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s) \\
&= \frac{(1+o_P(1))}{\sqrt{n}}\left(\theta_0-\widehat{\theta}\right)\sum_{t=1}^{n}\left(\sum_{s=1}^{n}G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right)e_t \\
&+ \frac{(1+o_P(1))}{\sqrt{n}}\sum_{t=1}^{n}\sum_{s=1}^{n}\left(r_n(x_s,z_s;\theta_0)-G_1(x_s,z_s;\theta_0)\left(\theta_0-\widehat{\theta}\right)\right)e_t \\
&\times L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s) \\
&\equiv \frac{(1+o_P(1))}{\sqrt{n}}\left(\theta_0-\widehat{\theta}\right)\cdot I_{1n}+\frac{(1+o_P(1))}{\sqrt{n}}\cdot I_{2n}. \qquad \text{(B.6)}
\end{aligned}
$$

In a similar fashion to the derivations in (A.11), (A.13) and (B.5), using Assumptions 2.1 and

2.2, we have as $n \to \infty$

$$
\begin{aligned}
E\left[I_{1n}^2\right] &= E\left[\sum_{t=1}^{n}\left(\sum_{s=1}^{n} G_1(x_s, z_s;\theta_0) L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right) \pi_1(x_s)\pi_2(z_s)\right) e_t\right]^2 \\
&= \sum_{t=1}^{n}\sum_{s=1}^{n} E\left[G_1^2(x_s, z_s;\theta_0) L_1^2\left(\frac{x_t - x_s}{h_1}\right) L_2^2\left(\frac{z_t - z_s}{h_2}\right) \pi_1^2(x_s)\pi_2^2(z_s)\right] \\
&\leq\ Cn\ h_1 h_2^d, \qquad\qquad (B.7)
\end{aligned}
$$

where Assumptions 2.3(i) and 2.4(i) have also been used in obtaining the last equation in (B.7).

Meanwhile, by Assumption 2.3(i) we have

$$
\begin{aligned}
|I_{2n}| &\leq \sum_{t=1}^{n}\sum_{s=1}^{n}\left| r_n(x_s, z_s;\theta_0) - G_1(x_s, z_s;\theta_0)\left(\theta_0 - \widehat{\theta}\right)\right| L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right) \\
&\times\ \pi_1(x_s)\pi_2(z_s)\,|e_t| \\
&\leq\ ||\widehat{\theta} - \theta_0||^2 \cdot \sum_{t=1}^{n}\sum_{s=1}^{n} G_2(x_s, z_s;\theta_0) L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right) \pi_1(x_s)\pi_2(z_s)|e_t| \\
&\equiv\ ||\widehat{\theta} - \theta_0||^2 \cdot I_{3n}, \qquad\qquad (B.8)
\end{aligned}
$$

in view of the fact that both $L_i(\cdot)$ and $G_2(x, z;\theta_0)$ are positive.

As in the derivations in (A.11), (A.13), (B.5) and (B.7), using Assumptions 2.1 and 2.2, we have as $n \to \infty$

$$
\begin{aligned}
E[I_{3n}] &= \sum_{t=1}^{n}\sum_{s=1}^{n} E\left(G_2(x_s, z_s;\theta_0) L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right) \pi_1(x_s)\pi_2(z_s) E\left[|e_t||\mathcal{F}_{n,t-1}\right]\right) \\
&\leq C n h_1 h_2^d. \qquad\qquad (B.9)
\end{aligned}
$$

Therefore, equations (B.6)–(B.9), along with Assumption 2.3(ii), imply that $\widehat{T}_{2n} = o_P(1)$. The same conclusion is true for $\widehat{T}_{3n}$. Therefore, we have shown that under H_0, as $n \to \infty$

$$
\begin{aligned}
\widehat{T}_n =&\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n} h_1 h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} \widehat{e}_s \widehat{e}_t \cdot K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) \\
&\times\ K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right) \pi_1(x)\pi_2(z) dz dx \\
=&\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n} h_1 h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n} e_s e_t \cdot K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) \\
&\times\ K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right) \pi_1(x)\pi_2(z) dz dx + o_P(1), \qquad\qquad (B.10)
\end{aligned}
$$

which completes the proof of Lemma A.1.

