The LLN and CLT for Kernel-Weighted U-Processes of Spatially Dependent Data with Applications to Nonparametric Model Specification Testing

Yiguo Sun*

Department of Economics and Finance, University of Guelph, Guelph, ON N1G2W1, Canada

February 13, 2018

In this paper we derive the law of large numbers (LLN) and central limit theorem (CLT) for second-order kernel-weighted U-processes when data exhibits spatial dependence. We then apply our theory to nonparametric specification test of functional form of exogenous regressors and/or of spatial weights in spatial autoregressive regression framework. We derive the limit results of our proposed test statistics and use Monte Carlo simulations to assess the finite sample performance of the test statistics.

Keywords: CLT; LLN; Spatial autoregressive models; U-processes.

JEL classification codes: C12; C14; C21

*E-mail address: yisun@uoguelph.ca. This research is supported by the Social Science and Humanities Research Council of Canada Insightful Grant 435-2016-0340.

1 Introduction

In this paper we consider the law of large numbers (or LLN) and central limit theorem (or CLT) for second-order kernel-weighted U-processes for spatially dependent random variables. Specifically, we consider

$$U_n = \frac{1}{n(n-1)|H|}(B_n V)' \mathbf{K}_H (B_n V), \tag{1.1}$$

where $V = [v_1, \ldots, v_n]'$ is an $n \times 1$ vector of i.i.d. random variables, B_n is an $n \times n$ non-stochastic matrix and its (i,j)th element equal to b_{ij}, $\mathbf{K}_H$ is an $n \times n$ matrix with zero diagonal elements and its (i,j)th element equal to $K_{H,ij} = K(H^{-1}(D_i - D_j))$ for $i \neq j$, and $D_i \in \mathcal{S}_D \subset R^q$ for $q \geq 1$ is an i.i.d. random variable across index i. Also, $K(\cdot)$ is a product kernel function and $H = \text{diag}\{h_1, \ldots, h_q\}$ is a $q \times q$ diagonal matrix of bandwidth and $|H| = \prod_{j=1}^{q} h_j$. In Section 2, we show that U_n can be rewritten as a linear combination of a symmetric second-, third-, and fourth-order U-processes of independent but not identically distributed random variables.

The study of the LLN for non-degenerate U statistics and CLT for degenerate U statistics has along history and can be well traced back to Hoeffding (1948); see Serfling (1980) and Lee (1990) for an excellent review of some LLN and CLT results and applications of U-statistics for estimation and hypothesis tests. Nasari (2012) derives the LLN for the non-degenerate U-statistics with non-stochastic weights for independent observations. The CLTs for degenerate U-statistics have been derived by De Jong (1987) and Khashimov (1988) for independent observations, Khashimov (1987) for m-dependent processes, Khashimov (1992) and Yoshihara (1989) for strictly stationary absolutely regular processes, Gao and Anh (2000) for strictly stationary α-mixing processes, and Gao and Hong (2008) for more general U-statistics for strictly stationary absolutely regular processes. In nonparametric kernel estimation and test literature, kernel-weighted U-statistics have been studied by Powell et al. (1989), Hall (1984), and Fan and Li (1996) for independent observations, Takahata and Yoshihara (1987) and Fan and Li (1999) for strictly stationary absolutely regular processes, and Kim et al. (2011) for strictly stationary α-mixing processes, where Powell et al. (1989) derive the LLN for non-degenerate kernel-weighted U-statistics while the others give the CLT for degenerate kernel-weighted U-statistics. Depending on the sparseness of B_n, the U-process defined in (1.1) introduces variant degrees of cross-sectional dependence and serial dependence to $B_n V$ for cross-sectional data and time series data, respectively. To the best of our knowledge, we have not found any existing publications studying the LLN and CLT of U-processes defined in (1.1).

In this paper, we fit our theory into spatial econometrics instead of time series framework because the model specification tests given below are considered in nonparametric spatial econometric framework. Spatial econometric models are seen popular applications in regional economic developments, real estate pricing, environmental economics, labor economics and so on when spatial spillover effects and/or spatial interaction exist; see, e.g., Case et al. (1993), Brett and Pinkse (2000), Brueckner (2003), Ertur and Koch (2007), Baltagi et al. (2008). It is standard practice in econometric modeling to search for correct specification of the relationship between covariates, which motivates intensive contribution to nonparametric model specification test literature. The spatial econometric modelling requires not only the correct specification of functional relationship between covariates, but also correct specification of how one spatial unit is connected to the other spatial units or the correct specification of spatial weights. In this paper we therefore will apply the LLN and CLT derived for U_n to test for model specification both in functional form of regressors and of the spatial weights.

Specifically, we start with Sun's (2016) *functional-coefficient spatial autoregressive model with nonparametric spatial weights*:

$$Y_i = \sum_{j \neq i} g(\mathcal{Z}_{ij}) Y_j + X_i' \theta(D_i) + u_i, i = 1, 2, \ldots, n, \tag{1.2}$$

where X_i is a $p \times 1$ vector, D_i is a continuous random vector of dimension q the same as defined above, both $g(\cdot)$ and $\theta(\cdot)$ are unknown measurable functions, and $\mathcal{Z}_{ij} > 0$, measures the geographic distance (or any other exogenous distance) between unit i and unit j and is treated as a nonstochastic variable. Assuming that $\{(X_i, D_i, u_i)\}$ is an independent sequence with $E(u_i|X_i, D_i) = 0$ almost surely for all i, she then estimates model (1.2) by a mixed kernel and sieve estimation method. The models considered by Su (2012), and Su and Jin (2010) and Zhang (2013) are all nested to model (1.2).

