ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

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

Specification Testing for Nonlinear Multivariate Cointegrating Regressions

Chaohua Dong, Jiti Gao, Dag Tjøstheim and Jiying Yin

February 2014

Working Paper 08/14

Specification Testing for Nonlinear Multivariate Cointegrating Regressions*

Chaohua Dong*, Jiti Gao*†, Dag Tjøstheim‡ and Jiying Yin*

Monash University and The University of Bergen‡*

February 19, 2014

Abstract

This paper considers a general model specification test for nonlinear multivariate cointegrating regressions where the regressor consists of a univariate integrated time series and a vector of stationary time series. The regressors and the errors are generated from the same innovations, so that the model accommodates endogeniety. 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, endogeneity, nonparametric kernel estimation, parametric model specification, time series.

JEL Classification: C12, C14, C22.

Abbreviated Title: Model Specification in Nonstationary Cointegration.

*The authors acknowledge comments and suggestions from participants at many 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.

†*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 the integrated time series regressor is univariate without additional regressors involved. A main reason for this is that 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 Chen, Gao and Li (2012), and Gao and Phillips (2013) 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, the nonparametric part of the model again contains only one univariate nonstationary regressor and no stationary regressors are involved in the same nonparametric part. 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. Section 5 below examines the suitability of linear models and applicability of nonlinear models. Additionally, endogeneity is naturally inherited in many economic variables, such as, supply and demand, and disposable income and expenditure consumption. In the estimation case, to the best of our knowledge, Wang and Phillips (2009b) were among the first investigating a nonparametric cointegrating regression for the case where the errors are defined by a functional form of innovations which also are building blocks of the regressors. For specification testing purposes, this paper shall establish some large small sample properties for a newly proposed test statistic under two types of endogeneity.

A main objective of using a parametric model specification is to find a best available parametric function to approximate an unknown nonparametric 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, x_t and z_t can be either mutually independent of each other or highly correlated, $\{e_t\}$ is a sequence of martingale differences, 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 $1 \le d \le 3$.

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^k for $k \ge 1$, and $\Delta_n(x, z)$ is a sequence of nonparametrically unknown departure functions.

Recent studies in the field of nonparametric model specification of integrated time series models include Marmer (2007), Kasparis (2008), Gao *et al* (2009a; 2009b), Hong and Phillips (2010), Kasparis (2010), Kasparis and Phillips (2012), Kasparis *et al* (2012), and Wang and Phillips (2012). 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. This paper takes the full feature of the integrated structure of $\{x_t\}$ into account and proposes a simple and powerful test for testing H_0 versus H_1. Both theoretical and empirical comparisons given in Sections 2–4, moreover, show that the proposed test is preferred to the test proposed in Gao *et al* (2009a) and Wang and Phillips (2012).

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, z_t and e_t can be highly correlated. 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. The proofs of the main results and the necessary lemmas are given in Appendix A. Appendix B gives a useful lemma and then the proof of the third lemma listed in Appendix A. The proof of the lemma listed in Appendix B and those of the first two lemmas of Appendix A are given in Appendix C.

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

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

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 $K_1(\cdot)$ is a univariate probability kernel function, K_2 is a d–dimensional probability kernel function, 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:

$$\widehat{q}(x,z) \;=\; \frac{1}{\sqrt{n} h_1 h_2^d} \sum_{t=1}^{n} K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) y_t \quad \text{and}$$

$$\widetilde{q}(x,z;\theta_0) = \frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^n K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)g(x_t,z_t;\theta_0). \quad (2.6)$$

This paper now proposes using a test statistic of the form

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

which is similar to the original proposal by Härdle and Mammen (1993) for the independent sample case, where $\widehat{\theta}$ is as defined in (2.2), $\pi_i(u)$ are both known probability weight functions satisfying $0 < \int_{-\infty}^{\infty}\pi_i^2(u)du < \infty$ for $i=1,2$, and $\widehat{p}(x,z) = \frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^n K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$. Note that throughout the rest of this paper, we use $\int_{-\infty}^{\infty}(\cdot)dz$ to denote $\int_{R^d}(\cdot)dz$ for notational simplicity.

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 $\widehat{L}_n$ under H_0. As shown in the proof of Theorem 2.1 given in Appendix A below, we have under H_0

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

where $\widehat{e}_t = y_t - g(x_t,z_t;\widehat{\theta})$, $L_i(v) = \int_{-\infty}^{\infty}K_i(u)K_i(u+v)du$ for $i=1,2$, and

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

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

Mainly because of the fact that $\widetilde{S}_{1n}$ already converges in distribution to a random variable, there is no need to standardise $\widetilde{L}_n$ by a random denominator of a quadratic form as has been done in Gao *et al* (2009a), although it is needed to normalise $\widetilde{S}_{1n}$ by a simple partial sum of the form $\widehat{\Sigma}_n = \frac{1}{n}\sum_{t=1}^n \widehat{e}_t^2\pi_2(z_t)$, which is a consistent estimator of $E[e_1^2\pi(z_1)]$. Moreover, existing literature (Gao *et al* 2009a; Wang and Phillips 2012) shows that it is much harder 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

$$\overline{M}_n = \frac{\widetilde{L}_n - \widetilde{S}_{1n}}{\widetilde{\sigma}_n} \tag{2.10}$$

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 random denominator of a quadratic form. 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. Throughout the rest of this paper, $\|\cdot\|$ denotes the conventional Euclidean norm.

Assumption 2.1. *Let $\{\varepsilon_t\}$ be a sequence of independent and identically distributed (i.i.d.) random errors with $E[\varepsilon_1] = 0$, $0 < E[\varepsilon_1^2] = \sigma_0^2 < \infty$ and $E\left[|\varepsilon_1|^{4+\delta}\right] < \infty$ for some $\delta > 0$. Let $\varphi(\cdot)$ be the characteristic function of ε_1 satisfying $|r|\ \varphi(r) \to 0$ as $r \to \infty$. Meanwhile, $\{\eta_t\}$ is another sequence of independent and identically distributed random variables with $E[\eta_1^2] < \infty$.*

(i) *Consider $x_t = x_{t-1} + u_t$. Let $u_t = \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$, where $\{\phi_i : i \geq 0\}$ is a sequence of real numbers such that $\sum_{j=0}^{\infty} \phi_j \neq 0$ and $\sum_{j=0}^{\infty} \phi_j^2 < \infty$.*

(ii) *z_t satisfies either (a) $z_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ where $\{\psi_j : j \geq 0\}$ is a sequence of vectors such that $\sum_{j=0}^{\infty} \psi_j \neq 0$, $\sum_{j=0}^{\infty} \|\psi_j\|^2 < \infty$ and $\sum_{j=k+1}^{\infty} \|\psi_j\|^2 = O\left(k^{-1}\right)$ as $k \to \infty$; or (b) $z_t = \Lambda(\varepsilon_{t-1}, \cdots, \varepsilon_{t-\tau}; \eta_t)$ where $\Lambda(\cdot, \cdot, \cdot)$ is a measurable vector function such that $E[\|z_1\|^{2+\delta}] < \infty$ for the same $\delta > 0$ as before, where $\tau \geq 1$ is a positive integer.*

(iii) *e_t is generated either by (a) $e_t = \sum_{j=0}^{\infty} \rho_j \varepsilon_{t-j}$ where $\sum_{j=0}^{\infty} \rho_j^2 < \infty$, $\rho := \sum_{j=0}^{\infty} \rho_j \neq 0$ and $\sum_{j=k+1}^{\infty} \rho_j^2 = O\left(k^{-1}\right)$ as $k \to \infty$; or by (b) $e_t = \Gamma(\varepsilon_{t-1}, \cdots, \varepsilon_{t-\tau}; \eta_t)$ with a measurable function $\Gamma(\cdot, \cdot, \cdot)$ such that $E[e_t] = 0$ and $E[e_t^{4+2\delta}] < \infty$ for all $t > \tau$; we define $e_t = 0$ for $t \leq \tau$. Here, τ and δ are the same as in (ii).*

(iv) *Let $p(z)$ be the marginal density function of z_1 and $p_\tau(x, y)$ be the joint density of $(z_{t+\tau}, z_t)$ for $\tau \geq 1$. Then $p(z)$ is continuous in z and $p_\tau(x, y)$ is continuous in (x, y) uniformly in $\tau \geq 1$. Let $q(v)$ be the marginal density function of e_1 and $q_\tau(x, y)$ be the joint density of $(e_{t+\tau}, e_t)$ for $\tau \geq 1$. Then $q(v)$ is continuous in v and $q_\tau(x, y)$ is continuous in (x, y) uniformly in $\tau \geq 1$.*

Assumption 2.2. *Let (e_t, z_t, u_t) be a vector of stationary and α–mixing time series with mixing coefficient $\alpha_{ezu}(k)$ satisfying $\sum_{k=0}^{\infty} \alpha_{ezu}^{\frac{\delta}{2+\delta}}(k) < \infty$.*

Assumption 2.3. (i) *$g(x, z; \theta)$ 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 $i, j = 1, 2$*

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

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* $\epsilon > 0$ *and as* $n \to \infty$*, we have* $P\left(\sqrt{n}h_1 h_2^d \left|\left|\widehat{\theta} - \theta_0\right|\right|^2 > \epsilon\right) \to 0$.

(iii) *Furthermore, for any small* $\eta > 0$*, we have* $P\left(R_{nj}\left|\left|\widehat{\theta} - \theta_0\right|\right|^2 > \eta\right) \to 0$ *as* $n \to \infty$*, where* $R_{nj} = \frac{1}{n}\sum_{t=1}^n \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ||G_j(\sqrt{t}\, x, z; \theta_0)||^2 \pi_2(z) f_{dt}(x, z)\, dxdz$*, in which* $f_{dt}(x, z)$ *denotes the joint density of* (x_t, z_t) *and* $\pi_2(z)$ *is the probability weight function.*

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

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

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

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

(iii) *Let* $\pi_1(\cdot)$ *and* $\pi_2(\cdot)$ *be known probability weight functions such that* $0 < \int_{-\infty}^{\infty} \pi_1^i(u)du < \infty$ *and* $0 < \int_{-\infty}^{\infty} \pi_2^i(z)p(z) < \infty$ *for* $i = 1, 2$*, and* $E\left[\left(e_1^2\pi_2(z_1)\right)^{2+\delta}\right] < \infty$ *for the same* $\delta > 0$ *as in Assumption 2.1.*

(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* $i = 1, 2$*, satisfying* $\int_{-\infty}^{\infty} |D_1(x)|^i dx < \infty$ *and* $\int_{-\infty}^{\infty} |D_2(z)|^i p(z)dz < \infty$ *for* $i = 1, 2$*, such that* $|\pi_1(y) - \pi_1(x)| \leq D_1(x) \cdot |y - x|$ *for any* $(x, y) \in \Omega_1(\epsilon) = \{(x, y) : |y - x| \leq \epsilon, x, y \in R^1\}$ *and* $|\pi_2(y) - \pi_2(x)| \leq D_2(x) \cdot ||y - x||$ *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-iii) imposes a stationarity structure on $\{e_t, u_t, z_t\}$. Note that they share the same building blocks $\{\varepsilon_t\}$, and so that they are flexible. For example, u_t and e_t are the same when $\phi_j = \rho_j$ for all j, u_t and e_t may be independent of each other given particular choice of their coefficients in the linear processes or particular form of the functional, u_t and e_t may be a component of z_t depending on ϕ_j and/or ρ_j is a component of ψ_j for all j, or u_t and e_t may be mutually independent of each other within the three processes. Such a setting gives a maximum applicability of the theoretical results below. Under Assumption 2.1(i-iii), Lemma B.1 listed in Appendix B below establishes a general and useful result, which implies the mutual independence between (z_t, z_s) and $\left(\frac{x_t}{\sqrt{t}}, \frac{x_s}{\sqrt{s}}\right)$ when $t, s \to \infty$ and $\frac{s}{t} \to 0$, even though x_t and z_t can be highly correlated. Assumption 2.1(iv) imposes some necessary and mild conditions on the marginal and joint density functions. Such conditions are very weak and easily verifiable. Assumption 2.2 imposes an α–mixing condition on (e_t, u_t, z_t). Such a condition is quite commonly used in the stationary time series case. When all e_t, u_t and z_t are linear processes, existing results (see, for example, Theorem 2.1 of Chanda 1974; Corollary 4 of Withers 1981) imply that the α–mixing condition is satisfied.

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). 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. However, 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) where e_t is independent of x_t or $(e_t,\mathcal{F}_t)$ is a martingale difference such that x_t is adapted with respect to $\mathcal{F}_{t-1}$. Assumption 2.3 (iii) involves G_1 and G_2 with the rate of convergence of $\widehat{\theta}-\theta_0$. This is not unreasonable, since similar assumptions have been used in the literature (see, for example, Theorems 5.1-5.2 of Park and Phillips 2001; Assumption 3.1 of Gao *et al* 2009b and Corollary 5.4 of Park and Phillips 2001). One particular example is $g(x_t,z_t;\theta_0)=\theta_0 x_t^2$, which implies $G_1(x_t,z_t;\theta_0)=x_t^2$ and $G_2(x_t,z_t;\theta_0)=0$. In this case, Assumption 2.3(iii) can easily be satisfied. 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 2.1–2.3(ii) and 2.4 hold. Then under* H_0

$$\begin{aligned}\widehat{L}_n &= \sqrt{nh_1^2h_2^{2d}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\widehat{q}(x,z)-\widetilde{q}\left(x,z;\widehat{\theta}\right)\right)^2\pi_1(x)\pi_2(z)dzdx\\ &\to_D C_1(K,\pi;\sigma_e^2)\cdot L_{B_u}(1,0)\quad \text{as } n\to\infty, \end{aligned}\tag{2.11}$$

where $L_{B_u}(r,0)=\lim_{\delta\to 0}\frac{1}{2\delta}\int_0^r I\left[|B_u(s)|<\delta\right]ds$ *is the local–time process associated with the Brownian motion* B_u*, which is the weak limit of* $U_n(r)$ *such that* $U_n(r)=\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}u_t\Longrightarrow_{\mathcal{D}} B_u(r)$ *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*

$$F_L(x)=P\left(L_{B_u}(1,0)\le x\right)=\begin{cases}2\Phi(x)-1, & x\ge 0,\\ 0, & x<0,\end{cases}\tag{2.12}$$

in which $\Phi(x)$ *is the cdf of* $N(0,1)$*, and* $C_1(K,\pi;\sigma_e^2)=\prod_{i=1}^2\int_{-\infty}^{\infty}K_i^2(v)dv\cdot E\left[e_1^2\ \pi_2(z_1)\right]$.

When $E\left[e_1^2\ \pi_2(z_1)\right]$ is unknown, it is estimated by $\widehat{\Sigma}_n=\frac{1}{n}\sum_{t=1}^n\left(y_t-g(x_t,z_t;\widehat{\theta})\right)^2\pi_2(z_t)$ under H_0. We then have the following corollary; its proof is given in Appendix A.

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

$$\frac{\widehat{L}_n}{\widehat{\Sigma}_n}\to_D C_2(K)\cdot L_{B_u}(1,0),\tag{2.13}$$

where $C_2(K)=\prod_{i=1}^2\int_{-\infty}^{\infty}K_i^2(v)dv$.

Note that Corollary 2.1 shows that the asymptotic distribution is proportional to $L_{B_u}(1, 0)$ that has a known distribution function given in (2.12). 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 natural to have 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 can be more powerful than the natural competitor under a sequence of local alternatives. 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.

We now consider an extended form of the test statistic proposed in Gao *et al* (2009a) and then used in Wang and Phillips (2012) as follows:

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

Let $\widehat{M}_{2n}^2 = 2\sum_{t=1}^n \sum_{s=1}^n K_1^2\left(\frac{x_t - x_s}{h_1}\right) K_2^2\left(\frac{z_t - z_s}{h_2}\right)\ \widehat{e}_s^2\ \widehat{e}_t^2$. Similarly to the derivations of equation (3.6) given in Appendix B below, we have under H_0:

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

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 $j = 1, 2$

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

In order to compare our test with $\widehat{M}_n(h_1, h_2)$, we need to further require

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta^2(x, z) p(z) dz dx < \infty. \tag{3.5}$$

Let $\widehat{M}_{1n} = \frac{1}{\sigma_{2n}} \sum_{t=1}^n \sum_{s=1}^n K_1\left(\frac{x_t - x_s}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2}\right) \ \widehat{e}_s \ \widehat{e}_t$. 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

$$\begin{aligned}
\widehat{M}_{1n} &= \frac{1}{\sigma_{2n}} \sum_{t=1}^n \sum_{s=1}^n K_1\left(\frac{x_t - x_s}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2}\right) \ \widehat{e}_s \ \widehat{e}_t \\
&= \frac{1}{\sigma_{2n}} \sum_{t=1}^n \sum_{s=1}^n K_1\left(\frac{x_t - x_s}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2}\right) \ e_s \ e_t \\
&\quad + \frac{1}{\sigma_{2n}} \sum_{t=1}^n \sum_{s=1}^n K_1\left(\frac{x_t - x_s}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2}\right) \ \Delta_n(x_s, z_s) \ \Delta_n(x_t, z_t) + o_P(1). \\
&\equiv M_{2n} + R_{2n} + o_P(1). \qquad (3.6)
\end{aligned}$$

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

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

It may then be shown that as $n \to \infty$

$$\begin{aligned}
E[R_{1n}] &= C_1(1 + o(1)) \cdot \delta_n^2 \cdot \sqrt{n} h_1 h_2^d, \\
E[R_{2n}] &= C_2(1 + o(1)) \cdot \delta_n^2 \cdot \sqrt{\sqrt{n} h_1 h_2^d} \qquad (3.7)
\end{aligned}$$

when $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta^2(x, z) p^2(z) \pi_1(x) \pi_2(z) dz dx < \infty$ and $\sqrt{n} h_1 h_2^d \to \infty$. An outline of the derivation of (3.7) is given in Appendix B.

