Testing Cross-sectional Dependence in Nonparametric Panel Data Models *

Liangjun Su, Yonghui Zhang
School of Economics, Singapore Management University, Singapore

September 12, 2010

Abstract

In this paper we propose a nonparametric test for cross-sectional contemporaneous dependence in large dimensional panel data models based on the L_2 distance between the pairwise joint density and the product of the marginals. The test can be applied to either raw observable data or residuals from local polynomial time series regressions for each individual to estimate the joint and marginal probability density functions of the error terms. In either case, we establish the asymptotic normality of our test statistic under the null hypothesis by permitting both the cross section dimension n and the time series dimension T to pass to infinity simultaneously and relying upon the Hoeffding decomposition of a two-fold U-statistic. We also establish the consistency of our test. We conduct a small set of Monte Carlo simulations to evaluate the finite sample performance of our test and compare it with that of Pesaran (2004) and Chen, Gao, and Li (2009).

JEL Classifications: C13, C14, C31, C33

Key Words: cross-sectional dependence; two-fold U-statistic; large dimensional panel; local polynomial regression; nonparametric test.

*Address Correspondence to: Liangjun Su, School of Economics, Singapore Management University, 90 Stamford Road, Singapore, 178903; E-mail: ljsu@smu.edu.sg, Phone: (+65) 6828 0386. The first author would like to acknowledge Peter Robinson for valuable discussions that motivate this paper. He also gratefully acknowledges the financial support from a research grant (#09-C244-SMU-008) from Singapore Management University.

1 Introduction

In recent years, there has been a growing literature on large dimensional panel data models with cross-sectional dependence. Cross-sectional dependence may arise due to spatial or spillover effects, or due to unobservable common factors. Much of the recent research on panel data has focused on how to handle cross-sectional dependence. There are two popular approaches in the literature: one is to assume that the individuals are spatially dependent, which gives rise to spatial econometrics; and the other is to assume that the disturbances have a factor structure, which gives rise to static or dynamic factor models. For a recent and comprehensive overview of panel data factor model, see the excellent monograph by Bai and Ng (2008).

Traditional panel data models typically assume observations are independent across individuals, which leads to immense simplification to the rules of estimation and inference. Nevertheless, if observations are cross-sectionally dependent, parametric or nonparametric estimators based on the assumption of cross-sectional independence may be inconsistent and statistical inference based on these estimators can generally be misleading. It has been well documented that panel unit root and cointegration tests based on the assumption of cross-sectional independence are generally inadequate and tend to lead to significant size distortions in the presence of cross-sectional dependence; see Chang (2002), Bai and Ng (2004, 2010), Bai and Kao (2006), and Pesaran (2007), among others. Therefore, it is important to test for cross-sectional independence before embarking on estimation and statistical inference.

Many diagnostic tests for cross-sectional dependence in parametric panel data model have been suggested. When the individuals are regularly spaced or ranked by certain rules, several statistics have been introduced to test for spatial dependence, among which the Moran-I test statistic is the most popular one. See Anselin (1988, 2001) and Robinson (2008) for more details. However, economic agents are generally not regularly spaced, and there does not exist a "spatial metric" that can measure the degree of spatial dependence across economic agents effectively. In order to test for cross-sectional dependence in a more general case, Breusch and Pagan (1980) develop a Lagrange multiplier (LM) test statistic to check the diagonality of the error covariance matrix in SURE models. Noticing that Breusch and Pagan's LM test is only effective if the number of time periods T is large relative to the number of cross sectional units n, Frees (1995) considers test for cross-sectional correlation in panel data models when n is large relative to T and show that both the Breusch and Pagan's and his test statistic belong to a general family of test statistics. Noticing that Breusch and Pagan's LM test statistic suffers from huge finite sample bias, Pesaran (2004) proposes a new test for cross-sectional dependence (CD) by averaging all pair-wise correlation coefficients of regression residuals. Nevertheless, Pesaran's CD test is not consistent against all global alternatives. In particular, his test has no power in detecting cross-sectional dependence when the mean of factor loadings is zero. Hence,

Ng (2006) employs spacing variance ratio statistics to test cross-sectional correlations, which is more robust and powerful than that of Pesaran (2004). Huang, Kao, and Urga (2008) suggest a copula-based tests for testing cross-sectional dependence of panel data models. Pesaran, Ullah, and Yamagata (2008) improve Pesaran (2004) by considering a bias adjusted LM test in the case of normal errors. Based on the concept of generalized residuals (e.g., Gourieroux et al. (1987)), Hsiao, Pesaran, and Pick (2009) propose a test for cross-sectional dependence in the case of non-linear panel data models. Interestingly, an asymptotic version of their test statistic can be written as the LM test of Breusch and Pagan (1980). Sarafidis, Yamagata, and Robertson (2009) consider tests for cross-sectional dependence in dynamic panel data models.

All the above tests are carried out in the parametric context. They can lead to meaningful interpretations if the parametric models or underlying distributional assumptions are correctly specified, and may yield misleading conclusions otherwise. To avoid the potential misspecification of functional form, Chen, Gao, and Li (2009, CGL hereafter) consider tests for cross-sectional dependence based on nonparametric residuals. Their test is a nonparametric counterpart of Pesaran's (2004) test. So it is constructed by averaging all pair-wise cross-sectional correlations and therefore, like Pesaran's (2004) test, it does not test for "*pair-wise independence*" but "*pair-wise uncorrelation*". It is well known that uncorrelation is generally different from independence in the case of non-Gaussianity or nonlinear dependence (e.g., Granger, Maasoumi, and Racine (2004)). There exist cases where testing for cross-sectional pair-wise independence is more appropriate than testing pair-wise uncorrelation.

Since Hoeffding (1948), there has developed an extensive literature on testing independence or serial independence. See Robinson (1991), Brock et al. (1996), Ahmad and Li (1997), Johnson and McClelland (1998), Pinkse (1998), Hong (1998, 2000), Hong and White (2005), among others. All these tests are based on some measure of deviations from independence. For example, Robinson (1991) and Hong and White (2005) base their tests for serial independence on the Kullback-Leibler information criterion, Ahmad and Li (1997) on an L_2 measure of the distance between the joint density and the product of the marginals, and Pinkse (1998) on the distance between the joint characteristic function and the product of the marginal characteristic functions. In addition, Neumeyer (2009) considers a test for independence between regressors and error term in the context of nonparametric regression. Su and White (2003, 2007, 2008) adopt three different methods to test for conditional independence. Except CGL, none of the above nonparametric tests are developed to test for cross-sectional independence in panel data model.

In this paper, we propose a nonparametric test for contemporary "*pair-wise cross-sectional independence*", which is based on the average of pair-wise L_2 distance between the joint density and the product of pair-wise marginals. Like CGL, we base our test on the residuals from local polynomial regressions. Unlike them, we are interested in the pair-wise independence of the

error terms so that our test statistic is based on the comparison of the joint probability density with the product of pair-wise marginal probability densities. We first consider the case where tests for cross-sectional dependence are conducted on raw data so that there is no parameter estimation error involved and then consider the case with parameter estimation error. For both cases, we establish the asymptotic normal distribution of our test statistic under the null hypothesis of cross-sectional independence when $n \to \infty$ and $T \to \infty$ simultaneously. We also show that the test is consistent against global alternatives.

The rest of the paper is organized as follows. Assuming away parameter estimation error, we introduce our testing statistic in Section 2 and study its asymptotic properties under both the null and the alternative hypotheses in Section 3. In Section 4 we study the asymptotic distribution of our test statistic when tests are conducted on residuals from heterogeneous nonparametric regressions. In Section 5 we provide a small set of Monte Carlo simulation results to evaluate the finite sample performance of our test. Section 6 concludes. All proofs are relegated to the appendix.

NOTATION. Throughout the paper we adopt the following notation and conventions. For a matrix A, we denote its transpose as A' and Euclidean norm as $\|A\| \equiv [\mathrm{tr}\,(AA')]^{1/2}$, where $\equiv$ means "is defined as". When A is a symmetric matrix, we use $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ to denote its minimum and maximum eigenvalues, respectively. The operator $\xrightarrow{p}$ denotes convergence in probability, and $\xrightarrow{d}$ convergence in distribution. Let $P_T^l \equiv T!/(T-l)!$ and $C_T^l \equiv T!/\left[(T-l)!l!\right]$ for integers $l \leq T$. We use $(n,T) \to \infty$ to denote the joint convergence of n and T when n and T pass to the infinity simultaneously.

2 Hypotheses and test statistics

To fix ideas and avoid distracting complications, we focus on testing pair-wise cross-sectional dependence in observables in this section and the next. The case of testing pair-wise cross-sectional dependence using unobservable error terms is studied in Section 4.

2.1 The hypotheses

Consider a nonparametric panel data model of the form

$$y_{it} = g_i\left(X_{it}\right) + u_{it},\ i = 1, 2, \ldots, n;\ t = 1, 2, \ldots, T, \tag{2.1}$$

where y_{it} is the dependent variable for individual i at time t, X_{it} is a $d \times 1$ vector of regressors in the ith equation, $g_i(\cdot)$ is an unknown smooth regression function, and u_{it} is a scalar random error term. We are interested in testing for the cross-sectional dependence in $\{u_{it}\}$. Since it seems impossible to design a test that can detect all kinds of cross-sectional dependence among $\{u_{it}\}$, as a starting point we focus on testing pair-wise cross-sectional dependence among them.

For each i, we assume that $\{u_{it}\}_{t=1}^{T}$ is a stationary time series process that has a probability density function (PDF) $f_i(\cdot)$. Let $f_{ij}(\cdot,\cdot)$ denote the joint PDF of u_{it} and u_{jt}. We can formulate the null hypothesis of pair-wise cross-sectional independence among $\{u_{it}, i=1,...,n\}$ as

$$H_0 : f_{ij}(u_{it}, u_{jt}) = f_i(u_{it}) f_j(u_{jt}) \text{ almost surely (a.s.) for all } i,\ j=1,\ldots,n, \text{ and } i \neq j. \tag{2.2}$$

That is, under H_0, u_{it} and u_{jt} are pair-wise independent for all $i \neq j$. The alternative hypothesis is

$$H_1 : \text{ the negation of } H_0. \tag{2.3}$$

2.2 The test statistic

For the moment, we assume that $\{u_{it}\}$ is observed and consider a test for the null hypothesis in (2.2). Alternatively, one can regard g_i's are identically zero in (2.1) and testing for potential cross-sectional dependence among $\{y_{it}\}$. The proposed test is based on the average pairwise L_2 distance between the joint density and the product of the marginal densities:

$$\Gamma_n = \frac{1}{n(n-1)} \sum_{1\le i\neq j\le n} \int\int \left[f_{ij}(u,v) - f_i(u) f_j(v)\right]^2 dudv, \tag{2.4}$$

where $\sum_{1\le i\neq j\le n}$ stands for $\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}$. Obviously, $\Gamma_n = 0$ under H_0 and is nonzero otherwise.

Since the densities are unknown to us, we propose to estimate them by the kernel method. That is, we estimate $f_i(u)$ and $f_{ij}(u,v)$ by

$$\begin{aligned} \widehat{f}_i(u) &\equiv T^{-1}\sum\nolimits_{t=1}^{T} h^{-1}k\left((u_{it}-u)/h\right), \text{ and} \\ \widehat{f}_{ij}(u,v) &\equiv T^{-1}\sum\nolimits_{t=1}^{T} h^{-2}k\left((u_{it}-u)/h\right)k\left((u_{jt}-v)/h\right), \end{aligned}$$

where h is a bandwidth sequence and $k(\cdot)$ is a symmetric kernel function. Note that we use the same bandwidth and (univariate or product of univariate) kernel functions in estimating both the marginal and joint densities, which can facilitate the asymptotic analysis to a great deal. Then a natural test statistic is given by

$$\widehat{\Gamma}_{1nT} = \frac{1}{n(n-1)} \sum_{1\le i\neq j\le n} \int\int \left[\widehat{f}_{ij}(u,v) - \widehat{f}_i(u)\widehat{f}_j(v)\right]^2 dudv. \tag{2.5}$$

Let $\overline{k}_{h,ts}^{i} \equiv h^{-1}\overline{k}\left((u_{it}-u_{is})/h\right)$, where $\overline{k}(\cdot) \equiv \int k(u)k(\cdot - u)du$ is the two-fold convolution of $k(\cdot)$. It is easy to verify that we can rewrite $\widehat{\Gamma}_{1nT}$ as follows:

$$\widehat{\Gamma}_{1nT} = \frac{1}{n(n-1)} \sum_{1\le i\neq j\le n} \left\{ \frac{1}{T^4} \sum_{1\le t,s,r,q\le T} \overline{k}_{h,ts}^{i}\left(\overline{k}_{h,ts}^{j} + \overline{k}_{h,rq}^{j} - 2\overline{k}_{h,tr}^{j}\right)\right\}, \tag{2.6}$$

where $\sum_{1\leq t,s,r,q\leq T} \equiv \sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}$.

The above statistic is simple to compute and offers a natural way to test H_0. Nevertheless, we propose a bias-adjusted test statistic, namely

$$\widehat{\Gamma}_{nT} = \frac{1}{n(n-1)} \sum_{1\leq i\neq j\leq n} \left\{ \frac{1}{P_T^4} \sum_{1\leq t\neq s\neq r\neq q\leq T} \overline{k}_{h,ts}^{i} \left(\overline{k}_{h,ts}^{j} + \overline{k}_{h,rq}^{j} - 2\overline{k}_{h,tr}^{j} \right) \right\}, \tag{2.7}$$

where $P_T^4 \equiv T!/[(T-4)!]$ and $\sum_{1\leq t\neq s\neq r\neq q\leq T}$ denotes the sum over all different arrangements of the distinct time indices t, s, r, and q. In effect, $\widehat{\Gamma}_{nT}$ removes the the "diagonal" (e.g. $t = s, r = q, t = r$) elements from $\widehat{\Gamma}_{1nT}$, thus reducing the bias of the statistic in finite samples. A similar idea has been used in Lavergne and Vuong (2000), Su and White (2007), and Su and Ullah (2009), to name just a few. We will show that, after being appropriately centered and scaled, $\widehat{\Gamma}_{nT}$ is asymptotically normally distributed under the null hypothesis of cross-sectional independence and some mild conditions.

3 Asymptotic distributions of the test statistic

In this section we first present a set of assumptions that are used in deriving the asymptotic null distribution of our test statistic. Then we study the asymptotic distribution of our test statistic under the null hypothesis and establish its consistency.

3.1 Assumptions

To study the asymptotic null distribution of the test statistic with observable "errors" $\{u_{it}\}$, we make the following assumptions.

Assumption A.1 (i) For each i, $\{u_{it}, t = 1, 2, ...\}$ is stationary and α-mixing with mixing coefficient $\{\alpha_i(\cdot)\}$ satisfying $\alpha_i(l) = O(\rho_i^l)$ for some $0 \leq \rho_i < 1$. Let $\overline{\rho} \equiv \max_{1\leq i\leq n} \rho_i$. We further require that $0 \leq \overline{\rho} < 1$.

(ii) For each i and $1 \leq l \leq 8$, the probability density function (PDF) $f_{i,t_1,...,t_l}$ of $(u_{it_1}, ..., u_{it_l})$ is bounded and satisfies a Lipschitz condition: $|f_{i,t_1,...,t_l}(u_1+v_1, \ldots, u_l+v_l) - f_{i,t_1,...,t_l}(u_1, \ldots, u_l)| \leq D_{i,t_1,...,t_l}(\mathbf{u})||\mathbf{v}||$, where $\mathbf{u} \equiv (u_1, ..., u_l)$, $\mathbf{v} \equiv (v_1, ..., v_l)$, and $D_{i,t_1,...,t_l}$ is integrable and satisfies the conditions that $\int_{\mathbb{R}^l} D_{i,t_1,...,t_l}(\mathbf{u})\, ||\mathbf{u}||^{2(1+\delta)} d\mathbf{u} < C_1$ and $\int_{\mathbb{R}^l} D_{i,t_1,...,t_l}(\mathbf{u})\, f_{i,t_1,...,t_l}(\mathbf{u}) d\mathbf{u} < C_1$ for some $C_1 < \infty$ and $\delta \in (0, 1)$. When $l = 1$, we denote the marginal PDF of u_{it} simply as f_i.

Assumption A.2 The kernel function $k : \mathbb{R} \to \mathbb{R}$ is a symmetric, continuous and bounded function such that $k(\cdot)$ is a γth order kernel: $\int k(u)\, du = 1$, $\int u^j k(u)\, du = 0$ for $j = 1, \ldots, \gamma - 1$, and $\int u^\gamma k(u)\, du = \kappa_\gamma < \infty$.

Assumption A.3 As $(n, T) \to \infty$, $h \to 0$, $nT^2h^2 \to \infty$, $nh^{\frac{1-\delta}{1+\delta}}/T \to 0$.

Remark 1. Assumption A.1(i) requires that $\{u_{it},\ t = 1,\ 2, ...\}$ be a stationary strong mixing process with geometric decay rate. This requirement on the mixing rate is handy for our asymptotic analysis but can be relaxed to the usual algebraic decay rate with more complications involved in the proof. It is also assumed in several early works for stationary β-mixing processes such as Fan and Li (1999), Li (1999), and Su and White (2008), and can be satisfied by many well-known processes such as linear stationary autoregressive moving average (ARMA) processes, and bilinear and nonlinear autoregressive processes. Here we only assume that the stochastic process is strong mixing, which is weaker than β-mixing. Assumption A.1(ii) assumes some standard smooth conditions on the PDF of $(u_{it_1}, ..., u_{it_l})$. Assumption A.2 imposes conditions on the kernel function which may or may not be a higher order kernel. The use of a higher order kernel typically aims at reducing the bias of kernel estimates, which is common in the nonparametric literature (see Robinson, 1988; Fan and Li, 1996; Li, 1999, and Su and White, 2008). Assumption A.3 imposes restrictions on the bandwidth, n, and T. These restrictions are weak and can be easily met in practice for a wide combinations of n and T. In addition, it is possible to have $n/T \to c \in [0, \infty]$ as $(n, T) \to \infty$.

By the proof of Theorem 3.1 below, one can relax Assumption A.1(i) to:

Assumption A.1(i*) For each i, $\{u_{it},\ t = 1,\ 2, ...\}$ is stationary and α-mixing with mixing coefficient $\alpha_i(\cdot)$. Let $\alpha(s) \equiv \max_{1 \le i \le n} \alpha_i(s)$. $\sum_{\tau=1}^{\infty} \alpha^{\frac{\delta}{1+\delta}}(\tau) \le C_2$ for some $C_2 < \infty$ and $\delta \in (0, 1)$. There exists $m \equiv m(n, T)$ such that

$$\max\left(n^{-1}T^4 h^{\frac{4}{1+\delta}}, T^4 h^{\frac{2(2+\delta)}{1+\delta}}, T^2 h^{\frac{2}{1+\delta}}\right) \alpha^{\frac{\delta}{1+\delta}}(m) \to 0 \tag{3.1}$$

and $\max\left(m^4 h^4, m^3 h^2\right) \to 0$ as $(n, T) \to \infty$.

For the result in Corollary 3.2 to hold, we further need m and $\alpha(\cdot)$ to meet the following condition.

Assumption A.1(i)** For the m and $\alpha(\cdot)$ defined in Assumption A.1(i*), they satisfy that $h^{\frac{2(1-\delta)}{1+\delta}} T^4 \alpha^{\frac{\delta}{1+\delta}}(m) + h^2 m^4 \to 0$ as $(n, T) \to \infty$.

Clearly, under Assumption A.1(i), we can take $m = \lfloor L \log T \rfloor$ (the integer part of $L \log T$) for a large positive constant L such that both Assumptions A.1(i*) and A.1(i**) are satisfied. For notational simplicity, we continue to apply Assumption A.1(i).

3.2 Asymptotic null distributions

To state our main results, we further introduce some notation. Let E_t denote expectation with respect to variables with time indexed by t only. For example, $E_t[\overline{k}^i_{h,ts}] \equiv \int \overline{k}^i_{h,ts} f_i(u_{is})\, du_{it}$, and $E_t E_s[\overline{k}^i_{h,ts}] \equiv \int \left[\int \overline{k}^i_{h,ts} f_i(u_{it})\, du_{is}\right] f_i(u_{it})\, du_{it}$. Let $\varphi_{i,ts} \equiv \overline{k}^i_{h,ts} - E_t[\overline{k}^i_{h,ts}] - E_s[\overline{k}^i_{h,ts}] +$

$E_t E_s[\overline{k}^i_{h,ts}]$. Define[1]

$$B_{nT} \equiv \frac{1}{n-1} \sum_{1\leq i\neq j\leq n} \frac{h}{T-1} \sum_{1\leq t\neq s\leq T} E\left[\varphi_{i,ts}\right] E\left[\varphi_{j,ts}\right], \text{ and} \tag{3.2}$$

$$\sigma^2_{nT} \equiv \frac{4h^2}{n(n-1)} \sum_{1\leq i\neq j\leq n} \frac{1}{T(T-1)} \sum_{1\leq t\neq s\leq T} \mathrm{Var}\left(\overline{k}^i_{h,ts}\right) \mathrm{Var}\left(\overline{k}^j_{h,ts}\right). \tag{3.3}$$

We establish the asymptotic null distribution of the $\widehat{\Gamma}_{nT}$ test statistic in the following theorem.

Theorem 3.1 *Suppose Assumptions A.1-A.3 hold. Then under the null of cross-sectional independence we have*

$$nTh\widehat{\Gamma}_{nT} - B_{nT} \xrightarrow{d} N\left(0, \sigma^2_0\right) \text{ as } (n,T) \to \infty,$$

where $\sigma^2_0 \equiv \lim_{(n,T)\to\infty} \sigma^2_{nT}$.

Remark 2. The proof of Theorem 3.1 is tedious and is relegated to Appendix A. The idea underlying the proof is simple but the details are quite involved. To see how complications arise, let $\gamma_{nT,ij} \equiv \gamma_{nT}(\mathbf{u}_i, \mathbf{u}_j) \equiv \frac{1}{P^4_T} \sum_{1\leq t\neq s\neq r\neq q\leq T} \overline{k}^i_{h,ts}(\overline{k}^j_{h,ts} + \overline{k}^j_{h,rq} - 2\overline{k}^j_{h,tr})$ where $\mathbf{u}_i \equiv (u_{i1}, ..., u_{iT})'$. Then we have $\widehat{\Gamma}_{nT} = \frac{1}{n(n-1)} \sum_{1\leq i\neq j\leq n} \gamma_{nT}(\mathbf{u}_i, \mathbf{u}_j)$. Clearly, for each pair (i,j) with $i \neq j$, $\gamma_{nT,ij}$ is a fourth order U-statistic along the time dimension, and by treating γ_{nT} as a kernel function, $\widehat{\Gamma}_{nT}$ can be regarded as a second order U-statistic along the individual dimension. To the best of our knowledge, there is no literature that treats such a *two-fold* U-statistic, and it is not clear in the first sight how one should pursue in order to yield a useful central limit theorem (CLT) for $\widehat{\Gamma}_{nT}$. Even though it seems apparent for us to apply the idea of Hoeffding decomposition, how to pursue it is still challenging.

In this paper, we first apply the Hoeffding decomposition on $\gamma_{nT,ij}$ for each pair (i,j) and demonstrate that $\gamma_{nT,ij}$ can be decomposed as follows

$$\gamma_{nT,ij} = 6G^{(2)}_{nT,ij} + 4G^{(3)}_{nT,ij} + G^{(4)}_{nT,ij}$$

where, for $l = 2, 3, 4$, $G^{(l)}_{nT,ij} \equiv \frac{1}{P^l_T} \sum_{1\leq t_1\neq ...\neq t_l\leq T} \vartheta^{(l)}_{ij}(Z_{ij,t_1}, ..., Z_{ij,t_l})$ is an l-th order degenerate U-statistic with kernel $\vartheta^{(l)}_{ij}$ being formerly defined in Appendix A, and $Z_{ij,t} \equiv (u_{it}, u_{jt})$. Then we can obtain the corresponding decomposition for $\widehat{\Gamma}_{nT}$:

$$\widehat{\Gamma}_{nT} = 6G^{(2)}_{nT} + 4G^{(3)}_{nT} + G^{(4)}_{nT}$$

where $G^{(l)}_{nT} \equiv \frac{1}{n(n-1)} \sum_{1\leq i\neq j\leq n} G^{(l)}_{nT,ij}$ for $l = 2, 3, 4$. Even though for each pair (i,j), $G^{(l)}_{nT,ij}$ is an l-th order degenerate U-statistic with kernel $\vartheta^{(l)}_{ij}$ along the time dimension under H_0,

[1] The notation can be greatly simplied under identical distributions across individuals. In this case, $B_{nT} = n(T-1)^{-1} h \sum_{1\leq t\neq s\leq n} [E(\varphi_{1,ts})]^2$, and $\sigma^2_{nT} = 4[T(T-1)]^{-1} h^2 \sum_{1\leq t\neq s\leq n} [\mathrm{Var}(\overline{k}^1_{h,ts})]^2$.

$G_{nT}^{(l)}$ is by no means an l-th order degenerate U-statistic along the individual dimension under H_0. Despite this, we can conjecture as usual that the dominant term in the decomposition of $\widehat{\Gamma}_{nT}$ is given by the first term $6G_{nT}^{(2)}$, and the other two terms $4G_{nT}^{(3)}$ and $G_{nT}^{(4)}$ are asymptotically negligible. So in the second step, we make a decomposition for $6G_{nT}^{(2)} - 6E[G_{nT}^{(2)}]$ and demonstrate that

$$nTh\left\{6G_{nT}^{(2)} - 6E[G_{nT}^{(2)}]\right\} = \sum_{1\leq i<j\leq n} w_{nT}\left(\mathbf{u}_i, \mathbf{u}_j\right) + o_P\left(1\right)$$

where $w_{nT}\left(\mathbf{u}_i, \mathbf{u}_j\right) \equiv \frac{4h}{nT}\sum_{1\leq t<s\leq T}\varphi_{i,ts}^c\varphi_{j,ts}^c$, and $\varphi_{i,ts}^c = \varphi_{i,ts} - E\left[\varphi_{i,ts}\right]$. Despite the fact that $w_{nT,ij} \equiv w_{nT}\left(\mathbf{u}_i, \mathbf{u}_j\right)$ is a non-degenerate second order U-statistic along the time dimension any more, $\sum_{1\leq i<j\leq n} w_{nT}\left(\mathbf{u}_i, \mathbf{u}_j\right)$ is a *degenerate* second order U-statistic along the individual dimension. The latter enables us to apply the de Jong's (1987) CLT for second order degenerate U-statistics with independent but non-identical observations. [Under the null hypothesis of cross-sectional independence $\mathbf{u}_i$'s are independent across i but not identically distributed.] The asymptotic variance of $\sum_{1\leq i<j\leq n} w_{nT}\left(\mathbf{u}_i, \mathbf{u}_j\right)$ is given by σ_0^2 defined in Theorem 3.1 and $6nThE[G_{nT}^{(2)}]$ delivers the asymptotic bias B_{nT} to be corrected from the final test statistic. In the third step, for $l = 3, 4$ we demonstrate $nThG_{nT}^{(l)} = o_P\left(1\right)$ by using the explicit formula of $\vartheta_{ij}^{(l)}$.

Remark 3. The asymptotic distribution in Theorem 3.1 is obtained by letting n and T pass to ∞ simultaneously. Phillips and Moon (1999) introduce three approaches to handle large dimensional panel, namely, sequential limit theory, diagonal path limit theory, and joint limit theory, and discuss relationships between the sequential and joint limit theory. As they remark, the joint limit theory generally requires stronger conditions to establish than the sequential or diagonal path convergence, and by the same token, the results are also stronger and may be expected to be relevant to a wider range of circumstances.