Proof of Lemma A.2. We only prove equation (A.4), as the proof of (A.3) follows similarly. Let

$$
\widehat{\varepsilon}_t(x, z) = \widehat{e}_t\ K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) \quad \text{and} \quad \widehat{e}_t = e_t + r_n(x_t, z_t;\theta_0) + \Delta_n(x_t, z_t) \qquad (B.11)
$$

under H_1, where $r_n(x,z;\theta_0) = g(x,z;\theta_0) - g(x,z;\widehat{\theta})$ and $\Delta_n(x,z) = \delta_n\Delta(x,z)$ is the same as defined in (3.3) and (3.4).

Then, we have under H_1:

$$\widehat{T}_n = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1h_2^d} \sum_{t=1}^n\sum_{s=1}^n \widehat{e}_s\widehat{e}_t \cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx$$
$$= \sum_{j=1}^4 \widehat{T}_{jn} + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1h_2^d} \sum_{t=1}^n\sum_{s=1}^n \Delta_n(x_s,z_s)\Delta_n(x_t,z_t) \cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx$$
$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1h_2^d} \sum_{t=1}^n\sum_{s=1}^n \Delta_n(x_s,z_s)r_n(x_t,z_t;\theta_0) \cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx$$
$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1h_2^d} \sum_{t=1}^n\sum_{s=1}^n \Delta_n(x_t,z_t)r_n(x_s,z_s;\theta_0) \cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx$$
$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1h_2^d} \sum_{t=1}^n\sum_{s=1}^n \Delta_n(x_t,z_t)e_s \cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx$$
$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n}h_1h_2^d} \sum_{t=1}^n\sum_{s=1}^n \Delta_n(x_s,z_s)e_t \cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \equiv \sum_{j=1}^4 \widehat{T}_{jn} + \sum_{k=5}^9 \widehat{T}_{kn}, \tag{B.12}$$

where $\widehat{T}_{jn}$ for $1 \le j \le 4$ are the same as in (B.3).

In view of the proof of (B.10) and equation (A.4), in order to complete the proof of Lemma A.2, we need only to deal with $\sum_{k=6}^9 \widehat{T}_{kn}$. We will show that under H_1:

$$\widehat{T}_{kn} = o_P\left(\delta_n^2\sqrt{n}h_1h_2^d\right) \quad \text{for } k = 6, \cdots, 9. \tag{B.13}$$

To complete the proof of (B.13), we need only to deal with $\widehat{T}_{6n}$ and $\widehat{T}_{9n}$. Using similar arguments to those used in the proofs of (B.4) and (B.5), we have as $n \to \infty$

$$\sum_{t=1}^n\sum_{s=1}^n E\left[||G_1(x_t,z_t;\theta_0)|| \cdot |Z(x_s,z_s)|\pi_1(x_s)\pi_2(z_s)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\right]$$
$$\le Cnh_1h_2^d, \tag{B.14}$$

which, along with $\frac{||\widehat{\theta}-\theta_0||\sqrt{\sqrt{n}h_1h_2^d}}{\delta_n\sqrt{\sqrt{n}h_1h_2^d}} = o_P(1)$ by Assumption 2.3(ii) and $\lim_{n\to\infty}\delta_n^2\sqrt{n}h_1h_2^d = \infty$, implies that $\widehat{T}_{6n} = o_P\left(\delta_n^2 nh_1h_2^d\right)$.