In view of Assumption 2.4(ii), equation (3.7) therefore implies that there is some $C_0 > 0$ such that

$$\frac{E[R_{1n}]}{E[R_{2n}]} = C_0 \sqrt{\sqrt{n} h_1 h_2^d} \to \infty. \qquad (3.8)$$

when $\sqrt{n} h_1 h_2^d \to \infty$.

Equations (3.8) indicates that $\widehat{L}_n(h_1, h_2)$ may be more powerful than $\widehat{M}_n(h_1, h_2)$ under a sequence of local departure functions that satisfy (3.3) and (3.4). As a matter of fact, $\widehat{L}_n(h_1, h_2)$ is indeed more powerful than $\widehat{M}_n(h_1, h_2)$. Let c_α be the α–level critical value of the limiting distribution of $\widehat{L}_n(h_1, h_2)$ and d_α be the α–level critical value of the limiting distribution of $\widehat{M}_n(h_1, h_2)$. Let $F_{n,i}(x)$ be the distributional function of $\widehat{L}_n(h_1, h_2)$ and $G_{n,i}(x)$

be the distributional function of $\widehat{M}_n(h_1, h_2)$ under H_i for $i = 0, 1$. Then, by Theorem 2.1 of this paper and an extension of Theorem 2.1 of Gao *et al* (2009b), we may have as $n \to \infty$

$$\begin{aligned} F_{n,1}(x) &= P\left(\widehat{L}_n \le x\right) = (1+o(1))P\left(L_{1n} \le x - r_{1n}\right) \\ &= (1+o(1))F_{n,0}(x - r_{1n}), \\ G_{n,1}(x) &= P\left(\widehat{M}_n \le x\right) = (1+o(1))P\left(M_{2n} \le x - r_{2n}\right) \\ &= (1+o(1))G_{n,0}(x - r_{2n}), \end{aligned} \tag{3.9}$$

where $r_{1n} = E\left[R_{1n}\right]$ and $r_{2n} = E\left[R_{2n}\right]$.

Then, the power functions of $\widehat{L}_n$ and $\widehat{M}_n$ can be respectively represented by

$$\begin{aligned} \beta_{l,n}(h) &= P\left(\widehat{L}_n > c_\alpha | H_1\right) = (1+o(1))\left(1 - F_{n,0}(c_\alpha - r_{1n})\right) \\ &= (1+o(1))\left(1 - F_l(c_\alpha - r_{1n})\right), \end{aligned} \tag{3.10}$$

$$\begin{aligned} \beta_{m,n}(h) &= P\left(\widehat{M}_n > d_\alpha | H_1\right) = (1+o(1))\left(1 - G_{n,0}(d_\alpha - r_{1n})\right) \\ &= (1+o(1))\left(1 - \Phi(d_\alpha - r_{2n})\right), \end{aligned} \tag{3.11}$$

where $F_l(\cdot)$ is the limit distribution of $\widehat{L}_n(h_1, h_2)$ as stated in Theorem 2.1 above, and $\Phi(\cdot)$ is the standard normal distribution as the limiting distribution of $\widehat{M}_n(h_1, h_2)$ being established in Theorem 2.1 of Gao *et al* (2009b). Further study of the power functions could be done when Edgeworth expansions of the power functions for the stationary stationary time series case become available. Proposition 3.1 of Chen *et al* (2011) have only established an Edgeworth expansion for the case where e_t is independent and identically distributed.

In view of the fact that $F_l(x) = 0$ for $x < 0$ and $r_{1n} \to \infty$, we have

$$\begin{aligned} \beta_{l,n}(h) - \beta_{m,n}(h) =& (1+o(1))\left(\Phi(d_\alpha - r_{2n}) - F_l(c_\alpha - r_{1n})\right) \\ =& (1+o(1)) \cdot \Phi(d_\alpha - r_{2n}) \ge 0. \end{aligned} \tag{3.12}$$

Our simulation study in Section 4 below supports the fact that $\widehat{L}_n(h_1, h_2)$ is indeed more powerful than $\widehat{M}_n(h_1, h_2)$ in the case where $\Delta_n(x, z) = \delta_n\ \Delta(x, z)$.

We then summarise the above discussion into the following proposition; its proof follows immediately from equation (3.10).

Proposition 3.1. *Suppose that Assumptions 2.1–2.4 hold. If, in addition, equations (3.3) and (3.4) are satisfied, then we have under* H_1

$$\lim_{n\to\infty} P\left(\widehat{L}_n > C\right) = 1, \tag{3.13}$$

for any positive constant $C > c_\alpha$*, where* c_α *is the* α*–level critical value of the limiting distribution of* $\widehat{L}_n(h_1, h_2)$*.*

It should be pointed out that there is a kind of trade–off between ensuring that $\widehat{L}_n$ is more powerful than $\widehat{M}_n$ and involving the weight functions $\pi_1(\cdot)$ and $\pi_2(\cdot)$ as well as requiring (3.4) and (3.5), in addition to requiring Assumption 2.4(ii). This is mainly because $\widehat{M}_n$ can be more powerful than $\widehat{L}_n$ when (3.4) is satisfied but (3.5) 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 Chen *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

$$\delta_n = \frac{\log(n)}{2n^{\frac{1}{8}}} \quad \text{and} \quad \Delta(x_t) = x_t^2, \tag{3.14}$$

where $x_t = x_{t-1} + u_t$ with $x_0 = o_P(\sqrt{n})$ 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

$$E\left[\Delta_n(x_n)\right] = \delta_n \, E[x_n^2] = \frac{n \log(n)}{n^{\frac{1}{8}}} \to \infty \tag{3.15}$$

even though $\delta_n \to 0$.

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 (3.5). 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 $\widehat{L}_n$ has power to pick up such 'small' departure and is more powerful than $\widehat{M}_n$ 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 $\widehat{L}_n$ is a more powerful test than $\widehat{M}_n$ under a sequence of local alternatives.

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:

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

Define the following test statistic:

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

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

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

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

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

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 Chen *et al* 2011). Instead, we propose using an estimation–based optimal bandwidths of the form:

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

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

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

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_{cv}, we actually calculated equation (4.3) over all possible intervals. Our computation indicates that H_{cv} is the shortest possible interval on which the CV function attains its smallest value.

Let $Q_n(h_1, h_2)$ denote either $\widehat{L}_{1n}$ or $\widehat{L}_{2n}$. 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 $\left(\widehat{h}_{1\text{cv}}, \widehat{h}_{2\text{cv}}\right)$. 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 $\widehat{m}(x,z;\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}}) = \frac{\sum_{s=1}^n K_1\left(\frac{x-x_s}{\widehat{h}_{1\mathrm{cv}}}\right)K_2\left(\frac{z-z_s}{\widehat{h}_{2\mathrm{cv}}}\right)y_s}{\sum_{s=1}^n K_1\left(\frac{x-x_s}{\widehat{h}_{1\mathrm{cv}}}\right)K_2\left(\frac{z-z_s}{\widehat{h}_{2\mathrm{cv}}}\right)}$. Let q_r be the asymptotic critical value of $Q_n\left(\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}}\right)$ 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: Let $\widehat{e}_t = \left(y_t - \widehat{m}(x_t,z_t;\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}})\right)$. Generate the bootstrap residuals $\{e_t^*\}$ by $e_t^* = \widehat{e}_t\eta_t^*$, where $\{\eta_t^*, 1\leq t\leq n\}$ is a sequence of i.i.d. random variables drawn from

$$P\left(\eta_1^* = -\frac{\sqrt{5}-1}{2}\right) = \frac{\sqrt{5}+1}{2\sqrt{5}} \text{ and } P\left(\eta_1^* = \frac{\sqrt{5}+1}{2}\right) = \frac{\sqrt{5}-1}{2\sqrt{5}}. \tag{4.4}$$

Step 2: Obtain $y_t^* = \widehat{m}(x_t,z_t;\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}})+e_t^*$. The resulting sample $\{(y_t^*,x_t,z_t), 1\leq t\leq n\}$ is called a bootstrap sample.

Step 3: Use the data set $\{(y_t^*,x_t,z_t), 1\leq t\leq n\}$ to re–estimate $m(x,z)$ and denote its estimate by $\widehat{m}^*(x,z;\widehat{h}_{1\mathrm{cv}}^*,\widehat{h}_{2\mathrm{cv}}^*)$, in which $\widehat{h}_{1\mathrm{cv}}^*$ and $\widehat{h}_{2\mathrm{cv}}^*$ are calculated based on the data $\{(y_t^*,x_t,z_t), 1\leq t\leq n\}$. Then calculate the test statistic $\widehat{Q}_n^*\left(\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}}\right)$, which is the corresponding version of $\widehat{Q}_n\left(\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}}\right)$ by replacing $\{(y_t,x_t,z_t)\}$ and $\widehat{m}(x,z;\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}})$ with $\{(y_t^*,x_t,z_t)\}$ and $\widehat{m}^*(x,z;\widehat{h}_{1\mathrm{cv}}^*,\widehat{h}_{2\mathrm{cv}}^*)$, respectively.

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

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

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

For each pair $\left(\widehat{h}_{1\mathrm{cv}}^*,\widehat{h}_{2\mathrm{cv}}^*\right)$, choose q_r^* such that $P^*\left(\widehat{Q}_n^*\left(\widehat{h}_{1\mathrm{cv}}^*,\widehat{h}_{2\mathrm{cv}}^*\right) > q_r^*\right) = r$ and estimate q_r by q_r^*. For $\widehat{L}_{1n}$ or $\widehat{L}_{2n}$, we approximate the asymptotic critical value z_r or l_r by z_r^* or l_r^*, respectively. For $M=250$, let also $f_{j\mathrm{cv}}^*$ denote the frequency of $\widehat{L}_{1n}(\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}};j) > z_r^*$ for $j=0,1$ under H_0 or H_1, and $g_{j\mathrm{cv}}^*$ denote the frequency of $\widehat{L}_{2n}(\widehat{h}_{1\mathrm{cv}},\widehat{h}_{2\mathrm{cv}};j) > l_r^*$ for $j=0,1$ under H_0 or H_1.

In the case where (e_t,u_t,z_t) is a vector of stationary time series, one should use a block bootstrap method. Since we only consider the case where (e_t,u_t,z_t) is a vector of i.i.d. random variables in Examples 4.1 and 4.2 below, the regression bootstrap with the choice of $\{\eta_t^*\}$ works well. 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 $\{(e_t, u_t, z_t) : 1 \le t \le n\}$ 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 (3.7). 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	0.006	0.027	0.009	0.052	0.007	0.019	0.008	0.054
300	0.009	0.034	0.011	0.046	0.009	0.034	0.012	0.049
500	0.008	0.045	0.009	0.053	0.006	0.045	0.007	0.053
H_1	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
100	0.562	0.591	0.590	0.631	0.763	0.815	0.812	0.871
300	0.601	0.642	0.642	0.691	0.814	0.867	0.858	0.924
500	0.674	0.713	0.721	0.762	0.877	0.919	0.913	0.972

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 $\{(e_t, u_t, z_t) : 1 \le t \le n\}$ 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 $\Delta_n(x, z) = \frac{\delta_n\ z^2}{\sqrt{1+x^2}}$ with $\delta_n = \frac{\log(n)}{2n^{\frac{1}{8}}}$.

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	0.008	0.032	0.012	0.048	0.006	0.032	0.009	0.049
300	0.011	0.044	0.009	0.054	0.008	0.029	0.012	0.054
500	0.009	0.039	0.011	0.049	0.012	0.043	0.008	0.046
H_1	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
100	0.536	0.572	0.602	0.641	0.731	0.759	0.787	0.821
300	0.584	0.619	0.658	0.689	0.777	0.801	0.830	0.878
500	0.631	0.676	0.722	0.755	0.835	0.862	0.893	0.929

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	0.007	0.025	0.016	0.057	0.008	0.026	0.015	0.057
300	0.006	0.037	0.009	0.048	0.011	0.032	0.008	0.048
500	0.011	0.029	0.012	0.052	0.009	0.043	0.011	0.052
H_1	$\widehat{L}_{1n}$		$\widehat{L}_{2n}$		$\widehat{L}_{1n}$		$\widehat{L}_{2n}$	
100	0.602	0.661	0.643	0.704	0.868	0.899	0.911	0.942
300	0.637	0.682	0.711	0.746	0.909	0.932	0.946	0.977
500	0.716	0.735	0.779	0.818	0.957	0.979	0.991	1.000

REMARK 4.1 (i) Tables 4.1 and 4.2 show that both $\widehat{L}_{1n}$ and $\widehat{L}_{2n}$ 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 $\widehat{L}_{2n}$ is more powerful than $\widehat{L}_{1n}$ 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 $\widehat{L}_{1n}$ and $\widehat{L}_{2n}$ 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 $\widehat{L}_{1n}$ or $\widehat{L}_{2n}$. Similar observations from their simulation study are given in Wang and Phillips (2012) for the univariate version of $\widehat{L}_{1n}$. Tables 4.1 and 4.2 clearly support that the asymptotic theory remains true for the case where there is a type of endogeneity between (x_t, z_t) and e_t as imposed in Assumption 2.1.

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 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.

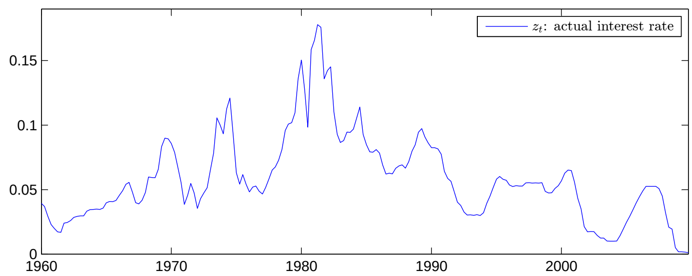

Figure 1a. Nominal interest rate

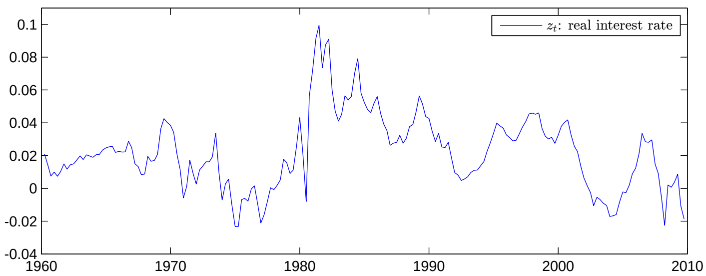

Figure 1b. Real interest rate

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.

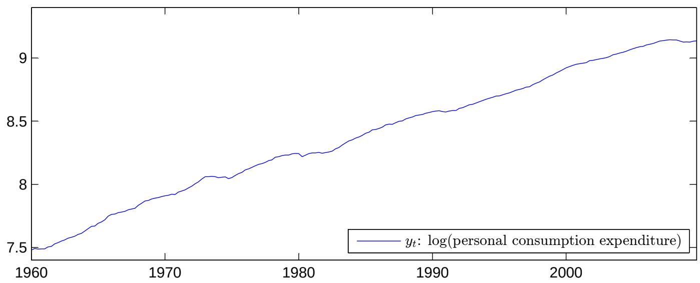

Figure 2a. Plot of c_t

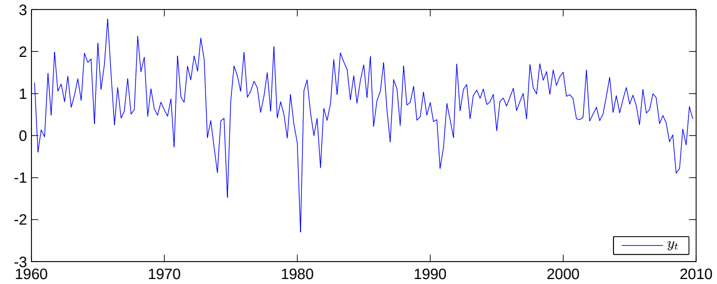

Figure 2b. Plot of $c_t - c_{t-1}$

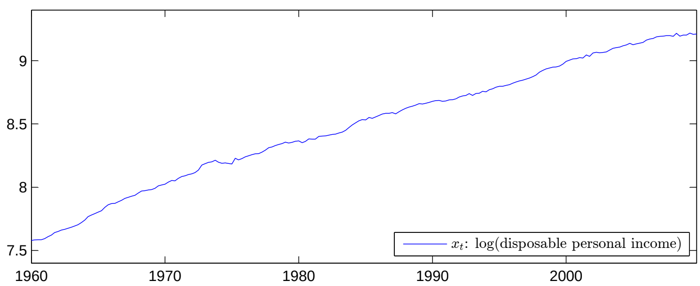

Figure 2c. Plot of I_t

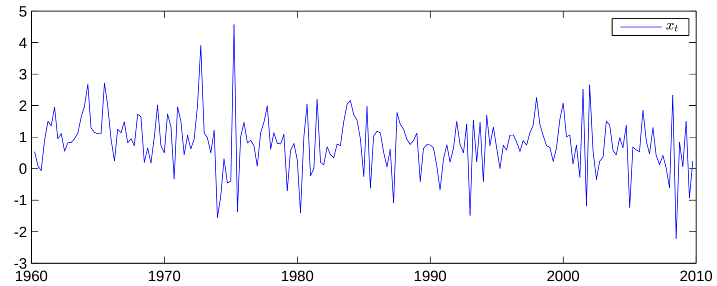

Figure 2d. Plot of $I_t - I_{t-1}$

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).