To implement the test, we require consistent estimates of σ_{nT}^2 and B_{nT}. Noting that

$$\begin{aligned}
\sigma_{nT}^2 &= \frac{4h^2}{n\left(n-1\right)T\left(T-1\right)}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t\neq s\leq T} E\left[\left(\overline{k}_{h,ts}^i\right)^2\right] E\left[\left(\overline{k}_{h,ts}^j\right)^2\right] + o\left(1\right) \\
&= \frac{4R\left(\overline{k}\right)^2}{n\left(n-1\right)T\left(T-1\right)}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t\neq s\leq T}\int f_{i,ts}\left(u,u\right)du\int f_{j,ts}\left(v,v\right)dv + o\left(1\right),
\end{aligned}$$

where $R\left(\overline{k}\right) \equiv \int \overline{k}\left(u\right)^2 du$, then we can estimate σ_{nT}^2 by

$$\widehat{\sigma}_{nT}^2 \equiv \frac{4R\left(\overline{k}\right)^2}{n\left(n-1\right)}\sum_{1\leq i\neq j\leq n}\frac{1}{T}\sum_{t=1}^{T}\widehat{f}_{ij,-t}\left(u_{it}, u_{jt}\right)$$

where $\widehat{f}_{ij,-t}\left(u_{it}, u_{jt}\right) \equiv (T-1)^{-1}\sum_{s=1,s\neq t}^{T} h^{-2}\, k\left(\left(u_{is}-u_{it}\right)/h\right) k\left(\left(u_{js}-u_{jt}\right)/h\right)$, i.e., $\widehat{f}_{ij,-t}(u_{it}, u_{jt})$ is the leave-one-out estimate of $f_{ij}(u_{it}, u_{jt})$. One can readily demonstrate $\widehat{\sigma}_{nT}^2$ is a con-

sistent estimate of σ_{nT}^2 under the null. Let

$$\widehat{B}_{nT} \equiv \frac{2}{T-1} \sum_{r=2}^{T} \frac{(T-r+1)\,h}{n-1} \sum_{1 \leq i \neq j \leq n} \widehat{E}\left[\varphi_{i,1r}\right] \widehat{E}\left[\varphi_{j,1r}\right],$$

where $\widehat{E}\left[\varphi_{i,1r}\right] \equiv (T-r+1)^{-1} \sum_{t=1}^{T-r+1} \overline{k}_{h,t,t+r-1}^{i} - T^{-1}(T-1)^{-1} \sum_{1 \leq t \neq s \leq T} \overline{k}_{h,ts}^{i}$. We establish the consistency of $\widehat{B}_{nT}$ for B_{nT} in Appendix B. Then we can define a feasible test statistic:

$$\widehat{I}_{nT} = \frac{nTh\widehat{\Gamma}_{nT} - \widehat{B}_{nT}}{\widehat{\sigma}_{nT}},$$

which is asymptotically distributed as standard normal under the null. We can compare $\widehat{I}_{nT}$ to the one-sided critical value z_α, the upper α percentile from the standard normal distribution, and reject the null if $\widehat{I}_{nT} > z_\alpha$. The following corollary formally establishes the asymptotic normal distribution of $\widehat{I}_{nT}$ under H_0

Corollary 3.2 *Suppose the conditions in Theorem 3.1 hold. Then we have*

$$\widehat{I}_{nT} \xrightarrow{d} N(0,1) \;\; as \;\; (n,T) \to \infty.$$

3.3 Consistency

To study the consistency of our test, we consider the nontrivial case where $\mu_A \equiv \lim_{n\to\infty} \Gamma_n > 0$, where

$$\Gamma_n \equiv \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \int\int \left[f_{ij}(u,v) - f_i(u) f_j(v)\right]^2 dudv.$$

We need to add the following assumption that takes into account cross-sectional dependence under the alternative.

Assumption A.4 For each pair (i,j) with $i \neq j$, the joint PDF f_{ij} of u_{it} and u_{jt} is bounded and satisfies a Lipschitz condition: $|f_{ij}(u_1+v_1, u_2+v_2) - f_{ij}(u_1,u_2)| \leq D_{ij}(u_1,u_2)||(v_1,v_2)||$, and D_{ij} is integrable uniformly in (i,j): $\int\int D_{ij}(u,v)\; f_{ij}(u,v)dudv < C_3$ for some $C_3 < \infty$.

The following theorem establishes the consistency of the test.

Theorem 3.3 *Suppose Assumptions A.1-A.4 hold and $\mu_A > 0$. Then under H_1, $P\left(\widehat{I}_{nT} > d_{nT}\right) \to 1$ for any sequence $d_{nT} = o_P(nTh)$ as $(n,T) \to \infty$.*

Remark 4. Theorem 3.3 indicates that under H_1 our test statistic $\widehat{I}_{nT}$ explodes at the rate nTh provided $\mu_A > 0$. This can occur if $f_{ij}(u,v)$ and $f_i(u) f_j(v)$ differ on a set of positive measure for a "large" number of pairs (i,j) where the number explodes to the infinity at rate n^2. It rules out the case where they differ on a set of positive measure only for a finite fixed number of pairs, or the case where the number of pairwise joint PDFs that differ from the

product of the corresponding marginal PDFs on a set of positive measure is diverging to infinity as $n \to \infty$ but at a slower rate than n^2. In either case, our test statistic $\widehat{I}_{nT}$ cannot explode to the infinity at the rate nTh, but can still be consistent. Specifically, as long as $\lambda_{nT}\Gamma_n \to \mu_A$ and $\lambda_{nT}/(nTh) \to 0$ as $(n,T) \to \infty$ for some diverging sequence $\{\lambda_{nT}\}$, our test is still consistent as $\widehat{I}_{nT}$ now diverges to infinite at rate $(nTh)/\lambda_{nT}$.

Remark 5. We have not studied the asymptotic local power property of our test. Unlike the CGL's test for cross-sectional uncorrelation, it is difficult for us to set up a desirable sequence of Pitman local alternatives that converge to the null at a certain rate and yet enable us to obtain the nontrivial asymptotic power property of our test. Once we deviate from the null hypothesis, all kinds of cross-sectional dependence can arise in the data, which makes the analysis complicated and challenging. See also the remarks in Section 6.

4 Tests based on residuals from nonparametric regressions

In this section, we consider tests for cross-sectional dependence among the unobservable error terms in the nonparametric panel data model (2.1). We must estimate the error terms from the data before conducting the test.

We assume that the regression functions $g_i(\cdot)$, $i = 1, \ldots, n$, are sufficiently smooth, and consider estimating them by the pth order local polynomial method ($p = 1$, 2, 3 in most applications). See Fan and Gijbels (1996) and Li and Racine (2007) for the advantage of local polynomial estimates over the local constant (Nadaraya-Watson) estimates. If $g_i(\cdot)$ has derivatives up to the pth order at a point x, then for any X_{it} in a neighborhood of x, we have

$$\begin{aligned} g_i(X_{it}) &= g_i(x) + \sum_{1 \le |\mathbf{j}| \le p} \frac{1}{\mathbf{j}!} D^{|\mathbf{j}|} g_i(x) (X_{it} - x)^{\mathbf{j}} + o(\|X_{it} - x\|^p) \\ &\equiv \sum_{0 \le |\mathbf{j}| \le p} \beta_{i,\mathbf{j}}(x; b) ((X_{it} - x)/b)^{\mathbf{j}} + o(\|X_{it} - x\|^p). \end{aligned}$$

Here, we use the notation of Masry (1996a, 1996b): $\mathbf{j} = (j_1, ..., j_d)$, $|\mathbf{j}| = \sum_{a=1}^d j_a$, $x^{\mathbf{j}} = \mathbf{\Pi}_{a=1}^d x_a^{j_a}$, $\sum_{0 \le |\mathbf{j}| \le p} = \sum_{l=0}^p \sum_{\substack{j_1=0 \\ j_1+\ldots+j_d=l}}^l \cdots \sum_{j_d=0}^l$, $D^{|\mathbf{j}|} g_i(x) = \frac{\partial^{|\mathbf{j}|} g_i(x)}{\partial^{j_1} x_1 \ldots \partial^{j_d} x_d}$, $\beta_{i,\mathbf{j}}(x; b) = \frac{b^{|\mathbf{j}|}}{\mathbf{j}!} D^{|\mathbf{j}|} g_i(x)$, where $\mathbf{j}! \equiv \mathbf{\Pi}_{a=1}^d j_a!$ and $b \equiv b(n,T)$ is a bandwidth parameter that controls how "close" X_{it} is from x. With observations $\{(y_{it}, X_{it})\}_{t=1}^T$, we consider choosing $\boldsymbol{\beta}_i$, the stack of $\beta_{i,\mathbf{j}}$ in a lexicographical order, to minimize the following criterion function

$$Q_T(x; \boldsymbol{\beta}_i) \equiv T^{-1} \sum_{t=1}^T \left(y_{it} - \sum_{0 \le |\mathbf{j}| \le p} \beta_{\mathbf{j}}((X_{it} - x)/b)^{\mathbf{j}} \right)^2 w_b(X_{it} - x), \tag{4.1}$$

where $w_b(x) = b^{-d} w(x/b)$, and w is a symmetric PDF on $\mathbb{R}^d$. The pth order local polynomial estimate of $g_i(x)$ is then defined as the minimizing concept in the above minimization problem.

Let $N_l \equiv (l+d-1)!/(l!(d-1)!)$ be the number of distinct d-tuples $\mathbf{j}$ with $|\mathbf{j}| = l$. It denotes the number of distinct l-th order partial derivatives of $g_i(x)$ with respect to x. Arrange the N_l d-tuples as a sequence in the lexicographical order (with highest priority to last position), so that $\phi_l(1) \equiv (0, 0, ..., l)$ is the first element in the sequence and $\phi_l(N_l) \equiv (l, 0, ..., 0)$ is the last element, and let ϕ_l^{-1} denote the mapping inverse to ϕ_l. Let $N \equiv \sum_{l=0}^{p} N_l$. Define $\mathbf{S}_{iT}(x)$ and $\mathbf{W}_{iT}(x)$ as a symmetric $N \times N$ matrix and an $N \times 1$ vector, respectively:

$$\mathbf{S}_{iT}(x) \equiv \begin{bmatrix} \mathbf{S}_{iT,0,0}(x) & \mathbf{S}_{iT,0,1}(x) & \cdots & \mathbf{S}_{iT,0,p}(x) \\ \mathbf{S}_{iT,1,0}(x) & \mathbf{S}_{iT,1,1}(x) & \cdots & \mathbf{S}_{iT,1,p}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_{iT,p,0}(x) & \mathbf{S}_{iT,p,1}(x) & \cdots & \mathbf{S}_{iT,p,p}(x) \end{bmatrix}, \quad \mathbf{W}_{iT}(x) \equiv \begin{bmatrix} \mathbf{W}_{iT,0}(x) \\ \mathbf{W}_{iT,1}(x) \\ \vdots \\ \mathbf{W}_{iT,p}(x) \end{bmatrix}$$

where $\mathbf{S}_{iT,j,k}(x)$ is an $N_j \times N_k$ submatrix with the (l, r) element given by

$$[\mathbf{S}_{iT,j,k}(x)]_{l,r} \equiv \frac{1}{T} \sum_{t=1}^{T} \left(\frac{X_{it} - x}{b} \right)^{\phi_j(l)+\phi_k(r)} w_b(X_{it} - x),$$

and $\mathbf{W}_{iT,j}(x)$ is an $N_j \times 1$ subvector whose r-th element is given by

$$[\mathbf{W}_{iT,j}(x)]_r \equiv \frac{1}{T} \sum_{t=1}^{T} y_{it} \left(\frac{X_{it} - x}{b} \right)^{\phi_j(r)} w_b(X_{it} - x).$$

Then we can denote the pth order local polynomial estimate of $g_i(x)$ as

$$\widetilde{g}_i(x) \equiv e_1' [\mathbf{S}_{iT}(x)]^{-1} \mathbf{W}_{iT}(x)$$

where $e_1 \equiv (1, 0, \cdots, 0)'$ is an $N \times 1$ vector.

For each $\mathbf{j}$ with $0 \leq |\mathbf{j}| \leq 2p$, let $\mu_{\mathbf{j}} \equiv \int_{\mathbb{R}^d} x^{\mathbf{j}} w(x) dx$. Define the $N \times N$ dimensional matrix $\mathbb{S}$ by

$$\mathbb{S} \equiv \begin{bmatrix} \mathbb{S}_{0,0} & \mathbb{S}_{0,1} & ... & \mathbb{S}_{0,p} \\ \mathbb{S}_{1,0} & \mathbb{S}_{1,1} & ... & \mathbb{S}_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{S}_{p,0} & \mathbb{S}_{p,1} & ... & \mathbb{S}_{p,p} \end{bmatrix}, \tag{4.2}$$

where $\mathbb{S}_{i,j}$ is an $N_i \times N_j$ dimensional matrix whose (l, r) element is $\mu_{\phi_i(l)+\phi_j(r)}$. Note that the elements of the matrix $\mathbb{S}$ are simply multivariate moments of the kernel w. For example, if $p = 1$, then

$$\mathbb{S} = \begin{bmatrix} \int w(x) dx & \int x' w(x) dx \\ \int x w(x) dx & \int x x' w(x) dx \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}_{1\times d} \\ \mathbf{0}_{d\times 1} & \int x x' w(x) dx \end{bmatrix},$$

where $\mathbf{0}_{a\times c}$ is an $a \times c$ matrix of zeros.

Let $\widetilde{u}_{it} \equiv y_{it} - \widetilde{g}_i(X_{it})$ for $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Define $\widetilde{\Gamma}_{nT}$, $\widetilde{B}_{nT}$, and $\widetilde{\sigma}^2_{nT}$ analogously to $\widehat{\Gamma}_{nT}$, $\widehat{B}_{nT}$, $\widehat{\sigma}^2_{nT}$ but with $\{u_{it}\}$ being replaced by $\{\widetilde{u}_{it}\}$. Then we can consider the following "feasible" test statistic

$$\widetilde{I}_{nT} \equiv \frac{nTh\widetilde{\Gamma}_{nT} - \widetilde{B}_{nT}}{\widetilde{\sigma}_{nT}}.$$

To demonstrate the asymptotic equivalence of $\widetilde{I}_{nT}$ and $\widehat{I}_{nT}$, we add the following assumptions.

Assumption A.5 (i) For each $i = 1, \ldots, n$, $\{X_{it}, t = 1, 2, ...\}$ is stationary and α-mixing with mixing coefficient $\{a_i(\cdot)\}$ satisfying $\sum_{j=1}^{\infty} j^{\kappa_0} a(j)^{\delta_0/(2+\delta_0)} < C_4$ for some $C_4 < \infty$, $\kappa_0 > \delta_0/(2+\delta_0)$, and $\delta_0 > 0$, where $a(j) \equiv \max_{1 \le i \le n} a_i(j)$.

(ii) For each $i = 1, \ldots, n$, the support $\mathcal{X}_i$ of X_{it} is compact on $\mathbb{R}^d$. The PDF p_i of X_{it} exists, is Lipschitz continuous, and is bounded away from zero on $\mathcal{X}_i$ uniformly in i : $\min_{1 \le i \le n} \inf_{x_i \in \mathcal{X}_i} p_i(x_i) > C_5$ for some $C_5 > 0$. The joint PDF of X_{it} and X_{is} is uniformly bounded for all $t \neq s$ by a constant that does not depend on i or $|t - s|$.

(iii) $\{u_{it}, i = 1, 2, \ldots, t = 1, 2, \ldots\}$ is independent of $\{X_{it}, i = 1, 2, \ldots, t = 1, 2, \ldots\}$.

Assumption A.6 (i) For each $i = 1, \ldots, n$, the individual regression function $g_i(\cdot)$, is $p+1$ times continuously partially differentiable.

(ii) The $(p+1)$-th order partial derivatives of g_i are Lipschitz continuous on $\mathcal{X}_i$.

Assumption A.7 (i) The kernel function $w : \mathbb{R}^d \to \mathbb{R}^+$ is a continuous, bounded, and symmetric PDF; $\mathbb{S}$ is positive definite (p.d.).

(ii) Let $\mathbf{w}(x) \equiv \|x\|^{2(2+\delta_0)p} w(x)$. $\mathbf{w}$ is integrable with respect to the Lebesgue measure.

(iii) Let $W_{\mathbf{j}}(x) \equiv x^{\mathbf{j}} w(x)$ for all d-tuples $\mathbf{j}$ with $0 \le |\mathbf{j}| \le 2p+1$. $W_{\mathbf{j}}(x)$ is Lipschitz continuous for $0 \le |\mathbf{j}| \le 2p+1$. For some $C_6 < \infty$ and $C_7 < \infty$, either $w(\cdot)$ is compactly supported such that $w(x) = 0$ for $\|x\| > C_6$, and $||W_{\mathbf{j}}(x) - W_{\mathbf{j}}(\widetilde{x})|| \le C_7 ||x - \widetilde{x}||$ for any x, $\widetilde{x} \in \mathbb{R}^d$ and for all $\mathbf{j}$ with $0 \le |\mathbf{j}| \le 2p+1$; or $w(\cdot)$ is differentiable, $\|\partial W_{\mathbf{j}}(x)/\partial x\| \le C_6$, and for some $\iota_0 > 1$, $|\partial W_{\mathbf{j}}(x)/\partial x| \le C_6 \|x\|^{-\iota_0}$ for all $\|x\| > C_7$ and for all $\mathbf{j}$ with $0 \le |\mathbf{j}| \le 2p+1$.

Assumption A.8 (i) The kernel function k is second order differentiable with first order derivative k' and second order derivative k''. Both $uk(u)$ and $uk'(u)$ tend to 0 as $|u| \to \infty$. (ii) For some $c_k < \infty$ and $A_k < \infty$, $|k''(u)| \le c_k$ and for some $\gamma_0 > 1$, $|k''(u)| \le c_k |u|^{-\gamma_0}$ for all $|u| > A_k$.

Assumption A.9 (i) Let $\eta \equiv T^{-1}b^{-d} + b^{2(p+1)}$. As $(n, T) \to \infty$, $Th^5 \to \infty$, $T^{3/2}b^d h^5 \to \infty$, and $nTh(\eta^2 + h^{-4}\eta^3 + h^{-8}\eta^4) \to 0$.

(ii) For the m defined in Assumption A.1(i*), $\max(nhmb^{2(p+1)}, nmT^{-1}b^{-d}, n^2T^{-4}m^6h^{-2}, n^2m^2h^{-2}b^{4(p+1)}, nhm^2/T, nh^{-3}m^3/T^2, m^3/T) \to 0$.

Remark 6 Assumptions A.5 (i)-(ii) are subsets of some standard conditions to obtain the uniform convergence of local polynomial regression estimates. Like CGL, we assume the

independence of $\{u_{it}\}$ and $\{X_{js}\}$ for all i, j, t, s in Assumptions A.5(iii), which will greatly facilitate our asymptotic analysis. Assumptions A.6 and A.7 are standard in the literature on local polynomial estimation. In particular, following Hansen (2008), the compact support of the kernel function w in Masry (1996b) can be relaxed as in Assumption A.7(iii). Assumption A.8 specifies more conditions on the kernel function k used in the estimation of joint and marginal densities of the error terms. They are needed because we need to apply Taylor expansions on functions associated with k. Assumption A.9 imposes further conditions on h, n, and T and their interaction with the smoothing parameter b and the order p of local polynomial used in the local polynomial estimation. If we relax the geometric α-mixing rate in Assumption A.1(i) to the algebraic rate, then we need to add the following condition on the bandwidth parameters, sample sizes, and the choices of m and p :

Assumption A.1(i*)** For the m, $\alpha(\cdot)$, and δ defined in Assumption A.1(i*), they also satisfy that

$$\max\left\{n^2T^2h^{-3-\frac{\delta}{1+\delta}}, T^2h^{-4-\frac{2\delta}{1+\delta}},\ T^2h^{-5-\frac{2\delta}{1+\delta}}b^{4(p+1)}\right\}\alpha^{\frac{\delta}{1+\delta}}(m) \to 0 \text{ as } (n,T)\to\infty.$$

Theorem 4.1 *Suppose Assumptions A.1-A.3 and A.5-A.9 hold. Then under the null of cross-sectional independence*

$$\widetilde{I}_{nT} \to N(0,1) \ \ as \ \ (n,T)\to\infty.$$

Remark 7. The above theorem establishes the asymptotic equivalence of $\widetilde{I}_{nT}$ and $\widehat{I}_{nT}$. That is, the test statistic $\widetilde{I}_{nT}$ that is based on the estimated residuals from heterogeneous local polynomial regressions is asymptotically equivalent to $\widehat{I}_{nT}$ that is constructed from the generally unobservable errors. If evidence suggests that the nonparametric regression relationships are homogeneous, i.e., $g_i(X_{it}) = g(X_{it})$ a.s. for some function g on $\mathbb{R}^d$ and for all i, then one can pool the cross section data together and estimate the homogeneous regression function g at a faster rate than estimating each individual regression function g_i by using the time series observations for cross section i only. In this case, we expect that the requirement on the relationship of n, T, h, b, and p becomes less stringent. Similarly, if $g_i(X_{it}) = \beta_{0i} + \beta_{1i}'X_{it}$ a.s. for some unknown parameters β_{0i} and β_{1i}, then we can estimate such parametric regression functions at the usual parametric rate $T^{-1/2}$, and it is easy to verify that the result in Theorem 4.1 continue to hold by using the residuals from time series parametric regressions for each individual.

The following theorem establishes the consistency of the test.

Theorem 4.2 *Suppose Assumptions A.1-A.9 hold and $\mu_A > 0$. Then under H_1, $P\left(\widetilde{I}_{nT} > d_{nT}\right) \to 1$ for any sequence $d_{nT} = o_P(nTh)$ as $(n,T)\to\infty$.*

The proof of the above theorem is almost identical to that of Theorem 3.3. The main difference is that one needs to apply Taylor expansions to show that $(nTh)^{-1}\widetilde{I}_{nT}$ is asymptotically equivalent to $(nTh)^{-1}\widehat{I}_{nT}$ under H_1. Remark 4 also holds for the test $\widetilde{I}_{nT}$.

5 Monte Carlo simulations

In this section, we conduct a small set of Monte Carlo simulations to evaluate the finite sample performance of our test and compare it with Pesaran's and CGL's tests for cross-sectional uncorrelation.

5.1 Data generating processes

We consider the following six data generating processes (DGPs) in our Monte Carlo study. DGPs 1-2 are for size study, and DGPs 3-6 are for power comparisons.

DGP 1:

$$y_{it} = \alpha_i + \beta_i X_{it} + u_{it},$$

where across both i and t, $X_{it} \sim$ IID $U(-3,3)$, $\alpha_i \sim$IID $U(0,1)$, $\beta_i \sim$ IID $N(0,1)$, and they are mutually independent of each other.

DGP 2:

$$y_{it} = (1+\theta_i)\exp(X_{it})/(1+\exp(X_{it})) + u_{it},$$

where across both i and t, $X_{it} \sim$ IID $U(-3,3)$, $\theta_i \sim$ IID $N(0,0.25)$, and they are mutually independent of each other.

In DGPs 1-2, we consider two kinds of error terms: (i) $u_{it} \sim$ IID $N(0,1)$ across both i and t and independent of $\{\alpha_i, \beta_i, X_{it}\}$; and (ii) $\{u_{it}\}$ is IID across i and an AR(1) process over t: $u_{it} = 0.5u_{i,t-1} + \varepsilon_{it}$, where $\varepsilon_{it} \sim$ IID $N(0,0.75)$ across both i and t and independent of $\{\alpha_i, \beta_i, X_{it}\}$. Clearly, there is no cross-sectional dependence in either case.

In terms of conditional mean specification, DGPs 3 and 5 are identical to DGP 1, and DGPs 4 and 6 are identical to DGP2. The only difference lies in the specification of the error term u_{it}. In DGPs 3-4, we consider the following single-factor error structure:

$$u_{it} = 0.5\lambda_i F_t + \varepsilon_{it} \tag{5.1}$$

where the factors F_t are IID $N(0,1)$, and the factor loadings λ_i are IID $N(0,1)$ and independent of $\{F_t\}$. We consider two configurations for ε_{it} : (i) ε_{it} are IID $N(0,1)$ and independent of $\{F_t, \lambda_i\}$, and (ii) $\varepsilon_{it} = 0.5\varepsilon_{it-1} + \eta_{it}$ where η_{it} are IID $N(0,0.75)$ across both i and t, and independent of $\{F_t, \lambda_i\}$.

In DGPs 5-6, we consider the following two-factor error structure:

$$u_{it} = 0.3\lambda_{1i}F_{1t} + 0.3\lambda_{2i}F_{2t} + \varepsilon_{it} \tag{5.2}$$

where both factors F_{1t} and F_{2t} are IID $N(0,1)$, λ_{1i} are IID $N(0,1)$, λ_{2i} are IID $N(0.5,1)$, F_{1t}, F_{2t}, λ_{1i}, and λ_{2i} are mutually independent of each other, and the error process $\{\varepsilon_{it}\}$ is specified as in DGPs 3-4 with two configurations.

5.2 Bootstrap

It is well known that the asymptotic normal distribution typically cannot approximate well the finite sample distribution of many nonparametric test statistics under the null hypothesis. In fact, the empirical level of these tests can be sensitive to the choice of bandwidths or highly distorted in finite samples. So we suggest using a bootstrap method to obtain the bootstrap p-values. Note that we need to estimate $E(\varphi_{ts})$ in B_{nT}, and that the dependence structure in each individual error process $\{u_{it}\}_{t=1}^{T}$ will affect the asymptotic distribution of our test under the null. Like Hsiao and Li (2001), we need to mimic the dependence structure over time. So we propose to apply the stationary bootstrap procedure of Politis and Romano (1994) to each individual i's residual series $\{\widetilde{u}_{it}\}_{t=1}^{T}$. The procedure goes as follows:

1. Obtain the local polynomial regression residuals $\widetilde{u}_{it} = Y_{it} - \widetilde{g}_i(\mathbf{x}_{it})$ for each i and t.
2. For each i, obtain the bootstrap time series sequence $\{u_{it}^*\}_{t=1}^{T}$ by the method of stationary bootstrap. [2]
3. Calculate the bootstrap test statistic $\widetilde{I}_{nT}^* = (nTh\widetilde{\Gamma}_{nT}^* - \widetilde{B}_{nT}^*)/\widetilde{\sigma}_{nT}^*$, where $\widetilde{\Gamma}_{nT}^*$, $\widetilde{B}_{nT}^*$ and $\widetilde{\sigma}_{nT}^*$ are defined analogously to $\widetilde{\Gamma}_{nT}$, $\widetilde{B}_{nT}$ and $\widetilde{\sigma}_{nT}$ but with $\widetilde{u}_{it}$ be replaced by u_{it}^*.
4. Repeat steps 1-3 for B times and index the bootstrap statistics as $\{\widetilde{I}_{nT,j}^*\}_{j=1}^{B}$. Calculate the bootstrap p-value $p^* \equiv B^{-1}\sum_{j=1}^{B}\mathbf{1}(\widetilde{I}_{nT,j}^* > \widetilde{I}_{nT})$ where $\mathbf{1}(\cdot)$ is the usual indicator function, and reject the null hypothesis of cross-sectional independence if p^* is smaller than the prescribed level of significance.

Note that we have imposed the null restriction of cross-sectional independence implicitly because we generate $\{u_{it}^*\}$ independently across all individuals. We conjecture that for sufficiently large B, the empirical distribution of $\{\widetilde{I}_{nT,j}^*\}_{j=1}^{B}$ is able to approximate the finite sample distribution of $\widetilde{I}_{nT}$ under the null hypothesis, but are not sure whether this can have any improvement over the asymptotic normal approximation. The theoretical justification for the validity of our bootstrap procedure goes beyond the scope of this paper.

5.3 Test results

We consider three tests of cross-sectional dependence in this section: Pesaran's CD test for cross-sectional dependence, CGL test for cross-sectional uncorrelation, and the $\widetilde{I}_{nT}$ test pro-

[2] A simple description of the resampling algorithm goes as follows. Let p be a fixed number in $(0,1)$. Let u_{i1}^* be picked at random from the original T residuals $\{\widetilde{u}_{i1}, ..., \widetilde{u}_{iT}\}$, so that $u_{i1}^* = \widetilde{u}_{iT_1}$, say, for some $T_1 \in \{1, ..., T\}$. With probability p, let u_{i2}^* be picked at random from the original T residuals $\{\widetilde{u}_{i1}, ..., \widetilde{u}_{iT}\}$; with probability $1-p$, let $u_{i2}^* = \widetilde{u}_{i,T_1+1}$ so that u_{i2}^* would be the "next" observation in the original residual series following $\widetilde{u}_{iT_1}$. In general, given that u_{it}^* is determined by the Jth observation $\widetilde{u}_{iJ}$ in the original residual series, let $u_{i,t+1}^*$ be equal to $\widetilde{u}_{i,J+1}$ with probability $1-p$ and be picked at random from the original T residuals with probability p. We set $p = T^{-1/3}$ in the simulations.