Meanwhile, similarly to the derivations in (A.13), in order to deal with $\widehat{T}_{9n}$, it suffices to show that as $n \to \infty$

$$\begin{aligned} & E\left[\delta_n \sum_{t=1}^n \left(\sum_{s=1}^n \Delta(x_s, z_s)\pi_1(x_s)\pi_2(z_s) L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right)\right) e_t\right]^2 \\ & \leq C\ \delta_n^2 n h_1 h_2^d, \end{aligned} \tag{B.15}$$

which follows similarly from the proof of (B.7). Equation (B.15) then implies that as $n \to \infty$

$$\widehat{T}_{9n} = O_P\left(\delta_n \sqrt{\sqrt{n} h_1 h_2^d}\right) = o_P\left(\delta_n^2 \sqrt{n} h_1 h_2^d\right), \tag{B.16}$$

which, along with the proofs of (B.3) and (B.12)–(B.16), shows that under H_1, as $n \to \infty$

$$\begin{aligned} \widehat{T}_n = & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n} h_1 h_2^d} \sum_{t=1}^n \sum_{s=1}^n \widehat{e}_s \widehat{e}_t \cdot K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) \\ & \times K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right) \pi_1(x)\pi_2(z) dz dx \\ \geq & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n} h_1 h_2^d} \sum_{t=1}^n \sum_{s=1}^n \Delta_n(x_s, z_s)\Delta_n(x_t, z_t) \cdot K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) \\ & \times K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right) \pi_1(x)\pi_2(z) dz dx + o_P(1), \end{aligned} \tag{B.17}$$

which completes the proof of Lemma A.2.

Proof of Lemma A.3. Let $\psi_3(x_t) = E[\psi_2(z_t)|x_t]$ and $\psi(x_t, z_t) = \psi_1(x_t)\left(\psi_2(z_t) - \psi_3(x_t)\right)$. To be able to use (2.11), we need to consider the case where t is large or not so large individually. Choose some positive integer $m = m(n)$ such that $\lim_{n\to\infty} \frac{m}{n} = 0$. Observe that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{t=1}^n \psi_1(x_t)\psi_2(z_t) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \psi_1(x_t)\psi_3(x_t) + \frac{1}{\sqrt{n}} \sum_{t=1}^n \psi(x_t, z_t) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^m \psi_1(x_t)\psi_2(z_t) + \frac{1}{\sqrt{n}} \sum_{t=m+1}^n \psi_1(x_t)\psi_3(x_t) + \frac{1}{\sqrt{n}} \sum_{t=m+1}^n \psi(x_t, z_t) \\ &\equiv J_{1n} + J_{2n} + J_{3n}. \end{aligned}$$

The following steps are to prove as $n \to \infty$

$$\begin{aligned} & J_{2n} \to_D L_{B_u}(1,0) \cdot \int_{-\infty} \int_{-\infty}^{\infty} \psi_1(x)\psi_2(z)p(z) dz dx, \\ & J_{in} \to_P 0 \quad \text{for } i = 1, 3. \end{aligned} \tag{B.18}$$

We first deal with the following term:

$$J_{1n} = \frac{1}{\sqrt{n}} \sum_{t=1}^m \psi_1(x_t)\psi_2(z_t). \tag{B.19}$$

Recall from Assumption 2.1(ii) that $f_t(\cdot)$ is the density function of x_t, $g_t(\cdot)$ is the density function of $\frac{x_t}{\sqrt{t}}$, $p_t(x|z)$ is the conditional density of x_t given $z_t = z$ and $q_t(\cdot|z)$ is the conditional density of $\frac{x_t}{\sqrt{t}}$ given $z_t = z$. By Assumption 2.1(ii), we then have