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.

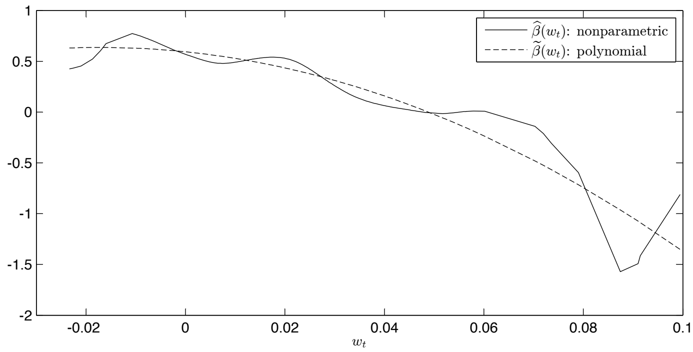

Figure 3. Plots of $\widehat{\beta}(w)$ and $\widetilde{\beta}(w)$

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

$$H_0: \ y_t = \theta_0 + \theta_1 x_t + \theta_2 x_t w_t + \theta_3 x_t w_t^2 + \theta_4 w_t + e_t \quad \text{versus} \tag{5.3}$$

$$H_1: \ y_t = \theta_0 + \theta_1 x_t + \theta_2 x_t w_t + \theta_3 x_t w_t^2 + \theta_4 w_t + \Delta_n(x_t, w_t) + e_t, \tag{5.4}$$

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

Let $\widehat{Q}_n = \widehat{L}_{1n}\left(\widehat{h}_{1\mathrm{cv}}, \widehat{h}_{2\mathrm{cv}}\right)$ or $\widehat{L}_{2n}\left(\widehat{h}_{1\mathrm{cv}}, \widehat{h}_{2\mathrm{cv}}\right)$. To check whether model (5.3) is appropriate, we propose using the following simulation procedure.

- **Step 1**: Let $g(x, w; \theta) = \theta_0 + \theta_1 x_t + \theta_2 x_t w_t + \theta_3 x_t w_t^2 + \theta_4 w_t$ and $\widehat{\theta}$ be the least squares estimator. Generate the bootstrap residuals $\{e_t^*\}$ by $e_t^* = \widehat{e}_t \eta_t^*$, where $\widehat{e}_t = \left(y_t - g(x_t, w_t; \widehat{\theta})\right)$, and $\{\eta_t^*, 1 \le t \le n\}$ is a sequence of i.i.d. random variables drawn from

$$P\left(\eta_t^* = -\frac{\sqrt{5}-1}{2}\right) = \frac{\sqrt{5}+1}{2\sqrt{5}} \quad \text{and} \quad P\left(\eta_t^* = \frac{\sqrt{5}+1}{2}\right) = \frac{\sqrt{5}-1}{2\sqrt{5}}. \tag{5.5}$$

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

- **Step 3**: Repeat Steps 1–2 $M = 250$ times, find the bootstrap distribution of $\widehat{Q}_n^*$ and then compute the the proportion of $\widehat{Q}_n < \widehat{Q}_n^*$ for model (5.3). This proportion is an approximate P–value of $\widehat{Q}_n$ in each case.

An application of the proposed tests $\widehat{L}_{1n}\left(\widehat{h}_{1\text{cv}}, \widehat{h}_{2\text{cv}}\right)$ and $\widehat{L}_{2n}\left(\widehat{h}_{1\text{cv}}, \widehat{h}_{2\text{cv}}\right)$, along with the proposed simulation procedure, 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:

$$y_t - y_{t-1} = \alpha_0 + \alpha_1(x_t - x_{t-1}) + \alpha_2 w_t + \varepsilon_t \tag{5.6}$$

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 set.

Let $\widetilde{y}_t = y_t - y_{t-1}$ and $\widetilde{x}_t = x_t - x_{t-1}$. In view of models (5.1)–(5.6), we propose using a first–order polynomial function of w_t to replace α_1 in model (5.6) and then compare the following model

$$\widetilde{y}_t = \beta_0 + (\beta_1 + \beta_2 w_t)\,\widetilde{x}_t + \beta_3 w_t + \eta_t \tag{5.7}$$

with a commonly used model of the form

$$\widetilde{y}_t = \gamma_0 + \gamma_1 \widetilde{x}_t + \gamma_2 w_t + \zeta_t. \tag{5.8}$$

The estimated versions of models (5.7) and (5.8) become respectively

$$\widetilde{y}_t = 0.554 + (0.345 - 0.366 w_t)\,\widetilde{x}_t + 0.116 w_t \quad \text{and} \tag{5.9}$$

$$\widetilde{y}_t = 0.558 + 0.341 \widetilde{x}_t - 0.199 w_t, \tag{5.10}$$

with the estimated standard deviations of the parameter estimates being between 0.0241 and 0.0632. An application of $\widehat{L}_{1n}(\widehat{h}_{1\text{cv}}, \widehat{h}_{2\text{cv}})$ to check whether either model (5.7) or model (5.8) 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 (5.7) 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 to investigate 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). Such an issue is left for future research.

7 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 (3.6) and (3.7), 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

This appendix gives three useful lemmas with the proofs of the first two lemmas being given in Appendix C below, while the proof of the third lemma is given in Appendix B below.

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

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

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 $n \to \infty$*

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

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

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \psi_1(x_t)\psi_2(z_t) \to_D L_{B_u}(1,0) \cdot E\left[e_1^2 \psi_2(z_1)\right] \cdot \int_{-\infty}^{\infty} \psi_1(x)dx, \tag{A.5}$$

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

Before the proof of Lemma A.3 is given in Appendix B below and the proofs of Lemmas A.1 and A.2 are given in Appendix C, 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 $n \to \infty$

$$\begin{aligned} S_{1n} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{n} h_1 h_2^d} \sum_{t=1}^{n} K_1^2\left(\frac{x_t - x}{h_1}\right) K_2^2\left(\frac{z_t - z}{h_2}\right) e_t^2 \right) \pi_1(x)\pi_2(z)dzdx \\ &\to_D L_{B_u}(1,0) \cdot \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)dvdu \right) \cdot \int_{-\infty}^{\infty} \pi_2(z)p(z)dz, \end{aligned} \tag{A.6}$$

$$\begin{aligned} S_{2n} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n} h_1 h_2^d} \sum_{t=1}^{n} \sum_{s=1,\neq t}^{n} e_s\, e_t \cdot K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) \\ &\times K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right) \pi_1(x)\pi_2(z)dzdx = o_P(1). \end{aligned} \tag{A.7}$$

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

$$\begin{aligned} S_{1n} &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)\pi_1(x_t - uh_1)\pi_2(z_t - vh_2)dudv \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)\pi_1(x_t)\pi_2(z_t)dudv \\ &+ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)\left(\pi_1(x_t - uh_1) - \pi_1(x_t)\right)\left(\pi_2(z_t - vh_2) - \pi_2(z_t)\right)dudv \\ &+ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)\left(\pi_1(x_t - uh_1) - \pi_1(x_t)\right)\pi_2(z_t)dvdu \\ &+ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)\pi_1(x_t)\left(\pi_2(z_t - vh_2) - \pi_2(z_t)\right)dudv \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \pi_1(x_t)\pi_2(z_t) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)dudv \\ &+ O(h_1 h_2^d) \cdot \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 D_1(x_t)D_2(z_t) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u||v|K_1^2(u)K_2^2(v)dvdu \\ &+ O(h_1) \cdot \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 D_1(x_t)\pi_2(z_t) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u|K_1^2(u)K_2^2(v)dvdu \\ &+ O(h_2^d) \cdot \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \pi_1(x_t)D_2(z_t) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |v|K_1^2(u)K_2^2(v)dvdu \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} e_t^2 \pi_1(x_t)\pi_2(z_t) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1^2(u)K_2^2(v)dudv + o_P(1) \end{aligned}$$

$$\to_D \quad L_{B_u}(1,0)\cdot\int_{-\infty}^{\infty}\pi_2(z)p(z)dz\cdot\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1^2(u)K_2^2(v)dudv, \tag{A.8}$$

which completes the proof of (A.6).

We then prove (A.7). Let

$$\begin{aligned} B(s,t) &\equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)K_1\left(\frac{x_s-x}{h_1}\right) \\ &\quad\times K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx, \\ S_{2n} &= \frac{2}{\sqrt{n}h_1h_2^d}\sum_{t=2}^{n}\sum_{s=1}^{t-1}B(s,t)e_se_t. \end{aligned} \tag{A.9}$$

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

$$\begin{aligned} B(s,t) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_1\left(\frac{x_t-x_s}{h_1}+\frac{x_s-x}{h_1}\right)K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_t-z_s}{h_2}+\frac{z_s-z}{h_2}\right) \\ &\quad\times\ K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx \\ &= h_1h_2^d\cdot\int_{-\infty}^{\infty}K_1\left(\frac{x_t-x_s}{h_1}+u\right)K_1\left(u\right)\pi_1(x_s-uh_1)du \\ &\quad\times\ \int_{-\infty}^{\infty}K_2\left(\frac{z_t-z_s}{h_2}+v\right)K_2\left(v\right)\pi_2(z_s-vh_2)dv \\ &= h_1h_2^d\ \pi_1(x_s)\ \pi_2(z_s)\cdot L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right) \\ &\quad+\ h_1h_2^d\cdot\int_{-\infty}^{\infty}K_1\left(\frac{x_t-x_s}{h_1}+u\right)K_1\left(u\right)\left(\pi_1(x_s-uh_1)-\pi_1(x_s)\right)du \\ &\quad\times\ \int_{-\infty}^{\infty}K_2\left(\frac{z_t-z_s}{h_2}+v\right)K_2\left(v\right)\left(\pi_2(z_s-vh_2)-\pi_2(z_s)\right)dv \\ &\quad+h_1h_2^d\pi_2(z_s)L_2\left(\frac{z_t-z_s}{h_2}\right)\int_{-\infty}^{\infty}K_1\left(\frac{x_t-x_s}{h_1}+u\right)K_1\left(u\right)\left(\pi_1(x_s-uh_1)-\pi_1(x_s)\right)du \\ &\quad+h_1h_2^d\pi_1(x_s)L_1\left(\frac{x_t-x_s}{h_1}\right)\int_{-\infty}^{\infty}K_2\left(\frac{z_t-z_s}{h_2}+v\right)K_2\left(v\right)\left(\pi_2(z_s-vh_2)-\pi_2(z_s)\right)dv \\ &\equiv h_1h_2^d\left(B_1(s,t)+B_2(s,t)+B_3(s,t)+B_4(s,t)\right), \end{aligned} \tag{A.10}$$

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

$$\begin{aligned} S_{2n} =& \frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}B_1(s,t)e_se_t+\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}B_2(s,t)e_se_t \\ &+\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}B_3(s,t)e_se_t+\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}B_4(s,t)e_se_t. \end{aligned} \tag{A.11}$$

Similar to the proof of Lemma A.1, we are able to utilise the α-mixing property for e_t, and making use of the joint density $f_{ts}(x,y,u,v)$ of (x_t,x_s,z_t,z_s), which then can be approximated by the product of their marginal densities due to Lemma B.1, we may have

$$\begin{aligned} &\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}E\left[|B_1(s,t)e_se_t|\right] \\ \leq& C\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(t-s)E\left[\pi_1(x_s)\pi_2(z_s)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\right] \end{aligned}$$

$$
\begin{aligned}
=&C\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(t-s)\frac{1}{\sqrt{ts}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\pi_1(y)\pi_2(v)L_1\left(\frac{x-y}{h_1}\right)L_2\left(\frac{u-v}{h_2}\right)\\
&\times f_{ts}\left(\frac{x}{\sqrt{t}},\frac{y}{\sqrt{s}},u,v\right)dxdydudv\\
=&Ch_1h_2^d\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(t-s)\frac{1}{\sqrt{ts}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\pi_1(x-h_1y)\pi_2(u-h_2v)L_1(y)L_2(v)\\
&\times f_{ts}\left(\frac{x}{\sqrt{t}},\frac{x-h_1y}{\sqrt{s}},u,u-h_2v\right)dxdydudv\\
\le&Ch_1h_2^d\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(t-s)\frac{1}{\sqrt{ts}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}L_1(y)L_2(v)dydv\\
&\times\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\pi_1(x-h_1y)\pi_2(u-h_2v)p(u)p(u-h_2v)dxdu\\
\le&Ch_1h_2^d\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{k=1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(k)\frac{1}{\sqrt{t(t-k)}}\int_{-\infty}^{\infty}L_1(y)dy\int_{-\infty}^{\infty}L_2(v)dv\int_{-\infty}^{\infty}\pi_1(x)dx\int_{-\infty}^{\infty}\pi_2(u)p(u)du\\
\le&Ch_1h_2^d=o(1),
\end{aligned}
$$

noting that $\int_{-\infty}^{\infty}L_i(v)dv=\left(\int_{-\infty}^{\infty}K_i(v)dv\right)^2<\infty$.

To handle the term involving $B_2(s,t)$, notice that by virtue of Assumption 2.4(iv), we have

$$
\begin{aligned}
|B_2(s,t)|\le&h_1h_2^dD_1(x_s)D_2(z_s)\int_{-\infty}^{\infty}K_1\left(\frac{x_t-x_s}{h_1}+u\right)K_1(u)|u|du\\
&\times\int_{-\infty}^{\infty}K_2\left(\frac{z_t-z_s}{h_2}+v\right)K_2(v)\|v\|dv.
\end{aligned}
$$

Then, similar to the above, we have

$$
\begin{aligned}
\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{s=1}^{t-1}E\left[|B_2(s,t)e_se_t|\right]\le&Ch_1^2h_2^{2d}\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\sum_{k=1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(k)\frac{1}{\sqrt{t(t-k)}}\int_{-\infty}^{\infty}K_1(y+u)K_1(u)|u|dudy\\
&\times\int_{-\infty}^{\infty}D_1(x)dx\int_{-\infty}^{\infty}K_2(z+v)K_2(v)\|v\|dvdz\int_{-\infty}^{\infty}D_2(z)p(z)dz=Ch_1^2h_2^{2d}=o(1).
\end{aligned}
$$

The same results can be derived similarly for the terms where $B_3(s,t)$ and $B_4(s,t)$ are involved. These show that $S_{2n}=o_P(1)$ and thus the proof is completed.