We are interested in testing the following null and alternative hypotheses:

$$H_0^1 : \Pr\{\theta(D_i) = \theta_0\} = 1 \text{ for some } \theta_0 \in \Theta \subset R^p \tag{1.3}$$

$$H_1^1 : \Pr\{\theta(D_i) = \theta\} < 1 \text{ for any } \theta \in \Theta \subset R^p \tag{1.4}$$

and

$$H_0^2 : g(z) = \rho_0 g_0(z) \text{ a.s. for some } \rho_0 \in [a_1, a_2] \subset R \text{ and } z \in R^+ \tag{1.5}$$

$$H_1^2 : g(z) \neq \rho g_0(z) \text{ for any } \rho \in [a_1, a_2] \subset R \text{ over some non-tempty interval} \tag{1.6}$$

where Θ is a compact subset of R^p, $R^+ = [0, \infty)$, "a.s." means almost surely, H_0^1 assumes constant parameters with unknown spatial weight function, H_0^2 assumes a functional coefficient model with a parametric spatial weight function, and the intersection of H_0^1 and H_0^2, denoted by $H_0^1 \cap H_0^2$, assumes a constant parameter vector and a parametric spatial weight function or a linear mixed regressive spatial autoregressive model. As H_0^2 includes the case of no spatial interactive relation as a special case, we do not separately mention the test for zero spatial interactive relation.

In the nonparametric spatial econometrics literature, Su (2012) estimates nonparametric spatial autoregressive (or SAR) models with spatial autoregressive errors, or SARAR models, when exogenous regressors enter into a regression model in nonparametric way and briefly mentioned that Wald type of test statistic can be constructed to test for zero spatial lag parameters. When the spatial weights are known, Malikov and Sun (2017) test parameter constancy of functional coefficient spatial autoregressive models, where the spatial lag parameter is allowed to be a functional of exogenous covariate, and hence the proposed residual-based test statistic and L_2-distance based test statistic take the test for zero spatial lag parameter as a special case. Also, Su and Qu (2017) test for functional form in nonparametric SAR models with *known* spatial weight matrix, and Kang and Li (2017) consider varying coefficient SAR model with constant spatial lag parameter, which is nested to Malikov and Sun (2017), and test for parameter constancy via a kernel-based likelihood ratio test. All these tests essentially test H_0^1 against H_1^1, while assuming known spatial weights.

Given a pre-determined spatial weight matrix, testing the existence of spatial dependence can be traced back to Moran (1950) and Sen (1976), where the data are i.i.d. under the null hypothesis. Kelejian and Prucha (2001) extend Moran's I test to allow independent but heteroskedastic errors under the null hypothesis in the framework of several limited dependent variable models and spatial autoregressive models with possibly spatial autoregressive errors. Applying to mixed regressive SAR, SARAR, and spatial error (or SE) models, Yang (2015) explains how to improve the power of the LM test of spatial dependence based on bootstrap critical values. Debarsy and Ertur (2010) construct both LM and LR test for spatial dependence in parametric fixed-effect panel data framework.

One key issure of spatial econometric modelling is to correctly specify the spatial weights which are used to describe how one spatial unit connects with another spatial unit. In practice, it is common that researchers pre-define a spatial weight function before estimating their model. Compared with the literature on nonparametric model specification test, the literature on testing spatial weight matrix is very thin. The available work includes Kelejian and Piras (2011) and Kelejian and Prucha's (2001) J-test that can be used to select from a pool of spatial weight matrix candidates. If the weight matrix pool does not include the true spatial weight matrix, one would not learn the true weight matrix by the application of their J test. Han and Lee (2013) apply J-test procedure to select between a SAR model and matrix exponential spatial specification (MESS) model, which, of course, also suffers the general critique of the J-test that it is possible that both parametric spatial structures may deviate from the true one.

To sum up, it is to our best knowledge that we have not found existing work that can be used to conduct model specification test under three cases: (i) testing for the functional form of $\theta(\cdot)$ and functional form of spatial weight function simultaneously, (ii) testing for the functional form of $\theta(\cdot)$ in the presence of nonparametric spatial weights; (iii) testing for the functional form of the spatial weight function while allowing nonparametric functional form of $\theta(\cdot)$. This paper therefore fills into this literature gap by introducing a test statistic to simultaneously and alternatively test for model specification in both the functional form of the regressor covariates and the functional form of the spatial weights. Following Zheng (1996), we therefore construct a residual-based test statistic:

$$T_n = \frac{1}{n^2 |H|} \sum_{i=1}^{n} \sum_{j \neq i} \hat{u}_i \hat{u}_j K\left(H^{-1}(D_i - D_j)\right)$$

where $\hat{u}_i$'s are the residuals calculated under the null hypothesis and T_n is a second-order U-statistic. Note that $\{\hat{u}_i\}$ will be cross-sectionally dependent although $\{u_i\}$ is an i.i.d. sequence. This thus motivates us to derive the LLN and CLT for a second-order U-statistic defined in (1.1).

The rest of the paper is organized as follows. In Section 2, we derive the LLN and CLT for U_n defined in (1.1). Section 3 provides the limit results of our proposed test statistic for three pairs of null and alternative hypotheses; i.e., $H_0^1 \cap H_0^2$ against $H_1^1 \cup H_1^2$, H_0^1 against H_1^1, and H_0^2 against H_1^2. In Section 4 we report a small Monte Carlo simulation to assess the finite sample performance of our proposed test statistic. Section 5 concludes. All the mathematical proofs are delayed to an Appendix.

For the convenience of readers, we summarize our notation here. (i) We denote a generic positive constant by M, which may take different values at different places. Also, I_n is the $n \times n$ identity matrix. (ii) We denote an eigenvalue of an $n \times n$ matrix A_n by $\lambda(A_n)$ and $\rho(A_n) = \max_{1 \leq i \leq n} |\lambda_i(A_n)|$, where $\lambda_1(A) \geq \ldots \geq \lambda_n(A)$ are the eigenvalues of a symmetric matrix A_n in non-increasing order. (iii) For an $m \times n$ matrix A, $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|$ is the column sum norm of matrix A. (iv) For any functions g and f, we denote $(fg)(d) = f(d)\,g(d)$.