Table 1: Finite sample rejection frequency for DGPs 1-2 (size study, nomial level 0.05)

DGP	n	T	(i) $u_{it} \sim$ IID $N(0,1)$			(ii) $u_{it} = 0.5u_{i,t-1} + \varepsilon_{it}$		
			P	CGL	SZ	P	CGL	SZ
1	25	25	0.040	0.044	0.054	0.092	0.060	0.082
		50	0.060	0.044	0.048	0.130	0.062	0.082
		100	0.056	0.058	0.064	0.126	0.080	0.066
	50	25	0.060	0.044	0.062	0.118	0.066	0.128
		50	0.070	0.052	0.080	0.112	0.076	0.074
		100	0.034	0.030	0.048	0.124	0.066	0.064
2	25	25	0.038	0.044	0.052	0.088	0.050	0.090
		50	0.056	0.062	0.060	0.122	0.062	0.082
		100	0.058	0.044	0.064	0.128	0.068	0.070
	50	25	0.054	0.042	0.058	0.076	0.078	0.120
		50	0.064	0.060	0.060	0.110	0.050	0.084
		100	0.038	0.052	0.052	0.108	0.068	0.060

Note: P, CGL, and SZ refer to Pesaran's, CGL's and our tests, respectively.

posed in this paper. To conduct our test, we need to choose kernels and bandwidths. To estimate the heterogeneous regression functions, we conduct a third-order local polynomial regression ($p = 3$) by choosing the second order Gaussian kernel and rule-of-thumb bandwidth: $b = s_X T^{-1/9}$ where s_X denotes the sample standard deviation of $\{X_{it}\}$ across i and t. To estimate the marginal and pairwise joint densities, we choose the second order Gaussian kernel and rule-of-thumb bandwidth $h = s_{\widetilde{u}} T^{-1/6}$, where $s_{\widetilde{u}}$ denotes the sample standard deviation of $\{\widetilde{u}_{it}\}$ across i and t. For the CGL test, we follow their paper and consider a local linear regression to estimate the conditional mean function by using the Gaussian kernel and choosing the bandwidth through the leave-one-out cross-validation method. For the Pesaran's test, we estimate the heterogeneous regression functions by using the linear model, and conduct his CD test based on the parametric residuals.

For all tests, we consider $n = 25$, 50, and $T = 25$, 50, 100. For each combination of n and T, we use 500 replications for the level and power study, and 200 bootstrap resamples in each replication.

Table 1 reports the finite sample level for Pesaran's CD test, the CGL test and our test (denoted as P, CGL, and SZ, respectively in the table). When the error terms u_{it} are IID across t, all three tests perform reasonably well for all combinations of n and T and both DGPs under investigation in that the empirical levels are close to the nominal level. When $\{u_{it}\}$ follows an AR(1) process along the time dimension, we find out the CGL test outperforms the Pesaran's test in terms of level performance: the latter test tends to have a large size

distortion which does not improve when either n or T increases. In contrast, our test can be oversized when n/T is not small (e.g., $n = 50$ and $T = 25$) so that the parameter estimation error plays a non-negligible role in the finite samples, but the level of our test improves quickly as T increases for fixed n.

Table 2 reports the finite sample power performance of all three tests for DGPs 3-6. For DGPs 3-4, we have a single-factor error structure. Noting that the factor loadings λ_i have zero mean in our setup, neither Pesaran's nor CGL's test has power in detecting cross-sectional dependence in this case. This is confirmed by our simulations. In contrast, our tests have power in detecting deviations from cross-sectional dependence. As either n or T increases, the power of our test increases. DGPs 5-6 exhibit a two-factor error structure where one of the two sequences of factor loadings have nonzero mean, and all three tests have power in detecting cross-sectional dependence. As either n or T increases, the powers of all three tests increase quickly and our test tends to more powerful than the Pesaran's and CGL's tests.

6 Concluding remarks

In this paper, we propose a nonparametric test for cross-sectional dependence in large dimensional panel. Our tests can be applied to both raw data and residuals from heterogenous nonparametric (or parametric) regressions. The requirement on the relative magnitude of n and T is quite weak in the former case, and very strong in the latter case in order to control the asymptotic effect of the parameter estimation error on the test statistic. In both cases, we establish the asymptotic normality of our test statistic under the null hypothesis of cross-sectional independence. The global consistency of our test is also established. Monte Carlo simulations indicate our test performs reasonably well in finite samples and has power in detecting cross-sectional dependence when the Pesaran's and CGL's tests fail.

We have not pursued the asymptotic local power analysis for our nonparametric test in this paper. It is well known that the study of asymptotic local power is rather difficult in nonparametric testing for serial dependence, see Tjøstheim (1996) and Hong and White (2005). Similar remark holds true for nonparametric testing for cross-sectional dependence. To analyze the local power of their test, Hong and White (2005) consider a class of locally j-dependent processes for which there exists serial dependence at lag j *only*, but j may grow to infinity as the sample size passes to infinity. It is not clear whether one can extend their analysis to our framework since there is no natural ordering along the individual dimensions in panel data models. In addition, it may not be advisable to consider a class of panel data models for which there exists cross-sectional dependence at pairwise level only: if any two of u_{it}, u_{jt}, and u_{kt} $(i \neq j \neq k)$ are dependent, they tend to be dependent on the other one also. Thus we conjecture that it is very challenging to conduct the asymptotic local power analysis for our nonparametric test.

Table 2: Finite sample rejection frequency for DGPs 3-6 (power study, nomial level 0.05)

DGP	n	T	(i) $\varepsilon_{it} \sim$ IID $N(0,1)$			(ii) $\varepsilon_{it} = 0.5\varepsilon_{it-1} + \eta_{it}$		
			P	CGL	SZ	P	CGL	SZ
3	25	25	0.040	0.046	0.446	0.092	0.052	0.590
		50	0.060	0.058	0.778	0.130	0.060	0.860
		100	0.056	0.074	0.950	0.126	0.038	0.984
	50	25	0.060	0.040	0.772	0.118	0.070	0.866
		50	0.070	0.060	0.972	0.112	0.074	0.992
		100	0.034	0.064	0.998	0.124	0.068	1.000
4	25	25	0.038	0.074	0.446	0.098	0.044	0.616
		50	0.056	0.052	0.772	0.206	0.066	0.858
		100	0.058	0.062	0.954	0.234	0.044	0.984
	50	25	0.054	0.046	0.772	0.148	0.086	0.870
		50	0.064	0.068	0.970	0.190	0.072	0.990
		100	0.038	0.062	0.998	0.270	0.068	1.000
5	25	25	0.326	0.248	0.208	0.410	0.304	0.418
		50	0.412	0.332	0.444	0.486	0.350	0.672
		100	0.584	0.446	0.740	0.594	0.424	0.910
	50	25	0.550	0.442	0.456	0.626	0.508	0.680
		50	0.720	0.620	0.812	0.754	0.640	0.918
		100	0.842	0.742	0.988	0.888	0.776	0.996
6	25	25	0.304	0.232	0.250	0.420	0.292	0.406
		50	0.428	0.330	0.424	0.488	0.348	0.634
		100	0.568	0.426	0.762	0.588	0.402	0.908
	50	25	0.548	0.454	0.424	0.624	0.516	0.662
		50	0.724	0.636	0.814	0.760	0.636	0.908
		100	0.838	0.746	0.980	0.888	0.794	1.000

Note: P, CGL, and SZ refer to Pesaran's, CGL's and our tests, respectively.

APPENDIX

Throughout this appendix, we use C to signify a generic constant whose exact value may vary from case to case. Recall $P_T^l \equiv T!/(T-l)!$ and $C_T^l \equiv T!/\left[(T-l)!l!\right]$ for integers $l \leq T$.

A Proof of Theorem 3.1

Recall $\varphi_{i,ts} \equiv \overline{k}_{h,ts}^i - E_t[\overline{k}_{h,ts}^i] - E_s[\overline{k}_{h,ts}^i] + E_tE_s[\overline{k}_{h,ts}^i]$ where $\overline{k}_{h,ts}^i \equiv \overline{k}_h\left(u_{it}-u_{is}\right)$ and E_s denotes expectation taken only with respect to variables indexed by time s, that is, $E_s(\overline{k}_{h,ts}^i) \equiv \int \overline{k}_h\left(u_{it}-u\right) f_i\left(u\right) du$. Let $c_{i,ts} \equiv E(\varphi_{i,ts})$, and $c_{ts} \equiv (n-1)^{-1}\sum_{i=1}^n c_{i,ts}$. We will frequently use the fact that for $t \neq s$,

$$c_{i,ts} \leq Ch^{-\frac{\delta}{1+\delta}} \alpha_i^{\frac{\delta}{1+\delta}}\left(|t-s|\right) \tag{A.1}$$

as by the law of iterated expectations, the triangle inequality, and Lemma E.2, we have $|c_{i,ts}| = |E[\overline{k}_{h,ts}^i] - E_tE_s[\overline{k}_{h,ts}^i]| = |E\{E[\overline{k}_{h,ts}^i|u_{it}] - E_s[\overline{k}_{h,ts}^i]\}| \leq E|E[\overline{k}_{h,ts}^i|u_{it}] - E_s[\overline{k}_{h,ts}^i]| \leq Ch^{-\frac{\delta}{1+\delta}} \alpha_i^{\frac{\delta}{1+\delta}}\left(|t-s|\right)$. Let $\alpha\left(j\right) \equiv \max_{1\leq i\leq n} \alpha_i\left(j\right)$. Let $m \equiv \lfloor L\log T\rfloor$ (the integer part of $L\log T$) where L is a large positive constant so that the conditions on m in Assumption A.1(i*) are all met by Assumption A.1(i). In addition, it is obvious that $\sum_{\tau=1}^{\infty} \alpha^{\frac{\delta}{1+\delta}}\left(\tau\right) = O\left(1\right)$ under Assumption A.1(i).

Let $Z_{ij,t} \equiv (u_{it}, u_{jt})$ and $\varsigma_{ij,tsrq} \equiv \varsigma\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}, Z_{ij,q}\right) = \overline{k}_{h,ts}^i(\overline{k}_{h,ts}^j + \overline{k}_{h,rq}^j - 2\overline{k}_{h,tr}^j)$. Let $\overline{\varsigma}_{ij,tsrq} \equiv \overline{\varsigma}\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}, Z_{ij,q}\right) \equiv \frac{1}{4!}\sum_{4!}\varsigma_{ij,tsrq}$, where $\sum_{4!}$ denotes summation over all 4! different permutations of (t,s,r,q). That is, $\overline{\varsigma}_{ij,tsrq}$ is a symmetric version of $\varsigma_{ij,tsrq}$ by symmetrizing over the four time indices and it is easy to verify that

$$\begin{aligned}
\bar{\varsigma}_{ij,tsrq} \quad = \quad & \frac{1}{12}\{\overline{k}_{h,ts}^i(2\overline{k}_{h,ts}^j + 2\overline{k}_{h,rq}^j - \overline{k}_{h,tr}^j - \overline{k}_{h,sr}^j - \overline{k}_{h,tq}^j - \overline{k}_{h,sq}^j) \\
& +\overline{k}_{h,tr}^i(2\overline{k}_{h,tr}^j + 2\overline{k}_{h,qs}^j - \overline{k}_{h,ts}^j - \overline{k}_{h,sr}^j - \overline{k}_{h,tq}^j - \overline{k}_{h,rq}^j) \\
& +\overline{k}_{h,tq}^i(2\overline{k}_{h,tq}^j + 2\overline{k}_{h,sr}^j - \overline{k}_{h,tr}^j - \overline{k}_{h,qr}^j - \overline{k}_{h,ts}^j - \overline{k}_{h,sq}^j) \\
& +\overline{k}_{h,sr}^i(2\overline{k}_{h,sr}^j + 2\overline{k}_{h,qt}^j - \overline{k}_{h,st}^j - \overline{k}_{h,rt}^j - \overline{k}_{h,sq}^j - \overline{k}_{h,rq}^j) \\
& +\overline{k}_{h,sq}^i(2\overline{k}_{h,sq}^j + 2\overline{k}_{h,rt}^j - \overline{k}_{h,st}^j - \overline{k}_{h,qt}^j - \overline{k}_{h,sr}^j - \overline{k}_{h,qr}^j) \\
& +\overline{k}_{h,rq}^i(2\overline{k}_{h,rq}^j + 2\overline{k}_{h,st}^j - \overline{k}_{h,rt}^j - \overline{k}_{h,qt}^j - \overline{k}_{h,rs}^j - \overline{k}_{h,qs}^j)\}.
\end{aligned} \tag{A.2}$$

Then we can write $\widehat{\Gamma}_{nT}$ as

$$\begin{aligned}
\widehat{\Gamma}_{nT} \quad &= \quad \frac{1}{n\left(n-1\right)} \sum_{1\leq i\neq j\leq n} \frac{1}{P_T^4} \sum_{1\leq t\neq s\neq r\neq q\leq T} \varsigma_{ij,tsrq} \\
&= \quad \frac{1}{n\left(n-1\right)} \sum_{1\leq i\neq j\leq n} \frac{1}{C_T^4} \sum_{1\leq t_1<t_2<t_3<t_4\leq T} \overline{\varsigma}_{ij,t_1t_2t_3t_4}.
\end{aligned} \tag{A.3}$$

Let $\theta_{ij} = E_1E_2E_3E_4\left[\overline{\varsigma}\left(Z_{ij,1}, Z_{ij,2}, Z_{ij,3}, Z_{ij,4}\right)\right]$ and $\bar{\varsigma}_{ij,c}\left(z_1,\ldots,z_c\right) = E_{c+1}\cdots E_4[\overline{\varsigma}(z_1,\ldots,z_c, Z_{ij,c+1},\ldots,Z_{ij,4})]$ for nonrandom $z_1,\ldots,z_c$ and $c = 1,2,3,4$. Let $\vartheta_{ij}^{(1)}\left(z_1\right) = \bar{\varsigma}_{ij,1}\left(z_1\right) - \theta_{ij}$ and $\vartheta_{ij}^{(c)}\left(z_1,\ldots,z_c\right) = \bar{\varsigma}_{ij,c}\left(z_1,\ldots,z_c\right) - \sum_{k=1}^{c-1}\sum_{(c,k)}\vartheta_{ij}^{(k)}(z_{t_1},\ldots,z_{t_k}) - \theta_{ij}$ for $c = 2,3,4$, where the sum $\sum_{(c,k)}$ is taken over all subsets $1 \leq t_1 < \cdots < t_k \leq c$ of $\{1,2,\ldots,c\}$. It is easy to verify that $\theta_{ij} = 0$, $\vartheta_{ij}^{(1)}\left(Z_{ij,t}\right) = 0$, and

$$\vartheta_{ij}^{(2)}\left(Z_{ij,t}, Z_{ij,s}\right) = \bar{\varsigma}_{ij,2}\left(Z_{ij,t}, Z_{ij,s}\right) = \frac{1}{6}\varphi_{i,ts}\varphi_{j,ts}. \tag{A.4}$$

Similarly, straightforward but tedious calculations show that

$$\begin{aligned}
& \vartheta_{ij}^{(3)}\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}\right) \\
= \; & \bar{\varsigma}_{ij,3}\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,t}, Z_{ij,s}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,t}, Z_{ij,r}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,s}, Z_{ij,r}\right) \\
= \; & -\frac{1}{12}\left[\varphi_{i,ts}\left(\varphi_{j,tr}+\varphi_{j,sr}\right)+\varphi_{i,tr}\left(\varphi_{j,ts}+\varphi_{j,sr}\right)+\varphi_{i,sr}\left(\varphi_{j,st}+\varphi_{j,rt}\right)\right] \qquad \text{(A.5)}
\end{aligned}$$

and

$$\begin{aligned}
& \vartheta_{ij}^{(4)}\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}, Z_{ij,q}\right) \\
= \; & \bar{\varsigma}\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}, Z_{ij,q}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,t}, Z_{ij,s}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,t}, Z_{ij,r}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,t}, Z_{ij,q}\right) \\
& -\bar{\varsigma}_{ij,2}\left(Z_{ij,s}, Z_{ij,r}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,s}, Z_{ij,q}\right) - \bar{\varsigma}_{ij,2}\left(Z_{ij,r}, Z_{ij,q}\right) - \bar{\varsigma}_{ij,3}\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}\right) \\
& -\bar{\varsigma}_{ij,3}\left(Z_{ij,t}, Z_{ij,s}, Z_{ij,q}\right) - \bar{\varsigma}_{ij,3}\left(Z_{ij,t}, Z_{ij,r}, Z_{ij,q}\right) - \bar{\varsigma}_{ij,3}\left(Z_{ij,s}, Z_{ij,r}, Z_{ij,q}\right) \\
= \; & \frac{1}{6}\left\{\varphi_{i,ts}\varphi_{j,rq}+\varphi_{i,tr}\varphi_{j,sq}+\varphi_{i,rq}\varphi_{j,ts}+\varphi_{i,sq}\varphi_{j,tr}+\varphi_{i,tq}\varphi_{j,sr}+\varphi_{i,sr}\varphi_{j,tq}\right\}, \qquad \text{(A.6)}
\end{aligned}$$

where (A.5) and (A.6) will be needed in the proofs of Propositions A.4 and A.5, respectively.

Let $G_{nT}^{(k)} \equiv \frac{1}{n(n-1)P_T^k}\sum_{1\leq i\neq j\leq n}\sum_{(T,k)}\vartheta_{ij}^{(k)}\left(Z_{ij,t_1},\ldots,Z_{ij,t_k}\right)$ for $k=1,2,3,4$, where $\sum_{(T,k)}$ denotes summation over all P_T^k permutations $(t_1,...,t_k)$ of distinct integers chosen from $\{1,2,...,T\}$ (See Lee (1990), Ch 1). Then by the Hoeffding decomposition, we have

$$\widehat{\Gamma}_{nT} = 6G_{nT}^{(2)} + 4G_{nT}^{(3)} + G_{nT}^{(4)}. \qquad \text{(A.7)}$$

Let $\overline{\Gamma}_{nT} \equiv 6G_{nT}^{(2)}$. Noting that $nThE(\overline{\Gamma}_{nT}) = \frac{2h}{(n-1)(T-1)}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s\leq T}E\left[\varphi_{i,ts}\varphi_{j,ts}\right] = B_{nT}$ under H_0, we complete the proof of the theorem by showing that: (i) $nTh[\overline{\Gamma}_{nT}-E\left(\overline{\Gamma}_{nT}\right)] \xrightarrow{d} N\left(0,\sigma_0^2\right)$, (ii) $nThG_{nT}^{(3)} = o_P(1)$, and (iii) $nThG_{nT}^{(4)} = o_P(1)$. These results are established respectively in Propositions A.1, A.4, and A.5 below.

Proposition A.1 $nTh\left[\overline{\Gamma}_{nT}-E\left(\overline{\Gamma}_{nT}\right)\right] \xrightarrow{d} N\left(0,\sigma_0^2\right)$.

Proof. Let $\varphi_{i,ts}^c \equiv \varphi_{i,ts} - E(\varphi_{i,ts})$. Then we have $\overline{\Gamma}_{nT} - E(\overline{\Gamma}_{nT}) = \overline{\Gamma}_{nT,1} + \overline{\Gamma}_{nT,2}$, where

$$\begin{aligned}
\overline{\Gamma}_{nT,1} &\equiv \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n}\frac{1}{C_T^2}\sum_{1\leq t<s\leq T}\varphi_{i,ts}^c\varphi_{j,ts}^c, \text{ and} \\
\overline{\Gamma}_{nT,2} &\equiv \frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n}\frac{1}{C_T^2}\sum_{1\leq t<s\leq T}\left\{\varphi_{i,ts}^c E\left[\varphi_{j,ts}\right]+\varphi_{j,ts}^c E\left[\varphi_{i,ts}\right]\right\}.
\end{aligned}$$

We prove the proposition by showing that

$$nTh\overline{\Gamma}_{nT,1} = \frac{nT}{(n-1)(T-1)}W_{nT} \xrightarrow{d} N\left(0,\sigma_0^2\right), \qquad \text{(A.8)}$$

and

$$nTh\overline{\Gamma}_{nT,2} = o_P(1), \qquad \text{(A.9)}$$

where $W_{nT} \equiv \sum_{1\leq i<j\leq n} w_{ij}$, $w_{ij} \equiv w_{nT,ij} \equiv w_{nT}\left(\mathbf{u}_i,\mathbf{u}_j\right) \equiv \frac{4h}{nT}\sum_{1\leq t<s\leq T}\varphi_{i,ts}^c\varphi_{j,ts}^c$, and $\mathbf{u}_i \equiv (u_{i1},....,u_{iT})'$. Noting that $nT/[(n-1)(T-1)] \to 1$, the proof is completed by Lemmas A.2-A.3 below. ■

Lemma A.2 $W_{nT} \xrightarrow{d} N\left(0, \sigma_0^2\right)$ *under* H_0.

Proof. W_{nT} is a second order degenerate U-statistic that is "clean" (i.e., $E\left[w_{nT}\left(\mathbf{u}_i, \mathbf{u}_j\right)|\mathbf{u}_i\right] = E\left[w_{nT}\left(\mathbf{u}_i, \mathbf{u}_j\right)|\mathbf{u}_j\right] = 0$ for $i \neq j$) under H_0, we can apply Proposition 3.2 of de Jong (1987) to prove (A.8) by showing that

$$\begin{aligned}
\overline{\sigma}_{nT}^2 &\equiv \operatorname{Var}\left(W_{nT}\right) = \sigma_{nT}^2 + o\left(1\right), && \text{(A.10)}\\
G_I &\equiv \sum_{1\leq i<j\leq n} E\left[w_{ij}^4\right] = o\left(1\right), && \text{(A.11)}\\
G_{II} &\equiv \sum_{1\leq i<j<k\leq n} E\left[w_{ij}^2 w_{ik}^2 + w_{ji}^2 w_{jk}^2 + w_{ki}^2 w_{kj}^2\right] = o\left(1\right), && \text{(A.12)}\\
G_{IV} &\equiv \sum_{1\leq i<j<k<l\leq n} E\left[w_{ij}w_{ik}w_{lj}w_{lk} + w_{ij}w_{il}w_{kj}w_{kl} + w_{ik}w_{il}w_{jk}w_{jl}\right] = o\left(1\right). && \text{(A.13)}
\end{aligned}$$

Step 1. Proof of (A.10). First, notice that

$$\begin{aligned}
\overline{\sigma}_{nT}^2 &= \frac{16h^2}{n^2T^2}\operatorname{Var}\left(\sum_{1\leq i<j\leq n}\sum_{1\leq t<s\leq T}\varphi_{i,ts}^c\varphi_{j,ts}^c\right)\\
&= \frac{16h^2}{n^2T^2}\sum_{1\leq i<j\leq n}\sum_{1\leq t_1<t_2\leq T,\ 1\leq t_3<t_4\leq T} E\left[\varphi_{i,t_1t_2}^c\varphi_{i,t_3t_4}^c\right]E\left[\varphi_{j,t_1t_2}^c\varphi_{j,t_3t_4}^c\right].
\end{aligned}$$

We consider three cases for the summation in the last expression: the number of distinct indices in $\{t_1, t_2, t_3, t_4\}$ are 4, 3, and 2, respectively, and use (a), (b), and (c) to denote these three cases in order. In cases (a)-(b), we can apply similar arguments to those used in the proof of (A.11) below and demonstrate the corresponding sum is $o\left(1\right)$. It follows that

$$\overline{\sigma}_{nT}^2 = \frac{16h^2}{n^2T^2}\sum_{1\leq i<j\leq n}\sum_{1\leq t<s\leq T}\operatorname{Var}\left(\varphi_{i,ts}^c\right)\operatorname{Var}\left(\varphi_{j,ts}^c\right) + o\left(1\right) = \sigma_{nT}^2 + o\left(1\right).$$

Step 2. Proof of (A.11). We prove a stronger result: $G_I = o(n^{-1})$ by showing that $\max_{1\leq i\neq j\leq n} G_{ijI} = o(n^{-3})$ where $G_{ijI} \equiv E(w_{ij}^4)$. For $i \neq j$, we have that under H_0,

$$G_{ijI} = \frac{256h^4}{n^4T^4}\sum_{1\leq t_{2k-1}<t_{2k}\leq T,\ k=1,2,3,4} E\left[\prod_{l=1}^4\varphi_{i,t_{2l-1}t_{2l}}^c\right]E\left[\prod_{l=1}^4\varphi_{j,t_{2l-1}t_{2l}}^c\right].$$

We consider five cases inside the summation: the number of distinct elements in $\{t_1, t_2, ..., t_8\}$ are 8, 7, 6, 5, and 4 or less. We use (A), (B), (C), (D), and (E) to denote these five cases, respectively, and denote the corresponding sum in G_{ijI} as $G_{ijI,A}$, $G_{ijI,B}$, $G_{ijI,C}$, $G_{ijI,D}$, and $G_{ijI,E}$, respectively (e.g., $G_{ijI,A}$ is defined as G_{ijI} but with the time indices restricted to case (A)).

For case (A), we consider two different subcases: (Aa) there exists $k_0 \in \{1, ..., 8\}$ such that, $|t_l - t_{k_0}| > m$ for all $l \neq k_0$; (Ab) all the other remaining cases. We use $G_{ijI,Aa}$ and $G_{ijI,Ab}$ to denote $G_{ijI,A}$ but with the time indices restricted to subcases (Aa) and (Ab), respectively. Let $1 \leq r_1 < ... < r_8 \leq T$ be the permutation of $t_1, ..., t_8$ in ascending order. Denote $A_i\left(r_1, ..., r_8\right) \equiv \prod_{l=1}^4\varphi_{i,t_{2l-1}t_{2l}}^c$. Then it is easy to see that $|E[A_j\left(r_1, ..., r_8\right)]| \leq C$ uniformly in j.