Proof of Corollary 2.1. To prove Corollary 2.1, in view of Theorem 2.1, it suffices to show that $\frac{1}{n}\sum_{t=1}^{n}\left(y_t-g(x_t,z_t;\widehat{\theta})\right)^2\pi_2(z_t)\to_P E\left[e_1^2\pi_2(z_1)\right]$. Note that, under H_0, $\widehat{e}_t=y_t-g(x_t,z_t;\widehat{\theta})=e_t+\widehat{r}_n(x_t,z_t;\theta_0)$, where $\widehat{r}_n(x_t,z_t;\theta_0):=g(x_t,z_t;\theta_0)-g(x_t,z_t;\widehat{\theta})$. Thus, we have

$$
\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\left(y_t-g(x_t,z_t;\widehat{\theta})\right)^2\pi_2(z_t)=&\frac{1}{n}\sum_{t=1}^{n}e_t^2\pi_2(z_t)+\frac{1}{n}\sum_{t=1}^{n}\widehat{r}_n^2(x_t,z_t;\theta_0)\pi_2(z_t)\\
&+\frac{2}{n}\sum_{t=1}^{n}\widehat{r}_n(x_t,z_t;\theta_0)e_t\pi_2(z_t).
\end{aligned}\tag{A.12}
$$

It follows from the stationarity and α-mixing property of (e_t,z_t) that $\frac{1}{n}\sum_{t=1}^{n}e_t^2\pi_2(z_t)\to_P E[e_1^2\pi_2(z_1)]$. Meanwhile, since $\widehat{\theta}\to_P\theta_0$, we need only to consider the case where $\|\widehat{\theta}-\theta_0\|<\epsilon$ holds in probability for some $\epsilon>0$ in what follows. By Assumption 2.3(i), $|\widehat{r}_n(x_t,z_t;\theta_0)|\le|G_1(x_t,z_t;\theta_0)^\tau(\widehat{\theta}-\theta_0)|+G_2(x_t,z_t;\theta_0)\|\widehat{\theta}-\theta_0\|^2$. Whence,

$$
\frac{1}{n}\sum_{t=1}^{n}\widehat{r}_n^2(x_t,z_t;\theta_0)\pi_2(z_t)\le\frac{1}{n}\sum_{t=1}^{n}|G_1(x_t,z_t;\theta_0)^\tau(\widehat{\theta}-\theta_0)|^2\pi_2(z_t)
$$

$$
\begin{aligned}
&+ \frac{1}{n}\sum_{t=1}^{n} G_2^2(x_t, z_t; \theta_0)\pi_2(z_t)\|\widehat{\theta} - \theta_0\|^4 \\
&+ \frac{2}{n}\sum_{t=1}^{n} \pi_2(z_t)|G_1(x_t, z_t; \theta_0)^\tau(\widehat{\theta} - \theta_0)|G_2(x_t, z_t; \theta_0)\|\widehat{\theta} - \theta_0\|^2 \\
&\leq \|\widehat{\theta} - \theta_0\|^2 \cdot \frac{1}{n}\sum_{t=1}^{n} \|G_1(x_t, z_t; \theta_0)\|^2\pi_2(z_t) + \|\widehat{\theta} - \theta_0)\|^4 \cdot \frac{1}{n}\sum_{t=1}^{n} G_2^2(x_t, z_t; \theta_0)\ \pi_2(z_t) \\
&+ \frac{2}{n}\sum_{t=1}^{n} \pi_2(z_t)|G_1(x_t, z_t; \theta_0)^\tau(\widehat{\theta} - \theta_0)|G_2(x_t, z_t; \theta_0) \cdot \|\widehat{\theta} - \theta_0\|^2 \equiv A_1 + A_2 + A_3. \qquad \text{(A.13)}
\end{aligned}
$$

Making use of the joint density $f_{dt}(x, z)$ of (x_t, z_t) by Lemma B.1, Assumption 2.3 (iii) implies $A_1 = o_P(1)$ and $A_2 = o_P(1)$. Since $A_1 = o_P(1)$ and $A_2 = o_P(1)$ imply that $A_3 = o_P(1)$, and hence $\frac{1}{n}\sum_{t=1}^{n} \widehat{r}_n(x_t, z_t; \theta_0)e_t = o_P(1)$ by Cauchy-Schwartz inequality. This finishes the proof.

Appendix B

This appendix gives the proof of Lemma A.3 and then the derivations of equations (3.6) and (3.7) are given in the last part of this appendix.

Let $x_{d,s} = \frac{x_s}{\sqrt{s}}$. Let $f_{st}(x, y, u, v)$, $f_{dt}(x)$, $f_{ds}(y)$, $p(u)$ and $p(v)$ be the joint and marginal density functions of $(x_{d,s}, x_{d,t}, z_s, z_t)$, $x_{d,s}$, $x_{d,t}$, z_s and z_t, respectively. Let $p_t(z|x)$ be the conditional density function of z_t given $x_{d,t}$. We then have Lemma B.1 below; its proof is given in Appendix C.

Lemma B.1. *Let Assumptions 2.1 and 2.2 hold.*

(1) Suppose z_t is assumed in Assumption 2.1(ii)(a). As $t > s$ and $s \to \infty$,

$$
\begin{aligned}
|f_{st}(x, y, u, v) - f_{dt}(x)f_{ds}(y)p(u)p(v)| &\leq C_0(1 + o(1)) \cdot \gamma_{st}(x, y, u, v) \\
&\times \exp\left(-\frac{1}{2}\left(x^2 + y^2 + u^2 + v^2\right)\right) \qquad \text{(B.1)}
\end{aligned}
$$

for some $C_0 > 0$, and $\gamma_{st}(x, y, u, v)$ is a sequence of positive continuous functions with $\gamma_{st}(x, y, u, v) = O\left(\max\left(\sqrt{\frac{s}{t}}, \frac{1}{\sqrt{s}}\right)\right)$ for given (x, y, u, v).

In addition, for some $C_1 > 0$ and $C_2 > 0$ and as $t \to \infty$

$$
|f_{dt}(x) - \phi(x)| \leq C_1 \cdot x^2 \cdot e^{-\frac{x^2}{2}} \cdot t^{-\frac{1}{2}}, \qquad \text{(B.2)}
$$

$$
|p_t(z|x) - p(z)| \leq C_2 \cdot (x^2 + z^2)\ e^{-\frac{z^2}{2}} \cdot t^{-\frac{1}{2}}. \qquad \text{(B.3)}
$$

(2) *Suppose that $z_t = \Lambda(\varepsilon_{t-1}, \cdots, \varepsilon_{t-\tau}; \eta_t)$ being assumed in Assumption 2.1(ii)(b). Then the conclusions of (i) remain true.*

(3) *Suppose that both z_t and e_t have the linear process forms in Assumption 2.1. Define $w_t = (z_t', e_t)'$. Then the conclusions in (1) still hold with z_t being replaced by w_t.*

(4) *Suppose that both z_t and e_t have the functional forms given in Assumption 2.1. Define $w_t = (z_t', e_t)'$. Then the conclusions in (2) still hold with z_t being replaced by w_t.*

Remark B.1: (i) Before proving this lemma, we point out that the results of Lemma B.1 is extendable to the situation where more variables are involved. For example, the joint density of $(x_{t_1}, \cdots, x_{t_4}, z_{t_1}, \cdots, z_{t_4})$ can be approximated by the product of their marginal densities with ρ_{st} being substituted by $\sqrt{t_4/t_1}$ for the case where $t_1 > \cdots > t_4$ and $t_4 \to \infty$. This may be shown similarly to the proof of Lemma B.1.

(ii) We also point out that Lemma B.1, along with the fact that $f_{st}(x,y)$, $f_{dt}(x)$, $f_{ds}(y)$ and $p_t(z|x)$ are all bounded uniformly in (t,s) (this may be shown in a similar way to the proof of Corollary 2.2 of Wang and Phillips 2009a), is applicable to derive various quantities, such as, $\frac{1}{n^2}\sum_{t=1}^n\sum_{s=1}^n f_{st}(x,y) = (1+o(1))\phi(x)\phi(y)$, $\frac{1}{n}\sum_{t=1}^n f_{dt}(x) = (1+o(1))\phi(x)$, $\frac{1}{n}\sum_{s=1}^n f_{ds}(y) = (1+o(1))\phi(y)$ and $\frac{1}{n}\sum_{t=2}^n p_t(z|x) = (1+o(1))p(z)$, where $\phi(x)$ denotes the density function of $U \sim N(0,1)$. The main steps are given as follows. Letting $m_n \to \infty$ and $\frac{m_n}{n} \to 0$, we have $\frac{1}{n}\sum_{t=1}^n f_{dt}(x) = \frac{1}{n}\sum_{t=1}^{m_n} f_{dt}(x) + \frac{1}{n}\sum_{t=m_n+1}^n (f_{dt}(x)-\phi(x)) + \frac{(n-m_n)}{n}\phi(x) \to \phi(x)$ as $n\to\infty$.

(iii) In addition, Lemma B.1 implies the mutual independence between $x_{d,t}$ and $(x_{d,s}, z_t, z_s, e_t, e_s)$ when $t,s\to\infty$ and $\frac{s}{t}\to 0$. As a consequence, it implies the mutual independence between (z_t, z_s) and $(x_{d,t}, x_{d,s})$, between (e_t, e_s) and $(x_{d,t}, x_{d,s})$, between $x_{d,t}$ and z_t and between $x_{d,t}$ and e_t when $t,s\to\infty$ and $\frac{s}{t}\to 0$. Additionally, (B.2) and (B.3) also imply the marginal densities $f_{dt}(x)$ and the conditional densities $q_t(z|x)$ of z_t given $x_{d,t}=x$ are uniformly bounded over all (t,x) and all (t,z,x), respectively.

Proof of Lemma A.3. Denote $w_t = e_t^2\ \psi_2(z_t)$ and $E_2 = E[w_t]$ for brevity. Note that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n e_t^2\psi_1(x_t)\psi_2(z_t) = E_2\frac{1}{\sqrt{n}}\sum_{t=1}^n\psi_1(x_t) + \frac{1}{\sqrt{n}}\sum_{t=1}^n\psi_1(x_t)[\psi_2(z_t)e_t^2 - E_2] \tag{B.4}$$

It follows that $\frac{1}{\sqrt{n}}\sum_{t=1}^n\psi_1(x_t) \to_D L_{B_u}(1,0)\int_{-\infty}^{\infty}\psi_1(x)dx$. Thus, to prove Lemma A.3, it suffices to show that

$$A_1 := \frac{1}{\sqrt{n}}\sum_{t=1}^n\psi_1(x_t)(w_t - E_2] = o_P(1). \tag{B.5}$$

Let m_n be a sequence such that $m_n\to\infty$ and $m_n^3/n\to 0$ when $n\to\infty$. We deal with A_1 as follows:

$$\begin{aligned}A_1 &= \frac{1}{\sqrt{n}}\sum_{t=1}^n\psi_1(x_t)[w_t - E_2] = \frac{1}{\sqrt{n}}\sum_{t=1}^{m_n}\psi_1(x_t)[w_t - E_2]\\ &+ \frac{1}{\sqrt{n}}\sum_{t=m_n+1}^n\psi_1(x_t)[w_t - E_2] = A_2 + A_3.\end{aligned} \tag{B.6}$$

We first show that $A_2 = o_P(1)$. Because $E|\psi_1(x_t)(w_t-E_2)| \le (E[\psi_1^2(x_t)]E[w_t-E_2]^2)^{1/2} = O(1)\ \sqrt{E[\psi_1^2(x_t)]}$ and

$$E\left[\psi_1^2(x_t)\right] = \int_{-\infty}^{\infty}\psi_1^2(\sqrt{t}x)f_{dt}(x)dx = \frac{1}{\sqrt{t}}\int_{-\infty}^{\infty}\psi_1^2(x)f_{dt}\left(\frac{x}{\sqrt{t}}\right)dx \le C\frac{1}{\sqrt{t}}\int_{-\infty}^{\infty}\psi_1^2(x)dx,$$

by uniformly boundedness of $f_{dt}(\cdot)$, which gives

$$E\left|A_2\right| \le O(1)\frac{1}{\sqrt{n}}\sum_{t=1}^{m_n}\frac{1}{\sqrt[4]{t}} = O(1)\frac{1}{\sqrt{n}}m_n^{5/4} = o(1).$$

Next, we shall show $A_3 = o_P(1)$. Observe that

$$\begin{aligned}E[A_3]^2 =& \frac{1}{n}\sum_{t=m_n+1}^n E\left(\psi_1^2(x_t)[w_t-E_2]^2\right)\\ &+ \frac{1}{n}\sum_{t=m_n+2}^n\sum_{s=m_n+1}^{t-1}E\left(\psi_1(x_t)[w_t-E_2]\psi_1(x_s)[w_s-E_2]\right)\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{n}\sum_{t=m_n+1}^{n}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_1^2(\sqrt{t}x)[w-E_2]^2 r(w)q_t(x|w)dxdw\\
&+\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E\left(\psi_1(x_t)\psi_1(x_s)E\{[w_t-E_2][w_s-E_2]|x_t,x_s\}\right)\\
=&\frac{1}{n}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_1^2(x)[w-E_2]^2 r(w)q_t\left(\frac{x}{\sqrt{t}}|w\right)dxdw\\
&+\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E\left(\psi_1(x_t)\psi_1(x_s)\widetilde{\psi}_w(x_s,x_t)\right)\\
=&\frac{1}{n}\sum_{t=m_n+1}^{n}\frac{1}{\sqrt{t}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_1^2(x)[w-E_2]^2 p_t\left(w|\frac{x}{\sqrt{t}}\right)f_{dt}\left(\frac{x}{\sqrt{t}}\right)dxdw\\
&+\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E\left(\psi_1(x_t)\psi_1(x_s)\widetilde{\psi}_w(x_s,x_t)\right)\\
\le&\frac{C}{\sqrt{n}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_1^2(x)[w-E_2]^2 r(w)dxdw+\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E|\psi_1(x_t)\psi_1(x_s)\widetilde{\psi}_w(x_s,x_t)|\\
=&o(1)+\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E|\psi_1(x_t)\psi_1(x_s)\widetilde{\psi}_z(x_s,x_t)|,
\end{aligned}
$$

where $r(w)$ denotes the marginal density of w_t, $p_t\left(w|\frac{x}{\sqrt{t}}\right)$ denotes the conditional density of w_t given $\frac{x_t}{\sqrt{t}}=\frac{x}{\sqrt{t}}$, $\widetilde{\psi}_w(x_s,x_t):=E\{[w_t-E_2][w_s-E_2]|x_t,x_s\}$ and using Lemma B.1, $p_t\left(w|\frac{x}{\sqrt{t}}\right)$ is approximated by $r(w)$ for large t. To deal with $\widetilde{\psi}_w(x_s,x_t)$, let $p_{st}(w|x,y)$ denote the conditional density of w_u given $x_t=x$ and $x_s=y$ and $q_{st}(w|x,y)$ denote the conditional density of w_u given $\frac{x_t}{\sqrt{t}}=\frac{x}{\sqrt{t}}$ and $\frac{x_s}{\sqrt{s}}=\frac{y}{\sqrt{s}}$. Using the density approximation in Lemma B.1 to imply $q_{st}\left(w|\frac{x}{\sqrt{t}},\frac{y}{\sqrt{s}}\right)\le 2r(w)$, we have for either $u=s$ or $u=t$

$$
\begin{aligned}
&E\left[|w_u-E_2|^{2+\delta}|x_t=x,x_s=y\right]=\int_{-\infty}^{\infty}|w_u-E_2|^{2+\delta}p_{st}(w|x,y)dz\\
=&\int_{-\infty}^{\infty}|w-E_2|^{2+\delta}q_{st}\left(w|\frac{x}{\sqrt{t}},\frac{y}{\sqrt{s}}\right)dw\le 2\int_{-\infty}^{\infty}|w-E_2|^{2+\delta}r(w)dz\\
=&2E\left[|w_u-E_2|^{2+\delta}\right]<\infty.
\end{aligned}
$$

We then apply some properties for the α–mixing condition assumed in Assumption 2.2 on $\{(u_t,z_t)\}$ (see, for example, Lemma A.1 of Gao 2007) to imply that for given (x_t,x_s) and $n\to\infty$

$$
\begin{aligned}
\left|\widetilde{\psi}_w(x_s,x_t)\right|\le&C\cdot\alpha^{\frac{\delta}{2+\delta}}(t-s)\left(E\left[|w_t-E_2|^{2+\delta}|(x_s,x_t)\right]\right)^{\frac{1}{2+\delta}}\\
&\times\left(E\left[|w_s-E_2|^{2+\delta}|(x_s,x_t)\right]\right)^{\frac{1}{2+\delta}}\le C\cdot\alpha^{\frac{\delta}{2+\delta}}(t-s),
\end{aligned}
$$

which gives

$$
\begin{aligned}
E[A_3]^2\le&o(1)+\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E|\psi_1(x_t)\psi_1(x_s)\widetilde{\psi}_w(x_s,x_t)|\\
\le&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(t-s)E|\psi_1(x_t)\psi_1(x_s)|
\end{aligned}
$$

$$
\begin{aligned}
=&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\alpha^{\frac{\delta}{2+\delta}}(t-s)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\psi_1(\sqrt{t}x)\psi_1(\sqrt{s}y)|f_{ts}(x,y)dxdy\\
=&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{\sqrt{t}}\frac{1}{\sqrt{s}}\alpha^{\frac{\delta}{2+\delta}}(t-s)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\psi_1(x)\psi_1(y)|f_{ts}\left(\frac{x}{\sqrt{t}},\frac{y}{\sqrt{s}}\right)dxdy\\
=&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{\sqrt{t}}\frac{1}{\sqrt{s}}\alpha^{\frac{\delta}{2+\delta}}(t-s)\\
=&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\sum_{k=1}^{t-(m_n+1)}\frac{1}{\sqrt{t}}\frac{1}{\sqrt{t-k}}\alpha^{\frac{\delta}{2+\delta}}(k)\\
\le&O(1)\frac{1}{n}\sum_{t=m_n+2}^{n}\frac{1}{\sqrt{t}}\sum_{k=1}^{\infty}\alpha^{\frac{\delta}{2+\delta}}(k)=O(1)\frac{1}{n}\sqrt{n}=o(1),
\end{aligned}
$$

where $f_{ts}(x,y)$ is the joint density of $(x_{d,t}, x_{d,s})$, which is uniformly bounded (see Remark of Lemma B.1). Therefore, the proof of Lemma A.3 is completed.