2 LLN and CLTs

This section will derive the LLN and CLTs for the second-order U-process, U_n, defined in (1.1). Applying simple algebra to (1.1) gives

$$\begin{aligned}
& n(n-1)|H|U_n \\
= & \sum_{i=1}^{n}\sum_{j\neq i}\sum_{l=1}^{n}\sum_{l'=1}^{n} b_{il}b_{jl'}v_l v_{l'} K_{H,ij} \\
= & \sum_{i=1}^{n}\sum_{j\neq i}\left[b_{ii}b_{ji}v_i^2 + b_{ij}b_{jj}v_j^2 + (b_{ii}b_{jj}+b_{ij}b_{ji})v_j v_i\right] K_{H,ij} \\
& + \sum_{i=1}^{n}\sum_{j\neq i}\sum_{l\neq i\neq j}\left[b_{il}b_{jl}v_l^2 + (b_{ii}b_{jl}+b_{il}b_{ji})v_i v_l + (b_{ij}b_{jl}+b_{il}b_{jj})v_j v_l\right] K_{H,ij} \\
& + \sum_{i=1}^{n}\sum_{j\neq i}\sum_{l\neq i\neq j}\sum_{l'\neq l\neq i\neq j} b_{il}b_{jl'}v_l v_{l'} K_{H,ij}. \qquad (2.1)
\end{aligned}$$

As U_n contains elements like v_i^2, we see that U_n will not be a degenerate U-process unless B_n is a diagonal matrix.

Denote $\chi_i = (D_i', v_i)'$, $p_1(\chi_i, \chi_j) = |H|^{-1}\left[b_{ii}b_{ji}v_i^2 + b_{ij}b_{jj}v_j^2 + (b_{ii}b_{jj}+b_{ij}b_{ji})v_j v_i\right] K_{H,ij}$, $p_2(\chi_i, \chi_j, \chi_l) = |H|^{-1}\left[b_{il}b_{jl}v_l^2 + (b_{jl}b_{ii}+b_{ji}b_{il})v_i v_l + (b_{ij}b_{jl}+b_{il}b_{jj})v_j v_l\right] K_{H,ij}$, and $p_3(\chi_i, \chi_j, \chi_l, \chi_{l'}) = |H|^{-1} b_{il}b_{jl'}v_l v_{l'} K_{H,ij}$, where $p_1(\chi_i, \chi_j)$ is a symmetric function, but $p_2(\chi_i, \chi_j, \chi_l)$ and $p_3(\chi_i, \chi_j, \chi_l, \chi_{l'})$ are asymmetric. We therefore denote $p_2^*(\chi_i, \chi_j, \chi_l) = \sum_{t=1}^{3!} p_2(\chi_{i_t}, \chi_{j_t}, \chi_{l_t})$ with (i_t, j_t, l_t) being any permutation of (i, j, l) and $p_3^*(\chi_i, \chi_j, \chi_l, \chi_{l'}) = \sum_{s=1}^{4!} p_3(\chi_{i_s}, \chi_{j_s}, \chi_{l_s}, \chi_{l_s'})$ with (i_s, j_s, l_s, l_s') being any permutation of (i, j, l, l'). Then, we have

$$\begin{aligned}
U_n = & \frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} p_1(\chi_i, \chi_j) + \frac{1}{n(n-1)}\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{l=j+1}^{n} p_2^*(\chi_i, \chi_j, \chi_l) \\
& + \frac{1}{n(n-1)}\sum_{i=1}^{n-3}\sum_{j=i+1}^{n-2}\sum_{l=j+1}^{n-1}\sum_{l'=l+1}^{n} p_3^*(\chi_i, \chi_j, \chi_l, \chi_{l'}) \\
= & U_{n1} + U_{n2} + U_{n3}, \qquad (2.2)
\end{aligned}$$

and we see that U_n equals a linear combination of second-, third-, and fourth-order U statistics of independent observations. The existing work mainly focuses on deriving the limit results for a U statistic of a specific order, and we are the very first to study the limit results of a linear combination of U statistics with several different orders.

Next, we denote

$$\hat{U}_{nj} = \sum_{i=1}^{n}\left[E(U_{nj}|\chi_i) - E(U_{nj})\right], \; j = 1, 2, 3$$

and

$$\hat{U}_n = \hat{U}_{n1} + \hat{U}_{n2} + \hat{U}_{n3}. \tag{2.3}$$

In order to derive the limit result for U_n we need the following regularity conditions.

Assumption 1. (i) $\{(D_i, v_i)\}$ is an i.i.d. sequence with finite variance and D_i does not depend on v_i and $E|v_i|^{2+\delta} < M$ for some $\delta > 0$; (ii) $\|B_n\|_1 + \|B_n'\|_1 < M$; (iii) $f(d)$, $m_1(d)$ and $m_2(d)$ are all Hölder continuous over $\mathcal{S}_D$, where $f(d)$ is the pdf of D_i, $m_1(d) = E(v_i|D_i = d)$, and $m_2(d) = E(v_i^2|D_i = d)$, and $\mathcal{S}_D$ is the support of $f(\cdot)$.

Assumption 2. (i) $K(u) = \prod_{l=1}^{p} k(u_l)$ and $k(\cdot)$ is a symmetric probability density function over a compact support $[-1, 1]$ and we denote $\kappa_{i,j} = \int K^i(u) u^j du$; (ii) $\|H\| \to 0$ and $n|H| \to \infty$ as $n \to \infty$, where $\|H\| = \sqrt{\sum_{l=1}^{q} h_l^2}$.