For subcase (Aa), without loss of generality (WLOG) we assume $t_{k_0} = t_1$. We consider two sub-subcases: (Aa1) $t_1 = r_1$, (Aa2) $t_1 = r_{l_0}$ for $l_0 \in \{2, ..., 7\}$. In subsubcase (Aa1), by splitting variables indexed by t_1 from those indexed by $t_2, \ldots, t_8$, we have by Lemma E.1 that

$$|E[A_i(r_1, ..., r_8)]| \leq \left|E\left\{E_{t_1}\left(\varphi^c_{i,t_1t_2}\right)\varphi^c_{i,t_3t_4}\varphi^c_{i,t_5t_6}\varphi^c_{i,t_7t_8}\right\}\right| + Ch^{-\frac{4\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m).$$

To bound the first term in the last expression, we apply Lemma E.2 to obtain

$$\begin{aligned}\left|E_{t_1}\left(\varphi^c_{i,t_1t_2}\right)\right| &= \left|E_{t_1}E_{t_2}(\overline{k}^i_{h,t_1t_2}) - E(\overline{k}^i_{h,t_1t_2})\right| = \left|E[E_{t_2}(\overline{k}^i_{h,t_1t_2}) - E(\overline{k}^i_{h,t_1t_2}|u_{it_1})]\right| \\ &\leq E\left|E_{t_2}(\overline{k}^i_{h,t_1t_2}) - E(\overline{k}^i_{h,t_1t_2}|u_{it_1})\right| \leq Ch^{-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m). \end{aligned} \tag{A.14}$$

Consequently, we have $|A_i(t_1, ..., t_8)| \leq Ch^{-\frac{4\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)$. In subsubcase (Aa2), noting that $t_2 \in \{r_{l_0+1}, ..., r_8\}$ we split first variables indexed by $r_1, ..., r_{l_0-1}$ from others and then variables indexed by $r_{l_0}(= t_1)$ from $\{r_{l_0+1}, ..., r_8\}$ to obtain

$$\begin{aligned}|E[A_i(r_1, ..., r_8)]| &\leq |E\{E_{1,...,l_0-1}[A_i(r_1, ..., r_8)]\}| + Ch^{-\frac{4\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m) \\ &\leq |E[E_{t_1}\{E_{1,...,l_0-1}[A_i(r_1, ..., r_8)]\}]| + Ch^{-\frac{3\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m) + Ch^{-\frac{4\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m).\end{aligned}$$

Now we can apply Fubini theorem and (A.14) to bound the first term in the last expression by $Ch^{-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)$. Consequently, we have $|E[A_i(r_1, ..., r_8)]| \leq Ch^{-\frac{4\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)$ uniformly i in case (Aa). It follows that

$$G_{ijI,Aa} \leq \frac{Ch^4}{n^4T^4}T^8h^{-\frac{4\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m) = O\left(n^{-4}T^4h^{\frac{4}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)\right) = o\left(n^{-3}\right), \tag{A.15}$$

where here and below $o\left(n^{-3}\right)$ holds uniformly in (i, j). In case (Ab), the number of terms in the summation for $G_{ijI,Ab}$ is of order $O\left(T^4m^4\right)$ and each term is uniformly bounded by a constant C. It follows that

$$G_{ijI,Ab} \leq \frac{Ch^4}{n^4T^4}T^4m^4 = O\left(n^{-4}h^4m^4\right) = o\left(n^{-3}\right). \tag{A.16}$$

Now, we consider case (B). WLOG we assume $t_8 = t_6$ and consider two subcases for the indices $\{t_1, ..., t_7\}$: (Ba) there exist two distinct integers $k_1, k_2 \in \{1, ..., 7\}$ such that $|t_l - t_{k_s}| > m$ for all $l \neq k_s$ and $s = 1, 2$; (Bb) all the other remaining cases. We use $G_{ijI,Ba}$ and $G_{ijI,Bb}$ to denote $G_{ijI,B}$ but with the time indices restricted to subcases (Ba) and (Bb), respectively. In case (Ba), at least one (say t_{k_1}) of the two time indices satisfying the condition in (Ba) is not t_6 so that we can apply the same argument as used in case (Aa) to obtain the bound for $G_{I,Ba}$ as

$$G_{ijI,Ba} \leq \frac{Ch^4}{n^4T^4}T^7h^{-\frac{4\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m) = O\left(n^{-4}T^3h^{\frac{4}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)\right) = o\left(n^{-3}\right). \tag{A.17}$$

In case (Bb), the number of terms in the summation for $G_{ijI,Bb}$ is of order $O\left(T^4m^3\right)$ and each term is uniformly bounded by a constant C. It follows that

$$G_{I,Bb} \leq \frac{Ch^4}{n^4T^4}T^4m^3 = O\left(n^{-4}h^4m^3\right) = o\left(n^{-3}\right). \tag{A.18}$$

For case (C), we consider two subcases for the indices $\{t_1, ..., t_8\}$: (Ca) there exists four distinct integers $k_1, k_2, k_3, k_4 \in \{1, ..., 8\}$ such that $|t_l - t_{k_s}| > m$ for all $l \neq k_s$ and $s = 1, 2, 3, 4$ (note that

some of the t_l indices coincide here so that the total number of distinct indices among $\{t_1, ..., t_8\}$ is six); (Cb) all the other remaining cases. We use $G_{ijI,Ca}$ and $G_{ijI,Cb}$ to denote $G_{I,C}$ but with the time indices restricted to subcases (Ca) and (Cb), respectively. In case (Ca) we can follow the same arguments as used in case (Aa) to bound $G_{ijI,Ca}$ as

$$G_{ijI,Ca} \leq \frac{Ch^4}{n^4T^4} T^6 h^{-\frac{2+4\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) = O\left(n^{-4}T^2 h^{\frac{2}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)\right) = o\left(n^{-3}\right). \tag{A.19}$$

In case (Cb), the number of terms in the summation for $G_{ijI,Cb}$ is of order $O\left(T^4 m^2\right)$ and each term is uniformly bounded by a constant Ch^{-2}. It follows that

$$G_{ijI,Cb} \leq \frac{Ch^4}{n^4T^4} T^4 m^2 h^{-2} = O\left(n^{-4}h^2m^2\right) = o\left(n^{-3}\right). \tag{A.20}$$

For case (D), we consider two subcases for the indices $\{t_1, ..., t_8\}$: (Da) for all distinct integers $k \in \{1, ..., 8\}$ such that $|t_l - t_k| > m$ for all $l \neq k$ with $t_l \neq t_k$; (Db) all the other remaining cases. We use $G_{ijI,Da}$ and $G_{ijI,Db}$ to denote $G_{ijI,D}$ but with the time indices restricted to subcases (Da) and (Db), respectively. In case (Da) we can follow the same arguments used in cases (Ca), (Ba), and (Aa) to bound $G_{ijI,Da}$ as

$$G_{ijI,Da} \leq \frac{Ch^4}{n^4T^4} T^5 h^{-\frac{2+4\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) = O\left(n^{-4}Th^{\frac{2}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)\right) = o\left(n^{-3}\right). \tag{A.21}$$

In case (Db), the number of terms in the summation for $G_{ijI,Db}$ is of order $O\left(T^4 m\right)$ and each term is uniformly bounded by Ch^{-2}. It follows that

$$G_{ijI,Db} \leq \frac{Ch^4}{n^4T^4} T^4 m h^{-2} = O\left(n^{-4}h^2m\right) = o\left(n^{-3}\right). \tag{A.22}$$

In case (E), it is straightforward to bound $G_{ijI,E}$ as

$$G_{ijI,E} \leq \frac{Ch^4}{n^4T^4}\left(T^4h^{-4} + T^3h^{-4} + T^2h^{-6}\right) = O\left(n^{-4} + n^{-4}T^{-2}h^{-2}\right) = o\left(n^{-3}\right). \tag{A.23}$$

In sum, combining (A.15)-(A.23) yields

$$\max_{1\leq i\neq j\leq n} G_{ijI} = o\left(n^{-3}\right). \tag{A.24}$$

Step 3. Proof of (A.12). By the Jensen inequality and (A.24), $G_{II} \leq \sum_{1\leq i<j<k\leq n}[\{E(w_{ij}^4) \times E(w_{ik}^4)\}^{1/2} + \{E(w_{ji}^4)E(w_{jk}^4)\}^{1/2} + \{E(w_{ki}^4)E(w_{kj}^4)\}^{1/2}] \leq \frac{n^3}{2}\max_{1\leq i\neq j\leq n} E(w_{ij}^4) = o(1)$.

Step 4. Proof of (A.13). Write $G_{IV} = \sum_{1\leq i<j<k<l\leq n}\{E[w_{ij}w_{ik}w_{lj}w_{lk}] + E[w_{ij}w_{il}w_{kj}w_{kl}] + E[w_{ik}w_{il}w_{jk}w_{jl}]\} \equiv G_{IV1} + G_{IV2} + G_{IV3}$. Recalling $w_{ij} \equiv \frac{4h}{nT}\sum_{1\leq t<s\leq T}\varphi_{i,ts}^c\varphi_{j,ts}^c$,

$$\begin{aligned} G_{IV1} &= \sum_{1\leq i_1<i_2<i_3<i_4\leq n} E\left[w_{i_1i_2}w_{i_1i_3}w_{i_4i_2}w_{i_4i_3}\right] \\ &= \frac{256h^4}{n^4T^4}\sum_{1\leq i_1<i_2<i_3<i_4\leq n}\ \sum_{1\leq t_{2k-1}<t_{2k}\leq T,\ k=1,2,3,4} E\left[\varphi_{i_1,t_1t_2}^c\varphi_{i_1,t_3t_4}^c\right]E\left[\varphi_{i_2,t_1t_2}^c\varphi_{i_2,t_5t_6}^c\right] \\ &\quad \times E\left[\varphi_{i_3,t_3t_4}^c\varphi_{i_3,t_7t_8}^c\right]E\left[\varphi_{i_4,t_5t_6}^c\varphi_{i_4,t_7t_8}^c\right]. \end{aligned}$$

Like in the analysis of G_I, we consider five cases inside the above summation: the number of distinct elements in $\{t_1, t_2, ..., t_8\}$ are 8, 7, 6, 5, and 4 or less. We continue to use (A), (B), (C), (D), and (E) to denote these five cases, respectively, and denote the corresponding sum in G_{IV1} as $G_{IV1,A}$, $G_{IV1,B}$, $G_{IV1,C}$, $G_{IV1,D}$, and $G_{IV1,E}$, respectively (e.g., $G_{IV1,A}$ is defined as G_{IV1} but with the time indices restricted to case (A)). For case (A), we consider two different subcases: (Aa) there exists $k_0 \in \{1, ..., 8\}$ such that, $|t_l - t_{k_0}| > m$ for all $l \neq k_0$; (Ab) all the other remaining cases. We use $G_{IV1,Aa}$ and $G_{IV1,Ab}$ to denote $G_{IV1,A}$ but with the time indices restricted to subcases (Aa) and (Ab), respectively. In case (Aa) we can follow the same argument as used in case (Aa) in Step 2 to bound $G_{IV1,Aa}$ as $G_{IV1,Aa} \leq \frac{Ch^4}{n^4T^4} n^4 T^8 h^{-\frac{2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) = O(T^4 h^{\frac{2(2+\delta)}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)) = o(1)$. In case (Ab), the number of terms in the summation for $G_{IV1,Ab}$ is of order $O\left(T^4 m^4\right)$ and each term is uniformly bounded by a constant C. It follows that $G_{IV1,Ab} \leq \frac{Ch^4}{n^4T^4} n^4 T^4 m^4 = O\left(h^4 m^4\right) = o(1)$.

For case (B), we consider two different subcases: (Ba) there exists $k_0 \in \{1, ..., 8\}$ such that, $|t_l - t_{k_0}| > m$ for all $l \neq k_0$ with $t_l \neq t_{k_0}$; (Bb) all the other remaining cases. For subcase (Ba), we consider only two representative subcases: (Ba1) $t_8 = t_1$ or $t_8 = t_2$, (Ba2) $t_8 = t_5$ or $t_8 = t_6$ since the other cases are analogous. For subsubcase (Ba1) WLOG we assume $t_8 = t_1$. Noting that all the four time indices in each of the four expectations $E[\varphi^c_{i_1,t_1t_2}\varphi^c_{i_1,t_3t_4}]$, $E[\varphi^c_{i_2,t_1t_2}\varphi^c_{i_2,t_5t_6}]$, $E[\varphi^c_{i_3,t_3t_4}\varphi^c_{i_3,t_7t_1}]$, and $E[\varphi^c_{i_4,t_5t_6}\varphi^c_{i_4,t_7t_1}]$ are different from each other, we can easily get the bound for $G_{IV1,B}$ (with the restriction $t_8 = t_1$) as $O(T^3 h^{\frac{2(2+\delta)}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)) = o(1)$. For subsubcase (Ba2) we assume $t_8 = t_5$ and consider bounding the following objects: $E[\varphi^c_{i_1,t_1t_2}\varphi^c_{i_1,t_3t_4}]$, $E[\varphi^c_{i_2,t_1t_2}\varphi^c_{i_2,t_5t_6}]$, $E[\varphi^c_{i_3,t_3t_4}\varphi^c_{i_3,t_7t_5}]$, and $E[\varphi^c_{i_4,t_5t_6}\varphi^c_{i_4,t_7t_5}]$. Note that the indices in the last expectation $E[\varphi^c_{i_4,t_5t_6}\varphi^c_{i_4,t_7t_5}]$ are not all distinct. Despite this, since all the four indices in each of the other three expectations are distinct, we can continue to bound $G_{IV1,B}$ (with the restriction $t_8 = t_5$) as $O(T^3 h^{\frac{2(2+\delta)}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)) = o(1)$. For subcase (Bb), it is easy to tell $G_{IV1,B}$ is bounded by $T^{-4}h^4 O\left(T^4 m^3\right) = O\left(h^4 m^3\right) = o(1)$. It follows that $G_{IV1,B} = o(1)$. For case (C), analogous to the study of case (C) in Step 2, we have

$$G_{IV1,C} = \frac{h^4}{T^4} O\left(T^6 h^{-1-\frac{2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) + T^4 m^2 h^{-1}\right) = O\left(T^2 h^{\frac{3+\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) + h^3 m^2\right) = o(1).$$

Similarly, in case (D) we have

$$G_{IV1,D} \leq \frac{h^4}{T^4} O\left(T^5 h^{-1-\frac{2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) + T^4 m h^{-1}\right) = O\left(T h^{\frac{3+\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) + h^3 m\right) = o(1).$$

In case (E), it is straightforward to bound $G_{IV1,E}$ as

$$G_{IV1,E} \leq \frac{Ch^4}{n^4T^4} n^4 \left(T^4 h^{-2} + T^3 h^{-3} + T^2 h^{-4}\right) = O\left(h^2 + T^{-1}h + T^{-2}\right) = o(1).$$

In sum, $G_{IV1} = o(1)$. Similarly we can show that $G_{IVs} = o(1)$ for $s = 2, 3$. ■

Lemma A.3 $nTh\overline{\Gamma}_{nT,2} = o_P(1)$.

Proof. Let $n_1 \equiv n - 1$ and $T_1 \equiv T - 1$. Recalling that $c_{i,ts} \equiv E\left(\varphi_{i,ts}\right)$ and $c_{ts} \equiv n_1^{-1} \sum_{i=1}^n c_{i,ts}$,

we have

$$\begin{aligned}
nTh\overline{\Gamma}_{nT,2} &= \frac{2h}{n_1}\sum_{1\leq j\neq i\leq n} T_1^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left[\varphi_{i,ts}^c c_{j,ts}+\varphi_{j,ts}^c c_{i,ts}\right] \\
&= \frac{2h}{n_1}\sum_{i=1}^{n}\sum_{j=1}^{n} T_1^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\left[\varphi_{i,ts}^c c_{j,ts}+\varphi_{j,ts}^c c_{i,ts}\right]-\frac{4h}{n_1}\sum_{i=1}^{n} T_1^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\varphi_{i,ts}^c c_{i,ts} \\
&= 4h\sum_{i=1}^{n} T_1^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\varphi_{i,ts}^c c_{ts}-\frac{4h}{n_1}\sum_{i=1}^{n} T_1^{-1}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\varphi_{i,ts}^c c_{i,ts} \\
&\equiv 4V_{1nT}-4V_{2nT}, \text{ say.}
\end{aligned}$$

We complete the proof by showing that $V_{1nT}=o_P(1)$ and $V_{2nT}=o_P(1)$. We only prove the first claim since the proof of the second one is similar.

Let $v_{i,t}\equiv\sum_{s=1}^{t-1}h^{1/2}\varphi_{i,ts}^c c_{ts}$ and $v_i\equiv T_1^{-1}\sum_{t=2}^{T}v_{i,t}$. Then we can write $V_{1nT}=h^{1/2}\sum_{i=1}^{n}v_i$. Note that $E(v_i)=0$ and $\{v_i\}_{i=1}^n$ are independently distributed under H_0, we have $E[(V_{1nT})^2]=h\sum_{i=1}^{n}\mathrm{Var}(v_i)$. For $\mathrm{Var}(v_i)$, we have

$$\mathrm{Var}(v_i)=E\left[\frac{1}{T_1}\sum_{t=2}^{T}v_{i,t}\right]^2=\frac{1}{T_1^2}\sum_{t=2}^{T}E\left[v_{i,t}^2\right]+\frac{2}{T_1^2}\sum_{t_1=3}^{T}\sum_{t_2=2}^{t_1-1}E\left[v_{i,t_1}v_{i,t_2}\right]\equiv V_{1i}+V_{2i}, \text{ say.}$$

For V_{1i}, we have

$$V_{1i}=\frac{h}{T_1^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}E\left[\varphi_{i,ts}^{c2}\right]c_{ts}^2+\frac{2h}{T_1^2}\sum_{t=3}^{T}\sum_{s=2}^{t-1}\sum_{r=1}^{s-1}E\left[\varphi_{i,ts}^c\varphi_{i,tr}^c\right]c_{ts}c_{tr}\equiv V_{1i,1}+V_{1i,2}, \text{ say.}$$

By (A.1) and Assumption A.1, $|c_{ts}|=|n_1^{-1}\sum_{i=1}^{n}E[\varphi_{i,ts}]|\leq Ch^{\frac{-\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(t-s)$. Thus uniformly in i

$$\begin{aligned}
V_{1i,1} &\leq \frac{C}{T_1^2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}h^{\frac{-2\delta}{1+\delta}}\alpha^{\frac{2\delta}{1+\delta}}(t-s)\max_{1\leq t\neq s\leq T}\left\{hE\left[\varphi_{i,ts}^{c2}\right]\right\} \\
&\leq \frac{C}{T_1}\max_{1\leq i\leq n}\max_{1\leq t\neq s\leq T}\left\{hE\left[\varphi_{i,ts}^{c2}\right]\right\}h^{\frac{-2\delta}{1+\delta}}\sum_{\tau=1}^{T-1}\alpha^{\frac{2\delta}{1+\delta}}(\tau)=O\left(T^{-1}h^{\frac{-2\delta}{1+\delta}}\right).
\end{aligned}$$

For $V_{1i,2}$, we have that uniformly in i

$$\begin{aligned}
|V_{1i,2}| &= \frac{2h}{T_1^2}\sum_{t=3}^{T}\sum_{s=2}^{t-1}\sum_{r=1}^{s-1}\left|E\left(\varphi_{i,ts}^c\varphi_{i,tr}^c\right)\right||c_{ts}||c_{tr}|\leq\frac{Chh^{\frac{-2\delta}{1+\delta}}}{T_1^2}\sum_{t=3}^{T}\sum_{s=2}^{t-1}\sum_{r=1}^{s-1}\alpha^{\frac{\delta}{1+\delta}}(t-s)\alpha^{\frac{\delta}{1+\delta}}(t-r) \\
&\leq \frac{Ch^{\frac{1-\delta}{1+\delta}}}{T_1}\sum_{\tau_1=1}^{\infty}\sum_{\tau_2=1}^{\infty}\alpha^{\frac{\delta}{1+\delta}}(\tau_1)\alpha^{\frac{\delta}{1+\delta}}(\tau_2)=O\left(T^{-1}h^{\frac{1-\delta}{1+\delta}}\right).
\end{aligned}$$

It follows that $V_{1i}=O(T^{-1}h^{\frac{-2\delta}{1+\delta}}+T^{-1}h^{\frac{1-\delta}{1+\delta}})$ uniformly in i.

For V_{2i}, we have

$$\begin{aligned}
V_{2i} &= \frac{2h}{T_1^2}\sum_{t_1=3}^{T}\sum_{t_2=2}^{t_1-1}\sum_{t_3=1}^{t_1-1}\sum_{t_4=1}^{t_2-1}E\left[\varphi_{i,t_1t_3}^c\varphi_{i,t_2t_4}^c\right]c_{t_1t_3}c_{t_2t_4} \\
&= \frac{4h}{T_1^2}\sum_{t_1=3}^{T}\sum_{t_2=2}^{t_1-1}\sum_{t_3=1}^{t_2-1}E\left[\varphi_{i,t_1t_3}^c\varphi_{i,t_2t_3}^c\right]c_{t_1t_3}c_{t_2t_3} \\
&\quad+\frac{2h}{T_1^2}\sum_{t_1=3}^{T}\sum_{t_2=2}^{t_1-1}\sum_{t_3=1,t_3\neq t_4,t_2}^{t_1-1}\sum_{t_4=1}^{t_2-1}E\left[\varphi_{i,t_1t_3}^c\varphi_{i,t_2t_4}^c\right]c_{t_1t_3}c_{t_2t_4}\equiv V_{2i,1}+V_{2i,2} \text{ say,}
\end{aligned}$$

where the first term is obtained when $t_3 = t_4$ or t_2 as $\varphi_{i,ts} = \varphi_{i,st}$. Following the analysis of $V_{1i,2}$, we can show that $|V_{2i,1}| \leq CT_1^{-1}h^{\frac{1-\delta}{1+\delta}} \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=1}^{\infty} \alpha^{\frac{\delta}{1+\delta}}(\tau_1)\,\alpha^{\frac{\delta}{1+\delta}}(\tau_2) = O(T^{-1}h^{\frac{1-\delta}{1+\delta}})$ uniformly in i. For $V_{2i,2}$, we consider three cases: (a) $1 \leq t_3 < t_4 < t_2 < t_1 \leq T$; (b) $1 \leq t_4 < t_3 < t_2 < t_1 \leq T$; (c) $1 \leq t_4 < t_2 < t_3 < t_1 \leq T$, and use $V_{2i,2a}$, $V_{2i,2b}$, and $V_{2i,2c}$ to denote the summation over these three cases of indices, respectively. In case (a), by separating variables indexed by t_3 from those indexed by t_4, t_2, and t_1 and Lemma E.1, we have

$$\left|E\left[\varphi^c_{i,t_1t_3}\varphi^c_{i,t_2t_4}\right]\right| \leq \left|E\left[E_{t_3}\left(\varphi^c_{i,t_1t_3}\right)\varphi^c_{i,t_2t_4}\right]\right| + Ch^{\frac{-2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(t_4-t_3) = Ch^{\frac{-2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(t_4-t_3),$$

where the equality follows from the fact that $E_s(\varphi^c_{i,ts}) = E_tE_s(\overline{k}^i_{h,ts}) - E(\overline{k}^i_{h,ts})$ is a constant and that $E(\varphi^c_{i,ts}) = 0$ for $t \neq s$. It follows that uniformly in i

$$\begin{aligned}
|V_{2i,2a}| &\leq \frac{2h}{T_1^2}\sum_{t_1=3}^{T}\sum_{t_2=2}^{t_1-1}\sum_{t_4=1}^{t_2-1}\sum_{t_3=1}^{t_4-1}\left|E\left[\varphi^c_{i,t_1t_3}\varphi^c_{i,t_2t_4}\right]\right||c_{t_1t_3}||c_{t_2t_4}| \\
&\leq \frac{Chh^{\frac{-4\delta}{1+\delta}}}{T_1^2}\sum_{t_1=3}^{T}\sum_{t_2=2}^{t_1-1}\sum_{t_3=1,t_3\neq t_4,t_2}^{t_1-1}\sum_{t_4=1}^{t_2-1}\alpha^{\frac{\delta}{1+\delta}}(t_4-t_3)\,\alpha^{\frac{\delta}{1+\delta}}(t_1-t_3)\,\alpha^{\frac{\delta}{1+\delta}}(t_2-t_4) \\
&\leq \frac{Ch^{\frac{1-3\delta}{1+\delta}}}{T_1}\sum_{\tau_3=1}^{\infty}\sum_{\tau_2=1}^{\infty}\sum_{\tau_1=1}^{\infty}\alpha^{\frac{\delta}{1+\delta}}(\tau_1)\,\alpha^{\frac{\delta}{1+\delta}}(\tau_2)\,\alpha^{\frac{\delta}{1+\delta}}(\tau_3) = O\left(T^{-1}h^{\frac{1-3\delta}{1+\delta}}\right).
\end{aligned}$$

By the same token, we can show that $|V_{2i,2\xi}| = O(T^{-1}h^{\frac{1-3\delta}{1+\delta}})$ uniformly in i for $\xi = b, c$. Hence $V_{2i,2} = O(T^{-1}h^{\frac{1-3\delta}{1+\delta}})$ and $V_{2i} = O(T^{-1}h^{\frac{1-\delta}{1+\delta}}) + O(T^{-1}h^{\frac{1-3\delta}{1+\delta}}) = O(T^{-1}h^{\frac{1-3\delta}{1+\delta}})$ uniformly in i. Consequently

$$E[(V_{1nT})^2] = h\sum_{i=1}^{n}(V_{1i}+V_{2i}) = O\left(nh\left(T^{-1}h^{\frac{-2\delta}{1+\delta}} + T^{-1}h^{\frac{1-3\delta}{1+\delta}}\right)\right) = O\left(nh^{\frac{1-\delta}{1+\delta}}/T\right) = o(1).$$

Then $V_{1nT} = o_P(1)$ by the Chebyshev inequality. ■

Proposition A.4 $nThG^{(3)}_{nT} = o_P(1)$.

Proof. By the definition of $G^{(3)}_{nT}$ and (A.5), we have

$$\begin{aligned}
-12nThG^{(3)}_{nT} &= \frac{-12nTh}{n(n-1)C_T^3}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s<r\leq T}\vartheta^{(3)}_{ij}(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}) \\
&= \frac{Th}{n_1C_T^3}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s<r\leq T}[\varphi_{i,ts}\varphi_{j,tr} + \varphi_{i,ts}\varphi_{j,sr} + \varphi_{i,tr}\varphi_{j,ts} + \varphi_{i,tr}\varphi_{j,sr} \\
&\quad + \varphi_{i,sr}\varphi_{j,st} + \varphi_{i,sr}\varphi_{j,rt}] \\
&\equiv U_{1nT} + U_{2nT} + U_{3nT} + U_{4nT} + U_{5nT} + U_{6nT}, \text{ say,}
\end{aligned}$$

where, e.g., $U_{1nT} \equiv \frac{Th}{n_1C_T^3}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s<r\leq T}\varphi_{i,ts}\varphi_{j,tr}$. It suffices to show that $U_{rnT} = o_P(1)$ for $r = 1, 2, ..., 6$.

For U_{1nT}, we have

$$\begin{aligned}
U_{1nT} &= \frac{Th}{n_1C_T^3}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s<r\leq T}\varphi^c_{i,ts}\varphi^c_{j,tr} + \frac{Th}{n_1C_T^3}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s<r\leq T}c_{i,ts}\varphi^c_{j,tr} \\
&\quad + \frac{Th}{n_1C_T^3}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s<r\leq T}\varphi^c_{i,ts}c_{j,tr} + \frac{Th}{n_1C_T^3}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t<s<r\leq T}c_{i,ts}c_{j,tr} \\
&\equiv U_{1nT,1} + U_{1nT,2} + U_{1nT,3} + U_{1nT,4}, \text{ say,}
\end{aligned}$$

where recall $\varphi_{i,ts}^c \equiv \varphi_{i,ts} - E(\varphi_{i,ts})$ and $c_{i,ts} \equiv E\left(\varphi_{i,ts}\right)$. We further decompose $U_{1nT,1}$ as follows

$$\begin{aligned} U_{1nT,1} &= \frac{Th}{n_1 C_T^3} \sum_{1\le i<j\le n} \sum_{1\le t<s<r\le T} \varphi_{i,ts}^c \varphi_{j,tr}^c + \frac{Th}{n_1 C_T^3} \sum_{1\le j<i\le n} \sum_{1\le t<s<r\le T} \varphi_{i,ts}^c \varphi_{j,tr}^c \\ &\equiv U_{1nT,1a} + U_{1nT,1b}. \end{aligned}$$

Noting that $E\left(U_{1nT,1a}\right) = 0$ under H_0, we have

$$\begin{aligned} \operatorname{Var}\left(U_{1nT,1a}\right) &= \frac{T^2h^2}{\left(n_1 C_T^3\right)^2} \sum_{1\le i_1<i_2\le n} \sum_{\substack{1\le t_1<t_2<t_3\le T \\ 1\le t_4<t_5<t_6\le T}} E\left[\varphi_{i_1,t_1t_2}^c \varphi_{i_2,t_1t_3}^c \varphi_{i_1,t_4t_5}^c \varphi_{i_2,t_4t_6}^c\right] \\ &= \frac{T^2h^2}{\left(n_1 C_T^3\right)^2} \sum_{1\le i_1<i_2\le n} \sum_{\substack{1\le t_1<t_2<t_3\le T \\ 1\le t_4<t_5<t_6\le T}} E\left[\varphi_{i_1,t_1t_2}^c \varphi_{i_1,t_4t_5}^c\right] E\left[\varphi_{i_2,t_1t_3}^c \varphi_{i_2,t_4t_6}^c\right]. \end{aligned}$$

Analogously to the proof of (A.13), we can show

$$\begin{aligned} \operatorname{Var}\left(U_{1nT,1a}\right) &\le \frac{CT^2h^2}{\left(n_1 C_T^3\right)^2} \left\{ n^2T^6 h^{\frac{-2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) + n^2T^3m^3 + n^2T^3h^{-2} \right\} \\ &= O\left(T^2 h^{\frac{2}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) + T^{-1}h^2m^3 + T^{-1} \right) = o(1). \end{aligned}$$

Hence $U_{1nT,1a} = o_P(1)$ by the Chebyshev inequality. Similarly, $U_{1nT,1b} = o_P(1)$. It follows that $U_{1nT,1} = o_P(1)$.

For $U_{1nT,2}$, write

$$\begin{aligned} U_{1nT,2} &= \frac{Th}{n_1 C_T^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{1\le t<s<r\le T} c_{j,ts} \varphi_{i,tr}^c - \frac{Th}{n_1 C_T^3} \sum_{i=1}^n \sum_{1\le t<s<r\le T} c_{i,ts} \varphi_{i,tr}^c \\ &= \frac{Th}{C_T^3} \sum_{i=1}^n \sum_{1\le t<s<r\le T} c_{ts} \varphi_{i,tr}^c - \frac{Th}{n_1 C_T^3} \sum_{i=1}^n \sum_{1\le t<s<r\le T} c_{i,ts} \varphi_{i,tr}^c \equiv U_{1nT,2a} - U_{1nT,2b}, \end{aligned}$$

where recall $c_{ts} \equiv n_1^{-1} \sum_{i=1}^n c_{i,ts}$. Noting $E\left(U_{1nT,2a}\right) = 0$, we have

$$\begin{aligned} \operatorname{Var}\left(U_{1nT,2a}\right) &= \frac{T^2h^2}{\left(C_T^3\right)^2} \sum_{i=1}^n \sum_{\substack{1\le t_1<t_2<t_3\le T \\ 1\le t_4<t_5<t_6\le T}} c_{t_1t_2} c_{t_4t_5} E\left[\varphi_{i,t_1t_3}^c \varphi_{i,t_4t_6}^c\right] \\ &= \frac{T^2h^2}{\left(C_T^3\right)^2} \sum_{i=1}^n \sum_{\substack{1\le t_1<t_2<t_3\le T, \\ 1\le t_4<t_5<t_6\le T, \\ t_1,\ldots,t_6 \text{ are all distinct}}} c_{t_1t_2} c_{t_4t_5} E\left[\varphi_{i,t_1t_3}^c \varphi_{i,t_4t_6}^c\right] + o(1) \\ &\le \frac{Ch^2 h^{\frac{-4\delta}{1+\delta}}}{T} \sum_{i=1}^n \sum_{\tau_3=1}^\infty \sum_{\tau_2=1}^\infty \sum_{\tau_1=1}^\infty \alpha^{\frac{\delta}{1+\delta}}(\tau_1) \alpha^{\frac{\delta}{1+\delta}}(\tau_2) \alpha^{\frac{\delta}{1+\delta}}(\tau_3) + o(1) \\ &= O\left(nh^{\frac{2(1-\delta)}{1+\delta}}/T \right) + o(1) = o(1). \end{aligned}$$

So $U_{1nT,2a} = o_P(1)$. By the same token $U_{1nT,2b} = o_P(1)$. Thus $U_{1nT,2} = o_P(1)$. Similarly we can

show that $U_{1nT,3} = o_P(1)$. For $U_{1nT,4}$, we have

$$\begin{aligned}
|U_{1nT,4}| &\leq \frac{Th}{n_1 C_T^3} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r\leq T} |c_{i,ts}| \, |c_{j,tr}| \\
&\leq \frac{CThh^{\frac{-2\delta}{1+\delta}}}{n_1 C_T^3} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r\leq T} \alpha^{\frac{\delta}{1+\delta}}(s-t)\, \alpha^{\frac{\delta}{1+\delta}}(r-t) \\
&\leq \frac{Cnh^{\frac{1-\delta}{1+\delta}}}{T} \sum_{\tau_1=1}^{\infty} \sum_{\tau_2=1}^{\infty} \alpha^{\frac{\delta}{1+\delta}}(\tau_1)\, \alpha^{\frac{\delta}{1+\delta}}(\tau_2) = O\left(nh^{\frac{1-\delta}{1+\delta}}/T\right) = o(1).
\end{aligned}$$

Consequently, $U_{1nT} = o_P(1)$. Analogously we can show that $U_{rnT} = o_P(1)$ for $r = 2, 3, ..., 6$. This completes the proof of the proposition. ■

Proposition A.5 $nThG_{nT}^{(4)} = o_P(1)$.