Derivations of (3.7) and (3.8). Similarly to the proof of Lemma A.2, under H_1, we have as $n \to \infty$

$$
\begin{aligned}
&\sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t-x_s}{h_1}\right)K_2\left(\frac{z_t-z_s}{h_2}\right)\ \widehat{e}_s\ \widehat{e}_t\\
&=\sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t-x_s}{h_1}\right)K_2\left(\frac{z_t-z_s}{h_2}\right)\ e_s\ e_t\\
&+\delta_n^2\sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t-x_s}{h_1}\right)K_2\left(\frac{z_t-z_s}{h_2}\right)\ \Delta(x_s,z_s)\ \Delta(x_t,z_t)+o_P(1)\\
&=\ \delta_n^2\cdot\sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t-x_s}{h_1}\right)K_2\left(\frac{z_t-z_s}{h_2}\right)\ \Delta(x_s,z_s)\ \Delta(x_t,z_t)+o_P(1)\\
&\equiv\ \sum_{t=1}^{n}\sum_{s=1}^{n}K_1\left(\frac{x_t-x_s}{h_1}\right)K_2\left(\frac{z_t-z_s}{h_2}\right)\ e_s\ e_t+\ \delta_n^2\cdot\Pi_n+o_P(1),
\end{aligned}
\tag{B.7}
$$

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

For $t > s$, let $p_{st}(x,y,u,v)$ and $q_{st}(x,y,u,v)$ denote the joint densities of $(x_t - x_s, x_s, z_t - z_s, z_s)$ and $\left(\frac{x_t-x_s}{\sqrt{t-s}}, \frac{x_s}{\sqrt{s}}, z_t - z_s, z_s\right)$, respectively. Straightforward derivations, similar to equation (C.4), imply that as $n \to \infty$

$$
\begin{aligned}
E\left[\Pi_n\right]=&2\sum_{t=2}^{n}\sum_{s=1}^{t-1}K_1\left(\frac{x_t-x_s}{h_1}\right)K_2\left(\frac{z_t-z_s}{h_2}\right)\ \Delta(x_s,z_s)\ \Delta(x_t,z_t)\\
=&\ 2\sum_{t=2}^{n}\sum_{s=1}^{t-1}\int\cdots\int K_1\left(\frac{x}{h_1}\right)K_2\left(\frac{u}{h_2}\right)\ \Delta(y,v)\ \Delta(x+y,u+v)p_{st}(x,y,u,v)dvdudydx\\
=&\ 2h_1h_2^d\sum_{t=2}^{n}\sum_{s=1}^{t-1}\frac{1}{\sqrt{(t-s)s}}\int\cdots\int K_1(x_1)K_2(u_1)\Delta(y,v)\Delta(y+x_1h_1,v+u_1h_2)\\
&\times q_{st}\left(\frac{x_1h_1}{\sqrt{t-s}},\frac{y}{\sqrt{s}},u,v\right)dvdudu_1dx_1=\ C_1(1+o(1))\ nh_1h_2^d
\end{aligned}
\tag{B.8}
$$

using Lemma B.1, where $C_1 > 0$ is some constant.

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

$$\sigma_{2n}^2 \equiv E\left[\sum_{t=1}^n \sum_{s=1}^n K_1^2\left(\frac{x_t - x_s}{h_1}\right) K_2^2\left(\frac{z_t - z_s}{h_2}\right) e_s^2 e_t^2\right] = C_2(1+o(1))\ n^{\frac{3}{2}} h_1 h_2^d \tag{B.9}$$

In a similar fashion to the derivations used in the proof of Lemma A.2, we have as $n \to \infty$ and for some $C_1 > 0$ and $C_2 > 0$

$$\begin{aligned}
&\sum_{t=2}^n \sum_{s=1}^{t-1} E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right) K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right)\right.\\
&\times \left.\Delta(x_s, z_s)\Delta(x_t, z_t)\right] \pi_1(x)\pi_2(z) dz dx\\
&= \sum_{t=2}^n \sum_{s=1}^{t-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\left[K_1\left(\frac{x_t - x_s}{h_1} + \frac{x_s - x}{h_1}\right) K_2\left(\frac{z_t - z_s}{h_2} + \frac{z_s - z}{h_2}\right)\right.\\
&\times \left. K_1\left(\frac{x_s - x}{h_1}\right) K_2\left(\frac{z_s - z}{h_2}\right) \Delta(x_s, z_s)\Delta(x_t, z_t)\right] \pi_1(x)\pi_2(z) dz dx\\
&= (1+o(1)) h_1 h_2^d \sum_{t=2}^n \sum_{s=1}^{t-1} E\left[L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right) \Delta(x_t, z_t)\Delta(x_s, z_s)\pi_1(x_s)\pi_2(z_s)\right]\\
&= (1+o(1)) h_1 h_2^d \sum_{t=2}^n \sum_{s=1}^{t-1} \int \cdots \int L_1\left(\frac{x}{h_1}\right) L_2\left(\frac{u}{h_2}\right)\ \Delta(y, v)\ \Delta(x+y, u+v)\pi_1(y)\pi_2(v)\\
&\times p_{st}(x, y, u, v) dv du dy dx\\
&= (1+o(1)) h_1^2 h_2^{2d} \sum_{t=2}^n \sum_{s=1}^{t-1} \frac{1}{\sqrt{(t-s)s}} \int \cdots \int L_1(x_1) L_2(u_1)\Delta(y, v)\Delta(y + x_1 h_1, v + u_1 h_2)\\
&\times q_{st}\left(\frac{x_1 h_1}{\sqrt{t-s}}, \frac{y}{\sqrt{s}}, u, v\right) dv du du_1 dx_1 = \ C_1(1+o(1))\ n h_1^2 h_2^{2d},
\end{aligned} \tag{B.10}$$

where we haves used Lemma B.1 and Remark B.1 in a similar way to the derivations in equation (B.8), and $L_i(u) = \int K_i(u+v)K_i(v)dv$ for $i = 1, 2$.

We therefore have as $n \to \infty$

$$\begin{aligned}
E\left[R_{2n}\right] &= \frac{C_2(1+o(1)) n h_1 h_2^d \delta_n^2}{n^{\frac{3}{4}}\sqrt{h_1 h_2^d}} = C_2(1+o(1))\ \delta_n^2\ \sqrt{\sqrt{n} h_1 h_2^d},\\
E\left[R_{1n}\right] &= C_1(1+o(1))\ \delta_n^2\ \sqrt{n} h_1 h_2^d,
\end{aligned} \tag{B.11}$$

which complete the derivations of (3.7) and (3.8).

Appendix C

This appendix gives the full proofs of Lemmas A.1– A.2 and B.1 listed in Appendices A and B above.

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:

$$\widehat{e}_t = y_t - g(x_t, z_t; \widehat{\theta}) = e_t + g(x_t, z_t; \theta_0) - g(x_t, z_t; \widehat{\theta}) \equiv e_t + r_n(x_t, z_t; \theta_0), \tag{C.1}$$

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

Define

$$\widehat{T}_n \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sqrt{n} h_1 h_2^d} \sum_{t=1}^n \sum_{s=1}^n \widehat{e}_s \widehat{e}_t \cdot K_1\left(\frac{x_t - x}{h_1}\right) K_2\left(\frac{z_t - z}{h_2}\right)$$

$$
\begin{aligned}
&\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx\\
=&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}e_se_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)\\
&\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx\\
&+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}e_sr_n(x_t,z_t;\theta_0)\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)\\
&\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx\\
&+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}e_tr_n(x_s,z_s;\theta_0)\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)\\
&\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx\\
&+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}r_n(x_t,z_t;\theta_0)r_n(x_s,z_s;\theta_0)\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)\\
&\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx\\
\equiv&\widehat{T}_{1n}+\widehat{T}_{2n}+\widehat{T}_{3n}+\widehat{T}_{4n}. \qquad\qquad (C.2)
\end{aligned}
$$

By Assumption 2.3(ii), we have for some $\epsilon>0$ and $\delta>0$, $P(\|\widehat{\theta}-\theta_0\|>\epsilon)<\delta$, 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, we then have

$$
\begin{aligned}
\widehat{T}_{4n}=&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}r_n(x_t,z_t;\theta_0)r_n(x_s,z_s;\theta_0)\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)\\
&\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx\\
\le&\|\widehat{\theta}-\theta_0\|^2\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}\|G_1(x_t,z_t;\theta_0)\|\cdot\|G_1(x_s,z_s;\theta_0)\|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\\
&\times K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx+R_n\\
=&\|\widehat{\theta}-\theta_0\|^2\frac{(1+o_P(1))}{\sqrt{n}}\sum_{t=1}^{n}\sum_{s=1}^{n}\|G_1(x_t,z_t;\theta_0)\|\cdot\|G_1(x_s,z_s;\theta_0)\|\pi_1(x_s)\pi_2(z_s)\\
&\times L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)+R_n\equiv Q_n+R_n, \qquad\qquad (C.3)
\end{aligned}
$$

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

Notice that in Q_n we only need to consider s and t large enough, $s>m_n$ and $t>m_n$ say, where $m_n\to\infty$ and $m_n^4/n\to 0$ as $n\to\infty$, otherwise we could control it as small as we wish by virtue of Lemma B.1.

Using the joint density $f_{st}(x,y,u,v)$ of $(x_{d,t},x_{d,s},z_t,z_s)$ which is approximated by the product of their marginal densities due to Lemma B.1 for s and t being large, we have as $n\to\infty$

$$
2\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E\left[\|G_1(x_t,z_t;\theta_0)\|\cdot\|G_1(x_s,z_s;\theta_0)\|\pi_1(x_s)\pi_2(z_s)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\right]
$$

$$=2\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{\sqrt{st}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\|G_1(x,u;\theta_0)\|\cdot\|G_1(y,v;\theta_0)\|\pi_1(y)\pi_2(v)L_1\left(\frac{x-y}{h_1}\right)L_2\left(\frac{u-v}{h_2}\right)$$
$$\times f_{st}\left(\frac{x}{\sqrt{t}},\frac{y}{\sqrt{s}},u,v\right)dxdydudv$$
$$\leq Ch_1h_2^d\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{\sqrt{st}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\|G_1(x,u;\theta_0)\|\cdot\|G_1(x-h_1y,u-h_2v;\theta_0)\|$$
$$\times\pi_1(x-h_1y)\pi_2(u-h_2v)L_1(y)L_2(v)p(u)p(u-h_2v)dxdydudv$$
$$=Cnh_1h_2^d\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\|G_1(x,u;\theta_0)\|^2\pi_1(x)\pi_2(u)p^2(u)dxdu\int_{-\infty}^{\infty}L_1(y)dy\int_{-\infty}^{\infty}L_2(v)dv=Cnh_1h_2^d, \tag{C.4}$$

due to Assumption 2.3(i) and $\int_{-\infty}^{\infty}L_i(a)da=\left(\int_{-\infty}^{\infty}K_i(a)da\right)^2<\infty$ for $i=1,2$, which, along with Assumption 2.3(ii) and equation (C.3), 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 (C.3), we have

$$\widehat{T}_{2n}=\frac{(1+o_P(1))}{\sqrt{n}}\sum_{t=1}^{n}\sum_{s=1}^{n}e_tr_n(x_s,z_s;\theta_0)\cdot L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)$$
$$=\frac{(1+o_P(1))}{\sqrt{n}}\left(\theta_0-\widehat{\theta}\right)\sum_{t=1}^{n}\left(\sum_{s=1}^{n}G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right)e_t$$
$$+\ \frac{(1+o_P(1))}{\sqrt{n}}\sum_{t=1}^{n}\sum_{s=1}^{n}\left(r_n(x_s,z_s;\theta_0)-G_1(x_s,z_s;\theta_0)\left(\theta_0-\widehat{\theta}\right)\right)e_t$$
$$\times\ L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)$$
$$\equiv\frac{(1+o_P(1))}{\sqrt{n}}\left(\theta_0-\widehat{\theta}\right)\cdot I_{1n}+\frac{(1+o_P(1))}{\sqrt{n}}\cdot I_{2n}. \tag{C.5}$$

Observe that

$$E\left[I_{1n}^2\right]=E\left[\sum_{t=1}^{n}\left(\sum_{s=1}^{n}G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right)e_t\right]^2$$
$$=\sum_{t=1}^{n}E\left(\left[\sum_{s=1}^{n}G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right]^2e_t^2\right)$$
$$+\,2\sum_{t_1=2}^{n}\sum_{t_2=1}^{t_1-1}\sum_{s_1=1}^{n}E\left[G_1(x_{s_1},z_{s_1};\theta_0)L_1\left(\frac{x_{t_1}-x_{s_1}}{h_1}\right)L_2\left(\frac{z_{t_1}-z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})e_{t_1}\right.$$
$$\left.\times\sum_{s_2=1}^{n}G_1(x_{s_2},z_{s_2};\theta_0)L_1\left(\frac{x_{t_2}-x_{s_2}}{h_1}\right)L_2\left(\frac{z_{t_2}-z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})e_{t_2}\right]$$
$$=\sum_{t=1}^{n}\sum_{s=1}^{n}E\left(\left[G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right]^2e_t^2\right)$$
$$+\,2\sum_{t=1}^{n}\sum_{s_1=2}^{n}E\left[e_t^2G_1(x_{s_1},z_{s_1};\theta_0)L_1\left(\frac{x_t-x_{s_1}}{h_1}\right)L_2\left(\frac{z_t-z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})\right.$$
$$\left.\times\sum_{s_2=1}^{s_1-1}G_1(x_{s_2},z_{s_2};\theta_0)L_1\left(\frac{x_t-x_{s_2}}{h_1}\right)L_2\left(\frac{z_t-z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})\right]$$

$$
\begin{aligned}
&+2\sum_{t_1=2}^{n}\sum_{t_2=1}^{t_1-1}\sum_{s=1}^{n}E\left[e_{t_1}e_{t_2}G_1^2(x_s,z_s;\theta_0)\pi_1^2(x_s)\pi_2^2(z_s)L_1\left(\frac{x_{t_1}-x_s}{h_1}\right)L_2\left(\frac{z_{t_1}-z_s}{h_2}\right)\right.\\
&\qquad\left.\times L_1\left(\frac{x_{t_2}-x_s}{h_1}\right)L_2\left(\frac{z_{t_2}-z_s}{h_2}\right)\right]\\
&+4\sum_{t_1=2}^{n}\sum_{t_2=1}^{t_1-1}\sum_{s_1=2}^{n}E\left[e_{t_1}e_{t_2}G_1(x_{s_1},z_{s_1};\theta_0)L_1\left(\frac{x_{t_1}-x_{s_1}}{h_1}\right)L_2\left(\frac{z_{t_1}-z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})\right.\\
&\qquad\left.\times\sum_{s_2=1}^{s_1-1}G_1(x_{s_2},z_{s_2};\theta_0)L_1\left(\frac{x_{t_2}-x_{s_2}}{h_1}\right)L_2\left(\frac{z_{t_2}-z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})\right]\\
&:=J_1+J_2+J_3+J_4,\quad\text{say.}
\end{aligned}
$$

As pointed before, in what follows we may only consider $t, s > m_n$ where $m_n \to \infty$ and $m_n^4/n \to 0$ as $n \to \infty$. Notice that

$$
\begin{aligned}
J_1=&\sum_{t=m_n+1}^{n}\sum_{s=m_n+1}^{n}E\left(\left[G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right]^2e_t^2\right)\\
=&\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E\left(\left[G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right]^2e_t^2\right)\\
&+\sum_{t=m_n+1}^{n-1}\sum_{s=t}^{n}E\left(\left[G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right]^2e_t^2\right)\\
:=&J_1'+J_1''.
\end{aligned}
$$

Note also that by Lemma B.1, the joint density $f_{ts}(x, y, u, v, w)$ of $(x_{d,t}, x_{d,s}, z_t, z_s, e_t)$ can be approximated by the product of their marginal densities. Hence,

$$
\begin{aligned}
J_1'=&\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}E\left(\left[G_1(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)\right]^2e_t^2\right)\\
=&\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[G_1(\sqrt{s}y,v;\theta_0)L_1\left(\frac{\sqrt{t}x-\sqrt{s}y}{h_1}\right)L_2\left(\frac{u-v}{h_2}\right)\pi_1(\sqrt{s}y)\pi_2(v)\right]^2\\
&\times w^2f_{ts}(x,y,u,v,w)dx\cdots dw\\
=&\sum_{t=m_n+2}^{n}\sum_{s=m_n+1}^{t-1}\frac{1}{\sqrt{ts}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[G_1(y,v;\theta_0)L_1\left(\frac{x-y}{h_1}\right)L_2\left(\frac{u-v}{h_2}\right)\pi_1(y)\pi_2(v)\right]^2\\
&\times w^2f_{ts}\left(\frac{x}{\sqrt{t}},\frac{y}{\sqrt{s}},u,v,w\right)dx\cdots dw\\
\le&O(1)nh_1h_2^d\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[G_1(y,v;\theta_0)L_1(a)L_2(b)\pi_1(y)\pi_2(v)\right]^2w^2p(v)\wp(w)dadbdydvdw\\
=&O(1)nh_1h_2^d
\end{aligned}
$$

by Assumption 2.3 (i) and $\int_{-\infty}^{\infty} L_i(a)da < \infty$ for $i = 1, 2$. Similarly, $J_1'' = O(1)nh_1h_2^d$.