$$
\begin{aligned}
E\left[|\psi_1(x_t)| \cdot |\psi_2(z_t)|\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |p_1(x)|\ |\psi_2(z)|\ p_t(x|z)p(z)dzdx \\
&= \frac{1}{\sqrt{t}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\psi_1(x)| \cdot |\psi_2(z)| q_t\left(\frac{x}{\sqrt{t}}|z\right) p(z)dzdx \\
&= \frac{1}{\sqrt{t}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\psi_1(x)| \cdot |\psi_2(z)| q_t\left(z|\frac{x}{\sqrt{t}}\right) g_t\left(\frac{x}{\sqrt{t}}\right) dzdx \\
&\leq \frac{C}{\sqrt{t}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\psi_1(x)|\ |\psi_2(z)|\ p(z)dzdx,
\end{aligned}
\tag{B.20}
$$

which implies as $n \to \infty$

$$
E\left[|J_{1n}|\right] \leq \frac{C\sqrt{m}}{\sqrt{n}}\frac{1}{\sqrt{m}}\sum_{t=1}^{m}\frac{1}{\sqrt{t}} \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\psi_1(x)|\ |\psi_2(z)| p(z)dzdx = O\left(\frac{\sqrt{m}}{\sqrt{n}}\right) = o(1). \tag{B.21}
$$

Thus, in order to complete the proof of Lemma A.3, in view of equations (B.18)–(B.21), it suffices to show that as $n \to \infty$

$$
\frac{1}{\sqrt{n}}\sum_{t=m+1}^{n} \psi_1(x_t)\psi_3(x_t) \to_D L_{B_u}(1,0)\int_{-\infty}^{\infty} \psi_1(x) \cdot \int_{-\infty}^{\infty} \psi_2(z)p(z)dz, \tag{B.22}
$$

$$
\frac{1}{\sqrt{n}}\sum_{t=m+1}^{n} \psi(x_t, z_t) \to_P 0. \tag{B.23}
$$

Observe that

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{t=m+1}^{n} \psi_1(x_t)\psi_3(x_t) &= \frac{1}{\sqrt{n}}\sum_{t=m+1}^{n} \psi_1(x_t)E[\psi_2(z_t)] \\
&+ \frac{1}{\sqrt{n}}\sum_{t=m+1}^{n} \psi_1(x_t)\left(E[\psi(z_t)|x_t] - E[\psi_2(z_t)]\right).
\end{aligned}
\tag{B.24}
$$

Similarly to the derivation in (2.11), by Lemma B.1, we have as $t \to \infty$

$$
\begin{aligned}
E[\psi_2(z_t)|x_t = x] &= \int_{-\infty}^{\infty} \psi_2(z)p_t(z|x)dz = \int_{-\infty}^{\infty} \psi_2(z) \cdot \frac{q_t\left(\frac{x}{\sqrt{t}}|z\right)}{q_t\left(\frac{x}{\sqrt{t}}\right)} p(z)dz \\
&\to \int_{-\infty}^{\infty} \psi_2(z) \cdot \frac{\phi(0)}{\phi(0)} p(z)dz = E[\psi_2(z_1)]
\end{aligned}
\tag{B.25}
$$

almost surely, where $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

In an analogous way to the derivations in (B.20) and (B.21), we have as $m \to \infty$

$$
\frac{1}{\sqrt{n}}\sum_{t=m+1}^{n} E\left[|\psi_1(x_t)|^i\right] = O(1) \quad \text{for } i = 1, 2, 3. \tag{B.26}
$$

Equation (B.26) then implies for large enough n

$$
\frac{1}{\sqrt{n}}\sum_{t=m+1}^{n} |\psi_1(x_t)|^i = O_P(1) \quad \text{for } i = 1, 2, 3. \tag{B.27}
$$

Therefore, in view of equations (B.24)–(B.27), we have as $m \to \infty$

$$\frac{1}{\sqrt{n}} \sum_{t=m+1}^{n} \psi_1(x_t)\psi_3(x_t) = \frac{1}{\sqrt{n}} \sum_{t=m+1}^{n} \psi_1(x_t)E[\psi_2(z_t)] + o_P(1)$$
$$\to_D \ L_{B_u}(1,0) \cdot \int_{-\infty}^{\infty} \psi_1(x)dx \cdot \int_{-\infty}^{\infty} \psi_2(z)p(z)dz, \tag{B.28}$$

in which we have used the standard result: $\frac{1}{\sqrt{k}} \sum_{t=1}^{k} \psi_1(x_t) \to_D L_{B_u}(1,0) \cdot \int_{-\infty}^{\infty} \psi_1(x)dx$ as $k \to \infty$.