Theorem 2.1 *Under Assumptions 1-2,* $\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} b_{ii}^2 = c_b > 0$*,* $m_1(d) \equiv 0$ *and* $m_4(d) = E(v_i^4|D_i = d)$ *is uniformly bounded over* $d \in \mathcal{S}_D$*, we obtain*

$$n\sqrt{|H|}U_n \xrightarrow{d} N\left(0, 2c_b^2\kappa_{2,0}E\left[\left(m_2^2 f\right)(D_1)\right]\right).$$

Theorem 2.2 *Under Assumptions 1-2 and* $m_1(d) \neq 0$ *over at least one non-empty subset of* $\mathcal{S}_D$*, we have*

$$U_n - E(U_n) - \hat{U}_n = O_p\left(\left(n\sqrt{|H|}\right)^{-1}\right) \text{ and } \sqrt{n}\hat{U}_n \xrightarrow{d} N\left(0, \sigma_\psi^2\right)$$

where $\sigma_\psi^2 > 0$ *is defined in the Appendix in Eq. (A.4).*

When $E(v_i|D_i) = 0$ a.s., U_n is still a non-degenerate U-process unless all the off-diagonal elements in B_n equal zero. The asymptotic normality result of $n\sqrt{|H|}U_n$ in Theorem 2.1 relies crucially on the fact that the row and column sum norm of B_n are bounded, which can occur when each row and column of B_n only contains a finite number of non-zero elements or B_n contains many non-zero but trivially small elements. In spatial econometric field, it means that each unit has a finite number of "neighbors" or that every spatial unit is connected with other units but the connections are very weak and almost ignorable as the number of spatial units increases. Theorem 2.2 shows that the law of large numbers hold for a general non-degenerate U-statistic, U_n, under regularity Assumptions 1-2, and $\sqrt{n}(U_n - E(U_n)) \xrightarrow{d} N\left(0, \sigma_\psi^2\right)$. The two theorems are to be used to show the limting results of our test statistics.

3 The Test Statistics

We firstly derive the reduced form of model (1.2) as

$$Y = (I_n - G_n)^{-1} \{\mathrm{mtk}(X, \theta(D)) + u\} \quad (2.4)$$

when $I_n - G_n$ is non-singular, where Y, $\mathrm{mtk}(X, \theta(D))$, and u are all $n \times 1$ vectors that stake up Y_i, $X_i'\theta(D_i)$, and u_i, respectively, and G_n is an $n \times n$ spatial weight matrix with zero diagonal elements and its (i, j)th element $g_{ij} = g(\mathcal{Z}_{ij})$ for $i \neq j$. It implies that

$$E(Y|X, D) = (I_n - G_n)^{-1} \mathrm{mtk}(X, \theta(D)) \text{ a.s.} \quad (2.5)$$

Below, we list some regularity assumptions imposed on the model.

Assumption 3. (i) $\{(X_i, D_i)\}$ is a sequence of i.i.d. random variables; (ii) $\{u_i\}$ is a sequence of i.i.d. errors with $E(u_i|X_i, D_i) = 0$, $E(u_i^2|X_i, D_i) = \sigma_u^2$ almost surely, and $E\left(|u_i|^{2+\delta}\right) < M$ for some $\delta > 0$; (iii) Assumption 1(iii) holds; (iv) $E\left(X_{i,l_2}^{j_1} X_{i,l_2}^{j_2}|D_i = d\right)$ is continuously differentiable over $\mathcal{S}_D$ for j_1, $j_2 = 1, 2$ and l_1, $l_2 = 1, \ldots, p$.

Assumption 4. (i) $\lim_{n\to\infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n I(\mathcal{Z}_{ij} \in S_{\mathcal{Z}}) < \infty = O(1)$ for any fixed bounded set $S_{\mathcal{Z}}$, where the indicator function $I(\mathcal{A}) = 1$ if event $\mathcal{A}$ occurs, and zero otherwise; (ii) there exist a positive integer N and a constant $c_G \in (0, 1)$ such that for all $n > N$, $\rho(G_n) \leq c_G < 1$, $\|G_n + G_n'\|_1 \leq M < \infty$ and $\left\|(I_n - G_n)^{-1} + (I_n - G_n')^{-1}\right\|_1 \leq M < \infty$.

Assumptions 3(i)-(ii) state that the spatial dependence only occurs to the dependent variable, not the regressors (X_i, D_i) and the error term u_i, and the i.i.d. error is assumed for exposition easiness of our test results and can be relaxed to allow heteroskedastic errors without loss of essence. The smoothness imposed by Assumptions 3(iii)-(iv) are standard in nonparametric setup. Assumption 4(i) implies that each spatial unit only has a finite number of neighbors, which is also imposed in Pinkse et al. (2002, Condition (iv), p.1122). This assumption holds if the spatial weight function has a compact support in the increasing domain asymptotics and is imposed to simplify the exposition of our assumptions and can be relaxed to Sun's (2016) conditions which allow a spatial unit to have an infinite number of neighbors, but the number of neighbors is ignorable relative to the sample size. In addition, the assumption of $\rho(G_n) < 1$ ensures the spatial stationarity of the dependent variable, and Assumption 4(i), $\rho(G_n) \leq c_G < 1$, and $\sup_z |g(z)| < 1$ imply that G_n and $(I_n - G_n)^{-1}$ have finite row and column sum norms. We refer readers to Sun (2016) for detailed discussion of Assumption 4(ii).

This section contains three subsections, where Section 2.1, Section 2.2 and Section 2.3 contain the assumptions to ensure a consistent test and derive the limiting distribution of the proposed test for the null hypothesis of $H_0^1 \cap H_0^2$, H_0^1, and H_0^2, respectively.