Proof. By the definition of $G_{nT}^{(4)}$ and (A.6), we have

$$\begin{aligned}
6nThG_{nT}^{(4)} &= \frac{6Th}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q\leq T} \vartheta_{ij}^{(4)}(Z_{ij,t}, Z_{ij,s}, Z_{ij,r}, Z_{ij,q}) \\
&= \frac{Th}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q\leq T} \{\varphi_{i,ts}\varphi_{j,rq} + \varphi_{i,tr}\varphi_{j,sq} + \varphi_{i,rq}\varphi_{j,ts} + \varphi_{i,sq}\varphi_{j,tr} \\
&\quad + \varphi_{i,tq}\varphi_{j,sr} + \varphi_{i,sr}\varphi_{j,tq}\} \equiv \sum_{l=1}^{6} Q_{lnT}, \text{ say,}
\end{aligned}$$

where e.g., $Q_{1nT} = \frac{Th}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q} \varphi_{i,ts}\varphi_{j,rq}$. It suffices to show $Q_{lnT} = o_P(1)$ for $l = 1, 2, ..., 6$. We only show that $Q_{1nT} = o_P(1)$ since the other cases are similar. Write

$$\begin{aligned}
Q_{1nT} &= \frac{Th}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q\leq T} \varphi_{i,ts}^c \varphi_{j,rq}^c + \frac{Th}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q\leq T} c_{i,ts}\varphi_{j,rq}^c \\
&\quad + \frac{Th}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q\leq T} \varphi_{i,ts}^c c_{j,rq} + \frac{Th}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q\leq T} c_{i,ts} c_{j,rq} \\
&\equiv Q_{1nT,1} + Q_{1nT,2} + Q_{1nT,3} + Q_{1nT,4}, \text{ say.}
\end{aligned}$$

Analogously to the determination of the probability orders of $U_{1nT,1}$, $U_{1nT,2}$, and $U_{1nT,3}$ in the proof of Proposition A.4, we can show that $Q_{1nT,s} = o_P(1)$ for $s = 1, 2, 3$. For $Q_{1nT,4}$, we have

$$|Q_{1nT,4}| \leq \frac{CThh^{-\frac{2\delta}{1+\delta}}}{n_1 C_T^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t<s<r<q\leq T} \alpha^{\frac{1}{1+\delta}}(s-t)\, \alpha^{\frac{1}{1+\delta}}(q-r) = O(nh^{\frac{1-\delta}{1+\delta}}/T) = o(1).$$

It follows that $Q_{1nT} = o_P(1)$. ■

B Proof of Corollary 3.2

Given Theorem 3.1, it suffices to show: (i) $\widehat{D}_{1nT} \equiv \widehat{\sigma}_{nT}^2 - \sigma_{nT}^2 = o_P(1)$, and (ii) $\widehat{D}_{2nT} \equiv \widehat{B}_{nT} - B_{nT} = o_P(1)$. For (i), we write

$$\begin{aligned}
\sigma_{nT}^2 &= \frac{4h^2}{n(n-1)T(T-1)} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t\neq s\leq T} E\left[\left(\overline{k}_{h,ts}^i\right)^2\right] E\left[\left(\overline{k}_{h,ts}^j\right)^2\right] + o(1) \\
&= \frac{4R\left(\overline{k}\right)^2}{n(n-1)T(T-1)} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t\neq s\leq T} \int f_{i,ts}(u,u)\,du \int f_{j,ts}(v,v)\,dv + o(1).
\end{aligned}$$

Then

$$\begin{aligned}
\widehat{D}_{1nT} &= \frac{4R\left(\overline{k}\right)^2}{n\left(n-1\right)}\sum_{1\leq i\neq j\leq n}\frac{1}{T}\sum_{t=1}^{T}\widehat{f}_{ij,-t}\left(u_{it},u_{jt}\right) \\
&\quad -\frac{4R\left(\overline{k}\right)^2}{n\left(n-1\right)T\left(T-1\right)}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t\neq s\leq T}\int f_{i,ts}\left(u,u\right)du\int f_{j,ts}\left(v,v\right)dv-o(1) \\
&= D_{1nT}-o\left(1\right).
\end{aligned}$$

where $D_{1nT}\equiv\frac{4R\left(\overline{k}\right)^2}{n(n-1)T(T-1)}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t\neq s\leq T}\{k_{h,ts}^{i}k_{h,ts}^{j}-\int f_{i,ts}\left(u,u\right)du\int f_{j,ts}\left(v,v\right)dv\}$. It is easy to show that $E\left(D_{1nT}\right)=O\left(h^{\gamma}\right)=o\left(1\right)$ and $\mathrm{Var}(D_{1nT})=o\left(1\right)$. Consequently, $\widehat{D}_{1nT}=o_P\left(1\right)$.

Now we show (ii). Noting that $B_{nT}=\frac{2h}{n_1}\sum_{1\leq i\neq j\leq n}\sum_{r=2}^{T}\frac{T-r+1}{T-1}E\left[\varphi_{i,1r}\right]E\left[\varphi_{j,1r}\right]$, we have

$$\begin{aligned}
\widehat{B}_{nT}-B_{nT} &= \frac{2h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}\{\widehat{E}\left[\varphi_{i,1r}\right]\widehat{E}\left[\varphi_{j,1r}\right]-E\left[\varphi_{i,1r}\right]E\left[\varphi_{j,1r}\right]\} \\
&= \frac{2h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}E\left[\varphi_{i,1r}\right]\left\{\widehat{E}\left[\varphi_{j,1r}\right]-E\left[\varphi_{j,1r}\right]\right\} \\
&\quad +\frac{2h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}\left\{\widehat{E}\left[\varphi_{i,1r}\right]-E\left[\varphi_{i,1r}\right]\right\}E\left[\varphi_{j,1r}\right] \\
&\quad +\frac{2h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}\left\{\widehat{E}\left[\varphi_{i,1r}\right]-E\left[\varphi_{i,1r}\right]\right\}\left\{\widehat{E}\left[\varphi_{j,1r}\right]-E\left[\varphi_{j,1r}\right]\right\} \\
&\equiv 2D_{2nT,1}+2D_{2nT,2}+2D_{2nT,3},\ \text{say.}
\end{aligned}$$

Recalling $c_{i,ts}\equiv E\left[\varphi_{i,ts}\right]$ and $c_{ts}\equiv n_1^{-1}\sum_{i=1}^{n}c_{i,ts}$, we have $D_{2nT,1}=h\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{i=1}^{n}c_{1r}\{\widehat{E}\left[\varphi_{i,1r}\right]-E\left[\varphi_{i,1r}\right]\}-\frac{h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{i=1}^{n}c_{i,1r}\left\{\widehat{E}\left[\varphi_{i,1r}\right]-E\left[\varphi_{j,1r}\right]\right\}\equiv D_{2nT,1a}-D_{2nT,1b}$, say. We only show that $D_{2nT,1a}=o_P\left(1\right)$ as the proof that $D_{2nT,1b}=o_P\left(1\right)$ is analogous. Noting that

$$\widehat{E}\left[\varphi_{i,1r}\right]-E\left[\varphi_{i,1r}\right]=\frac{1}{T-r+1}\sum_{t=1}^{T-r+1}\{\overline{k}_{h,t,t+r-1}^{i}-E[\overline{k}_{h,t,t+r-1}^{i}]\}-\frac{1}{C_T^2}\sum_{1\leq t<s\leq T}\{\overline{k}_{h,ts}^{i}-E_tE_s[\overline{k}_{h,ts}^{i}]\}, \tag{B.1}$$

we have

$$\begin{aligned}
D_{2nT,1a} &= h\sum_{i=1}^{n}\sum_{r=2}^{T}\overline{c}_{1r}\frac{1}{T_r+1}\sum_{t=1}^{T_r+1}\left\{\overline{k}_{h,t,t+r-1}^{i}-E[\overline{k}_{h,t,t+r-1}^{i}]\right\} \\
&\quad -h\sum_{i=1}^{n}\sum_{r=2}^{T}\overline{c}_{1r}\frac{1}{C_T^2}\sum_{1\leq t<s\leq T}\left\{\overline{k}_{h,ts}^{i}-E_tE_s[\overline{k}_{h,ts}^{i}]\right\} \\
&\equiv D_{2nT,1a1}-D_{2nT,1a2},\ \text{say,}
\end{aligned} \tag{B.2}$$

where $\overline{c}_{1r} \equiv c_{1r}\left(T-r+1\right)/\left(T-1\right)$ and $T_r \equiv T-r$. Noting that $E\left(D_{2nT,1a1}\right)=0$, we have

$$\begin{aligned}
&\operatorname{Var}\left(D_{2nT,1a1}\right)\\
=\;& h^2\sum_{i=1}^{n}\sum_{r_1=2}^{T}\overline{c}_{1r_1}\sum_{r_2=2}^{T}\overline{c}_{1r_2}\frac{1}{(T_{r_1}+1)(T_{r_2}+1)}\sum_{t=1}^{T_{r_1}+1}\sum_{s=1}^{T_{r_2}+1}\operatorname{Cov}\left(\overline{k}^{i}_{h,t,t+r_1-1},\overline{k}^{i}_{h,s,s+r_2-1}\right)\\
=\;& h^2\sum_{i=1}^{n}\sum_{r_1=2}^{T}\overline{c}_{1r_1}\sum_{r_2=2}^{T}\overline{c}_{1r_2}\frac{1}{(T_{r_1}+1)(T_{r_2}+1)}\sum_{t=1}^{T_{r_1}+1}\sum_{s=1,s\neq t,s\neq t+r_1-r_2}^{T_{r_2}+1}\operatorname{Cov}\left(\overline{k}^{i}_{h,t,t+r_1-1},\overline{k}^{i}_{h,s,s+r_2-1}\right)\\
&+o\left(1\right). \qquad \text{(B.3)}
\end{aligned}$$

We consider three cases for the summation in the last expression: (a) $t<t+r_1-1<s<s+r_2-1$ or $s<s+r_2-1<t<t+r_1-1$, (b) $t<s<s+r_2-1<t+r_1-1$ or $s<t<t+r_1-1<s+r_2-1$, and (c) $t<s<t+r_1-1<s+r_2-1$ or $s<t<s+r_2-1<t+r_1-1$, and use VD_{2nTa}, VD_{2nTb}, and VD_{2nTc} denote the summation in (B.3) corresponding these three cases, respectively. In case (a) we can apply the fact that $\sum_{r=2}^{T}\overline{c}_{1r}\leq Ch^{-\frac{\delta}{1+\delta}}$ and the Davydov inequality to obtain $VD_{2nTa}\leq Cnh^{2-\frac{4\delta}{1+\delta}}/T=O(nh^{\frac{2(1-\delta)}{1+\delta}}/T)=o\left(1\right)$. In case (b), WLOG we assume $t<s<s+r_2-1<t+r_1-1$. Then we apply Lemma E.1 by first separating t from $(s,s+r_2-1,t+r_1-1)$ and then separating $t+r_1-1$ from $(s,s+r_2-1)$ to obtain

$$\begin{aligned}
&\left|\operatorname{Cov}\left(\overline{k}^{i}_{h,t,t+r_1-1},\ \overline{k}^{i}_{h,s,s+r_2-1}\right)\right|\\
=\;& \left|E\left\{\left[\overline{k}^{i}_{h,t,t+r_1-1}-E(\overline{k}^{i}_{h,t,t+r_1-1})\right]\left[\overline{k}^{i}_{h,s,s+r_2-1}-E(\overline{k}^{i}_{h,s,s+r_2-1})\right]\right\}\right|\\
\leq\;& \left|E\left\{E_t\left[\overline{k}^{i}_{h,t,t+r_1-1}-E(\overline{k}^{i}_{h,t,t+r_1-1})\right]\left[\overline{k}^{i}_{h,s,s+r_2-1}-E(\overline{k}^{i}_{h,s,s+r_2-1})\right]\right\}\right|+Ch^{-\frac{2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}\left(s-t\right)\\
\leq\;& Ch^{-\frac{2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}\left(t+r_1-s-r_2\right)+Ch^{-\frac{2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}\left(s-t\right).
\end{aligned}$$

Then we have

$$\begin{aligned}
&h^2\sum_{i=1}^{n}\sum_{r_1=2}^{T}\overline{c}_{1r_1}\sum_{r_2=2}^{T}\frac{\overline{c}_{1r_2}}{(T_{r_1}+1)(T_{r_2}+1)}\sum_{\substack{t=1\\ t<s<s+r_2-1<t+r_1-1}}^{T_{r_1}+1}\sum_{s=1,s\neq t,s\neq t+r_1-r_2}^{T_{r_2}+1}\left|\operatorname{Cov}\left(\overline{k}^{i}_{h,t,t+r_1-1},\overline{k}^{i}_{h,s,s+r_2-1}\right)\right|\\
\leq\;& Mh^{\frac{2}{1+\delta}}\sum_{i=1}^{n}\sum_{r_1=2}^{T}\sum_{r_2=2}^{T}\frac{\overline{c}_{1r_1}\overline{c}_{1r_2}}{(T_{r_1}+1)(T_{r_2}+1)}\sum_{\substack{t=1\\ t<s<s+r_2-1<t+r_1-1}}^{T_{r_1}+1}\sum_{s=1,s\neq t,s\neq t+r_1-r_2}^{T_{r_2}+1}\left\{\alpha^{\frac{\delta}{1+\delta}}\left(t+r_1-s-r_2\right)+\alpha^{\frac{\delta}{1+\delta}}\left(s-t\right)\right\}\\
=\;& O\left(nh^{\frac{2(1-\delta)}{1+\delta}}/T\right)=o\left(1\right).
\end{aligned}$$

It follows that $VD_{2nTb}=o\left(1\right)$. Similarly, we have $VD_{2nTc}=o\left(1\right)$. Hence $\operatorname{Var}(D_{2nT,1a1})=o\left(1\right)$ and $D_{2nT,1a1}=o_P\left(1\right)$ by the Chebyshev inequality.

To study $D_{2nT,1a2}$ in (B.2), let $\chi_{i,ts}\equiv\overline{k}^{i}_{h,ts}-E_tE_s[\overline{k}^{i}_{h,ts}]$, and $\chi^{c}_{i,ts}\equiv\chi_{i,ts}-E\left(\chi_{i,ts}\right)$. Noting that $\left|E\left(\chi_{i,ts}\right)\right|\leq Ch^{\frac{-\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}\left(|s-t|\right)$, we can readily show that $D_{2nT,1a2}=\overrightarrow{D}_{2nT,1a2}+o_P\left(1\right)$, where $\overrightarrow{D}_{2nT,1a2}=h\sum_{i=1}^{n}\sum_{r=2}^{T}\overline{c}_{1r}\frac{1}{C_T^2}\sum_{1\leq t<s\leq T}\chi^{c}_{i,ts}$. By construction, $E(\overrightarrow{D}_{2nT,1a2})=0$ and

$$\begin{aligned}
E\left[\left(\overrightarrow{D}_{2nT,1a2}\right)^2\right] &= h^2\sum_{i=1}^{n}\sum_{r_1=2}^{T}\overline{c}_{1r_1}\sum_{r_2=2}^{T}\overline{c}_{1r_2}\frac{1}{\left(C_T^2\right)^2}\sum_{1\leq t_1<t_2\leq T}\sum_{1\leq t_3<t_4\leq T}E\left(\chi^{c}_{i,t_1t_2}\chi^{c}_{i,t_3t_4}\right)\\
&\leq Cnh^{\frac{2}{1+\delta}}/T=o\left(1\right).
\end{aligned}$$

Consequently, $\overrightarrow{D}_{2nT,1a2} = o_P(1)$ and $D_{2nT,1a2} = o_P(1)$. Hence $D_{2nT,1a} = o_P(1)$. Analogously $D_{2nT,1b} = o_P(1)$ and hence $D_{2nT,1} = o_P(1)$.

By the same token we can show that $D_{2nT,2} = o_P(1)$. To show $D_{2nT,3} = o_P(1)$, by (B.1) we can decompose $D_{2nT,3}$ as follows

$$\begin{aligned}
D_{2nT,3} &= \frac{h}{n_1}\sum_{r=1}^{T_1}\frac{T_r}{T_1}\sum_{1\le i\ne j\le n}^{n}\frac{1}{T_r^2}\sum_{t=1}^{T_r}\sum_{s=1}^{T_r}\left(\overline{k}^i_{h,t,t+r} - E[\overline{k}^i_{h,t,t+r}]\right)\left(\overline{k}^j_{h,s,s+r} - E[\overline{k}^j_{h,s,s+r}]\right) \\
&\quad -\frac{h}{n_1}\sum_{r=1}^{T_1}\frac{T_r}{T_1}\sum_{1\le i\ne j\le n}^{n}\frac{1}{T_rC_T^2}\sum_{t_1=1}^{T_r}\sum_{1\le t_2<t_3\le T}\left(\overline{k}^i_{h,t_1,t_1+r} - E[\overline{k}^i_{h,t_1,t_1+r}]\right)\chi_{j,t_2t_3} \\
&\quad -\frac{h}{n_1}\sum_{r=1}^{T_1}\frac{T_r}{T_1}\sum_{1\le i\ne j\le n}^{n}\frac{1}{T_rC_T^2}\sum_{t_1=1}^{T_r}\sum_{1\le t_2<t_3\le T}\chi_{i,t_2t_3}\left(\overline{k}^j_{h,t_1,t_1+r} - E[\overline{k}^j_{h,t_1,t_1+r}]\right) \\
&\quad +\frac{h}{n_1}\sum_{r=1}^{T_1}\frac{T_r}{T_1}\sum_{1\le i\ne j\le n}^{n}\frac{1}{(C_T^2)^2}\sum_{1\le t_1<t_2\le T}\sum_{1\le t_3<t_4\le T}\chi_{i,t_1t_2}\chi_{j,t_3t_4} \\
&\equiv D_{2nT,3a} - D_{2nT,3b} - D_{2nT,3c} + D_{2nT,3d}, \text{ say}
\end{aligned}$$

It suffices to show $D_{2nT,3\xi} = o_P(1)$ for $\xi = a, b, c$, and d. We only sketch the proof of $D_{2nT,3d} = o_P(1)$ since the other cases are simpler. First, note that $D_{2nT,3d} = \overrightarrow{D}_{2nT,3d} + o_P(1)$ by a simple application of Lemma E.1, where $\overrightarrow{D}_{2nT,3d} = \frac{2h}{n_1(C_T^2)^2}\sum_{r=1}^{T_1}\frac{T_r}{T_1}\sum_{1\le i<j\le n}\sum_{1\le t_1<t_2\le T, 1\le t_3<t_4\le T}\chi^c_{i,t_1t_2}\chi^c_{j,t_3t_4}$. Second, noting that $E(\overrightarrow{D}_{2nT,3d}) = 0$, we can write

$$E\left[\left(\overrightarrow{D}_{2nT,3d}\right)^2\right] = \frac{16h^2}{(n_1)^2(C_T^2)^4}\left(\sum_{r=1}^{T_1}\frac{T_r}{T_1}\right)^2\sum_{1\le i<j\le n}\sum_{\substack{1\le t_1<t_2\le T\\ 1\le t_3<t_4\le T}}\sum_{\substack{1\le t_5<t_6\le T,\\ 1\le t_7<t_8\le T}} E\left[\chi^c_{i,t_1t_2}\chi^c_{i,t_3t_4}\right]E\left[\chi^c_{j,t_5t_6}\chi^c_{j,t_7t_8}\right].$$

Now, following the same arguments as used in the proof of (A.13) and applying Lemmas E.1 and E.2 repeatedly, we can show that $E[(\overrightarrow{D}_{2nT,3d})^2] = O(h^{\frac{2(1-\delta)}{1+\delta}}T^4\alpha^{\frac{\delta}{1+\delta}}(m) + h^2m^4) = o(1)$. Hence $\overrightarrow{D}_{2nT,3d} = o_P(1)$. This completes the proof of the corollary.

C Proof of Theorem 3.3

It suffices to show that under H_1,(i) $\widehat{\Gamma}_{nT} = \mu_A + o_P(1)$, (ii) $(nTh)^{-1}\widehat{B}_{nT} = o_P(1)$, and (iii) $\widehat{\sigma}^2_{nT} = \sigma^2_A + o_P(1)$, because then $(nTh)^{-1}\widehat{I}_{nT} = \frac{\widehat{\Gamma}_{nT}}{\widehat{\sigma}_{nT}} - \frac{(nTh)^{-1}\widehat{B}_{nT}}{\widehat{\sigma}_{nT}} \xrightarrow{p} \frac{\mu_A}{\sigma_A} > 0$. Using the expression of $\widehat{\Gamma}_{nT}$ in (2.7), we can easily show that $E[\widehat{\Gamma}_{nT}] = \mu_A + o(1)$ and $\mathrm{Var}(\widehat{\Gamma}_{nT}) = o(1)$. Then (i) follows by the Chebyshev inequality. Next, it is easy to show that $(nTh)^{-1}\widehat{B}_{nT} = O_P\left(T^{-1}\right) = o_P(1)$ and thus (ii) follows. Lastly one can show (iii) by the Chebyshev inequality.

D Proof of Theorem 4.1

Let $\widetilde{\Gamma}_{1nT}$, $\widetilde{\Gamma}_{nT}$, $\widetilde{B}_{nT}$, and $\widetilde{\sigma}^2_{nT}$ be analogously defined as $\widehat{\Gamma}_{1nT}$, $\widehat{\Gamma}_{nT}$, $\widehat{B}_{nT}$, and $\widehat{\sigma}^2_{nT}$ but with $\{u_{it}\}$ being replaced by the residuals $\{\widetilde{u}_{it}\}$ in their definitions. We prove the theorem by showing that: (i) $nTh(\widetilde{\Gamma}_{nT} - \widehat{\Gamma}_{nT}) = o_P(1)$; (ii) $\widetilde{\sigma}^2_{nT} = \widehat{\sigma}^2_{nT} + o_P(1)$; and (iii) $\widetilde{B}_{nT} - \widehat{B}_{nT} = o_P(1)$.

To show (i), let $\widehat{\Delta}_{nT} \equiv \widehat{\Gamma}_{1nT} - \widehat{\Gamma}_{nT}$ and $\widetilde{\Delta}_{nT} \equiv \widetilde{\Gamma}_{1nT} - \widetilde{\Gamma}_{nT}$. By straightforward but tedious calculations, we have $\widehat{\Delta}_{nT} = \widehat{\Delta}_{nT,1} + \widehat{\Delta}_{nT,2}$, where $\widehat{\Delta}_{nT,1} = R\left(\overline{k}\right)\left[\frac{1}{Th^2}(1+\frac{1}{T}) - \frac{2}{nTh}\sum_{i=1}^{n}\int \widehat{f}_i^2\left(u\right)du\right]$, and

$$\begin{aligned}
\widehat{\Delta}_{nT,2} &= \frac{1}{n\left(n-1\right)}\sum_{1\leq i\neq j\leq n}\left(\frac{1}{T^2}-\frac{1}{P_T^2}+\frac{6}{P_T^4}+\frac{2}{P_T^3}\right)\sum_{1\leq t\neq s\leq T}\overline{k}_{h,ts}^{i}\overline{k}_{h,ts}^{j} \\
&\quad +\frac{1}{n\left(n-1\right)}\sum_{1\leq i\neq j\leq n}\left(\frac{2}{P_T^3}-\frac{2}{T^3}+\frac{4}{P_T^4}\right)\sum_{1\leq t\neq s,t\neq r\leq T}\overline{k}_{h,ts}^{i}\overline{k}_{h,tr}^{j} \\
&\quad +\frac{1}{n\left(n-1\right)}\sum_{1\leq i\neq j\leq n}\left(\frac{1}{T^4}-\frac{1}{P_T^4}\right)\sum_{1\leq t\neq s,r\neq q\leq T}\overline{k}_{h,ts}^{i}\overline{k}_{h,rq}^{j}. \qquad \text{(D.1)}
\end{aligned}$$

Similarly, $\widetilde{\Delta}_{nT} = \widetilde{\Delta}_{nT,1} + \widetilde{\Delta}_{nT,2}$, where $\widetilde{\Delta}_{nT,1}$ and $\widetilde{\Delta}_{nT,2}$ are analogously defined as $\widehat{\Delta}_{nT,1}$ and $\widehat{\Delta}_{nT,2}$ but with $\{u_{it}\}$ being replaced by $\{\widetilde{u}_{it}\}$ in their definitions. It follows that $nTh(\widetilde{\Gamma}_{nT} - \widehat{\Gamma}_{nT}) = nTh(\widetilde{\Gamma}_{1nT} - \widehat{\Gamma}_{1nT}) - nTh(\Gamma\widetilde{\Delta}_{nT,1} - \widehat{\Delta}_{nT,1}) - nTh(\widetilde{\Delta}_{nT,2} - \widehat{\Delta}_{nT,2})$. We prove (i) by establishing that: (i1) $nTh(\widetilde{\Gamma}_{1nT} - \widehat{\Gamma}_{1nT}) = o_P\left(1\right)$, (i2) $nTh(\widetilde{\Delta}_{nT,1} - \widehat{\Delta}_{nT,1}) = o_P\left(1\right)$, and (i3) $nTh(\widetilde{\Delta}_{nT,2} - \widehat{\Delta}_{nT,2}) = o_P\left(1\right)$, respectively in Propositions D.1, D.4 and D.5 below.

For (ii), we have

$$\begin{aligned}
\widetilde{\sigma}_{nT}^2 - \widehat{\sigma}_{nT}^2 &= \frac{4R\left(\overline{k}\right)^2}{n\left(n-1\right)T}\sum_{1\leq i\neq j\leq n}\sum_{t=1}^{T}\left[\widetilde{f}_{ij,-t}\left(\widetilde{u}_{it},\widetilde{u}_{jt}\right) - \widehat{f}_{ij,-t}\left(u_{it},u_{jt}\right)\right] \\
&= \frac{4R\left(\overline{k}\right)^2}{n\left(n-1\right)T\left(T-1\right)}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t\neq s\leq T}\left[k_h\left(\widetilde{u}_{it}-\widetilde{u}_{is}\right)k_h\left(\widetilde{u}_{jt}-\widetilde{u}_{js}\right) - k_{h,ts}^{i}k_{h,ts}^{j}\right] \\
&= \frac{4R\left(\overline{k}\right)^2}{n\left(n-1\right)T\left(T-1\right)}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t\neq s\leq T}\{h^{-2}k_{h,ts}^{i}k_{j,ts}'\left(\Delta u_{jt}-\Delta u_{js}\right) \\
&\quad +h^{-2}k_{h,ts}^{j}k_{i,ts}'\left(\Delta u_{it}-\Delta u_{is}\right)\} + o_P\left(1\right),
\end{aligned}$$

where $\widetilde{f}_{ij,-t}$ is analogously defined as $\widehat{f}_{ij,-t}$ with $\{u_{it}\}$ being replaced by $\{\widetilde{u}_{it}\}$, $k_{i,ts}' \equiv k'\left(\left(u_{it}-u_{is}\right)/h\right)$ and $\Delta u_{it} \equiv \widetilde{u}_{it} - u_{it}$. Then following the proof of Lemma D.3 below, one can readily show that the dominant term in the last expression is $o_P\left(1\right)$ by the Chebyshev inequality.