To show (B.23), observe that

$$E\left[\sum_{t=m+1}^{n} \psi(x_t, z_t)\right]^2 = \sum_{t=m+1}^{n} E\left[\psi^2(x_t, z_t)\right] + 2 \sum_{t=m+2}^{n} \sum_{s=m+1}^{t-1} E\left[\psi(x_t, z_t)\psi(x_s, z_s)\right]$$
$$= \sum_{t=m+1}^{n} E\left[\psi_1^2(x_t)\left(\psi_2(z_t) - E[\psi_2(z_t)|x_t]\right)^2\right]$$
$$+ 2 \sum_{t=m+2}^{n} \sum_{s=m+1}^{t-1} E\left[\psi_1(x_t)\psi_1(x_s)\left(\psi_2(z_t) - E[\psi_2(z_t)|x_t]\right)\left(\psi_2(z_s) - E[\psi_2(z_s)|x_s]\right)\right]$$
$$\equiv \ J_{4n} + J_{5n}. \tag{B.29}$$

Analogously to the derivation in (B.20), by Assumption 2.1(ii) again we have for $i = 1, 2, 3$

$$E\left[|\psi_2(z_t)|^i | x_t = x\right] = \int_{-\infty}^{\infty} |\psi_2(z)|^i p_t(z|x)dz = \int_{-\infty}^{\infty} |\psi_2(z)|^i q_t\left(z|\frac{x}{\sqrt{t}}\right)dz$$
$$\leq \ 2\int_{-\infty}^{\infty} |\psi_2(z)|^i p(z)dz = 2E\left[|\psi_2(z_t)|^i\right], \tag{B.30}$$

which, along with (B.26), implies as $n \to \infty$

$$J_{4n} = \sum_{t=m+1}^{n} E\left[\psi_1^2(x_t)\left(\psi_2(z_t) - E[\psi_2(z_t)|x_t]\right)^2\right]$$
$$= \ \sum_{t=m+1}^{n} E\left[\psi_1^2(x_t)\left(E\left[\psi_2^2(z_t)|x_t\right] - \left(E[\psi_2(z_t)|x_t]\right)^2\right)\right]$$
$$\leq \ \sum_{t=m+1}^{n} E\left[\psi_1^2(x_t)E\left[\psi_2^2(z_t)|x_t\right]\right] \leq 2\sum_{t=m+1}^{n} E\left[\psi_1^2(x_t)\right] \cdot E\left[\psi_2^2(z_t)\right]$$
$$= \ O\left(\sqrt{n}\right). \tag{B.31}$$

To deal with J_{5n}, we apply some properties for the α–mixing condition assumed in Assumption 2.1(i) (see, for example, Lemma A.1 of Gao 2007) to imply that as $n \to \infty$

$$\sum_{t=m+2}^{n} \sum_{s=m+1}^{t-1} |E\left[\psi(x_t, z_t)\psi(x_s, z_s)\right]| \leq C \sum_{t=m+2}^{n} \sum_{s=m+1}^{t-1} \alpha^{\frac{\delta}{2+\delta}}(t-s)$$
$$\times \ \left(E\left[|\psi(x_s, z_s)|^{2+\delta}\right]\right)^{\frac{1}{2+\delta}} \cdot \left(E\left[|\psi(x_t, z_t)|^{2+\delta}\right]\right)^{\frac{1}{2+\delta}}$$
$$\leq \ C \sum_{t=m+2}^{n} \sum_{s=m+1}^{t-1} \alpha^{\frac{\delta}{2+\delta}}(t-s) \cdot \left(\frac{1}{\sqrt{s}}\right)^{\frac{1}{2+\delta}} \cdot \left(\frac{1}{\sqrt{t}}\right)^{\frac{1}{2+\delta}}$$
$$\leq \ C \sum_{s=m+1}^{n} \sum_{k=1}^{n} \alpha^{\frac{\delta}{2+\delta}}(k) \cdot \left(\frac{1}{\sqrt{s}}\right)^{\frac{2}{2+\delta}} = O\left(n^{\frac{1+\delta}{2+\delta}}\right) = o(n), \tag{B.32}$$