3.1 The null of $H_0^1 \cap H_0^2$

In this subsection, our null hypothesis of interest is $H_0^1 \cap H_0^2$, under which model (1.2) becomes

$$Y_i = \rho_0 \sum_{j \neq i} g_0\left(\mathcal{Z}_{ij}\right) Y_j + X_i' \theta_0 + u_i, i = 1, 2, \ldots, n \tag{2.6}$$

where Assumption 4 holds if $|\rho_0| < 1/\rho\left(G_{n,0}\right)$, and $G_{n,0}$ is an $n \times n$ spatial weight matrix with zero diagonal elements and its (i, j)th element equal to $g_{0,ij} = g_0\left(\mathcal{Z}_{ij}\right)$ for $i \neq j$. So, $G_n = \rho_0 G_{n,0}$ in model (2.6).

Let $\left(\hat{\rho}, \hat{\theta}'\right)$ be either the QMLE or the GMM estimator of model (2.6). As Lee (2004, 2007) derives the consistency and limiting results of the QMLE and GMM estimator of the above mixed regressive SAR model, it is reasonable for us to impose the following assumption.

Assumption 5. (i) Under $H_0^1 \cap H_0^2$, $\left(\hat{\rho}, \hat{\theta}'\right) = \left(\rho_0, \theta_0'\right) + O_p\left(n^{-1/2}\right)$; (ii) If $H_0^1 \cap H_0^2$ fails to hold, there exists a finite vector (ρ, θ') such that $\left(\hat{\rho}, \hat{\theta}'\right) = (\rho, \theta') + O_p\left(n^{-1/2}\right)$.

The limiting result of our test statistic is given below.

Theorem 3.1 *Under Assumptions 2-5, under $H_0^1 \cap H_0^2$, we have*

$$J_n = n\sqrt{|H|} T_n / \sqrt{\hat{\sigma}_n^2} \xrightarrow{d} N(0, 1),$$

where

$$\hat{\sigma}_n^2 = \frac{2}{n^2 |H|} \sum_{i=1}^{n} \sum_{j \neq i} \hat{u}_i^2 \hat{u}_j^2 K^2\left(H^{-1}\left(D_i - D_j\right)\right) \xrightarrow{p} 2\sigma_u^4 \kappa_{2,0} E\left[f\left(D_1\right)\right];$$

otherwise, we have $\Pr\{J_n > M_n\} \to 1$ *as $n \to \infty$ for any non-stochastic, positive sequence $M_n = o\left(n\sqrt{|H|}\right)$.*

In the Appendix we show that $n\sqrt{|H|} J_n$ converges to a positive constant if $H_0^1 \cap H_0^2$ fails to hold, see the results derived in Lemma . Hence, J_n is a one-sided test statistic. We reject $H_0^1 \cap H_0^2$ if J_n is greater than the critical value c_α at the significance level of α, where $\Pr\left(Z \geq c_\alpha\right) = \alpha$ and Z denotes a standard normal random variable.

3.2 H_0^1 vs. H_1^1

In this subsection, our null hypothesis of interest is H_0^1, under which model (1.2) becomes

$$Y_i = \sum_{j \neq i} g\left(\mathcal{Z}_{ij}\right) Y_j + X_i' \theta_0 + u_i, i = 1, 2, \ldots, n, \tag{2.7}$$

where the spatial weight function $g(\cdot)$ is unknown and is to be estimated with θ_0 simultaneously via the sieve estimation method studied by Pinske et al. (2002) and Sun (2016). Specifically, we approximate the unknown spatial weight $g(z)$ by a series expansion $g_0^*(z) = \sum_{l=1}^{L_n} \alpha_{0,l}\phi_l(z)$, where $\{\phi_j(z), j = 1, 2, \ldots\}$ is a sequence of orthonormal basis functions on $L_2[0, \infty)$, and $L_n \to \infty$ as $n \to \infty$. Denoting $\boldsymbol{\phi}_{L_n}(z) = [\phi_1(z), \ldots, \phi_{L_n}(z)]'$ and $\Phi_{\omega,i} = \sum_{j \neq i} \boldsymbol{\phi}_{L_n}(\mathcal{Z}_{ij}) \otimes \omega_j$, we then approximate the spatial lag term, $\sum_{j \neq i} g(\mathcal{Z}_{ij}) Y_j$, by

$$\sum_{j \neq i} g_0^*(\mathcal{Z}_{ij}) Y_j = \sum_{l=1}^{L_n} \alpha_{0,l} \sum_{j \neq i} \phi_l(\mathcal{Z}_{ij}) Y_j \equiv \alpha_0' \Phi_{Y,i}, \tag{2.8}$$

where "$\otimes$" is the Kronecker product, and $\alpha_0 = [\alpha_{0,1}, \ldots, \alpha_{0,L_n}]'$ is an $L_n \times 1$ vector. Applying these approximation procedures to (2.7) gives

$$Y_i \approx \alpha_0' \Phi_{Y,i} + X_i' \theta_0 + u_i = V_{n,i}' \gamma_0 + u_i, \tag{2.9}$$

where both $V_{n,i} = [\Phi_{Y,i}', X_i']'$ and $\gamma_0 \equiv [\alpha_0', \theta_0']'$ are $m_n \times 1$ vectors with $m_n = p + L_n$.

Let $Q_n = [Q_{1n}, X]$ be an $n \times d$ instrument matrix and V_n be the $n \times (p + L_n)$ data matrix, where Q_{1n} contains linearly independent instruments taken from $\Phi_{X,i}$, and the ith row vector of V_n equals $V_{n,i}'$. As in Pinske et al.'s (2002), letting $\hat{\gamma} = [\hat{\alpha}', \hat{\theta}']'$ be the 2SLS estimator of γ_0, we can estimate $g(z)$ by $\hat{g}(z) = \hat{\alpha}' \boldsymbol{\phi}_{L_n}(z)$. As both Pinske et al. (2002) and Sun (2016) derive the consistency and limiting results for model (1.2) although under different conditions and the paper focuses on hypothesis test, it is reasonable to directly assume the consistency of the estimator given above and we refer readers to the two papers for details.