For (iii), letting $\widetilde{E}\left[\varphi_{i,1r}\right]$ be analogously defined as $\widehat{E}\left[\varphi_{i,1r}\right]$ but with $\{u_{it}\}$ being replaced by $\{\widetilde{u}_{it}\}$, we have

$$\begin{aligned}
\widetilde{B}_{nT} - \widehat{B}_{nT} &= \frac{h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}\{\widetilde{E}\left[\varphi_{i,1r}\right]\widetilde{E}\left[\varphi_{j,1r}\right] - \widehat{E}\left[\varphi_{i,1r}\right]\widehat{E}\left[\varphi_{j,1r}\right]\} \\
&= \frac{h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}\widehat{E}\left[\varphi_{i,1r}\right]\left\{\widetilde{E}\left[\varphi_{j,1r}\right] - \widehat{E}\left[\varphi_{j,1r}\right]\right\} \\
&\quad +\frac{h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}\left\{\widetilde{E}\left[\varphi_{i,1r}\right] - \widehat{E}\left[\varphi_{i,1r}\right]\right\}\widehat{E}\left[\varphi_{j,1r}\right] \\
&\quad +\frac{h}{n_1}\sum_{r=2}^{T}\frac{T-r+1}{T-1}\sum_{1\leq i\neq j\leq n}\left\{\widetilde{E}\left[\varphi_{i,1r}\right] - \widehat{E}\left[\varphi_{i,1r}\right]\right\}\left\{\widetilde{E}\left[\varphi_{j,1r}\right] - \widehat{E}\left[\varphi_{j,1r}\right]\right\} \\
&\equiv D_{nT,1} + D_{nT,2} + D_{nT,3}, \text{ say.}
\end{aligned}$$

Analogously to the proofs of Lemmas D.2-D.3 below, we can use the expression $\widetilde{E}\left[\varphi_{j,1r}\right]-\widehat{E}\left[\varphi_{j,1r}\right]$ $=\frac{1}{T-r}\sum_{t=1}^{T-r}\{\overline{k}_h\left(\widetilde{u}_{it}-\widetilde{u}_{i,t+r}\right)-\overline{k}_{h,t,t+r}^{i}\}-\frac{1}{C_T^2}\sum_{1\leq t<s\leq T}\{\overline{k}_h\left(\widetilde{u}_{it}-\widetilde{u}_{is}\right)-\overline{k}_{h,ts}^{i}\}$, the Taylor expansions, and the Chebyshev inequality to show that $D_{nT,s}=o_P\left(1\right)$ for $s=1,2,3$.

Proposition D.1 $nTh(\widetilde{\Gamma}_{1nT}-\widehat{\Gamma}_{1nT})=o_P\left(1\right).$

Proof. Noting that $x^2-y^2=\left(x-y\right)^2+2\left(x-y\right)y$, we have

$$\begin{aligned}\widetilde{\Gamma}_{1nT}-\widehat{\Gamma}_{1nT} &= \frac{1}{n\left(n-1\right)}\sum_{1\leq i\neq j\leq n}\int R_{ij}\left(u,v\right)^2 dudv \\ &\quad+\frac{2}{n\left(n-1\right)}\sum_{1\leq i\neq j\leq n}\int R_{ij}\left(u,v\right)\left[\widehat{f}_{ij}\left(u,v\right)-\widehat{f}_i\left(u\right)\widehat{f}_j\left(v\right)\right]dudv\equiv\Gamma_{nT,1}+\Gamma_{nT,2},\end{aligned}$$

where $R_{ij}\left(u,v\right)\equiv\widetilde{f}_{ij}\left(u,v\right)-\widetilde{f}_i\left(u\right)\widetilde{f}_j\left(v\right)-\widehat{f}_{ij}\left(u,v\right)+\widehat{f}_i\left(u\right)\widehat{f}_j\left(v\right)$, $\widetilde{f}_i$ and $\widetilde{f}_{ij}$ are analogously defined as $\widehat{f}_i$ and $\widehat{f}_{ij}$ with $\{u_{it},u_{jt}\}_{t=1}^T$ being replaced by $\{\widetilde{u}_{it},\widetilde{u}_{jt}\}_{t=1}^T$. Expanding $k_h\left(\widetilde{u}_{it}-u\right)=h^{-1}k\left(\left(\widetilde{u}_{it}-u\right)/h\right)$ in a Taylor series around $u_{it}-u$ with an integral remainder term, we have

$$k_h\left(\widetilde{u}_{it}-u\right)=h^{-1}k_{it}\left(u\right)+h^{-2}k_{it}'\left(u\right)\Delta u_{it}+h^{-2}\Delta u_{it}\int_0^1 k_{it}^{+}\left(u,\lambda\right)d\lambda, \tag{D.2}$$

where $\Delta u_{it}\equiv\widetilde{u}_{it}-u_{it}$, $k_{it}\left(u\right)\equiv k\left(\left(u_{it}-u\right)/h\right)$, $k_{it}'\left(u\right)\equiv k'\left(\left(u_{it}-u\right)/h\right)$, $k_{it}^{+}\left(u,\lambda\right)\equiv k'((u_{it}-u+\lambda\Delta u_{it})/h)-k_{it}'\left(u\right)$, and k' denotes the first order derivative of k. It follows that

$$\begin{aligned}R_{ij}\left(u,v\right) &= \frac{1}{T}\sum_{t=1}^{T}\left[k_h\left(\widetilde{u}_{it}-u\right)k_h\left(\widetilde{u}_{jt}-v\right)-k_h\left(u_{it}-u\right)k_h\left(u_{jt}-v\right)\right] \\ &\quad-\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[k_h\left(\widetilde{u}_{it}-u\right)k_h\left(\widetilde{u}_{js}-v\right)-k_h\left(u_{it}-u\right)k_h\left(u_{js}-v\right)\right]=\sum_{r=1}^{8}R_{rij}\left(u,v\right),\end{aligned}$$

where

$$\begin{aligned}R_{1ij}\left(u,v\right) &\equiv \frac{1}{T^2h^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[k_{jt}\left(v\right)-k_{js}\left(v\right)\right]k_{it}'\left(u\right)\Delta u_{it}, \\ R_{2ij}\left(u,v\right) &\equiv \frac{1}{T^2h^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[k_{it}\left(u\right)-k_{is}\left(u\right)\right]k_{js}'\left(v\right)\Delta u_{js}, \\ R_{3ij}\left(u,v\right) &\equiv \frac{1}{T^2h^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[k_{jt}\left(v\right)-k_{js}\left(v\right)\right]\Delta u_{it}\int_0^1 k_{it}^{+}\left(u,\lambda\right)d\lambda, \\ R_{4ij}\left(u,v\right) &\equiv \frac{1}{T^2h^3}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[k_{it}\left(u\right)-k_{is}\left(u\right)\right]\Delta u_{jt}\int_0^1 k_{jt}^{+}\left(v,\lambda\right)d\lambda, \\ R_{5ij}\left(u,v\right) &\equiv \frac{1}{T^2h^4}\sum_{t=1}^{T}\sum_{s=1}^{T}k_{it}'\left(u\right)\Delta u_{it}\left[k_{jt}'\left(v\right)\Delta u_{jt}-k_{js}'\left(v\right)\Delta u_{js}\right], \\ R_{6ij}\left(u,v\right) &\equiv \frac{1}{T^2h^4}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[k_{jt}'\left(v\right)\Delta u_{jt}-k_{js}'\left(v\right)\Delta u_{js}\right]\Delta u_{it}\int_0^1 k_{it}^{+}\left(u,\lambda\right)d\lambda, \\ R_{7ij}\left(u,v\right) &\equiv \frac{1}{T^2h^4}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[k_{it}'\left(u\right)\Delta u_{it}-k_{is}'\left(u\right)\Delta u_{is}\right]\Delta u_{jt}\int_0^1 k_{jt}^{+}\left(v,\lambda\right)d\lambda, \\ R_{8ij}\left(u,v\right) &\equiv \frac{1}{T^2h^4}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\Delta u_{it}\int_0^1 k_{it}^{+}\left(u,\lambda\right)d\lambda-\Delta u_{is}\int_0^1 k_{is}^{+}\left(u,\lambda\right)d\lambda\right]\Delta u_{jt}\int_0^1 k_{jt}^{+}\left(v,\lambda\right)d\lambda.\end{aligned}$$

By the C_r inequality, it suffices to prove the theorem by showing that:

$$R_{rnT} \equiv \frac{Th}{n_1} \sum_{1 \leq i \neq j \leq n} \int R_{rij}(u,v)^2 \, dudv = o_P(1) \text{ for } r = 1, 2, ..., 8, \tag{D.3}$$

and

$$S_{rnT} \equiv \frac{Th}{n_1} \sum_{1 \leq i \neq j \leq n} \int R_{rij}(u,v) \left[\widehat{f}_{ij}(u,v) - \widehat{f}_i(u) \widehat{f}_j(v)\right] dudv = o_P(1) \text{ for } r = 1, 2, ..., 8. \tag{D.4}$$

We prove (D.3) in Lemma D.2 below and (D.4) in Lemma D.3 below. ■

To proceed, let $\boldsymbol{\tau}((X_{it} - x)/b)$ be the stack of $((X_{it} - x)/b)^{\mathbf{j}}$, $0 \leq |\mathbf{j}| \leq p$, in the lexicographical order such that we can write $\mathbf{S}_{iT}(x) = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\tau}((X_{it} - x)/b) \boldsymbol{\tau}((X_{it} - x)/b)' w_b(X_{it} - x)$. Let $\mathbf{V}_{iT}(x) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{v}_{it}(x) u_{it}$, and $\mathbf{B}_{iT}(x) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{v}_{it}(x) g_i(X_{it}) - g_i(x)$, where $\mathbf{v}_{it}(x) \equiv \boldsymbol{\tau}((X_{it} - x)/b) w_b(X_{it} - x)$. By Masry (1996b), we have $\sup_{x \in \mathcal{X}_i} ||\mathbf{B}_{iT}(x)|| = O_P\left(b^{p+1}\right)$, $\sup_{x \in \mathcal{X}_i} ||\mathbf{V}_{iT}(x)|| = O_P(T^{-1/2} b^{-d/2} \sqrt{\log T})$, and $\sup_{x \in \mathcal{X}_i} ||\mathbf{S}_{iT}(x) - f_i(x)\mathbb{S}|| = O_P(b + T^{-1/2} b^{-d/2} \sqrt{\log T})$, where $\mathbb{S}$ is defined in (4.2). Following Chen, Gao, and Li (2009, Lemma A.1), we can show that

$$\max_{1 \leq i \leq n} \sup_{x \in \mathcal{X}_i} ||\mathbf{S}_{iT}(x) - f_i(x)\mathbb{S}|| = o_P(1). \tag{D.5}$$

Then by the Slutsky lemma and Assumptions A.5(ii) and A.7(i), we have

$$\max_{1 \leq i \leq n} \sup_{x \in \mathcal{X}_i} [\lambda_{\min}(\mathbf{S}_{iT}(x))]^{-1} = \left[\min_{1 \leq i \leq n} \min_{x \in \mathcal{X}_i} f_i(x)\right]^{-1} [\lambda_{\min}(\mathbb{S})]^{-1} + o_P(1). \tag{D.6}$$

By the standard variance and bias decomposition, we have

$$\begin{aligned} u_{it} - \widetilde{u}_{it} &= \widehat{g}_i(X_{it}) - g_i(X_{it}) = e_1'[\mathbf{S}_{iT}(X_{it})]^{-1} \mathbf{V}_{iT}(X_{it}) + e_1'[\mathbf{S}_{iT}(X_{it})]^{-1} \mathbf{B}_{iT}(X_{it})] \\ &\equiv \mathbb{V}_{it} + \mathbb{B}_{it}. \end{aligned} \tag{D.7}$$

Let

$$\eta_{i,ts} \equiv e_1'[\mathbf{S}_{iT}(X_{it})]^{-1} \mathbf{v}_{is}(X_{it}). \tag{D.8}$$

We frequently need to evaluate terms associated with $\eta_{i,ts}$ and $\mathbb{B}_{it}$:

$$\frac{1}{n}\sum_{i=1}^{n} \left(\frac{1}{T^2} \sum_{1 \leq t,s \leq T} |\eta_{i,ts}|\right)^q = O_P(1), \; q = 1, 2, 3, \tag{D.9}$$

$$\frac{1}{n}\sum_{i=1}^{n} \left(\frac{1}{T^3} \sum_{1 \leq t,s,r \leq T} |\eta_{i,ts}\eta_{i,tr}|\right)^q = O_P(1), \; q = 1, 2, \tag{D.10}$$

and

$$\frac{1}{n}\sum_{i=1}^{n} \left(\frac{1}{T}\sum_{t=1}^{T} |\mathbb{B}_{it}|\right)^q = O_P\left(b^{q(p+1)}\right), \; q = 1, 2, 3, 4. \tag{D.11}$$

(D.9) and (D.10) can be proved by using (D.6) and the Markov inequality. For (D.11), we first need to apply the fact that $[\mathbf{S}_{iT}(X_{it})]^{-1} \mathbf{S}_{iT}(X_{it}) = I_N$ and expanding $g_i(X_{is})$ in a Taylor series around X_{it} with an integral remainder to obtain

$$\mathbb{B}_{it} = e_1'[\mathbf{S}_{iT}(X_{it})]^{-1} \frac{1}{T}\sum_{s=1}^{T} \mathbf{v}_{is}(X_{it}) \Delta_{is}(X_{it})$$

where $\Delta_{is}(x) \equiv g_i(X_{is}) - g_i(x) - \sum_{|\mathbf{j}|=1}^{p} \frac{1}{\mathbf{j}!} D^{\mathbf{j}} g_i(x)(X_{is}-x)^{\mathbf{j}} = \sum_{|\mathbf{j}|=p+1} \frac{1}{\mathbf{j}!} D^{\mathbf{j}} g_i(x)(X_{is}-x)^{\mathbf{j}} + (p+1) \sum_{|\mathbf{j}|=p+1} \frac{1}{\mathbf{j}!}(X_{is}-x)^{\mathbf{j}} \int \left[\left(D^{\mathbf{j}} g_i\right)(x+\lambda(X_{is}-x)) - D^{\mathbf{j}} g_i(x)\right](1-\lambda)^p d\lambda$. Then we can apply (D.6), the dominated convergence theorem, and the Markov inequality to show that (D.11) holds. Let $\mathbb{X} \equiv \{X_{it}, i=1,...,n, t=1,...,T\}$ and $E^{\mathbb{X}}(\cdot)$ denote expectation conditional on X.

Lemma D.2 $R_{rnT} \equiv \frac{Th}{n_1} \sum_{1 \le i \ne j \le n} \int R_{rij}(u,v)^2 \, dudv = o_P(1)$ *for* $r = 1, 2, ..., 8$.

Proof. We only prove the lemma for the cases where $r = 1$, 3, 5, 6, and 8 as the other cases can be proved analogously. By (D.7) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
R_{1nT} &\le \frac{2}{n_1 T^3 h^5} \sum_{1 \le i \ne j \le n} \int \left[\sum_{1 \le t \ne s \le T}^{T} [k_{jt}(v) - k_{js}(v)] k'_{it}(u) \mathbb{V}_{it} \right]^2 dudv \\
&\quad + \frac{2}{n_1 T^3 h^5} \sum_{1 \le i \ne j \le n} \int \left[\sum_{1 \le t \ne s \le T}^{T} [k_{jt}(v) - k_{js}(v)] k'_{it}(u) \mathbb{B}_{it} \right]^2 dudv \\
&= \frac{2}{n_1 T^5 h^3} \sum_{1 \le i \ne j \le n} \sum_{1 \le t_1 \ne t_2 \le T} \sum_{t_3=1}^{T} \sum_{1 \le t_4 \ne t_5 \le T} \sum_{t_6=1}^{T} \kappa_{j,t_1 t_2 t_4 t_5} \overline{k'}_{i,t_1 t_4} u_{it_3} u_{it_6} \eta_{i,t_1 t_3} \eta_{i,t_4 t_6} \\
&\quad + \frac{2}{n_1 T^3 h^3} \sum_{1 \le i \ne j \le n} \sum_{1 \le t_1 \ne t_2 \le T} \sum_{1 \le t_3 \ne t_4 \le T} \kappa_{j,t_1 t_2 t_3 t_4} \overline{k'}_{i,t_1 t_3} \mathbb{B}_{it_1} \mathbb{B}_{it_3} \\
&\equiv 2R_{1nT,1} + 2R_{1nT,2},
\end{aligned}$$

where $\overline{k'}$ is a two-fold convolution of k', and

$$\kappa_{j,tsrq} \equiv \overline{k}_{j,tr} - \overline{k}_{j,tq} - \overline{k}_{j,sr} + \overline{k}_{j,sq}. \tag{D.12}$$

Noting that $R_{1nT,r}$, $r = 1, 2$, are nonnegative, it suffices to prove $R_{1nT,r} = o_P(1)$ by showing that $E^{\mathbb{X}}[R_{1nT,r}] = o_P(1)$ by the conditional Markov inequality. For $R_{1nT,1}$, we can easily verify that $E^{\mathbb{X}}[R_{1nT,1}] = \overrightarrow{R}_{1nT,1} + o_P(1)$, where

$$\begin{aligned}
\overrightarrow{R}_{1nT,1} &\equiv \frac{1}{n_1 T^5 h^3} \sum_{1 \le i \ne j \le n} \sum_{t_1,t_2,t_3 \text{ are distinct}} \sum_{t_4,t_5,t_6 \text{ are distinct}} E(\kappa_{j,t_1 t_2 t_4 t_5}) \\
&\quad \times E(\overline{k'}_{i,t_1 t_4} u_{it_3} u_{it_6}) \eta_{i,t_1 t_3} \eta_{i,t_4 t_6}.
\end{aligned} \tag{D.13}$$

We consider two different cases for the time indices $\{t_1, ..., t_6\}$ in the above summation: (a) for at least four different k's in $\{1, ..., 6\}$, $|t_l - t_k| > m$ for all $l \ne k$; (b) all the other remaining cases. We use $\overrightarrow{R}_{1nT,1a}$ and $\overrightarrow{R}_{1nT,1b}$ to denote $\overrightarrow{R}_{1nT,1}$ when the summation over the time indices are restricted to these two cases, respectively. In case (a) we can apply Lemmas E.1 and E.2 repeatedly and show that either $\left|h^{-1} E(\kappa_{j,t_1 t_2 t_4 t_5})\right| \le C h^{\frac{-\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)$ or $\left|h^{-1} E(\overline{k'}_{i,t_1 t_4} u_{it_3} u_{it_6})\right| \le C h^{\frac{-\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)$ must hold. It follows that

$$\begin{aligned}
\overrightarrow{R}_{1nT,1a} &\le \frac{C h^{-\frac{\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)}{n_1 T^5 h} \sum_{1 \le i \ne j \le n} \sum_{t_1,t_2,t_3 \text{ are distinct}} \sum_{t_4,t_5,t_6 \text{ are distinct}} \left|\eta_{i,t_1 t_3} \eta_{i,t_4 t_6}\right| \\
&\le C n T h^{-\frac{1+2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) \left[n^{-1} \sum_{i=1}^{n} \left(T^{-2} \sum_{1 \le t,s \le T} \left|\eta_{i,ts}\right| \right)^2 \right] \\
&= O_P\left(n T h^{-\frac{1+2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) \right) = o_P(1),
\end{aligned}$$

where we have used the result in (D.9). In case (b) noting that we have $O(n^2T^4m^2)$ terms in the summation in (D.13) and $h^{-1}E\left(\kappa_{j,t_1t_2t_4t_5}\right)$ and $h^{-3}E(\overline{k'}_{i,t_1t_4}u_{it_3}u_{it_6})$ are bounded uniformly in all indices (as $\overline{k'}$ behaves like a second order kernel by Lemma E.3), we can apply (D.9) and show that $\overrightarrow{R}_{1nT,1b} = O_P\left(nhm^2/T\right) = o_P\left(1\right)$.

For $R_{1nT,2}$, we can show that $E^{\mathbb{X}}[R_{1nT,2}] = \overrightarrow{R}_{1nT,2} + o_P\left(1\right)$, where

$$\overrightarrow{R}_{1nT,2} = \frac{1}{n_1T^3h^3}\sum_{1\leq i\neq j\leq n}\sum_{t_1,t_2,t_3,t_4 \text{ are distinct}} E\left(\kappa_{j,t_1t_2t_3t_4}\right)E(\overline{k'}_{i,t_1t_3})\mathbb{B}_{it_1}\mathbb{B}_{it_3}.$$

We consider two cases for the time indices $\{t_1,...,t_4\}$ in the above summation: (a) for all k's in $\{1,...,4\}$, $|t_l - t_k| > m$ for all $l \neq k$; (b) all the other remaining cases. We use $\overrightarrow{R}_{1nT,2a}$, and $\overrightarrow{R}_{1nT,2b}$ to denote $\overrightarrow{R}_{1nT,2}$ when the summation over the time indices are restricted to these cases, respectively. In case (a) we can use the fact that $\left|h^{-1}E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right| \leq Ch^{\frac{-\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}\left(m\right)$, the fact that $h^{-1}E(\overline{k'}_{i,t_1t_3}) \leq Ch^2$ (by Lemma E.3) and (D.11) to obtain $\overrightarrow{R}_{1nT,2a} = O_P(nTh^{\frac{1}{1+\delta}}b^{2(p+1)}\alpha^{\frac{\delta}{1+\delta}}\left(m\right)) = o_P\left(1\right)$. In case (b), note that $E\left(\kappa_{j,t_1t_2t_3t_4}\right)$ cannot be bounded by a term proportional to $h^{\frac{-\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}\left(m\right)$ in the cases where one of the index-pair $\{(t_1,t_3), (t_1,t_4), (t_2,t_3), (t_2,t_4)\}$ has elements that do not fall from each other at least m-apart. But we can apply the fact that $\left|h^{-1}E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right| \leq C$, $|h^{-1}E(\overline{k'}_{i,t_1t_3})| \leq Ch^2$, and (D.11) to obtain $\overrightarrow{R}_{1nT,2b} = O_P(nmhb^{2(p+1)}) = o_P\left(1\right)$. Hence we have $E^{\mathbb{X}}[R_{1nT,2}] = o_P\left(1\right)$. Consequently, $R_{1nT} = o_P\left(1\right)$.

For R_{3nT}, write

$$\begin{aligned} R_{3nT} &= \frac{1}{n_1T^3h^4}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\kappa_{j,t_1t_2t_3t_4}\Delta u_{it_1}\Delta u_{it_3} \\ &\quad \times\int\int_0^1\int_0^1 k^+_{it_1}\left(u,\lambda_1\right)d\lambda_1 k^+_{it_3}\left(u,\lambda_2\right)d\lambda_2 du. \end{aligned}$$

As argued by Hansen (2008, pp.740-741), under Assumption A.8 there exists an integrable function k^* such that

$$\left|k^+_{it}\left(u,\lambda\right)\right| = \left|k'\left(\left(u_{it}-u+\lambda\Delta u_{it}\right)/h\right) - k'_{it}\left(u\right)\right| \leq \lambda h^{-1}\left|\Delta u_{it}\right|k^*\left(\left(u_{it}-u\right)/h\right). \tag{D.14}$$

It follows that

$$\begin{aligned} E^{\mathbb{X}}\left(R_{3nT}\right) &\leq \frac{1}{4n_1T^3h^5}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\left|E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right|E^{\mathbb{X}}\{\overline{k^*}_{i,t_1t_3}\left(\Delta u_{it_1}\right)^2\left(\Delta u_{it_3}\right)^2\} \\ &\leq \frac{1}{n_1T^3h^5}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\left|E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right|E^{\mathbb{X}}\{\overline{k^*}_{i,t_1t_3}[\mathbb{V}^2_{it_1}\mathbb{V}^2_{it_3} + \mathbb{B}^2_{it_1}\mathbb{B}^2_{it_3} \\ &\quad +\mathbb{V}^2_{it_1}\mathbb{B}^2_{it_3} + \mathbb{B}^2_{it_1}\mathbb{V}^2_{it_3}]\} \\ &\equiv ER_{3nT,1} + ER_{3nT,2} + ER_{3nT,3} + ER_{3nT,4}, \end{aligned}$$

where $\overline{k^*}_{i,ts} \equiv \overline{k^*}\left(\left(u_{it}-u_{is}\right)/h\right)$ and $\overline{k^*}$ is the two-fold convolution of k^*. It is easy to show that $ER_{3nT,1} = \overrightarrow{ER}_{3nT,1} + o_P\left(1\right)$, where $\overrightarrow{ER}_{3nT,1} = \frac{1}{n_1T^7h^5}\sum_{1\leq i\neq j\leq n}\sum_{t_1,...,t_8 \text{ are all distinct}}\left|E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right|$ $E(\overline{k^*}_{i,t_1t_3}u_{it_5}u_{it_6}u_{it_7}u_{it_8})\,\eta_{i,t_1t_5}\eta_{i,t_1t_6}\eta_{i,t_3t_7}\eta_{i,t_3t_8}$. We consider two cases for the time indices $\{t_1,...,t_8\}$ in the last summation: (a) for at least 4 distinct k's in $\{1,...,8\}$, $|t_l - t_k| > m$ for all $l \neq k$; (b) all the other remaining cases. We use $ER_{3nT,1a}$, and $ER_{3nT,1b}$ to denote $ER_{3nT,1}$ when the summation over

the time indices are restricted to these cases, respectively. In case (a), we have $\left|h^{-1}E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right| \leq Ch^{\frac{-\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)$ or $\left|h^{-1}E(\overline{k^*}_{i,t_1t_3}u_{it_5}u_{it_6}u_{it_7}u_{it_8})\right| \leq Ch^{\frac{-\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)$, and thus by (D.10)

$$\begin{aligned}
|ER_{3nT,1a}| &\leq \frac{CTh^{\frac{-\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)}{h^3}\sum_{i=1}^{n}\left\{\frac{1}{T^3}\sum_{1\leq t_1,t_5,t_6\leq T}\left|\eta_{i,t_1t_5}\eta_{i,t_1t_6}\right|\right\}^2 \\
&\leq O_P\left(nTh^{-3-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)\right) = o_P(1).
\end{aligned}$$

In case (b), noting that $h^{-1}\left|E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right| \leq C$ and $h^{-1}\left|E(\overline{k^*}_{i,t_1t_3}u_{it_5}u_{it_6}u_{it_7}u_{it_8})\right| \leq C$, we have by (D.10)

$$|ER_{3nT,1b}| \leq \frac{m^3}{T^2h^3}\sum_{j=1}^{n}\left\{\frac{1}{T^3}\sum_{1\leq t_1,t_5,t_6\leq T}\left|\eta_{j,t_1t_5}\eta_{j,t_1t_6}\right|\right\}^2 = O_P\left(nm^3h^{-3}/T^2\right) = o_P(1).$$

Consequently $ER_{3nT,1} = o_P(1)$. Next, it is easy to show that $ER_{3nT,2} = \overrightarrow{ER}_{3nT,2} + o_P(1)$, where $\overrightarrow{ER}_{3nT,2} = \frac{1}{n_1T^3h^5}\sum_{1\leq i\neq j\leq n}\sum_{t_1,\ldots,t_4 \text{ are all distinct}}\left|E\left(\kappa_{j,t_1t_2t_3t_4}\right)\right| E(\overline{k^*}_{i,t_1t_3})\mathbb{B}^2_{it_1}\mathbb{B}^2_{it_3}$. Then we can show that

$$\overrightarrow{ER}_{3nT,2} = O_P\left(nTh^{-3-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)\,b^{4(p+1)} + nTh^{-3}b^{4(p+1)}\right) = o_P(1).$$

Hence $ER_{3nT,2} = o_P(1)$. Similarly, we can show that $ER_{3nT,r} = o_P(1)$ for $r = 3, 4$.

For R_{5nT}, note that

$$\begin{aligned}
R_{5nT} &\leq 2n_1^{-1}Th\sum_{1\leq i\neq j\leq n}\int\int\left[\frac{1}{Th^4}\sum_{t=1}^{T}k'_{it}(u)\,\Delta u_{it}k'_{jt}(v)\,\Delta u_{jt}\right]^2 dudv \\
&\quad +2n_1^{-1}Th\sum_{1\leq i\neq j\leq n}\int\int\left[\frac{1}{T^2h^4}\sum_{1\leq t,s\leq T}k'_{it}(u)\,\Delta u_{it}k'_{js}(v)\,\Delta u_{js}\right]^2 dudv \equiv R_{5nT,1} + R_{5nT,2}.
\end{aligned}$$

By (D.9) and (D.11) and the fact that $\overline{k'}$ behaves like second order kernel (see Lemma E.3), we can show that

$$\begin{aligned}
E^{\mathbb{X}}(R_{5nT,1}) &= \frac{2}{n_1Th^5}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1,t_2\leq T}E^{\mathbb{X}}\left[\overline{k'}_{i,t_1t_2}\Delta u_{it_1}\Delta u_{it_2}\right]E^{\mathbb{X}}\left[\overline{k'}_{j,t_1t_2}\Delta u_{jt_1}\Delta u_{jt_2}\right] \\
&= O_P\left(nTh\left(T^{-2}b^{-2d} + b^{4(p+1)}\right)\right) = o_P(1).
\end{aligned}$$

It follows that $R_{5nT,1} = o_P(1)$. By the same token, $R_{5nT,2} = o_P(1)$. Consequently $R_{5nT} = o_P(1)$.