where $\delta > 0$ is the same as involved in Assumption 2.1(i), and we have used the fact that

$$E\left[|\psi(x_s, z_s)|^{2+\delta}\right] = E\left[|\psi_1(x_t)|^{2+\delta} E\left[|\psi_2(z_t) - E[\psi_2(z_t)|x_t]|^{2+\delta}|x_t\right]\right]$$
$$\leq\ 4E\left[|\psi_1(x_t)|^{2+\delta}\right] \cdot E\left[|\psi_2(z_t)|^{2+\delta}\right] = O\left(\frac{1}{\sqrt{t}}\right)$$

by equations (B.26) and (B.27) as well as Assumptions 2.1(ii) and 2.4(iii).

In view of (B.18)–(B.32), in order to complete the proof of Lemma A.3, it suffices to show that as $n \to \infty$

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\psi_1(x_t)\psi_2(z_t)e_t^2 = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\psi_1(x_t)\psi_2(z_t)E[e_t^2] + \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\psi_1(x_t)\psi_2(z_t)\left(e_t^2 - E\left[e_t^2\right]\right)$$
$$= \frac{\sigma_e^2}{\sqrt{n}}\sum_{t=1}^{n}\psi_1(x_t)\psi_2(z_t) + o_P(1), \tag{B.33}$$

which follows from

$$\frac{1}{n}\left[\sum_{t=1}^{n}\psi_1(x_t)\psi_2(z_t)\left(e_t^2 - E\left[e_t^2\right]\right)\right]^2 = \frac{1}{n}\sum_{t=1}^{n}E\left[\psi_1(x_t)\psi_2(z_t)\left(e_t^2 - E\left[e_t^2\right]\right)\right]^2$$
$$+\frac{2}{n}\sum_{t=2}^{n}\sum_{s=1}^{t-1}E\left[\psi_1(x_t)\psi_1(x_s)\psi_2(z_t)\psi_2(z_s)\left(e_t^2 - E\left[e_t^2\right]\right)\left(e_s^2 - E\left[e_s^2\right]\right)\right]$$
$$= \frac{1}{n}\sum_{t=1}^{n}E\left[\psi_1^2(x_t)E\left(\psi_2^2(z_t)|x_t\right)\right] \cdot E\left[\left(e_t^2 - E\left[e_t^2\right]\right)^2\right]$$
$$\leq \frac{2}{n}\sum_{t=1}^{n}E\left[\psi_1^2(x_t)\right] \cdot E\left[\psi_2^2(z_t)\right] \cdot E\left[\left(e_t^2 - E\left[e_t^2\right]\right)^2\right] = O\left(\frac{1}{\sqrt{n}}\right) = o(1)$$

by Assumption 2.2 as well as equations (B.26) and (B.27) again.

Equations (B.29)–(B.33) then complete the proof of (B.23). Therefore, the proof of Lemma A.3 is completed.