For J_2, notice that

$$
\begin{aligned}
J_2=&\sum_{t=1}^{n}\sum_{s_1=2}^{n}\sum_{s_2=1}^{s_1-1}E\left[e_t^2G_1(x_{s_1},z_{s_1};\theta_0)L_1\left(\frac{x_t-x_{s_1}}{h_1}\right)L_2\left(\frac{z_t-z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})\right.\\
&\left.\times G_1(x_{s_2},z_{s_2};\theta_0)L_1\left(\frac{x_t-x_{s_2}}{h_1}\right)L_2\left(\frac{z_t-z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})\right]
\end{aligned}
$$

$$
\begin{aligned}
=&\sum_{t=3}^{n}\sum_{s_1=2}^{t-1}\sum_{s_2=1}^{s_1-1} E\left[e_t^2 G_1(x_{s_1}, z_{s_1};\theta_0) L_1\left(\frac{x_t - x_{s_1}}{h_1}\right) L_2\left(\frac{z_t - z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})\right.\\
&\left.\times G_1(x_{s_2}, z_{s_2};\theta_0) L_1\left(\frac{x_t - x_{s_2}}{h_1}\right) L_2\left(\frac{z_t - z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})\right]\\
&+\sum_{t=1}^{n-2}\sum_{s_1=t}^{n}\sum_{s_2=1}^{s_1-1} E\left[e_t^2 G_1(x_{s_1}, z_{s_1};\theta_0) L_1\left(\frac{x_t - x_{s_1}}{h_1}\right) L_2\left(\frac{z_t - z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})\right.\\
&\left.\times G_1(x_{s_2}, z_{s_2};\theta_0) L_1\left(\frac{x_t - x_{s_2}}{h_1}\right) L_2\left(\frac{z_t - z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})\right].
\end{aligned}
$$

Because for these two terms the following calculation is the same, we only consider the first term. Once again, we only need to consider the case where t, s_1, s_2 are large. As pointed out in Remark of Lemma B.1, we may have an approximation of the joint density $f_{ts_1s_2}$ of $(x_{d,t}, x_{d,s_1}, x_{d,s_2}, z_{s_1}, z_{s_2}, e_t)$ by the product of their marginal densities. Thus, the first term is bounded in absolute value by

$$
\begin{aligned}
&\sum_{t=m_n+3}^{n}\sum_{s_1=m_n+2}^{t-1}\sum_{s_2=m_n+1}^{s_1-1} E\left[e_t^2 |G_1(x_{s_1}, z_{s_1};\theta_0)| L_1\left(\frac{x_t - x_{s_1}}{h_1}\right) L_2\left(\frac{z_t - z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})\right.\\
&\left.\times |G_1(x_{s_2}, z_{s_2};\theta_0)| L_1\left(\frac{x_t - x_{s_2}}{h_1}\right) L_2\left(\frac{z_t - z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})\right]\\
=&\sum_{t=m_n+3}^{n}\sum_{s_1=m_n+2}^{t-1}\sum_{s_2=m_n+1}^{s_1-1}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|G_1(\sqrt{s_1}x'', z'';\theta_0)| L_1\left(\frac{\sqrt{t}x' - \sqrt{s_1}x''}{h_1}\right) L_2\left(\frac{z' - z''}{h_2}\right)\\
&\pi_1(\sqrt{s_1}x'')\pi_2(z'')|G_1(\sqrt{s_2}x''', z''';\theta_0)| L_1\left(\frac{\sqrt{t}x' - \sqrt{s_2}x'''}{h_1}\right) L_2\left(\frac{z' - z'''}{h_2}\right)\\
&\times \pi_1(\sqrt{s_2}x''')\pi_2(z''')u^2 f_{ts_1s_2}(x', \cdots, z''', u)dx'\cdots dz'''du\\
=&\sum_{t=m_n+3}^{n}\sum_{s_1=m_n+2}^{t-1}\sum_{s_2=m_n+1}^{s_1-1}\frac{1}{\sqrt{t}}\frac{1}{\sqrt{s_1}}\frac{1}{\sqrt{s_2}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|G_1(x'', z'';\theta_0)| L_1\left(\frac{x' - x''}{h_1}\right) L_2\left(\frac{z' - z''}{h_2}\right)\\
&\pi_1(x'')\pi_2(z'')|G_1(x''', z''';\theta_0)| L_1\left(\frac{x' - x'''}{h_1}\right) L_2\left(\frac{z' - z'''}{h_2}\right)\\
&\times \pi_1(x''')\pi_2(z''')u^2 f_{ts_1s_2}\left(\frac{x'}{\sqrt{t}}, \frac{x''}{\sqrt{s_1}}, \frac{x'''}{\sqrt{s_2}}, z', z''z''', u\right)dx'\cdots dz'''du.
\end{aligned}
$$

Under the change of variables, $\frac{x'-x''}{h_1} = v_1$, $\frac{x'-x'''}{h_1} = v_2$, $\frac{z'-z''}{h_2} = v_3$ and $\frac{z'-z'''}{h_2} = v_4$, the integral becomes

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\|G_1(x'', z'';\theta_0)\| L_1\left(\frac{x' - x''}{h_1}\right) L_2\left(\frac{z' - z''}{h_2}\right)\pi_1(x'')\pi_2(z'')\|G_1(x''', z''';\theta_0)\| L_1\left(\frac{x' - x'''}{h_1}\right)\\
&\times L_2\left(\frac{z' - z'''}{h_2}\right)\pi_1(x''')\pi_2(z''')u^2 f_{ts_1s_2}\left(\frac{x'}{\sqrt{t}}, \frac{x''}{\sqrt{s_1}}, \frac{x'''}{\sqrt{s_2}}, z', z'', z''', u\right)dx'\cdots dz'''du\\
=&h_1^2h_2^{2d}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\|G_1(x' - h_1v_1, z' - h_2v_3;\theta_0)\| L_1(v_1)L_2(v_3)\pi_1(x' - h_1v_1)\pi_2(z' - h_2v_3)\\
&\times\|G_1(x' - h_1v_2, z' - h_2v_4;\theta_0)\| L_1(v_2)L_2(v_4)\pi_1(x' - h_1v_2)\pi_2(z' - h_2v_4)u^2\\
&\times f_{ts_1s_2}\left(\frac{x'}{\sqrt{t}}, \frac{x' - h_1v_1}{\sqrt{s_1}}, \frac{x' - h_1v_2}{\sqrt{s_2}}, z', z' - h_2v_3, z' - h_2v_4, u\right)dx'dz'dv_1\cdots dv_4du\\
\le&Ch_1^2h_2^{2d}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\|G_1(x' - h_1v_1, z' - h_2v_3;\theta_0)\| L_1(v_1)L_2(v_3)\pi_1(x' - h_1v_1)\pi_2(z' - h_2v_3)\\
&\times\|G_1(x' - h_1v_2, z' - h_2v_4;\theta_0)\| L_1(v_2)L_2(v_4)\pi_1(x' - h_1v_2)\pi_2(z' - h_2v_4)\\
&\times p(z')p(z' - h_2v_3)p(z' - h_2v_4)dx'dz'dv_1\cdots dv_4
\end{aligned}
$$

$$
\begin{aligned}
=&Ch_1^2h_2^{2d}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}L_1(v_1)L_2(v_3)L_1(v_2)L_2(v_4)dv_1\cdots dv_4\\
&\times\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\|G_1(x'-h_1v_1,z'-h_2v_3;\theta_0)\|\pi_1(x'-h_1v_1)\pi_2(z'-h_2v_3)p(z'-h_2v_3)\\
&\times\|G_1(x'-h_1v_2,z'-h_2v_4;\theta_0)\|\pi_1(x'-h_1v_2)\pi_2(z'-h_2v_4)p(z'-h_2v_4)dx'dz'\\
\le&Ch_1^2h_2^{2d}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}L_1(v_1)L_2(v_3)L_1(v_2)L_2(v_4)dv_1\cdots dv_4\\
&\times\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\|G_1(x'-h_1v_1,z'-h_2v_3;\theta_0)\|^2\pi_1^2(x'-h_1v_1)\pi_2^2(z'-h_2v_3)p^2(z'-h_2v_3)dx'dz'\right)^{1/2}\\
&\times\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\|G_1(x'-h_1v_2,z'-h_2v_4;\theta_0)\|^2\pi_1^2(x'-h_1v_2)\pi_2^2(z'-h_2v_4)p^2(z'-h_2v_4)dx'dz'\right)^{1/2}\\
=&Ch_1^2h_2^{2d}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}L_1(v_1)L_2(v_3)L_1(v_2)L_2(v_4)dv_1\cdots dv_4\\
&\times\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\|G_1(x,z;\theta_0)\|^2\pi_1^2(x)\pi_2^2(z)p^2(z)dxdz\right)^{1/2}\\
&\times\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\|G_1(x,z;\theta_0)\|^2\pi_1^2(x)\pi_2^2(z)p^2(z)dxdz\right)^{1/2}\\
=&Ch_1^2h_2^{2d}\left(\int_{-\infty}^{\infty}L_1(u)du\int_{-\infty}^{\infty}L_2(v)dv\right)^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\|G_1(x,z;\theta_0)\|^2\pi_1^2(x)\pi_2^2(z)p^2(z)dxdz\\
=&O(h_1^2h_2^{2d}),
\end{aligned}
$$

by Assumption 2.3(i) and $\int_{-\infty}^{\infty}L_i(a)da<\infty$ for $i=1,2$. Hence, $J_2=O(\sqrt{n}^3h_1^2h_2^{2d})$.

We are next about to tackle J_4. Since if J_4 is calculated, it is easy to deal with J_3. To this end, we first partition J_4 into 6 parts according as (1) $t_1>t_2>s_1>s_2$, (2) $t_1>s_1>t_2>s_2$, (3) $t_1>s_1>s_2>t_2$ and (4)-(6) that are the cases (1)-(3) but exchange the positions of t_i's and s_i's mutually. Due to similarity, in what follows we only consider the first part, namely, $t_1>t_2>s_1>s_2$.

Second, consider the conditional expectation of $E[|e_t|^{2+\delta}|x_{t_i},x_{s_i},z_{t_i},z_{s_i},i=1\text{ and }2]$ for $t=t_1,t_2$ where δ is given by Assumption 2.1. Since the conditional density of e_t given $x_{t_i},x_{s_i},z_{t_i},z_{s_i},i=1$ and 2 can be approximated by the density of e_t, the conditional expectation is confined by $CE[|e_t|^{2+\delta}]$ where C is some constant. This entails the application of α-mixing property for e_t. More precisely,

$$|E[e_{t_1}e_{t_2}|x_{t_i},x_{s_i},z_{t_i},z_{s_i},i=1\text{ and }2]|\le C\alpha^{\frac{\delta}{2+\delta}}(t_1-t_2).$$

Third, we shall only consider the case where t_i and s_i are larger than m_n for $i=1,2$. Whence, we have

$$
\begin{aligned}
|J_4|\le&O(1)\sum_{t_1=m_n+4}^{n}\sum_{t_2=m_n+3}^{t_1-1}\sum_{s_1=m_n+2}^{t_2-1}\sum_{s_2=m_n+1}^{s_1-1}\alpha^{\frac{\delta}{2+\delta}}(t_1-t_2)E\Big|G_1(x_{s_1},z_{s_1};\theta_0)\\
&\times L_1\left(\frac{x_{t_1}-x_{s_1}}{h_1}\right)L_2\left(\frac{z_{t_1}-z_{s_1}}{h_2}\right)\pi_1(x_{s_1})\pi_2(z_{s_1})G_1(x_{s_2},z_{s_2};\theta_0)\\
&\times L_1\left(\frac{x_{t_2}-x_{s_2}}{h_1}\right)L_2\left(\frac{z_{t_2}-z_{s_2}}{h_2}\right)\pi_1(x_{s_2})\pi_2(z_{s_2})\Big|.
\end{aligned}
$$

Utilising the joint density of $(x_{t_i},x_{s_i},z_{t_i},z_{s_i},i=1,2)$, which is approximated by the product of their marginal densities (see Remark of Lemma B.1), gives

$$|J_4|\le O(1)\sum_{t_1=m_n+4}^{n}\sum_{t_2=m_n+3}^{t_1-1}\sum_{s_1=m_n+2}^{t_2-1}\sum_{s_2=m_n+1}^{s_1-1}\alpha^{\frac{\delta}{2+\delta}}(t_1-t_2)\frac{1}{\sqrt{t_1t_2s_1s_2}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}$$

$$
\begin{aligned}
&\times \Bigg| G_1(w_3, u_3; \theta_0) L_1\left(\frac{w_1 - w_3}{h_1}\right) L_2\left(\frac{u_1 - u_3}{h_2}\right) \pi_1(w_3)\pi_2(u_3) \\
&\times G_1(w_4, u_4; \theta_0) L_1\left(\frac{w_2 - w_4}{h_1}\right) L_2\left(\frac{u_2 - u_4}{h_2}\right) \pi_1(w_4)\pi_2(u_4)\Bigg| \\
&\times f_{t_1 t_2 s_1 s_2}\left(\frac{w_1}{\sqrt{t_1}}, \frac{w_2}{\sqrt{t_2}}, \frac{w_3}{\sqrt{s_1}}, \frac{w_4}{\sqrt{s_2}}, u_1, u_2, u_3, u_4\right) dw_1 \cdots du_4 \\
=& O(1) h_1^2 h_2^{2d} \sum_{t_1=m_n+4}^{n} \sum_{t_2=m_n+3}^{t_1-1} \sum_{s_1=m_n+2}^{t_2-1} \sum_{s_2=m_n+1}^{s_1-1} \alpha^{\frac{\delta}{2+\delta}}(t_1 - t_2) \frac{1}{\sqrt{t_1 t_2 s_1 s_2}} \\
&\times \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |G_1(w_1 - h_1 w_3, u_1 - h_2 u_3; \theta_0)| L_1(w_3) L_2(u_3) \pi_1(w_1 - h_1 w_3) \\
&\times \pi_2(u_1 - h_2 u_3) |G_1(w_2 - h_1 w_4, u_2 - h_2 u_4; \theta_0)| L_1(w_4) L_2(u_4) \pi_1(w_2 - h_1 w_4) \pi_2(u_2 - h_2 u_4) \\
&\times f_{t_1 t_2 s_1 s_2}\left(\frac{w_1}{\sqrt{t_1}}, \frac{w_2}{\sqrt{t_2}}, \frac{w_1 - h_1 w_3}{\sqrt{s_1}}, \frac{w_2 - h_1 w_4}{\sqrt{s_2}}, u_1, u_2, u_1 - h_2 u_3, u_2 - h_2 u_4\right) dw_1 \cdots du_4 \\
\leq& O(1) h_1^2 h_2^{2d} \sum_{t_1=m_n+4}^{n} \sum_{t_2=m_n+3}^{t_1-1} \sum_{s_1=m_n+2}^{t_2-1} \sum_{s_2=m_n+1}^{s_1-1} \alpha^{\frac{\delta}{2+\delta}}(t_1 - t_2) \frac{1}{\sqrt{t_1 t_2 s_1 s_2}} \\
&\times \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L_1(w_3) L_2(u_3) L_1(w_4) L_2(u_4) dw_3 dw_4 du_3 du_4 \\
&\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G_1(w_1 - h_1 w_3, u_1 - h_2 u_3; \theta_0)| \pi_1(w_1 - h_1 w_3) \pi_2(u_1 - h_2 u_3) p(u_1) p(u_1 - h_2 u_3) dw_1 du_1 \\
&\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G_1(w_2 - h_1 w_4, u_2 - h_2 u_4; \theta_0)| \pi_1(w_2 - h_1 w_4) \pi_2(u_2 - h_2 u_4) p(u_2) p(u_2 - h_2 u_4) dw_2 du_2 \\
\leq& O(1) h_1^2 h_2^{2d} \sum_{t_1=m_n+4}^{n} \sum_{t_2=m_n+3}^{t_1-1} \sum_{s_1=m_n+2}^{t_2-1} \sum_{s_2=m_n+1}^{s_1-1} \alpha^{\frac{\delta}{2+\delta}}(t_1 - t_2) \frac{1}{\sqrt{t_1 t_2 s_1 s_2}} \\
&\times \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L_1(w_3) L_2(u_3) L_1(w_4) L_2(u_4) dw_3 dw_4 du_3 du_4 \\
&\times \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |G_1(x, z; \theta_0)| \pi_1(x) \pi_2(z) p(z) dx dz\right)^2 \\
=& O(1) h_1^2 h_2^{2d} \sum_{t_1=m_n+4}^{n} \sum_{t_2=m_n+3}^{t_1-1} \sum_{s_1=m_n+2}^{t_2-1} \sum_{s_2=m_n+1}^{s_1-1} \alpha^{\frac{\delta}{2+\delta}}(t_1 - t_2) \frac{1}{\sqrt{t_1 t_2 s_1 s_2}} \\
=& O(1) h_1^2 h_2^{2d} \sum_{t_1=m_n+4}^{n} \sum_{t_2=m_n+3}^{t_1-1} \frac{\sqrt{t_2}}{\sqrt{t_1}} \alpha^{\frac{\delta}{2+\delta}}(t_1 - t_2) \\
=& O(1) h_1^2 h_2^{2d} \sum_{t_1=m_n+4}^{n} \sum_{k=1}^{t_1-(m_n+3)} \frac{\sqrt{t_1 - k}}{\sqrt{t_1}} \alpha^{\frac{\delta}{2+\delta}}(k) \leq O(1) h_1^2 h_2^{2d} n.
\end{aligned}
$$

It follows that $I_{1n} = O_P(n^{3/4} h_1 h_2^d)$, implying that $\frac{(1+o_P(1))}{\sqrt{n}}(\theta_0 - \widehat{\theta}) \cdot I_{1n} = o_P(1)$ due to Assumption 2.3(ii).