For R_{6nT}, write $R_{6ij}(u,v) = \frac{1}{Th^4}\sum_{t=1}^{T}k'_{jt}(v)\,\Delta u_{jt}\Delta u_{it}\int_0^1 k^+_{it}(u,\lambda)\,d\lambda - \frac{1}{T^2h^4}\sum_{t=1}^{T}\sum_{s=1}^{T}k'_{js}(v)\,\Delta u_{js}\Delta u_{it}\int_0^1 k^+_{it}(u,\lambda)\,d\lambda \equiv R_{6ij,1}(u,v) - R_{6ij,2}(u,v)$. Define $R_{6nT,1}$ and $R_{6nT,2}$ analogously as R_{6nT} but with $R_{6ij}(u,v)$ being replaced by $R_{6ij,1}(u,v)$ and $R_{6ij,2}(u,v)$, respectively. Then

$$R_{6nT,1} = \frac{1}{n_1Th^6}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t,s\leq T}\overline{k'}_{j,ts}\Delta u_{jt}\Delta u_{js}\Delta u_{it}\Delta u_{is}\int\int_0^1 k^+_{it}(u,\lambda_1)\,d\lambda_1\int_0^1 k^+_{is}(u,\lambda_2)\,d\lambda_2du$$

Using (D.9) and (D.11), we have

$$\begin{aligned}
E^{\mathbb{X}}\left(R_{6nT,1}\right) &= \frac{1}{n_1 T h^6}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t,s\leq T} E^{\mathbb{X}}\left[\overline{k'}_{j,ts}\Delta u_{jt}\Delta u_{js}\right] \\
&\quad \times E^{\mathbb{X}}\left[\Delta u_{it}\Delta u_{is}\int\int_0^1 k_{it}^{+}\left(u,\lambda_1\right)d\lambda_1\int_0^1 k_{is}^{+}\left(u,\lambda_2\right)d\lambda_2 du\right] \\
&\leq \frac{1}{4n_1 T h^7}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t,s\leq T}\left|E^{\mathbb{X}}\left[\overline{k'}_{j,ts}\Delta u_{jt}\Delta u_{js}\right]\right| E^{\mathbb{X}}\{\overline{k^*}_{i,ts}\left(\Delta u_{it}\Delta u_{is}\right)^2\} \\
&= O_P\left(nTh^{-3}\left(T^{-3}b^{-3d}+b^{6(p+1)}\right)\right)=o_P\left(1\right).
\end{aligned}$$

Similarly, we can show that $E^{\mathbb{X}}\left(R_{8nT,1}\right)=O_P\left(nTh^{-7}\left(T^{-4}b^{-4d}+b^{8(p+1)}\right)\right)=o_P\left(1\right)$. ■

Lemma D.3 $S_{rnT}\equiv\frac{Th}{n_1}\sum_{1\leq i\neq j\leq n}\int R_{rij}\left(u,v\right)\left[\widehat{f}_{ij}\left(u,v\right)-\widehat{f}_i\left(u\right)\widehat{f}_j\left(v\right)\right]dudv=o_P\left(1\right)$ *for* $r=1,2,...,8$.

Proof. We only prove the lemma for the cases where $r=1$, 3, and 5 as the other cases can be proved analogously. Decompose

$$\begin{aligned}
S_{1nT} &= \frac{1}{T^3 n_1 h^4}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\int\int\left[k_{jt_1}\left(v\right)-k_{jt_2}\left(v\right)\right]\left[k_{jt_3}\left(v\right)-k_{jt_4}\left(v\right)\right] \\
&\quad \times k'_{it_1}\left(u\right)k_{it_3}\left(u\right)\Delta u_{it_1}dudv \\
&= \frac{1}{T^3 n_1 h^2}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\kappa_{j,t_1t_2t_3t_4}k_{i,t_1t_3}^{+}\Delta u_{it_1} \\
&= \frac{1}{T^4 n_1 h^2}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\sum_{t_5=1}^{T}\kappa_{j,t_1t_2t_3t_4}k_{i,t_1t_3}^{+}u_{it_5}\eta_{i,t_1t_5} \\
&\quad +\frac{1}{T^3 n_1 h^2}\sum_{1\leq i\neq j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\kappa_{j,t_1t_2t_3t_4}k_{i,t_1t_3}^{+}\mathbb{B}_{it_1} \\
&\equiv S_{1nT,1}+S_{1nT,2},
\end{aligned}$$

where $k_{i,ts}^{+}\equiv k^{+}\left(\left(u_{it}-u_{is}\right)/h\right)\equiv h^{-1}\int k'_{it}\left(u\right)k_{is}\left(u\right)du$, and $\kappa_{j,tsrq}$ is defined in (D.12). To show $S_{1nT,1}=o_P\left(1\right)$, we can first show that $S_{1nT,1}=\overrightarrow{S}_{1nT,1}+o_P\left(1\right)$, where $\overrightarrow{S}_{1nT,1}$ is analogously defined as $S_{1nT,1}$ but with all distinct time indices inside the summation. Second, we can decompose $\overrightarrow{S}_{1nT,1}$ as $\overrightarrow{S}_{1nT,11}+\overrightarrow{S}_{1nT,12}$ where $\overrightarrow{S}_{1nT,11}$ is analogously defined as $\overrightarrow{S}_{1nT,1}$ but with only $i<j$ terms in the summation and $\overrightarrow{S}_{1nT,12}\equiv\overrightarrow{S}_{1nT,1}-\overrightarrow{S}_{1nT,11}$. Let $e_{i,tsr}\equiv k_{i,ts}^{+}u_{ir}$, $e_{i,tsr}^{c}\equiv e_{i,tsr}-E\left(e_{i,tsr}\right)$, and $\kappa_{j,t_1t_2t_3t_4}^{c}\equiv\kappa_{j,t_1t_2t_3t_4}-E(\kappa_{j,t_1t_2t_3t_4})$. Then we can decompose $\overrightarrow{S}_{1nT,11}$ as follows

$$\begin{aligned}
\overrightarrow{S}_{1nT,11} &= \frac{1}{T^4 n_1 h^2}\sum_{1\leq i\neq j\leq n}\sum_{t_1,\ldots,t_5 \text{ are all distinct}}\kappa_{j,t_1t_2t_3t_4}e_{i,t_1t_3t_5}\eta_{i,t_1t_5}+o_P\left(1\right) \\
&= \frac{1}{T^4 n_1 h^2}\sum_{1\leq i\neq j\leq n}\sum_{t_1,\ldots,t_5 \text{ are all distinct}}\{\kappa_{j,t_1t_2t_3t_4}^{c}e_{i,t_1t_3t_5}^{c}\eta_{i,t_1t_5}+\kappa_{j,t_1t_2t_3t_4}^{c}E\left(e_{i,t_1t_3t_5}\right)\eta_{i,t_1t_5} \\
&\quad +E(\kappa_{j,t_1t_2t_3t_4})e_{i,t_1t_3t_5}^{c}\eta_{i,t_1t_5}+E(\kappa_{j,t_1t_2t_3t_4})E\left(e_{i,t_1t_3t_5}\right)\eta_{i,t_1t_5}\}+o_P\left(1\right) \\
&\equiv \overrightarrow{S}_{1nT,111}+\overrightarrow{S}_{1nT,112}+\overrightarrow{S}_{1nT,113}+\overrightarrow{S}_{1nT,114}+o_P\left(1\right).
\end{aligned}$$

For $\overrightarrow{S}_{1nT,111}$, we have

$$\begin{aligned}
E^{\mathbb{X}}\left[\left(\overrightarrow{S}_{1nT,111}\right)^2\right] &= \frac{1}{T^8 n_1^2 h^4}\sum_{1\leq i<j\leq n}\sum_{t_1,\ldots,t_5 \text{ are all distinct}}\sum_{t_6,\ldots,t_{10} \text{ are all distinct}}\eta_{i,t_1t_5}\eta_{i,t_6t_{10}} \\
&\quad \times E\left[e_{i,t_1t_3t_5}^{c}e_{i,t_6t_8t_{10}}^{c}\right]E[\kappa_{j,t_1t_2t_3t_4}^{c}\kappa_{j,t_6t_7t_8t_9}^{c}].
\end{aligned}$$

We consider two cases for the time indices $\{t_1, ..., t_{10}\}$: (a) for at least six different k's, $|t_l - t_k| > m$ for all $l \neq k$; (b) all the other remaining cases. We use $ES_{1,111a}$ and $ES_{1,111b}$ to denote the summation corresponding to these two cases, respectively. In the first case, $ES_{1,111a} \leq CT^2 h^{\frac{-2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m) \sum_{i=1}^n \{T^{-2} \sum_{1\leq t\neq s\leq T} |\eta_{i,ts}|\}^2 = O_P(T^2 h^{\frac{-2\delta}{1+\delta}} \alpha^{\frac{\delta}{1+\delta}}(m)) = o_P(1)$. In the second case,

$$ES_{1,111b} \leq CT^{-1} n_1^{-1} m^3 \sum_{i=1}^n \left(T^{-2} \sum_{1\leq t\neq s\leq T} |\eta_{i,ts}| \right)^2 = O_P\left(m^3/T\right) = o_P(1).$$

It follows that $\overrightarrow{S}_{1nT,111} = o_P(1)$. Analogously, we can show that $\overrightarrow{S}_{1nT,11r} = o_P(1)$ for $r = 2, 3, 4$. So $\overrightarrow{S}_{1nT,11} = o_P(1)$. Also $\overrightarrow{S}_{1nT,12} = o_P(1)$ by the same argument. Thus $\overrightarrow{S}_{1nT,1} = o_P(1)$ and $S_{1nT,1} = o_P(1)$. Analogously, we can show that $S_{1nT,2} = o_P(1)$. Consequently, $S_{1nT} = o_P(1)$.

For S_{3nT}, we have

$$\begin{aligned}
S_{3nT} &= \frac{1}{n_1 T^3 h^4} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t_1\neq t_2\leq T} \sum_{1\leq t_3\neq t_4\leq T} \int [k_{jt_1}(v) - k_{jt_2}(v)] [k_{jt_3}(v) - k_{jt_4}(v)] dv \Delta u_{it_1} \\
&\quad \times \int k_{it_3}(u) \int_0^1 k_{it_1}^+(u,\lambda) d\lambda du \\
&= \frac{1}{n_1 T^3 h^3} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t_1\neq t_2\leq T} \sum_{1\leq t_3\neq t_4\leq T} \kappa_{j,t_1t_2t_3t_4} \Delta u_{it_1} \int k_{it_3}(u) \int_0^1 k_{it_1}^+(u,\lambda) d\lambda du \\
&= \frac{1}{n_1 T^3 h^3} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t_1,t_2\leq T} \sum_{1\leq t_3,t_4\leq T} E[\kappa_{j,t_1t_2t_3t_4}] \Delta u_{it_1} \int k_{it_3}(u) \int_0^1 k_{it_1}^+(u,\lambda) d\lambda du \\
&\quad + \frac{1}{n_1 T^3 h^3} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t_1,t_2\leq T} \sum_{1\leq t_3,t_4\leq T} \kappa_{j,t_1t_2t_3t_4}^c \Delta u_{it_1} \int k_{it_3}(u) \int_0^1 k_{it_1}^+(u,\lambda) d\lambda du \\
&\equiv S_{3nT,1} + S_{3nT,2}.
\end{aligned}$$

Noting that $h^{-1}\kappa_{j,t_1t_2t_3t_4} = \varphi_{j,t_1t_3} - \varphi_{j,t_1t_4} - \varphi_{j,t_2t_3} + \varphi_{j,t_2t_4}$, we can decompose $S_{3nT,r} = S_{3nT,r1} - S_{3nT,r2} - S_{3nT,r3} + S_{3nT,r4}$, where $S_{3nT,r1}$, $S_{3nT,r2}$, $S_{3nT,r3}$, and $S_{3nT,r4}$ are defined analogously as $S_{3nT,r}$ with $E[\kappa_{j,t_1t_2t_3t_4}]$ (for $r = 1$) or $\kappa_{j,t_1t_2t_3t_4}^c$ (for $r = 2$) being respectively replaced by $hE[\varphi_{j,t_1t_3}]$, $hE[\varphi_{j,t_1t_4}]$, $hE[\varphi_{j,t_2t_3}]$, and $hE[\varphi_{j,t_2t_4}]$ (for $r = 1$), or by $h\varphi_{j,t_1t_3}^c$, $h\varphi_{j,t_1t_4}^c$, $h\varphi_{j,t_2t_3}^c$, and $h\varphi_{j,t_2t_4}^c$ (for $r = 2$). WLOG we prove $S_{3nT,r} = o_P(1)$ by showing that $S_{3nT,r1} = o_P(1)$ for $r = 1, 2$. For $S_{3nT,11}$, noting that

$$\left| \int k_{is}(u) \int_0^1 k_{it}^+(u,\lambda) d\lambda du \right| \leq \frac{1}{2} |\Delta u_{it}| h^{-1} \int |k((u_{is} - u)/h)| k^*((u_{it} - u)/h) du = \frac{1}{2} |\Delta u_{it}| k_{i,ts}^{\ddagger},$$

where $k_{i,ts}^{\ddagger} \equiv k^{\ddagger}((u_{it} - u_{is})/h)$ and $k^{\ddagger}(u) \equiv \int k^*(u - v) |k(v)| dv$, we have

$$\begin{aligned}
|S_{3nT,11}| &\leq \frac{1}{2n_1 Th} \sum_{1\leq i\neq j\leq n} \sum_{1\leq t,s\leq T} \left|E(\varphi_{j,ts})\right| (\Delta u_{it})^2 k_{i,ts}^{\ddagger} \\
&= \frac{1}{2n_1 Th} \left\{ \sum_{1\leq i\neq j\leq n} \sum_{|t-s|\geq m} + \sum_{1\leq i\neq j\leq n} \sum_{0<|t-s|<m} \right\} \left|E(\varphi_{j,ts})\right| (\Delta u_{it})^2 k_{i,ts}^{\ddagger} \\
&\equiv S_{3nT,11a} + S_{3nT,11b}.
\end{aligned}$$

By the fact that $\left|E(\varphi_{j,ts})\right| \leq Ch^{-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(|t-s|)$ (see (A.1)), we have

$$\begin{aligned} S_{3nT,11a} &\leq Ch^{-1-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)\sum_{i=1}^{n}\sum_{t=1}^{T}(\Delta u_{it})^2 k_{i,ts}^{\ddagger} \\ &= h^{-1-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)\, O_P\left(nT\left(T^{-1}b^{-d}+b^{2(p+1)}\right)\right) = o_P(1). \end{aligned}$$

For $S_{3nT,11b}$ we can apply (D.9) and (D.11) and the Markov inequality to show that $S_{3nT,11b} = O_P(nm(T^{-1}b^{-d}+b^{2(p+1)})) = o_P(1)$. It follows that $S_{3nT,11} = o_P(1)$.

For $S_{3nT,21}$, write $S_{3nT,21} = \frac{1}{n_1Th^2}\{\sum_{1\leq i<j\leq n}+\sum_{1\leq j<i\leq n}\}\sum_{1\leq t_1,t_2\leq T}\varphi_{j,t_1t_2}^c\Delta u_{it_1}\int k_{it_2}(u)\int_0^1 k_{it_1}^{+}(u,\lambda)d\lambda du \equiv S_{3nT,211}+S_{3nT,212}$. Note that $E^{\mathbb{X}}[S_{3nT,211}]=0$, and $E^{\mathbb{X}}\left[(S_{3nT,211})^2\right] = S_3 + o_P(1)$, where

$$\begin{aligned} S_3 &\equiv \frac{1}{(n_1Th^2)^2}\sum_{1\leq i_1\neq i_2<j\leq n}\sum_{1\leq t_1,t_2\leq T}\sum_{1\leq t_3,t_4\leq T}E\left(\varphi_{j,t_1t_2}^c\varphi_{j,t_3t_4}^c\right) \\ &\quad\times E^{\mathbb{X}}\left[\Delta u_{i_1t_1}\int k_{i_1t_2}(u)\int_0^1 k_{i_1t_1}^{+}(u,\lambda)\,d\lambda du\right]E^{\mathbb{X}}\left[\Delta u_{i_2t_3}\int k_{i_2t_4}(u)\int_0^1 k_{i_2t_3}^{+}(u,\lambda)\,d\lambda du\right] \\ &\leq \frac{1}{4(n_1Th^2)^2}\sum_{1\leq i_1\neq i_2<j\leq n}\sum_{1\leq t_1,t_2\leq T}\sum_{1\leq t_3,t_4\leq T}\left|E\{\varphi_{j,t_1t_2}^c\varphi_{j,t_3t_4}^c\}\right|E^{\mathbb{X}}\left[(\Delta u_{i_1t_1})^2 k_{i_1,t_1t_2}^{\ddagger}\right] \\ &\quad\times E^{\mathbb{X}}\left[(\Delta u_{i_2t_3})^2 k_{i_2,t_3t_4}^{\ddagger}\right]. \end{aligned}$$

It is easy to show that the dominant term on the r.h.s. of the last equation is given by $\overline{S}_3 = \left(n_1Th^2\right)^{-2}\sum_{1\leq i_1\neq i_2<j\leq n}\sum_{t_1,t_2,t_3,t_4 \text{ are all distinct}}\left|E(\varphi_{j,t_1t_2}^c\varphi_{j,t_3t_4}^c)\right|E^{\mathbb{X}}[(\Delta u_{i_1t_1})^2k_{i_1,t_1t_2}^{+}]E^{\mathbb{X}}[(\Delta u_{i_2t_3})^2\,k_{i_2,t_3t_4}^{+}]$. We consider two cases for the time indices $\{t_1,...,t_4\}$ in the last summation: (a) there exists at least an integer $k\in\{1,...,4\}$, $|t_l-t_k|>m$ for all $l\neq k$; (b) all the other remaining cases. We use $\overline{S}_{3a}$, and $\overline{S}_{3b}$ to denote $\overline{S}_3$ when the summation over the time indices are restricted to these cases, respectively. In case (a), WLOG we assume that t_1 lies at least m-apart from $\{t_2,t_3,t_4\}$. Then by Lemma E.1, $E\{\varphi_{j,t_1t_2}^c\varphi_{j,t_3t_4}^c\} \leq \left|E\{E_{t_1}(\varphi_{j,t_1t_2}^c)\varphi_{j,t_3t_4}^c\}\right| + Ch^{\frac{-2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m) = Ch^{\frac{-2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)$ as $E_{t_1}(\varphi_{j,t_1t_2}^c)$ is nonrandom.

$$\begin{aligned} \overline{S}_{3a} &\leq \frac{Ch^{\frac{-2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)}{n_1T^2h^2}\left\{\sum_{i=1}^{n}\sum_{1\leq t_1\neq t_2\leq T}E^{\mathbb{X}}\left\{(\Delta u_{it_1})^2h^{-1}k_{i,t_1t_2}^{\ddagger}\right\}\right\}^2 \\ &= nT^2h^{-2-\frac{2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)\,O_P\left(T^{-2}b^{-2d}+b^{4(p+1)}\right) = o_P(1). \end{aligned}$$

In case (b), noting that the total number of terms in the summation is of order $O\left(n^3T^2m^2\right)$, we can easily obtain $\left|\overline{S}_{3b}\right| = O\left(nm^2h^{-2}\right)O_P\left(T^{-2}b^{-2d}+b^{4(p+1)}\right) = O_P\left(nm^2h^{-2}(T^{-2}b^{-2d}+b^{4(p+1)})\right) = o_P(1)$. Consequently $S_3 = o_P(1)$ and $S_{3nT,211} = o_P(1)$ by the conditional Chebyshev inequality.

Next we study S_{5nT}. Write $S_{5nT} = \frac{Th}{n_1}\left(\sum_{1\leq i<j\leq n}+\sum_{1\leq j<i\leq n}\right)\int\int R_{5ij}(u,v)\left[\widehat{f}_{ij}(u,v)-\widehat{f}_i(u)\,\widehat{f}_j(v)\right] dudv \equiv S_{5nT,1}+S_{5nT,2}$. It suffices to show that $S_{5nT,1}=o_P(1)$ and $S_{5nT,2}=o_P(1)$. We only prove

the former claim as the latter one can be proved analogously. It is easy to show that

$$\begin{aligned} S_{5nT,1} &= \frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\int\int\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T} k'_{it_1}(u)\,\Delta u_{it_1}[k'_{jt_1}(v)\,\Delta u_{jt_1}-k'_{jt_2}(v)\,\Delta u_{jt_2}] \\ &\quad \times k_{jt_3}(v)\,[k_{it_3}(u)-k_{it_4}(u)]\,dudv \\ &= \frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4})\Delta u_{it_1}(k^{\dagger}_{j,t_1t_3}\Delta u_{jt_1}-k^{\dagger}_{j,t_2t_3}\Delta u_{jt_2}) \\ &= \overrightarrow{S}_{5nT,1}+o_P(1), \end{aligned}$$

where $k^{\dagger}_{i,ts}\equiv k^{\dagger}((u_{it}-u_{is})/h)$, $k^{\dagger}(u)\equiv\int k'(u-v)\,k(v)\,dv$,

$$\overrightarrow{S}_{5nT,1}=\frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{t_1\ldots t_4 \text{ are all distinct}}(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4})\Delta u_{it_1}(k^{\dagger}_{j,t_1t_3}\Delta u_{jt_1}-k^{\dagger}_{j,t_2t_3}\Delta u_{jt_2}),$$

and the $o_P(1)$ terms arises when the cardinality of the set $\{t_1,t_2,t_3,t_4\}$ is 3 or 2. In particular, by the standard bias-variance decomposition (for Δu_{it_1} and Δu_{jt_2}) and the conditional Chebyshev inequality, we can show that

$$\begin{aligned} &\frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{\substack{t_1\neq t_2,t_3\neq t_4\\ \#\{t_1\ldots t_4\}=3 \text{ or } 2}}(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4})\Delta u_{it_1}(k^{\dagger}_{j,t_1t_3}\Delta u_{jt_1}-k^{\dagger}_{j,t_2t_3}\Delta u_{jt_2}) \\ = \; & O_P\left(h^{-5}\left(T^{-1}+T^{-3/2}b^{-d}\right)+nhb^{2(p+1)}\right)=o_P(1). \end{aligned}$$

Decompose $\overrightarrow{S}_{5nT,1}=\overrightarrow{S}_{5nT,11}+\overrightarrow{S}_{5nT,12}$, where

$$\begin{aligned} \overrightarrow{S}_{5nT,11} &\equiv \frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{t_1\ldots t_4 \text{ are all distinct}}(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4})\Delta u_{it_1}(k^{\dagger}_{j,t_1t_3}-k^{\dagger}_{j,t_2t_3})\Delta u_{jt_1}, \text{ and} \\ \overrightarrow{S}_{5nT,12} &\equiv \frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{t_1\ldots t_4 \text{ are all distinct}}(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4})\Delta u_{it_1}k^{\dagger}_{j,t_2t_3}(\Delta u_{jt_1}-\Delta u_{jt_2}). \end{aligned}$$

We prove $\overrightarrow{S}_{5nT,1}=o_P(1)$ by showing that $\overrightarrow{S}_{5nT,11}=o_P(1)$ and $\overrightarrow{S}_{5nT,12}=o_P(1)$. We only prove the former claim as the latter can be proved analogously. Let

$$\mathcal{S}(A,B)\equiv\frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{1\leq t_1\neq t_2\leq T}\sum_{1\leq t_3\neq t_4\leq T}\left(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4}\right)A_{it_1}\left(k^{\dagger}_{j,t_1t_3}-k^{\dagger}_{j,t_2t_3}\right)B_{jt_2}.$$

By (D.7), we have $\overrightarrow{S}_{5nT,11}=\mathcal{S}(\Delta u,\Delta u)=\mathcal{S}(\mathbb{V},\mathbb{V})+\mathcal{S}(\mathbb{B},\mathbb{B})+\mathcal{S}(\mathbb{V},\mathbb{B})+\mathcal{S}(\mathbb{B},\mathbb{V})$. It suffices to show that each term in the last expression is $o_P(1)$.