Derivation of (3.5). Similarly to the proof of Lemma A.2, under H_1, we have as $n \to \infty$

$$\sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t - x_s}{h_1}\right)K_2\left(\frac{z_t - z_s}{h_2}\right)\ \widehat{e}_s\ \widehat{e}_t$$
$$= \sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t - x_s}{h_1}\right)K_2\left(\frac{z_t - z_s}{h_2}\right)\ e_s\ e_t$$
$$+\delta_n^2\sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t - x_s}{h_1}\right)K_2\left(\frac{z_t - z_s}{h_2}\right)\ \Delta(x_s, z_s)\ \Delta(x_t, z_t) + o_P(1)$$
$$\geq\ \delta_n^2 \cdot \sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t - x_s}{h_1}\right)K_2\left(\frac{z_t - z_s}{h_2}\right)\ \Delta(x_s, z_s)\ \Delta(x_t, z_t) + o_P(1)$$
$$\equiv\ \delta_n^2 \cdot Q_n(h_1, h_2) + o_P(1), \tag{B.34}$$

where $Q_n(h_1, h_2) = \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n}K_1\left(\frac{x_t - x_s}{h_1}\right)K_2\left(\frac{z_t - z_s}{h_2}\right)\ \Delta(x_s, z_s)\ \Delta(x_t, z_t)$.

Straightforward derivations as in equations (A.13)–(A.15), imply that as $n \to \infty$

$$E\left[Q_n(h_1, h_2)\right] = C_1(1 + o(1))\ nh_1h_2^d \tag{B.35}$$

for some $C_1 > 0$.

Similarly, we may show that as $n \to \infty$

$$\sigma_{2n}^2 \equiv E\left[\sum_{t=1}^n \sum_{s=1}^n K_1^2\left(\frac{x_t - x_s}{h_1}\right) K_2^2\left(\frac{z_t - z_s}{h_2}\right) e_s^2 e_t^2\right] = C_2(1 + o(1))\ n^{\frac{3}{2}} h_1 h_2^d \tag{B.36}$$

Equations (B.34)–(B.36) thus complete an outline of the derivation of (3.5).

Derivation of (3.6). In view of the proof of Lemma A.2, in order to complete the derivation of (3.6), it suffices to show that as $n \to \infty$ and for some $C_2 > 0$

$$\begin{aligned} &\frac{1}{\sqrt{n} h_1 h_2} \sum_{t=1}^n \sum_{s=1}^n E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right)\right. \\ &\times\ \left.\Delta_n(x_s, z_s)\Delta_n(x_t, z_t)\right] \pi_1(x)\pi_2(z) dz dx \\ &=\ C_2(1 + o(1))\ \sqrt{n} h_1 h_2^d, \end{aligned} \tag{B.37}$$

which follows similarly to the derivations in equations (A.13)–(A.15). Thus, we omit these details. Such details are available upon request, however.

REFERENCES

Cai, Z., Li, Q., Park, J., 2009. Functional–coefficient models for nonstationary time series data. Journal of Econometrics 148, 101–113.

Campbell, J., Lo, A. W., MacKinlay, A. C., 1997. The Econometrics of Financial Markets. Princeton University Press, New Jersey.

Campbell, J., Mankiw, N. G., 1990. Consumption, income and interest rates: reinterpreting the time series evidence. NBER Macroeconomics Series at http://ideas.repec.org/h/nbr/nberch/10965.html.

Chanda, K. C., 1974. Strong mixing properties of linear stochastic processes. Journal of Applied Probability 11, 401–408.

Chang, Y., Park, J. and Phillips, P. C. B., 2001. Nonlinear econometric models with cointegrated and deterministically trending regressors. Econometrics Journal 4, 1–36.

Chen, J., Gao, J., Li, D., 2012. Estimation in semiparametric regression with nonstationary regressors. Bernoulli 18, 678–702

Dam, B. K., 1998. Approximation orders in the conditional central limit theorem for weakly dependent random variables. *Acta Mathematica Vietnamica* 23, 347–358.

Denker, M., Gordin, M., 2003. On conditional central limit theorems for stationary processes," In Lecture Notes–Monograph Series, Vol. 41, *Probability, Statistics and Their Applications: Papers in Honor of Rabi Bhattacharya*, pp. 133–152.

Faff, R. W., Brooks, R. D., 1998. Time–varying beta risk for Australian industry portfolios: an exploratory analysis. Journal of Business Finance and Accounting 25, 721–745.