Assumption 6. (i) For any L_n, there exists a constant vector α_0 such that

$$\max_{1 \leq i \leq n} \sum_{j \neq i} \left| g(\mathcal{Z}_{ij}) - \alpha_0^T \boldsymbol{\phi}_{L_n}(\mathcal{Z}_{ij}) \right| = O\left(L_n^{-\xi}\right)$$

for some $\xi \geq 2$. (ii) $\max_l \sup_{z \in R^+} |\phi_l(z)| \leq M$.

Assumption 7. (i) Under H_0^1, $\|\hat{\alpha} - \alpha_0\| = O_p(\vartheta_n)$, $\left\|\hat{\theta} - \theta_0\right\| = O_p\left(n^{-1/2}\right)$, and $\sup_{z \in R^+} |\hat{g}(z) - g(z)| = O_p\left(L_n^{-\xi} + L_n^{1/2}\vartheta_n\right)$, where $\vartheta_n = L_n^{-\xi} + \sqrt{L_n/n}$; (ii) Under H_1^1, there exists an $L_n \times 1$ vector $\alpha = [\alpha_1, \ldots, \alpha_{L_n}]$ and a $p \times 1$ vector, θ, such that $\|\hat{\alpha} - \alpha\| = O_p(\vartheta_n)$, $\left\|\hat{\theta} - \theta\right\| = O_p\left(n^{-1/2}\right)$, and $\sup_{z \in R^+} |\hat{g}(z) - g^*(z)| = O_p\left(L_n^{1/2}\vartheta_n\right)$, where $g^*(z) = \sum_{l=1}^{L_n} \alpha_l \phi_l(z)$.

Assumption 8. (i) $L_n \to \infty$ and $L_n^2/n \to 0$ as $n \to \infty$; (ii) $L_n^2 |H| = o(1)$.

Theorem 3.2 *Under Assumptions 2-4 and 6-8, we have, under* H_0^1,

$$J_n = n\sqrt{|H|} T_n / \sqrt{\hat{\sigma}_n^2} \xrightarrow{d} N(0, 1),$$

where

$$\hat{\sigma}_n^2 = \frac{2}{n^2 |H|} \sum_{i=1}^{n} \sum_{j \neq i} \hat{u}_i^2 \hat{u}_j^2 K^2 \left(H^{-1} \left(D_i - D_j \right) \right) \xrightarrow{p} 2\sigma_u^4 \kappa_{2,0} E \left[f \left(D_1 \right) \right];$$

under H_1^1*, we have* $\Pr \{ J_n > M_n \} \to 1$ *as* $n \to \infty$ *for any non-stochastic, positive sequence* $M_n = o \left(n \sqrt{|H|} \right)$.

Again, J_n is a one-sided test as it converges to a positive constant under H_1^1 as shown in the Appendix, see the results given in Lemma A.10. Theorem **??** states that the nonparametric test for functional form of regressors does not require known spatial weights. Therefore, our work complements the functional form specification tests of Kang and Li (2017), Malikov and Sun (2017) and Su and Qu (2017) by allowing unknown spatial weights.

3.3 H_0^2 vs. H_1^2

In this subsection, our null hypothesis of interest is H_0^2, under which model (1.2) becomes

$$Y_i = \rho_0 \sum_{j \neq i} g_0 \left(\mathcal{Z}_{ij} \right) Y_j + X_i' \theta_0 \left(D_i \right) + u_i, i = 1, 2, \ldots, n, \tag{4.1}$$

where the spatial weight function $g_0(\cdot)$ is known and ρ_0 and $\theta_0(\cdot)$ are to be estimated. This model is nested to the varying coefficient SAR model considered in Malikov and Sun (2017) who allow both $\rho(\cdot)$ and $\theta_0(\cdot)$ depending on D_i. Once we calculate the respective kernel estimate, $\hat{\rho}(d)$ and $\hat{\theta}(d)$, of $\rho(d)$ and $\theta_0(d)$, following Malikov and Sun (2017), we can construct the estimator for ρ_0 by $\hat{\rho} = n^{-1} \sum_{i=1}^{n} \hat{\rho}(D_i)$. As the limit result for the estimator is available, letting $H_0 = \text{diag}\{h_{01}, ..., h_{0q}\}$ be the bandwidth matrix used to calculate the estimator, we impose the following assumptions.

Assumption 9. (i) Under H_0^2, $\hat{\rho} - \rho_0 = O_p\left(n^{-1/2}\right)$ and $\sup_{d \in \mathcal{S}_D} \left\| \hat{\theta}(d) - \theta_0(d) \right\| = O_p\left(\|H_0\|^2 + \sqrt{\ln n / (n |H_0|)} \right)$; (ii) Under H_1^2, there exists a constant ρ and a smooth curve $\theta(\cdot)$ such that $\hat{\rho} - \rho = O_p\left(n^{-1/2}\right)$, and $\sup_{d \in \mathcal{S}_D} \left\| \hat{\theta}(d) - \theta(d) \right\| = O_p\left(\|H_0\|^2 + \sqrt{\ln n / (n |H_0|)} \right)$.

Assumption 10. (i) $\|H_0\| \to 0$, $n|H_0| \to \infty$, $\|H\| \to 0$, and $n|H| \to \infty$ as $n \to \infty$; (ii) $n\sqrt{|H|} \|H_0\|^4 \to 0$ and $\sqrt{|H|} |H_0|^{-1} \ln n \to 0$ as $n \to \infty$.