First, we consider $\mathcal{S}(\mathbb{V},\mathbb{V})$. It is easy to verify that

$$\mathcal{S}(\mathbb{V},\mathbb{V})=S_1+o_P(1)$$

where

$$S_1\equiv\frac{1}{n_1T^5h^5}\sum_{1\leq i<j\leq n}\sum_{t_1\ldots t_6 \text{ are distinct}}\left(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4}\right)u_{it_5}\left(k^{\dagger}_{j,t_1t_3}-k^{\dagger}_{j,t_2t_3}\right)u_{jt_6}\eta_{i,t_1t_5}\eta_{i,t_2t_6}.$$

Let $\varphi^{\dagger}_{i,ts}\equiv k^{\dagger}_{i,ts}-E_t(k^{\dagger}_{i,ts})-E_s(k^{\dagger}_{i,ts})+E_tE_s(k^{\dagger}_{i,ts})$. Then $k^{\dagger}_{j,t_1t_3}-k^{\dagger}_{j,t_1t_4}=\varphi^{\dagger}_{j,t_1t_3}-\varphi^{\dagger}_{j,t_1t_4}+E_{t_1}(k^{\dagger}_{j,t_1t_3})-E_{t_1}(k^{\dagger}_{j,t_1t_4})$ and $k^{\dagger}_{j,t_1t_3}-k^{\dagger}_{j,t_2t_3}=\varphi^{\dagger}_{j,t_1t_3}-\varphi^{\dagger}_{j,t_2t_3}+E_{t_3}(k^{\dagger}_{j,t_1t_3})-E_{t_3}(k^{\dagger}_{j,t_2t_3})$. With these we can de-

compose S_1 as follows:

$$\begin{aligned} S_1 &= \frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{t_1...t_6 \text{ are distinct}}\{[\varphi^{\dagger}_{i,t_1t_3}-\varphi^{\dagger}_{i,t_1t_4}][\varphi^{\dagger}_{j,t_1t_3}-\varphi^{\dagger}_{j,t_2t_3}] \\ &\quad +[\varphi^{\dagger}_{i,t_1t_3}-\varphi^{\dagger}_{i,t_1t_4}][E_{t_3}(k^{\dagger}_{j,t_1t_3})-E_{t_3}(k^{\dagger}_{j,t_2t_3})]+[E_{t_1}(k^{\dagger}_{i,t_1t_3})-E_{t_1}(k^{\dagger}_{i,t_1t_4})][\varphi^{\dagger}_{j,t_1t_3}-\varphi^{\dagger}_{j,t_2t_3}] \\ &\quad +[E_{t_1}(k^{\dagger}_{i,t_1t_3})-E_{t_1}(k^{\dagger}_{i,t_1t_4})][E_{t_3}(k^{\dagger}_{j,t_1t_3})-E_{t_3}(k^{\dagger}_{j,t_2t_3})]\}u_{it_5}u_{jt_6}\eta_{i,t_1t_5}\eta_{j,t_2t_6} \\ &\equiv S_{11}+S_{12}+S_{13}+S_{14}, \text{ say,} \end{aligned} \tag{D.15}$$

where the definitions of S_{1r}, $r=1,2,3,4$, are self-evident. We further decompose S_{11} as follows:

$$\begin{aligned} S_{11} &= \frac{1}{n_1T^5h^5}\sum_{1\leq i<j\leq n}\sum_{t_1...t_6 \text{ are distinct}}\{\varphi^{\dagger}_{i,t_1t_3}\varphi^{\dagger}_{j,t_1t_3}-\varphi^{\dagger}_{i,t_1t_3}\varphi^{\dagger}_{j,t_2t_3}-\varphi^{\dagger}_{i,t_1t_4}\varphi^{\dagger}_{j,t_1t_3} \\ &\quad +\varphi^{\dagger}_{i,t_1t_4}\varphi^{\dagger}_{j,t_2t_3}\}u_{it_5}\varphi^{\dagger}_{j,t_1t_3}u_{jt_6}\eta_{i,t_1t_5}\eta_{j,t_2t_6} \\ &\equiv S_{111}-S_{112}-S_{113}+S_{114} \end{aligned}$$

To analyze S_{111}, let $A_{i_1j_1,i_2j_2}(t_1,...,t_{10})\equiv\varphi^{\dagger}_{i_1,t_1t_3}u_{i_1t_4}\varphi^{\dagger}_{j_1,t_1t_3}u_{j_1t_5}\eta_{i_1,t_1t_4}\eta_{j_1,t_2t_5}\varphi^{\dagger}_{i_2,t_6t_8}u_{i_2t_9}\varphi^{\dagger}_{j_2,t_6t_8}u_{j_2t_{10}}$ $\eta_{i_2,t_6t_9}\eta_{j_2,t_7t_{10}}$. Then

$$\begin{aligned} & E^{\mathbb{X}}\left[(S_{111})^2\right] \\ &= \frac{1}{(n_1T^4h^5)^2}\sum_{1\leq i_1<j_1\leq n}\sum_{1\leq i_2<j_2\leq n}\sum_{t_1...t_5 \text{ are distinct}}\sum_{t_6...t_{10} \text{ are distinct}}E^{\mathbb{X}}\left[A_{i_1j_1,i_2j_2}(t_1,...,t_{10})\right] \\ &= \frac{1}{(n_1T^4h^5)^2}\sum_{\substack{1\leq i_1<j_1\leq n,1\leq i_2<j_2\leq n,\\ i_1,i_2,j_1,j_2 \text{ are all distinct}}}\sum_{t_1...t_5 \text{ are distinct}}\sum_{t_6...t_{10} \text{ are distinct}}E^{\mathbb{X}}\left[A_{i_1j_1,i_2j_2}(t_1,...,t_{10})\right] \\ &\quad +\frac{1}{(n_1T^4h^5)^2}\sum_{\substack{1\leq i_1<j_1\leq n,1\leq i_2<j_2\leq n,\\ \#\{i_1,i_2,j_1,j_2\}=3}}\sum_{t_1...t_5 \text{ are distinct}}\sum_{t_6...t_{10} \text{ are distinct}}E^{\mathbb{X}}\left[A_{i_1j_1,i_2j_2}(t_1,...,t_{10})\right] \\ &\quad +\frac{1}{(n_1T^4h^5)^2}\sum_{1\leq i<j\leq n}\sum_{t_1...t_5 \text{ are distinct}}\sum_{t_6...t_{10} \text{ are distinct}}E^{\mathbb{X}}\left[A_{ij,ij}(t_1,...,t_{10})\right] \\ &\equiv ES_{111,1}+ES_{111,2}+ES_{111,3}, \end{aligned}$$

We prove $E^{\mathbb{X}}[(S_{111})^2]=o_P(1)$ by showing that $ES_{111,r}=o_P(1)$ for $r=1,3$ as one can analogously show that $ES_{111,2}=o_P(1)$. Write $ES_{111,1}$ as

$$\begin{aligned} ES_{111,1} &= \frac{1}{(n_1T^4h^5)^2}\sum_{\substack{1\leq i_1<j_1\leq n,1\leq i_2<j_2\leq n,\\ i_1,i_2,j_1,j_2 \text{ are all distinct}}}\sum_{t_1...t_5 \text{ are distinct}}\sum_{t_6...t_{10} \text{ are distinct}}E\left(\varphi^{\dagger}_{i_1,t_1t_3}u_{i_1t_4}\right) \\ &\quad \times E\left(\varphi^{\dagger}_{j_1,t_1t_3}u_{j_1t_5}\right)E\left(\varphi^{\dagger}_{i_2,t_6t_8}u_{i_2t_9}\right)E\left(\varphi^{\dagger}_{j_2,t_6t_8}u_{j_2t_{10}}\right)\eta_{i_1,t_1t_4}\eta_{j_1,t_2t_5}\eta_{i_2,t_6t_9}\eta_{j_2,t_7t_{10}} \end{aligned}$$

Let $\mathcal{G}_1\equiv\{t_1,t_3,t_4\}$, $\mathcal{G}_2\equiv\{t_1,t_3,t_5\}$, $\mathcal{G}_3\equiv\{t_6,t_8,t_9\}$, and $\mathcal{G}_4\equiv\{t_6,t_8,t_{10}\}$. We consider two cases: (a) there exists at least one time index that belongs to either one of these four groups and lies at least m-apart from all other indices within the same group, (b) all the other remaining cases. Noting that we can bound $|E(\varphi^{\dagger}_{i_1,t_1t_3}u_{i_1t_4})E(\varphi^{\dagger}_{j_1,t_1t_3}u_{j_1t_5})\ E(\varphi^{\dagger}_{i_2,t_6t_8}u_{i_2t_9})E(\varphi^{\dagger}_{j_2,t_6t_8}u_{j_2t_{10}})|$ by $Ch^{7-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)$ in case (a) and by Ch^8 in case (b), and the total number of terms in the summation is of order $O\left(n^4T^4m^6\right)$ in case (b), we can readily obtain $ES_{111,1}=O_P(n^2T^2h^{-3-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m)+n^2T^{-4}m^6h^{-2})=o_P(1)$. So $ES_{111,1}=o_P(1)$.

For $ES_{111,3}$, we have

$$\begin{aligned} ES_{111,3} &= \frac{1}{(n_1T^4h^5)^2}\sum_{1\leq i<j\leq n}\sum_{t_1...t_5 \text{ are distinct}}\sum_{t_6...t_{10} \text{ are distinct}} E\left[\varphi^{\dagger}_{i,t_1t_3}u_{it_4}\varphi^{\dagger}_{i,t_6t_8}u_{it_9}\right] \\ &\quad \times E\left[\varphi^{\dagger}_{j,t_1t_3}u_{jt_5}\varphi^{\dagger}_{j,t_6t_8}u_{jt_{10}}\right]\eta_{i,t_1t_4}\eta_{j,t_2t_5}\eta_{i,t_6t_9}\eta_{j,t_7t_{10}}. \end{aligned}$$

Let $\mathcal{G}_5 \equiv \{t_1, t_3, t_4, t_6, t_8, t_9\}$, $\mathcal{G}_6 \equiv \{t_1, t_3, t_5, t_6, t_8, t_{10}\}$ and $\mathcal{G} \equiv \mathcal{G}_5 \cup \mathcal{G}_6$. We can consider five cases: the number of distinct time indices in $\mathcal{G}$ are 8, 7, 6, 5, and 4, respectively, and use (a)-(e) to denote these five cases in order. Also, we use $ES_{111,3\xi}$ to denote $ES_{111,3}$ when the time indices in the summation are restricted to these five cases in order for $\xi = a, ..., e$. Following the arguments used in the analysis of $S_{111,1}$, we can show that $ES_{111,3a} = O_P(T^2h^{-4-\frac{2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m) + T^{-4}m^6h^{-2}) = o_P(1)$. Similarly we can show that $ES_{111,3\xi} = o_P(1)$ for $\xi = b, c, d$. For $ES_{111,3e}$, noting that the sets $\{t_1, t_3, t_4, t_5\}$ and $\{t_6, t_8, t_9, t_{10}\}$ must coincide, we have $|ES_{111,3e}| = O_P\left(T^{-2}h^{-8}\right) = o_P(1)$. Hence $ES_{111,3} = o_P(1)$, and we have shown that $E^{\mathbb{X}}[(S_{111})^2] = o_P(1)$, implying that $S_{111} = o_P(1)$. Similarly, we can show that $S_{11r} = o_P(1)$ for $r = 2, 3, 4$. It follows that $S_{11} = o_P(1)$.

For S_{12} defined in (D.15), we decompose it as follows:

$$\begin{aligned} S_{12} &= \frac{1}{n_1T^5h^5}\sum_{1\leq i<j\leq n}\sum_{t_1...t_6 \text{ are distinct}}[\varphi^{\dagger}_{i,t_1t_3}-\varphi^{\dagger}_{i,t_1t_4}]u_{it_5}[E_{t_3}(k^{\dagger}_{j,t_1t_3})-E_{t_3}(k^{\dagger}_{j,t_2t_3})]u_{jt_6}\eta_{i,t_1t_5}\eta_{j,t_2t_6} \\ &= \frac{1}{n_1T^5h^5}\sum_{1\leq i<j\leq n}\sum_{t_1...t_6 \text{ are distinct}}\{\varphi^{\dagger}_{i,t_1t_3}[E_{t_3}(k^{\dagger}_{j,t_1t_3})-c^{\dagger}_j]-\varphi^{\dagger}_{i,t_1t_4}[E_{t_3}(k^{\dagger}_{j,t_1t_3})-c^{\dagger}_j] \\ &\quad -\varphi^{\dagger}_{i,t_1t_3}u_{it_5}[E_{t_3}(k^{\dagger}_{j,t_2t_3})-c^{\dagger}_j]+\varphi^{\dagger}_{i,t_1t_4}[E_{t_3}(k^{\dagger}_{j,t_2t_3})-c^{\dagger}_j]\}u_{it_5}u_{jt_6}\eta_{i,t_1t_5}\eta_{j,t_2t_6} \\ &\equiv S_{121}-S_{122}-S_{123}+S_{124}, \end{aligned}$$

where $c^{\dagger}_j \equiv E_tE_s(k^{\dagger}_{j,ts})$. Analogously to the analysis of S_{111}, we can show $E^{\mathbb{X}}[(S_{12r})^2] = o_P(1)$ for $r =$1, 2, 3, 4. It follows that $S_{12} = o_P(1)$. By the same token, $S_{113} = o_P(1)$. For S_{114}, we have

$$\begin{aligned} S_{14} &= \frac{1}{n_1T^5h^5}\sum_{1\leq i<j\leq n}\sum_{t_1...t_6 \text{ are distinct}}\{[E_{t_1}(k^{\dagger}_{i,t_1t_3})-c^{\dagger}_i][E_{t_3}(k^{\dagger}_{j,t_1t_3})-c^{\dagger}_j] \\ &\quad -[E_{t_1}(k^{\dagger}_{i,t_1t_3})-c^{\dagger}_i][E_{t_3}(k^{\dagger}_{j,t_2t_3})-c^{\dagger}_j]-[E_{t_1}(k^{\dagger}_{i,t_1t_4})-c^{\dagger}_i][E_{t_3}(k^{\dagger}_{j,t_1t_3})-c^{\dagger}_j] \\ &\quad +[E_{t_1}(k^{\dagger}_{i,t_1t_4})-c^{\dagger}_i][E_{t_3}(k^{\dagger}_{j,t_2t_3})-c^{\dagger}_j]\}u_{it_5}u_{jt_6}\eta_{i,t_1t_5}\eta_{j,t_2t_6} \\ &\equiv S_{141}-S_{142}-S_{143}+S_{144}. \end{aligned}$$

Then we can show that $E^{\mathbb{X}}[(S_{14r})^2] = o_P(1)$ for $r =$1, 2, 3, 4. It follows that $S_{14} = o_P(1)$. Hence we have shown that $\mathcal{S}(\mathbb{V}, \mathbb{V}) = S_1 + o_P(1) = o_P(1)$.

Now, we consider $\mathcal{S}(\mathbb{B}, \mathbb{B})$. We have

$$\begin{aligned} \mathcal{S}(\mathbb{B}, \mathbb{B}) &= \frac{1}{n_1T^3h^5}\sum_{1\leq i<j\leq n}\sum_{t_1...t_4 \text{ are distinct}}\{(\varphi^{\dagger}_{i,t_1t_3}-\varphi^{\dagger}_{i,t_1t_4})(\varphi^{\dagger}_{j,t_1t_3}-\varphi^{\dagger}_{j,t_2t_3}) \\ &\quad +(\varphi^{\dagger}_{i,t_1t_3}-\varphi^{\dagger}_{i,t_1t_4})E_{t_3}(k^{\dagger}_{j,t_1t_3}-k^{\dagger}_{j,t_2t_3})+E_{t_1}(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4})(\varphi^{\dagger}_{j,t_1t_3}-\varphi^{\dagger}_{j,t_2t_3}) \\ &\quad +E_{t_1}(k^{\dagger}_{i,t_1t_3}-k^{\dagger}_{i,t_1t_4})E_{t_3}(k^{\dagger}_{j,t_1t_3}-k^{\dagger}_{j,t_2t_3})\}\mathbb{B}_{it_1}\mathbb{B}_{jt_2} \\ &\equiv S_{21}+S_{22}+S_{23}+S_{24}, \text{ say.} \end{aligned}$$

Write $S_{21} = \frac{1}{n_1 T^3 h^5} \sum_{1 \leq i < j \leq n} \sum_{t_1 ... t_4 \text{ are distinct}} \{\varphi^{\dagger}_{i,t_1t_3}\varphi^{\dagger}_{j,t_1t_3} - \varphi^{\dagger}_{i,t_1t_3}\varphi^{\dagger}_{j,t_2t_3} - \varphi^{\dagger}_{i,t_1t_4}\varphi^{\dagger}_{j,t_1t_3} + \varphi^{\dagger}_{i,t_1t_4} \times \varphi^{\dagger}_{j,t_2t_3}\}\mathbb{B}_{it_1}\mathbb{B}_{jt_2} \equiv S_{211} - S_{212} - S_{213} + S_{214}$. It is easy to show that S_{211} dominates S_{21r} for $r = 2, 3, 4$ and

$$E^{\mathbb{X}}\left[(S_{211})^2\right] = O_P(n^2T^2h^{-3-\frac{\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m) + n^2m^2h^{-2} + T^2h^{-5-\frac{2\delta}{1+\delta}}\alpha^{\frac{\delta}{1+\delta}}(m))\left(b^{4(p+1)}\right) = o_P(1).$$

Hence $S_{211} = o_P(1)$ and $S_{21} = o_P(1)$. Similarly, by decomposing $E_{t_1}(k^{\dagger}_{i,t_1t_3} - k^{\dagger}_{i,t_1t_4})$ as $[E_{t_1}(k^{\dagger}_{i,t_1t_3}) - c^{\dagger}_i] - [E_{t_1}(k^{\dagger}_{i,t_1t_4}) - c^{\dagger}_i]$ and $E_{t_3}(k^{\dagger}_{j,t_1t_3} - k^{\dagger}_{j,t_2t_3})$ as $[E_{t_3}(k^{\dagger}_{j,t_1t_3}) - c^{\dagger}_j] - [E_{t_3}(k^{\dagger}_{j,t_2t_3}) - c^{\dagger}_j]$, we can show $S_{2r} = o_P(1)$ for $r = 2, 3, 4$ by the conditional Chebyshev inequality. Consequently, $\mathcal{S}(\mathbb{B}, \mathbb{B}) = o_P(1)$. Analogously, we can show that $\mathcal{S}(\mathbb{V}, \mathbb{B}) = o_P(1)$ and $\mathcal{S}(\mathbb{B}, \mathbb{V}) = o_P(1)$. It follows that $S_{5nT,1} = o_P(1)$. ■

Proposition D.4 $nTh(\widetilde{\Delta}_{nT,1} - \widehat{\Delta}_{nT,1}) = o_P(1)$.

Proof. By the definitions of $\widehat{\Delta}_{nT,1}$ and $\widetilde{\Delta}_{nT,1}$, we have $-nTh(\widetilde{\Delta}_{nT,1} - \widehat{\Delta}_{nT,1})/\left[2R\left(\overline{k}\right)\right] = \sum_{i=1}^n \int [\widetilde{f}_i^2(u) - \widehat{f}_i^2(u)]du \equiv U_{1nT} + 2U_{2nT}$, where $U_{1nT} \equiv \sum_{i=1}^n \int [\widetilde{f}_i(u) - \widehat{f}_i(u)]^2 du$, and $U_{2nT} \equiv \sum_{i=1}^n \int [\widetilde{f}_i(u) - \widehat{f}_i(u)]\widehat{f}_i(u)\,du$. Then it is straightforward to show that $U_{1nT} = o_P(1)$ and $U_{2nT} = o_P(1)$ by arguments similar to but simpler than those used in the proof of Proposition D.1. ■

Proposition D.5 $nTh(\widetilde{\Delta}_{nT,2} - \widehat{\Delta}_{nT,2}) = o_P(1)$.

Proof. Let $\widehat{\Delta}_{nT,21}$, $\widehat{\Delta}_{nT,22}$, and $\widehat{\Delta}_{nT,23}$ denote the three terms on the right hand side of (D.1). Define $\widetilde{\Delta}_{nT,21}$, $\widetilde{\Delta}_{nT,22}$, and $\widetilde{\Delta}_{nT,23}$ analogously with the estimated residuals replacing the unobservable error terms. Then it suffices to show that $nTh(\widetilde{\Delta}_{nT,2r} - \widehat{\Delta}_{nT,2r}) = o_P(1)$ for $r = 1, 2, 3$. Each of them can be proved by the use of Taylor expansions and Chebyshev inequality. We omitted the details to save space. ■

E Some technical lemmas

This appendix presents some technical lemmas that are used in proving the main results.

Lemma E.1 *Let $\{W_t\}$ be a strong (α-) mixing process with mixing coefficient $\alpha(t)$. For any integer $l > 1$ and integers $(t_1, ..., t_l)$ such that $1 \leq t_1 < t_2 < \cdots < t_l$, let θ be a Borel measurable function such that*

$$\int |\theta(w_1, ..., w_l)|^{1+\delta} dF^{(1)}(w_1, ..., w_j)\, dF^{(2)}(w_{j+1}, ..., w_l) \leq M$$

for some $\delta > 0$ and $M > 0$, where $F^{(1)} = F_{t_1,...,t_j}$ and $F^{(2)} = F_{t_{j+1},...,t_l}$ are the distribution functions of $(W_{t_1}, ..., W_{t_j})$ and $(W_{t_{j+1}}, ..., W_{t_l})$, respectively. Let F denote the distribution function of $(W_{t_1}, ..., W_{t_l})$. Then

$$\left|\int \theta(w_1, ..., w_l)\, dF(w_1, ..., w_l) - \int \theta(w_1, ..., w_l)\, dF^{(1)}(w_1, ..., w_j)\, dF^{(2)}(w_{j+1}, ..., w_l)\right|$$
$$\leq 4M^{1/(1+\delta)}\alpha(t_{j+1} - t_j)^{\delta/(1+\delta)}.$$

Proof. See Lemma 2.1 of Sun and Chiang (1997). ■

Lemma E.2 *Let* $\{W_t\}$, θ, δ, *and* M *be defined as above. Let* $V_1 \equiv (W_{t_1}, ..., W_{t_j})$ *and* $V_2 \equiv (W_{t_{j+1}}, ..., W_{t_l})$. *Then* $E|E[\theta(V_1, V_2)|V_1] - \Theta(V_1)| \leq 4M^{1/(1+\delta)}\alpha\left(t_{j+1} - t_j\right)^{\delta/(1+\delta)}$, *where* $\Theta(v_1) \equiv E[\theta(v_1, V_2)]$.

Proof. See Yoshihara (1989) who proved the above lemma for β-mixing processes by using an inequality in Yoshihara (1976). The analogous result holds for α-mixing processes by using the Davydov inequality or Lemma E.1. ■

Let $k : \mathbb{R} \to \mathbb{R}$ be a differentiable kernel function, and k' be its first derivative. Define $\overline{k}(v) \equiv \int k(u)\, k(v-u)\, du$, $\overline{k'}(v) \equiv \int k'(u)\, k'(v-u)\, du$, and $k^+(v) \equiv \int k'(u)\, k(v-u)\, du$. The following lemma states some properties of $\overline{k}$, $\overline{k'}$, and k^+ that are used in the proof of our main results.

Lemma E.3 *Suppose* $k : \mathbb{R} \to \mathbb{R}$ *is a symmetric differential* γ*-th order kernel function such that* $\lim_{v\to\infty} v^l k(v) = 0$ *for* $l = 0, 1$. *Then*

(i) $\int \overline{k}(v)\, dv = 1$, $\int \overline{k}(v)\, v^l dv = 0$ *for* $l = 1, \ldots \gamma - 1$, *and* $\int \overline{k}(v)\, v^\gamma dv = 2\kappa_\gamma$ *where* $\kappa_\gamma = \int k(u)\, u^\gamma du$;

(ii) $\int \overline{k'}(v)\, v^l dv = 0$ *for* $l = 0, 1$ *and* $\int \overline{k'}(v)\, v^2 dv = 2$;

(iii) $\int k^+(v)\, dv = 0$, *and* $\int v k^+(v)\, dv = -1$.

Proof. (i) $\int \overline{k}(v)\, dv = \int\int k(u)\, k(v-u)\, dudv = \int k(u)\, du \int k(s)\, ds = 1$,
$\int \overline{k}(v)\, v^l dv = \sum_{s=0}^{l} C_l^s \int k(u)\, u^s du \int k(t)\, t^{l-s} dt = 0$ for $l = 1, \ldots, \gamma - 1$, and
$\int \overline{k}(v)\, v^\gamma dv = \sum_{s=0}^{\gamma} C_\gamma^s \int k(u)\, u^s du \int k(t)\, t^{\gamma-s} dt = 2\int k(u)\, du \int k(t)\, t^\gamma dt = 2\kappa_\gamma$.

(ii) $\int \overline{k'}(v)\, dv = \int\int k'(u)\, k'(v-u)\, dudv = \int k'(u)\, du \int k'(s)\, ds = 0$,
$\int \overline{k'}(v)\, vdv = 2\int k'(u)\, udu \int k'(t)\, dt = 0$ by the fact $\int k'(u)\, du = 0$, and
$\int \overline{k'}(v)\, v^2 dv = \int\int k'(u)\, k'(t) \left(u^2 + 2ut + t^2\right) dudt = 2\left[\int k'(u)\, udu\right]^2 = 2$.

(iii) $\int k^+(v)\, dv = \int k'(u) \int k(u-v)\, dvdu = \int k'(u)\, du = 0$, and
$\int v k^+(v)\, dv = \int k(u)\, k'(s)\, (s+u)\, dsdu = \int k'(s)\, sds + \int k'(s)\, ds \int uk(u)\, du = -1$. ■

References

Ahmad, I. A., Li, Q., 1997. Testing independence by nonparametric kernel method. Statistics & Probability Letter 34, 201-210.

Anselin, L., 1998. Spatial Econometrics: Methods and Models. Dorddrecht: Kluwer Academic Publishers.

Anselin, L., 2001. Spatial econometrics. B. Baltagi (eds.), A Companion to Theoretical Econometrics. Blackwell, Oxford.

Bai, J., Kao, C., 2006. On the estimation inference of a panel cointegration model with cross-sectional dependence. In: Baltagi, Badi (Ed.), Contributions to Economic Analysis. Elsevier, pp. 3-30.

Bai, J., Ng, S., 2004. A PANIC attack on unit roots and cointegration. Econometrica 72, 1127-1177.

Bai, J., Ng, S. 2008. Large Dimensional Factor Analysis. Foundations and Trends in Econometrics, Vol. 3, No. 2, 89-163.

Bai,J., Ng, S., 2010. Panel data unit root test with cross-section dependence: a further investigation. Econometric Theory 26, 1088-1114.

Breusch, T. S., Pagan, A. R., 1980. The Lagranger multiplier test and its application to model specifications in econometrics. Review of Economic Studies 47, 239-253.

Brock, W., Dechert, W., Scheinkman, J., LeBaron, B., 1996. A test for independence based on the correlation dimension. Econometric Reviews 15. 197-235.

Chang, Y., 2002. Nonlinear IV unit root tests in panels with cross-sectional dependency. Journal of Econometrics 110, 261-292,

Chen, J., Gao, J., Li, D., 2009. A new diagnostic test for cross-section uncorrelation in nonparametric panel data models. Working paper Series No. 0075, University of Adelaide.

de Jong, P., 1987. A central limit theorem for generalized quadratic forms. Probability Theory and Related Fields 75, 261-277.

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

Fan, Y., Li, Q., 1999. Central limit theorem for degenerate U-statistics of absolutely regular processed with applications to model specification testing. Journal of Nonparametric Statistics 10, 245-271.

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and Its Applications. Chapman & Hall, London.

Frees, E. W., 1995. Assessing cross-sectional correlation in panel data. Journal of Econometrics 69, 393-414.

Gourieroux, C., Monfort, A., Renault, E., Trognon, A., 1987. Generalised residuals. Journal of Econometrics 34, 5-32.

Granger, C., Maasoumi, E., Racine, J. S., 2004. A dependence metric for possibly nonlinear time series. Journal of Time Series Analysis 25, 649-669.

Hansen, B. E., 2008. Uniform convergence rates for kernel estimation with dependent data. Econometric Theory 24, 726-748.

Hoeffding, W., 1948. A nonparametric test of independence. The Annals of Mathematical Statistics 19, 546-557.

Hong, Y., 1998. Testing for Pairwise Serial Independence via the empirical distribution function. Journal of the Royal Statistical Society: Series B, 60, part 2, 429-453.

Hong, Y., 2000. Generalized Spectral Tests for Serial Independence. Journal of the Royal Statistical Society: Series B, 60, part 3, 557-574.

Hong, Y., White, H., 2005. Asymptotic Distribution Theory for Nonparametric Entropy Measures of Serial Independence. Econometrica 73, 837-901.

Hsiao, C., Li, Q., 2001. A consistent test for conditional heteroskedasticity in time series regression models. Econometric Theory 17, 188-221.

Hsiao, C., Pesaran, M. H., Pick, A., 2009. Diagnostic tests of cross section independence for nonlinear panel data models. IZA discussion paper No. 2756.

Huang, H., Kao, C., Urga, G., 2008. Copula-based tests for cross-sectional independence in panel models. Economics Letters 100, 24-228.

Johnson, D., McClelland, R., 1998. A general dependence test and applications. Journal of Applied Econometrics 13, 627-644.

Lavergne, P., Vuong, Q., 2000. Nonparametric significance testing. Econometric Theory 16, 576-601.

Lee, A. J., 1990. U-statistics: Theory and Practice. Marcel Dekker, Inc.: New York and Basel.

Li, Q., 1999. Consistent model specification tests for time series econometric models. Journal of Econometrics 92, 101-147.

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

Masry, E., 1996a. Multivariate regression estimation: Local polynomial fitting for time series. Stochastic processes and Their Application 65, 81-101.

Masry, E., 1996b. Multivariate local polynomial regression for time series:Uniform strong consistency rates. Journal of Time Series Analysis 17, 571-599.

Neumeyer, N., 2009. Testing independence in nonparametric regression. Journal of Multivariate Analysis 100, 1551-1566.

Ng, S., 2006. Testing cross section correlation in panel data using spacing. Journal of Business and Economic Statistics 24, 12-23.

Pesaran, M. H., 2007. A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics 22, 265-312.

Pesaran, M. H., 2004. General diagnostic tests for cross section dependence in panels. Cambridge Working Paper in Economics No. 0435.

Pesaran, M. H., Ullah, A., Yamagata, T., 2008. A bias adjusted LM test of error cross section independence. Econometrics Journal 11, 105-127.

Phillips, P. C. B., Moon, H., 1999. Linear regression limit theory for nonstationary panel data. Econometrica 67, 1057-1111.

Pinkse, J., 1998. A consistent nonparametric test for serial independence. Journal of Econometrics 84, 205-231.

Politis, D., Romano, J. 1994. The stationary bootstrap. Journal of the American Statistical Association 89, 1303-1313.

Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica 56, 931-954.

Robinson, P. M., 1991. Consistent nonparametric entropy-based testing. Review of Economic Studies 58, 437-453.

Robinson, P.M., 2008. Correlation testing in time series, spatial and cross-sectional data. Journal of Econometrics 147, 5-16.

Sarafidis, V., Yamagata, T., Robertson, D., 2009. A test of cross section dependence for a linear dynamic panel model with regressors. Journal of Econometrics 148, 149-161.

Su, L., Ullah, A., 2009. Testing conditional uncorrelatedness. Journal of Business and Economic Statistics 27, 18-29.

Su, L., White, H., 2003. Testing conditional independence via empirical likelihood. Discussion Paper, Department of Economics, UCSD.

Su, L., White, H., 2007. Consistent characteristic function-based test for conditional independence. Journal of Econometrics 141, 807-834.

Su, L., White, H., 2008, Nonparametric Hellinger metric test for conditional independence. Econometric Theory 24, 829-864.

Sun, S. and Chiang, C-Y., 1997. Limiting behavior of the perturbed empirical distribution functions evaluated at U-statistics for strongly mixing sequences of random variables. Journal of Applied Mathematics and Stochastic Analysis 10, 3-20.

Tjøstheim, D., 1996. Measures and tests of independence: a survey. Statistics 28, 249-284.

Yoshihara, K., 1976. Limit behavior of U-statistics for stationary, absolutely regular processes. Zeitschrift für Wahrscheinlichskeitstheorie und Verwandte Gebiete 35, 237-252.

Yoshihara, K., 1989. Limiting behavior of generalized quadratic forms generated by absolutely regular processes. In P. Mandl and M. Hušková (eds.), Proceedings of the Fourth Prague Symposium on Asymptotic Statistics, pp. 539-547. Charles University Press.