Meanwhile, by Assumption 2.3(i) we have

$$
\begin{aligned}
|I_{2n}| \leq& \sum_{t=1}^{n} \sum_{s=1}^{n} \left| r_n(x_s, z_s; \theta_0) - G_1(x_s, z_s; \theta_0)\left(\theta_0 - \widehat{\theta}\right) \right| L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right) \\
&\times \pi_1(x_s) \pi_2(z_s) |e_t| \\
\leq& ||\widehat{\theta} - \theta_0||^2 \cdot \sum_{t=1}^{n} \sum_{s=1}^{n} G_2(x_s, z_s; \theta_0) L_1\left(\frac{x_t - x_s}{h_1}\right) L_2\left(\frac{z_t - z_s}{h_2}\right) \pi_1(x_s) \pi_2(z_s) |e_t| \\
\equiv& ||\widehat{\theta} - \theta_0||^2 \cdot I_{3n}, \qquad \text{(C.6)}
\end{aligned}
$$

in view of the fact that both $L_i(\cdot)$ and $G_2(x,z;\theta_0)$ are positive. Moreover, similar to the calculation of J_1', we have

$$E[I_{3n}] = \sum_{t=1}^{n}\sum_{s=1}^{n} E\left[G_2(x_s,z_s;\theta_0)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\pi_1(x_s)\pi_2(z_s)|e_t|\right]$$
$$\leq\ Cnh_1h_2^d, \tag{C.7}$$

which implies that $\frac{(1+o_P(1))}{\sqrt{n}}\cdot I_{2n} = \frac{(1+o_P(1))\cdot||\widehat{\theta}-\theta_0||^2}{\sqrt{n}}\cdot I_{3n} = o_P(1)$ in view of Assumption 2.3 (ii) and hence, $\widehat{T}_{2n} = o_P(1)$. The same conclusion is true for $\widehat{T}_{3n}$. Therefore, we have shown under H_0 as $n\to\infty$

$$\widehat{T}_n = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}e_se_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx + o_P(1) \tag{C.8}$$

This 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

$$\widehat{\varepsilon}_t(x,z) = \widehat{e}_t\ K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)\quad\text{and}\quad \widehat{e}_t = e_t + r_n(x_t,z_t;\theta_0)+\Delta_n(x_t,z_t) \tag{C.9}$$

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:

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

$$+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}\Delta_n(x_s,z_s)e_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx\equiv\sum_{j=1}^{4}\widehat{T}_{jn}+\sum_{k=5}^{9}\widehat{T}_{kn}, \tag{C.10}$$

where $\widehat{T}_{jn}$ for $1\le j\le 4$ are the same as in (C.2).

In view of the proof of (C.8) 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:

$$\widehat{T}_{kn}=o_P\left(\delta_n^2\sqrt{n}h_1h_2^d\right)\quad\text{for } k=6,\cdots,9. \tag{C.11}$$

To complete the proof of (C.11), we need only to deal with $\widehat{T}_{6n}$ and $\widehat{T}_{9n}$. Using similar arguments to those used in the proofs of (C.3) and (C.4), we have as $n\to\infty$

$$\sum_{t=1}^{n}\sum_{s=1}^{n}E\left[||G_1(x_t,z_t;\theta_0)||\cdot|Z(x_s,z_s)|\pi_1(x_s)\pi_2(z_s)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\right]\le Cnh_1h_2^d, \tag{C.12}$$

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 $\widehat{T}_{6n}=o_P\left(\delta_n^2nh_1h_2^d\right)$.

Meanwhile, in order to deal with $\widehat{T}_{9n}$, it suffices to show that as $n\to\infty$

$$E\left[\delta_n\sum_{t=1}^{n}\left(\sum_{s=1}^{n}\Delta(x_s,z_s)\pi_1(x_s)\pi_2(z_s)L_1\left(\frac{x_t-x_s}{h_1}\right)L_2\left(\frac{z_t-z_s}{h_2}\right)\right)e_t\right]^2\le C\ \delta_n^2n^{3/2}h_1^2h_2^{2d}, \tag{C.13}$$

which follows similarly from the proof of $I_{1n}=o_P(1)$ in Lemma A.1, taking into account the condition for $\Delta(x,z)$ in (3.4) and (3.5). Equation (C.13) then implies that as $n\to\infty$

$$\widehat{T}_{9n}=O_P\left(\delta_nn^{1/4}h_1h_2^d\right)=o_P\left(\delta_n^2\sqrt{n}h_1h_2^d\right), \tag{C.14}$$

which, along with the proofs of (C.2) and (C.10)–(C.14), shows that under H_1, as $n\to\infty$

$$\widehat{T}_n=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}\widehat{e}_s\widehat{e}_t\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx$$
$$\ge\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{n}h_1h_2^d}\sum_{t=1}^{n}\sum_{s=1}^{n}\Delta_n(x_s,z_s)\Delta_n(x_t,z_t)\cdot K_1\left(\frac{x_t-x}{h_1}\right)K_2\left(\frac{z_t-z}{h_2}\right)$$
$$\times\ K_1\left(\frac{x_s-x}{h_1}\right)K_2\left(\frac{z_s-z}{h_2}\right)\pi_1(x)\pi_2(z)dzdx+o_P(1), \tag{C.15}$$

which completes the proof of Lemma A.2.

Proof of Lemma B.1: We consider without loss of generality the case of $d=1$. Otherwise, we consider $v_t=D^\tau z_t$, where D is a d–dimensional vector of real numbers.

(1) It follows from Assumption 2.1(i-ii) that we have

$$u_t=\sum_{j=0}^{\infty}\phi_j\varepsilon_{t-j}\quad\text{and}\quad z_t=\sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}, \tag{C.16}$$

where $\{\psi_j\}$ satisfies the same conditions that are imposed on $\psi_j = (\psi_{j1}, \cdots, \psi_{jd})^\tau$ in Assumption 2.1(i). To simplify the notation, we assume without loss of generality that $\text{var}(u_t) = \text{var}(z_t) \equiv 1$. In addition, we avoid involving s and t as the indices. For $t > s$, let $Y_1 = \frac{\sum_{i=1}^t u_i}{\sqrt{t}}$, $Y_2 = \frac{\sum_{j=1}^s u_j}{\sqrt{s}}$, $Y_3 = z_t$ and $Y_4 = z_s$.

Assumption 2.1(ii)(a) implies that we have

$$z_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \sum_{i=-\infty}^{t} \psi_{t-i}\varepsilon_i = \sum_{i=-\infty}^{0} \psi_{t-i}\varepsilon_i + \sum_{j=1}^{t} \psi_{t-j}\varepsilon_j. \tag{C.17}$$

Similarly, we have

$$u_t = \sum_{j=0}^{\infty} \phi_j \varepsilon_{t-j} = \sum_{i=-\infty}^{t} \phi_{t-i}\varepsilon_i = \sum_{i=-\infty}^{0} \phi_{t-i}\varepsilon_i + \sum_{j=1}^{t} \phi_{t-j}\varepsilon_j. \tag{C.18}$$

We now define $Y_{1t} = \frac{\sum_{i=1}^t \phi_{t-i}\varepsilon_i}{\sqrt{t}}$ and $Y_{2s} = \frac{\sum_{j=1}^s \phi_{s-j}\varepsilon_j}{\sqrt{s}}$.

In view of equations (C.16)–(C.18), we have the following decompositions: for $t > s$

$$\begin{aligned}
Y_{1t} &= \frac{1}{\sqrt{t}} \sum_{i=1}^{t} \phi_{t-i}\varepsilon_i = \frac{1}{\sqrt{t}} \sum_{i=1}^{s} \phi_{t-i}\varepsilon_i + \frac{1}{\sqrt{t}} \sum_{i=s+1}^{t} \phi_{t-i}\varepsilon_i, \quad Y_{2s} = \frac{1}{\sqrt{s}} \sum_{i=1}^{s} \phi_{s-j}\varepsilon_i, \\
Y_{3t} &= \sum_{j=0}^{t} \psi_j \varepsilon_{t-j} = \sum_{j=1}^{s} \psi_{t-j}\varepsilon_j + \sum_{j=s+1}^{t} \psi_{t-j}\varepsilon_j, \\
Y_{4s} &= \sum_{l=0}^{s} \psi_l \varepsilon_{s-l} = \sum_{l=1}^{s} \psi_{s-l}\varepsilon_l, \\
U_{1t} &= \sum_{i=-\infty}^{0} \phi_{t-i}\varepsilon_i \text{ and } U_{2t} = \sum_{j=-\infty}^{0} \psi_{t-j}\varepsilon_j,
\end{aligned} \tag{C.19}$$

which imply $Y_1 = Y_{1t} + \frac{1}{\sqrt{t}} U_{1t}$, $Y_2 = Y_{2s} + \frac{1}{\sqrt{s}} U_{1s}$, $Y_3 = Y_{3t} + U_{2t}$ and $Y_4 = Y_{4s} + U_{2s}$.

For given $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, we have

$$\lambda_1 Y_1 + \lambda_2 Y_2 + \lambda_3 Y_3 + \lambda_4 Y_4 = \sum_{i=1}^{s} a_{st}(i)\varepsilon_i + \sum_{j=s+1}^{t} b_{st}(j)\varepsilon_j + W_{st}, \tag{C.20}$$

where $a_{st}(i) = \frac{\lambda_1 \phi_{t-i}}{\sqrt{t}} + \frac{\lambda_2 \phi_{s-i}}{\sqrt{s}} + \lambda_3\ \psi_{t-i} + \lambda_4\ \psi_{s-i}$, $b_{st}(j) = \frac{\lambda_1 \phi_{t-j}}{\sqrt{t}} + \lambda_3\ \psi_{t-j}$ and $W_{st} = \frac{\lambda_1 U_{1t}}{\sqrt{t}} + \frac{\lambda_2 U_{1s}}{\sqrt{s}} + \lambda_3 U_{2t} + \lambda_4 U_{2s}$.

Let $\varphi(\cdot)$ be the characteristic function of ε_t. The logarithm of the characteristic function of $(Y_{1t}, Y_{2s}, Y_{3t}, Y_{4s})$ is then given by

$$\begin{aligned}
&\log\left(\Psi_{st}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)\right) = \log\left(\prod_{i=1}^{s} \varphi(a_{st}(i) \prod_{j=s+1}^{t} \varphi(b_{st}(j)) \cdot \exp\left(iW_{st}\right)\right) \\
&= \sum_{i=1}^{s} \log(\varphi(a_{st}(i))) + \sum_{j=s+1}^{t} \log(\varphi(b_{st}(j))) + \log\left(\exp\left(iW_{st}\right)\right) \\
&= \frac{-(1+o(1))}{2} \left(\sum_{i=1}^{s} a_{st}^2(i) + \sum_{j=s+1}^{t} b_{st}^2(j) + E\left[W_{st}^2\right]\right) \\
&= \frac{-(1+o(1))}{2} \sum_{i=1}^{s} \left(\frac{\lambda_1 \phi_{t-i}}{\sqrt{t}} + \frac{\lambda_2 \phi_{s-i}}{\sqrt{s}} + \lambda_3\ \psi_{t-i} + \lambda_4\ \psi_{s-i}\right)^2
\end{aligned}$$

$$+\frac{-(1+o(1))}{2}\sum_{j=s+1}^{t}\left(\frac{\lambda_1\phi_{t-j}}{\sqrt{t}}+\lambda_3\ \psi_{t-j}\right)^2+\frac{-(1+o(1))}{2}E\left[W_{st}^2\right], \tag{C.21}$$

which implies that there is some continuous function $\rho_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4)=O\left(\max\left(\sqrt{\frac{s}{t}},\frac{1}{\sqrt{s}}\right)\right)$ such that

$$\Psi_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4)=(1+o(1))\exp\left(-\frac{1}{2}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_4^2\right)\right)\exp\left(-\rho_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4)\right) \tag{C.22}$$

as $t>s$ and $s\to\infty$, where we have used the following facts: $\gamma(t-s)=E[v_sv_t]=o\left(\sqrt{\frac{s}{t}},\frac{1}{\sqrt{s}}\right)$ by Assumption 2.1 for $v_t=u_t$ or z_t, $\sum_{k=0}^{\infty}\psi_k^2=E\left[v_1^2\right]=1=E[z_1^2]=\sum_{j=0}^{\infty}\phi_j^2$,

$$\sum_{i=1}^{s}\psi_{t-i}^2=\sum_{j=t-s}^{t-1}\psi_j^2=O\left(\frac{s-1}{(t-s)(t-1)}\right)\quad\text{and}\quad\sum_{l=s+1}^{t}\psi_{t+l}^2=O\left(\frac{t-s}{(t+s+1)(2t+1)}\right)$$

for $t>s$ and $s\to\infty$, and

$$\begin{aligned}
E\left[W_{st}^2\right] &= E\left(\frac{\lambda_1U_{1t}}{\sqrt{t}}+\frac{\lambda_2U_{1s}}{\sqrt{s}}+\lambda_3U_{2t}+\lambda_4U_{2s}\right)^2=\frac{\lambda_1^2}{t}E\left[U_{1t}^2\right]\\
&+\frac{\lambda_2^2}{s}E\left[U_{1s}^2\right]+\lambda_3^2E[U_{2t}^2]+\lambda_4^2E[U_{2s}^2]+\frac{2\lambda_1\lambda_2}{\sqrt{st}}E[U_{1t}U_{1s}]+\frac{2\lambda_1\lambda_3}{\sqrt{t}}E[U_{1t}U_{2t}]\\
&+\frac{2\lambda_1\lambda_4}{\sqrt{t}}E[U_{1t}U_{2s}]+\frac{2\lambda_3\lambda_2}{\sqrt{s}}E[U_{2t}U_{1s}]+\frac{2\lambda_4\lambda_2}{\sqrt{s}}E[U_{1s}U_{2s}]+\frac{2\lambda_3\lambda_4}{\sqrt{s}}E[U_{2t}U_{2s}]\\
&= \frac{\lambda_1^2}{t}\sum_{j=-\infty}^{0}\phi_{t-j}^2+\frac{\lambda_2^2}{s}\sum_{j=-\infty}^{0}\phi_{s-j}^2+\lambda_3^2\sum_{j=-\infty}^{0}\psi_{t-j}^2+\lambda_4^2\sum_{j=-\infty}^{0}\psi_{s-j}^2\\
&+\frac{2\lambda_1\lambda_2}{\sqrt{st}}\sum_{k=-\infty}^{0}\phi_{s-k}\phi_{t-k}+\frac{2\lambda_1\lambda_3}{\sqrt{t}}\sum_{k=-\infty}^{0}\phi_{s-k}\psi_{t-k}+\frac{2\lambda_1\lambda_4}{\sqrt{t}}\sum_{k=-\infty}^{0}\phi_{t-k}\psi_{s-k}\\
&+\frac{2\lambda_3\lambda_2}{\sqrt{s}}\sum_{k=-\infty}^{0}\phi_{s-k}\psi_{t-k}+\frac{2\lambda_4\lambda_2}{\sqrt{s}}\sum_{k=-\infty}^{0}\phi_{s-k}\psi_{s-k}+\frac{2\lambda_3\lambda_4}{\sqrt{s}}\sum_{k=-\infty}^{0}\psi_{s-k}\psi_{t-k}\\
&= O\left(s^{-1}\right)\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_4^2+\rho_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4)\right)(1+o(1)).
\end{aligned} \tag{C.23}$$

Meanwhile, the product of the marginal characteristic functions of Y_1, Y_2, Y_3 and Y_4 is given by

$$\Psi_t(\lambda_1)\Psi_s(\lambda_2)\Psi_{3t}(\lambda_3)\Psi_{4s}(\lambda_4)=(1+o(1))\exp\left(-\frac{1}{2}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_4^2\right)\right), \tag{C.24}$$

which implies as $t-s\to\infty$

$$\begin{aligned}
&\Psi_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4)-\Psi_t(\lambda_1)\Psi_s(\lambda_2)\Psi_{3t}(\lambda_3)\Psi_{4s}(\lambda_4)\\
&= (1+o(1))\exp\left(-\frac{1}{2}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_4^2\right)\right)\\
&\times\left(\exp\left(-\frac{\lambda_{st}}{2}\left(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\lambda_4+\lambda_2\lambda_3+\lambda_2\lambda_4+\lambda_3\lambda_4\right)\right)-1\right)\\
&= \frac{-(1+o(1))}{2}\exp\left(-\frac{1}{2}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_4^2\right)\right)\cdot\rho_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4),
\end{aligned} \tag{C.25}$$

using a Taylor expansion of the form $e^x-1=x\left(1+\frac{x}{2!}+\frac{x^2}{3!}+\cdots\right)$ as $x\to 0$

Recall $f_{st}(x,y,u,v)$, $f_{dt}(x)$, $f_{ds}(y)$, $p(u)$ and $p(v)$ be the joint and marginal density functions of (Y_1,Y_2,Y_3,Y_4), Y_1, Y_2, Y_3 and Y_4, respectively.