Theorem 3.3 *Under Assumptions 2-4(ii) and 9-10, we have, under* H_0^2,

$$J_n = n\sqrt{|H|} T_n / \sqrt{\hat{\sigma}_n^2} \xrightarrow{d} N(0,1),$$

where

$$\hat{\sigma}_n^2 = \frac{2}{n^2 |H|} \sum_{i=1}^{n} \sum_{j \neq i} \hat{u}_i^2 \hat{u}_j^2 K^2 \left(H^{-1} \left(D_i - D_j \right) \right) \xrightarrow{p} 2\sigma_u^4 \kappa_{2,0} E \left[f \left(D_1 \right) \right];$$

under H_1^2*, we have* $\Pr\{J_n > M_n\} \to 1$ *as* $n \to \infty$ *for any non-stochastic, positive sequence* $M_n = o\left(n\sqrt{|H|}\right)$*.*

To the best of our knowledge, this is the first paper to formally test the functional form of spatial weight function. And, Theorem 3.3 states that we can test the spatial weight function while allowing other regressors enter into the SAR model in semiparametric way. Again, J_n is a one-sided test statistic as it converges to a positive constant under the alternative hypothesis as shown in the Appendix, see Lemma A.14.

4 Monte Carlo simulations

In this section, we evaluate the finite sample performance of our proposed test statistics by generating our data from the following four data generating mechanisms:

$$\begin{aligned}
\text{DGP}_1 &: \quad Y_i = 1 + \rho_0 \sum_{j \neq i} g_0(\mathcal{Z}_{ij}) Y_j + X_i + u_i \\
\text{DGP}_2 &: \quad Y_i = 1 + \rho_0 \sum_{j \neq i} g_0(\mathcal{Z}_{ij}) Y_j + \sin(\pi D_i/2) + X_i\left(1 - D_i^2\right) + u_i \\
\text{DGP}_3 &: \quad Y_i = 1 + \sum_{j \neq i} g(\mathcal{Z}_{ij}) Y_j + X_i + u_i \\
\text{DGP}_4 &: \quad Y_i = 1 + \sum_{j \neq i} g(\mathcal{Z}_{ij}) Y_j + \sin(\pi D_i/2) + X_i\left(1 - D_i^2\right) + u_i,
\end{aligned}$$

for $i = 1, 2, \ldots, n$, where we randomly draw $\{u_i\}$ from $N(0, 0.25)$ distribution, $\{(X_i, D_i)\}$ is randomly drawn from bivariate uniform random numbers over $[0,1] \times [0,1]$ with a correlation equal to $\rho = 0.5$ according to the Plackett distribution. For variable $\mathcal{Z}_{i,j}$ we randomly generate a_i and b_i from $U[0, R_n]$ with $R_n = [0.5n]$ and calculate $\mathcal{Z}_{i,j}$ as the Euclidean distance between (a_i, b_i) and (a_j, b_j). For each n, we only sample $\{\mathcal{Z}_{i,j}\}$ once so that $\{\mathcal{Z}_{i,j}\}$ is non-stochastic. By design, DGP_1 and DGP_2 have parametric spatial weight function, while DGP_1 and DGP_3 have constant parameters for intercept term and constant slope parameter in front of X. For DGP_1 and DGP_2, we define the spatial weight function as the popularly used inverse power function, $g_0(z) = c_0 z^{-2}$ for $z > 0$ and $c_0 > 0$, where we set ρ_0 and c_0 such that $\rho(G_n) = 0.1$.[1] Also, we define an exponential weight function $g(z) = c_0 0.5 \exp(-z/0.5)$ for $z > 0$ in DGP_3 and DGP_4. The sample size equals 100, 200, and 400, and the number of Monte Carlo replications equals 500.

[1] We also calculate rejection rates when $\rho(G_n) = 0.85$. The results for $\rho(G_n) = 0.85$ are similar to those for $\rho(G_n) = 0.1$ and are omitted to save space. The omitted simulation results are available upon request from the author.

We report the percentage rejection rates of testing $H_0^1 \cap H_0^2$ against $H_1^1 \cup H_1^2$ in Table 1 under the null hypothesis and Table 2 under the alternative hypotheses. We see that $H_0^1 \cap H_0^2$ holds for DGP_1 and $H_1^1 \cup H_1^2$ for DGP_2-DGP_4. For this case, we set $h = c_1 \hat{\sigma}_d n^{-1/5}$, where $\hat{\sigma}_d$ is the sample standard deviation of $\{D_i\}$, and $c_1 = 0.8$, 1, and 1.2 for robustness check and we report the rejection rates when the critical values are taken from the standard normal table and are generated from wild bootstrap method with 199 bootstrap replications. The proof for the validity of the bootstrap method is omitted as it has been shown in Su and Qu (2017) who assume known spatial weight matrix under both null and alternative hypotheses. The columns below N(0,1) give the percentage rejection rates when the standard normal critical values are used, while the columns below "Bootstrap" give the percentage rejection rates when the bootstrap critical values are used. Our simulation results indicate that the proposed test is consistent but the estimated sizes are underestimated even when sample size increases when the asymptotic critical values are used to make conclusions and that applying the bootstrap critical values can improve the performance of our test. In addition, the test results are robust to the choice of bandwidth choices. These observations are consistent with the findings from nonparametric kernel test literature.

The percentage rejection rates of testing H_0^1 against H_1^1 are given in Table 3. These pairs of null and alternative hypotheses test the functional form of $\theta(\cdot)$ without parametric specification of the spatial weight function, so the null hypothesis holds true for both DGP_1 and DGP_2, and DGP_3 and DGP_4 are models under the alternative hypothesis. It is no surprise that estimated sizes are the same for DGP_1 and DGP_2 and the estimated powers are the same for DGP_3 and DGP_4. We therefore only report the estimated size for DGP_1 and estimated power for DGP_3 in Table 3. The smoothing parameter h is set up the same way as above and we use Laguerre orthonormal basis functions with $L_n = \max([n^{0.05}], 2)$. Again, both asymptotic critical values and wild bootstrap critical values are used, and applying the bootstrap critical values improve the test performance under both null and alternative hypotheses.