Fan, J., Yao, Q., 2003. Nonlinear Time Series: Nonparametric and Parametric Methods. Springer, New York.

Gao, J., 2007. Nonlinear Time Series: Semiparametric and Nonparametric Methods. Chapman & Hall, London.

Gao, J., Gijbels, I., 2008. Bandwidth selection in nonparametric kernel testing. Journal of the American Statistical Association 484, 1584-1594.

Gao, J., King, M. L., 2012. An improved nonparametric unit–root test. Working paper available at http://www.jitigao.com/page1006.aspx.

Gao, J., King, M. L., Lu, D., Tjøstheim, D., 2009a. Specification testing in nonlinear and nonstationary time series autoregression. Annals of Statistics 37, 3893–3928.

Gao, J., King, M. L., Lu, D., Tjøstheim, D., 2009b. Nonparametric specification testing for nonlinear time series with nonstationarity. Econometric Theory 25, 1869–1892.

Gao, J., Phillips, P. C. B., 2011. Semiparametric estimation in multivariate nonstationary time series models. Available at http://www.buseco.monash.edu.au/ebs/pubs/wpapers/2011/wp17-11.pdf.

Gylfason, T., 1981. Interest rates, inflation and the aggregate consumption function. Review of Economics and Statistics 63, 233–245.

Hahm, J., Steigerwald, D. G., 1999. Consumption adjustment under time–varying income uncertainty. Review of Economics and Statistics 81, 32–40.

Härdle, W., Mammen, E., 1993. Comparing nonparametric versus parametric regression fits. Annals of Statistics 21, 1926–1947.

Hong, S. H., Phillips, P. C. B., 2010. Testing linearity in cointegrating relations with an application to purchasing power parity. Journal of Business and Economic Statistics 28, 96–114.

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

Karlsen, H. A., Myklebust, T., Tjøstheim, D., 2007. Nonparametric estimation in a nonlinear cointegration type model. Annals of Statistics 35, 252–299.

Kasparis, I. 2008. Detection of functional form misspecification in cointegrating relations. Econometric Theory 24, 1373-1403.

Li, D., Gao, J., Chen, J., Lin, Z., 2011. Nonparametric specification testing in nonlinear and nonstationary time series models. Working paper available at http://www.jitigao.com/page1006.aspx.

Li, Q., Racine, J., 2007. Nonparametric Econometrics: Theory and Practice. Princeton University Press, Princeton.

Myklebust, T., Karlsen, H., Tjøstheim, D., 2012. Null recurrent unit root processes. Econometric Theory 28, 1–41.

Peligrad, M., Utev, S., 2006. Central limit theorem for stationary linear processes. Annals of Probability 34, 16081622.

Phillips, P. C. B., Park, J., 1998. Nonstationary density estimation and kernel autoregression. Cowles Foundation Discussion Paper, No. 1181, Yale University.

Rényi, A., 1958. On mixing sequences of sets. Acta Mathematica Hungarica 9, 215–228.

Wang, Q. Y., Phillips, P. C. B., 2009a. Asymptotic theory for local time density estimation and nonparametric cointegrating regression. Econometric Theory 25, 710-738.

Wang, Q. Y., Phillips, P. C. B., 2009b. Structural Nonparametric Cointegrating Regression. Econometrica 77, 1901–1948.

Wang, Q. Y., Phillips, P. C. B., 2011. Asymptotic theory for zero energy functionals with nonparametric regression application. Econometric Theory 27, 235–259.

Wang, Q. Y., Phillips, P. C. B., 2012. Specification testing for nonlinear cointegrating regression. Annals of Statistics 40, 727–758.

Withers, C. S., 1981. Conditions for linear processes to be strong mixing. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 57, 509–534.

Xiao, Z., 2009. Functional–coefficient cointegration models. *Journal of Econometrics* 152, 81–92.