With the choice of $t > s$ and $s \to \infty$, equation (C.25) implies

$$\begin{aligned}
& f_{st}(x,y,u,v) - f_{dt}(x)f_{ds}(y)p(u)p(v) \\
&= \frac{1}{(2\pi)^4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(\lambda_1 x+\lambda_2 y+\lambda_3 u+\lambda_4 v)} \\
&\times \left(\Psi_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4) - \Psi_t(\lambda_1)\Psi_s(\lambda_2)\Psi_{3t}(\lambda_3)\Psi_{4s}(\lambda_4)\right)\, d\lambda_4\, d\lambda_3\, d\lambda_2\, d\lambda_1 \\
&= C\ (1+o(1))\gamma_{st}(x,y,u,v)\cdot \exp\left(-\frac{1}{2}\left(x^2+y^2+u^2+v^2\right)\right), \qquad \text{(C.26)}
\end{aligned}$$

where $\gamma_{st}(x,y,u,v)$ is a continuous function in (x,y,u,v) with $\gamma_{st}(x,y,u,v) = O\left(\max\left(\frac{s}{\sqrt{t}}, \frac{1}{\sqrt{s}}\right)\right)$ for given (x,y,u,v). This therefore completes the proof of equation (B.1) of Lemma B.1(1).

Meanwhile, the proof of equation (B.2) of Lemma B.1(i) follows similarly and easily from the derivations of equation (C.26). To prove equation (B.3) of Lemma B.1(1), denote $q_t(x,z)$ is the joint density function of $(x_{d,t}, z_t)$, and recall that $f_{dt}(x)$ and $q_t(x|z)$ are the marginal density of $x_{d,t}$ and the conditional density of $x_{d,t}$ given z_t, respectively. Also, recall that $p_t(z|x)$ is the conditional density function of z_t given $x_{d,t}$. The proof of equation (C.26) then implies

$$\begin{aligned}
|p_t(z|x) - p(z)| &= \left|\frac{q_t(x,z) - f_{dt}(x)p(z)}{f_{dt}(x)}\right| = \left|\frac{q_t(x,z) - f_{dt}(x)p(z)}{f_{dt}(x) - \phi(x) + \phi(x)}\right| \\
&\le C_2\ (x^2+z^2)\ e^{-\frac{z^2}{2}}\cdot t^{-\frac{1}{2}}, \qquad \text{(C.27)}
\end{aligned}$$

which completes the proof of (B.3), where $C_2 > 0$ is some constant and $\phi(\cdot)$ denotes the density function of the standard normal random variable, $U \sim N(0,1)$.

(2) Let us prove Lemma B.1 for the case where $\{z_t\}$ is generated by the functional form. Recall the notation of u_t and z_t for the case of $d = 1$. We then introduce the following notation: for $t > s + \tau - 1$

$$\begin{aligned}
Y_{1t} &= \frac{1}{\sqrt{t}}\left(\sum_{i=1}^{s-\tau-1}\phi_{t-i}\varepsilon_i + \sum_{j=s-\tau}^{s-1}\phi_{t-j}\varepsilon_j + \phi_{t-s}\varepsilon_s + \sum_{k=s+1}^{t-\tau-1}\phi_{t-k}\varepsilon_k + \sum_{l=t-\tau}^{t-1}\phi_{t-l}\varepsilon_l + \phi_0\varepsilon_t\right), \\
Y_{2s} &= \frac{1}{\sqrt{s}}\left(\sum_{i=1}^{s-\tau-1}\phi_{s-i}\varepsilon_i + \sum_{j=s-\tau}^{s-1}\phi_{s-j}\varepsilon_j + \phi_0\varepsilon_s\right), \quad U_{1t} = \sum_{i=-\infty}^{0}\phi_{t-i}\varepsilon_i, \\
Y_3 &= \lambda(\varepsilon_{t-1},\cdots,\varepsilon_{t-\tau};\eta_t), \quad Y_4 = \lambda(\varepsilon_{s-1},\cdots,\varepsilon_{s-\tau};\eta_s), \qquad \text{(C.28)}
\end{aligned}$$

which imply $Y_1 = Y_{1t} + \frac{1}{\sqrt{t}}U_{1t}$ and $Y_2 = Y_{2s} + \frac{1}{\sqrt{s}}U_{1s}$.

For given $(\lambda_1,\lambda_2,\lambda_3,\lambda_4)$, we have

$$\begin{aligned}
\lambda_1 Y_1 + \lambda_2 Y_2 + \lambda_3 Y_3 + \lambda_4 Y_4 &= \sum_{i=1}^{s-\tau-1} a_{st}(i)\varepsilon_i + \sum_{j=s-\tau}^{s-1} b_{st}(j)\varepsilon_j + \lambda_4 Y_4 \qquad \text{(C.29)} \\
&+ \sum_{k=s+1}^{t-\tau-1} c_{st}(k)\varepsilon_k + \sum_{l=t-\tau}^{t-1} d_{st}(l)\varepsilon_l + \lambda_3 Y_{3t} + f_{st}\varepsilon_s + g_{st}\varepsilon_t + W_{st}
\end{aligned}$$

where $a_{st}(i) = \frac{\lambda_1}{\sqrt{t}}\phi_{t-i} + \frac{\lambda_2}{\sqrt{s}}\phi_{s-i}$, $b_{st}(j) = \frac{\lambda_1}{\sqrt{t}}\phi_{t-j} + \frac{\lambda_2}{\sqrt{s}}\phi_{s-j}$, $c_{st}(k) = \frac{\lambda_1}{\sqrt{t}}\phi_{t-k}$, $d_{st}(l) = \frac{\lambda_1}{\sqrt{t}}\phi_{t-l}$, $f_{st} = \frac{\lambda_1}{\sqrt{t}}\phi_{t-s} + \frac{\lambda_2}{\sqrt{s}}\phi_0$, $g_{st} = \frac{\lambda_1}{\sqrt{t}}\phi_0$, and $W_{st} = \frac{\lambda_1}{\sqrt{t}}\sum_{i=-\infty}^{0}\phi_{t-i}\varepsilon_i + \frac{\lambda_2}{\sqrt{s}}\sum_{j=-\infty}^{0}\phi_{s-j}\varepsilon_j$.

Similarly to the derivations in (C.21), the characteristic function of (Y_1, Y_2, Y_3, Y_4) is given by

$$\Psi_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4) = E\left[\exp\left(i\left(\lambda_1 Y_1 + \lambda_2 Y_2 + \lambda_3 Y_3 + \lambda_4 Y_4\right)\right)\right]$$

$$= \prod_{i=1}^{s-\tau-1} \varphi(a_{st}(i)) \prod_{k=s+1}^{t-\tau-1} \varphi(c_{st}(k)) \cdot \psi_{1st} \cdot \psi_{2st} \cdot \varphi(f_{st}) \cdot \varphi(g_{st}) \cdot \phi_{st}, \tag{C.30}$$

where $\varphi(\cdot)$ is the characteristic function of ε_1, $\psi_{1st} = E\left[\exp\left(iU_{st}\right)\right]$, $\psi_{2st} = E\left[\exp\left(iV_{st}\right)\right]$, $U_{st} = \sum_{j=s-\tau}^{s-1} b_{st}(j)\varepsilon_j + \lambda_4 Y_4$, $V_{st} = \sum_{l=t-\tau}^{t-1} d_{st}(l)\varepsilon_l + \lambda_3 Y_3$, and $\phi_{st} = E\left[\exp\left(iW_{st}\right)\right]$.

Using the result that $e^x = 1 + x\left(1 + \frac{x}{2!} + \cdots\right)$ as $x \to 0$, we have as $t, s \to \infty$

$$\begin{aligned}
E\left[\exp\left(iU_{st}\right)\right] &= (1+o(1))\left(1 + i\ E[U_{st}] - \frac{1}{2}E\left[U_{st}^2\right]\right) = (1+o(1))\left(1 - \frac{1}{2}E\left[U_{st}^2\right]\right) \\
&= (1+o(1))\left(1 - \frac{1}{2}\left(E\left[\sum_{j=s-\tau}^{s-1} b_{st}(j)\varepsilon_j\right]^2 + \lambda_4^2 E[Y_4^2] + 2E\left[\left(\sum_{j=s-\tau}^{s-1} b_{st}(j)\varepsilon_j\right) Y_4\right]\right)\right) \\
&= (1+o(1))\left(1 - \frac{1}{2}\left(\sum_{j=s-\tau}^{s-1} b_{st}^2(j) + \lambda_4^2 E[Y_4^2] + 2\sum_{j=s-\tau}^{s-1} b_{st}(j) E\left[\varepsilon_j Y_4\right]\right)\right), \\
E\left[\exp\left(iV_{st}\right)\right] &= (1+o(1))\left(1 + i\ E[V_{st}] - \frac{1}{2}E\left[V_{st}^2\right]\right) = (1+o(1))\left(1 - \frac{1}{2}E\left[V_{st}^2\right]\right) \\
&= (1+o(1))\left(1 - \frac{1}{2}\left(E\left[\sum_{k=t-\tau}^{t-1} d_{st}(k)\varepsilon_k\right]^2 + \lambda_3^2 E[Y_3^2] + 2E\left[\left(\sum_{k=t-\tau}^{t-1} d_{st}(k)\varepsilon_k\right) Y_3\right]\right)\right) \\
&= (1+o(1))\left(1 - \frac{1}{2}\left(\sum_{k=t-\tau}^{t-1} d_{st}^2(k) + \lambda_3^2 E[Y_3^2] + 2\sum_{k=t-\tau}^{t-1} d_{st}(k) E\left[\varepsilon_k Y_3\right]\right)\right),
\end{aligned} \tag{C.31}$$

where we have used $E\left[\left(\sum_{j=s-\tau}^{s-1} \phi_{s-j}\varepsilon_j\right)^2\right] \leq C < \infty$ and $E\left[\left(\sum_{k=t-\tau}^{t-1} \phi_{t-k}\varepsilon_k\right)^2\right] \leq C < \infty$.

Similarly to equation (C.23), we have as $t > s$ and $s \to \infty$

$$\begin{aligned}
E\left[\exp\left(iW_{st}\right)\right] &= (1+o(1))\left(1 + iE\left[W_{st}\right] - \frac{1}{2}E\left[W_{st}^2\right]\right) = (1+o(1))\left(1 - \frac{1}{2}E\left[W_{st}^2\right]\right) \\
&= (1+o(1))\left(1 - \frac{1}{2}\left(\frac{\lambda_1^2}{t}\sum_{i=-\infty}^{0} \phi_{t-i}^2 + \frac{\lambda_2^2}{s}\sum_{j=-\infty}^{0} \phi_{s-j}^2 + \frac{2\lambda_1\lambda_2}{\sqrt{ts}}\sum_{j=-\infty}^{0} \phi_{t-j}\phi_{s-j}\right)\right).
\end{aligned} \tag{C.32}$$

Using the dominated convergence theorem and $\sum_{i=0}^{\infty} \phi_i^2 = 1$, equations (C.30)–(C.32) therefore imply that there is a continuous function $\lambda_{st}(\lambda_1, \cdots, \lambda_4) = O\left(\max\left(\frac{1}{\sqrt{s}}, \sqrt{\frac{s}{t}}\right)\right)$ such that as $t > s$ and $s \to \infty$

$$\begin{aligned}
\log\left(\Psi_{st}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)\right) &= \sum_{i=1}^{s-\tau-1} \log(\varphi(a_{st}(i))) + \log(\varphi(f_{st})) + \log(\varphi(g_{st})) \\
&+ \sum_{j=s+1}^{t-\tau-1} \log(\varphi(c_{st}(j)) + \log(\psi_{1st}) + \log(\psi_{2st}) + \log(\phi_{st}) \\
&= -\frac{1}{2}(s-\tau-1)\ \left(\frac{\lambda_1}{\sqrt{t}} + \frac{\lambda_2}{\sqrt{s}}\right)^2 - \frac{1}{2}\left(\frac{\lambda_1}{\sqrt{t}}\right)^2 (1+o(1)) - \frac{1}{2}\ (t-s-\tau-1)\ \left(\frac{\lambda_1}{\sqrt{t}}\right)^2 (1+o(1)) \\
&- \frac{1}{2}\ \left(\frac{\lambda_1}{\sqrt{t}} + \frac{\lambda_2}{\sqrt{s}}\right)^2 (1+o(1)) - \left(\frac{\lambda_3^2}{2} + \frac{\lambda_4^2}{2}\right)\left(1 + O\left(\frac{1}{\sqrt{t}}\right) + O\left(\frac{1}{\sqrt{s}}\right)\right) \\
&= -\frac{1}{2}\left(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 - \lambda_{st}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)\right)(1+o(1)),
\end{aligned} \tag{C.33}$$

which, similarly to equation (C.25), implies

$$\Psi_{st}\left(\lambda_1, \lambda_2, \lambda_3, \lambda_4\right) - \Psi_1(\lambda_1)\Psi_2(\lambda_2)\Psi_3(\lambda_3)\Psi_4(\lambda_4)$$

$$
\begin{aligned}
&= (1+o(1))\exp\left(-\frac{1}{2}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_4^2\right)\right) \\
&\times\left(\exp\left(-\frac{\gamma_{st}}{2}\left(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\lambda_4+\lambda_2\lambda_3+\lambda_2\lambda_4+\lambda_3\lambda_4\right)\right)-1\right) \\
&= \frac{-(1+o(1))}{2}\exp\left(-\frac{1}{2}\left(\lambda_1^2+\lambda_2^2+\lambda_3^2+\lambda_4^2\right)\right)\cdot\lambda_{st}(\lambda_1,\lambda_2,\lambda_3,\lambda_4), \qquad (C.34)
\end{aligned}
$$

where $\lambda_{st}(\lambda_1,\cdots,\lambda_4)=O\left(\max\left(\frac{1}{\sqrt{s}},\sqrt{\frac{s}{t}}\right)\right)$ for given $(\lambda_1,\cdots,\lambda_4)$, and $\Psi_1(\lambda_1)$, $\Psi_2(\lambda_2)$, $\Psi_3(\lambda_3)$ and $\Psi_4(\lambda_4)$ denote the characteristic functions of Y_1, Y_2, Y_3 and Y_4, respectively.

Using (C.34) and then the derivations of (C.26)–(C.27), the rest of the proof of Lemma B.1(ii) follows similarly. Furthermore, given that the conditions on e_t, $w_t=(z_t',e_t)'$ can be regarded as a new z_t with one more dimension. Thus, Lemma B.1(3-4) hold from the proof of Lemma B.1(1-2).

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.

Chen, J., Gao, J., Li, D. and Lin, Z., 2011. Nonparametric specification testing in nonlinear and nonstationary time series models: theory and practice. Available at http://ssrn.com/abstract=2235356.

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., 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., 2013. Semiparametric estimation in triangular system equations with nonstationarity. Journal of Econometrics 176, 5979.

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.

Kasparis, I. 2010. The Bierens test for certain nonstationary models. Journal of Econometrics 158, 221–230.

Kasparis, I., Andreou, E., Phillips, P. C. B., 2012. Nonparametric predictive regression. Cowles Foundation Discussion Paper No. 1878.

Kasparis, I., Phillips, P. C. B., 2012. Dynamic misspecification in nonparametric regression. Journal of Econometrics 168, 270–284.

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

Marmer, V., 2007. Nonlinearity, nonstationarity, and spurious forecasts. Journal of Econometrics 142, 1-27.

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

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

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.