When testing H_0^2 against H_1^2, we use $h_0 = 1.06\hat{\sigma}_d n^{-r_0}$ to estimate the unknown curve and use $h = 1.06\hat{\sigma}_d n^{-r}$ to construct the test statistic. Then, Assumption 10 implies that $r \in (0.4, 1)$ and $(1 - r/2)/4 < r_0 < r/2$. We then set $r_0 = 0.5$ and $r = 0.5$, 0.55, and 0.66 for robustness check. Also, as the estimation procedure can be very time consuming with quadratic moment conditions used in the nonparametric GMM estimation, we only use the localized linear moment conditions or the nonparametric 2SLS estimation method to speed up our calculation. The spatial weight function under H_0^2 equals $\rho_0 g_0(\cdot)$, while we set $g(z) = \exp(-z)$ under the alternative hypothesis. We report the estimated size when the

Table 1: Estimated sizes under $H_0^1 \cap H_0^2$

		N(0,1)			Bootstrap		
n	$\alpha \backslash c_1$	0.8	1	1.2	0.8	1	1.2
100	1	0.6	0.4	0.2	1	1.2	0.8
	5	2.4	1.6	1.2	5.8	5.2	5
	10	3.6	3.4	3	11.8	11.2	11.4
	20	8	6	4.4	24.4	23.2	23.4
200	1	1	0.6	0.6	1.4	1.2	1.6
	5	1.6	1.8	1.6	5	5.6	5
	10	3.2	2.8	2.4	11	11.4	11.4
	20	8.2	7.6	5.8	22.4	23.4	22.6
400	1	0.8	0.6	0.4	0.6	0.8	1.2
	5	2.2	2.2	2	4.8	5.8	5.6
	10	3.6	3.4	3.8	12.4	11.2	11.2
	20	8	7.2	6.2	21.8	22.8	23.2

data is generated from DGP_1 and estimated power when the data is generated from DGP_3 in Table 4. Table 4 supports the consistency of the proposed test statistic, but the power can be very low for small sample. Also, the bootstrap test results are very similar to the asymptotic test results and are therefore omitted. Our simulation experiences indicate that the residual-based test is more difficult and less powerful to test for the functional form of the spatial weights than to test for the unknown coefficient curve of regressors, which is consistent with our theory derived in the Appendix. In the Appendix, the leading term of the three cases we considered are given in (A.15), (A.21), and (A.25) under their alternative hypotheses, where the mean weighted squared differences between $X_i'\theta(D_i)$ and $X_i'\theta$ generate the power of our test when testing $H_0^1 \cap H_0^2$ against $H_1^1 \cup H_1^2$ and H_0^1 against H_1^1, while the power of our test when testing H_0^2 against H_1^2 can be seriously subdued when the spatial weights under the null are very close to the spatial weights under the alternative hypothesis.

Overall, we observe that the small simulation experiments do support our theories derived in the paper.

5 Conclusion

This paper derives the LLN and CLTs for second-order U statistics for spatially dependent data and applies the derived LLN and CLTs to show that the residual-based test statistic can be used to test the functional form of regressors and/or spatial weight function in semiparametric spatial autoregressive framework. It is the first paper to formally test the

Table 2: Estimated powers when $H_0^1 \cap H_0^2$ is the null hypothesis

		DGP$_2$						DGP$_4$					
		N(0,1)			Bootstrap			N(0,1)			Bootstrp		
n	$\alpha \backslash c_1$	0.8	1	1.2	0.8	1	1.2	0.8	1	1.2	0.8	1	1.2
100	1	52.4	52.8	50.2	54.2	54.8	57.6	52.4	52.4	50.2	54.2	55	57.6
	5	66.6	65.8	64.4	78.8	80	80.8	66.6	65.6	64.8	78.6	80.8	81.2
	10	73.4	73.2	73.2	86.6	88.4	87.8	73.8	73	72.8	86.6	88.6	87.6
	20	82.4	82.4	81.8	92.6	93.6	94.8	82	82.4	81.6	92	93.6	94.6
200	1	91	90.8	90.8	92	92.2	92.2	90.8	91	91.2	91.8	91.6	92
	5	95.2	95.6	95.6	97.2	98	98	95.4	95.4	95.4	97.2	97.8	98.2
	10	96.6	96.6	97	98.6	98.8	98.8	96.6	96.6	96.6	98.8	98.8	98.8
	20	98	98.2	98.6	99.8	99.6	100	97.6	98.4	98.6	99.8	99.6	100
400	1	100	100	100	99.8	99.8	99.8	100	100	100	99.8	99.8	99.8
	5	100	100	100	100	100	100	100	100	100	100	100	100
	10	100	100	100	100	100	100	100	100	100	100	100	100
	20	100	100	100	100	100	100	100	100	100	100	100	100

Table 3: Estimated rejection rates when H_0^1 is the null hypothesis

		DGP$_1$						DGP$_2$					
		N(0,1)			Bootstrap			N(0,1)			Bootstrap		
n	$\alpha \backslash c_1$	0.8	1	1.2	0.8	1	1.2	0.8	1	1.2	0.8	1	1.2
100	1	0.8	0.6	0.4	1.2	1.4	0.8	51.2	51.6	49.8	53.6	55.2	58.4
	5	2.2	1.8	1.4	5	5.2	5.4	65.6	65	63.2	77.4	79	80.4
	10	3.4	3.4	2.8	10.8	11.4	10.8	72	72.6	71.4	84.6	85.6	87.6
	20	7.8	6	5	22.6	23.8	23.6	80.2	80.6	79.2	91.8	92.6	93.6
200	1	1	1	0.4	1.6	1	1.4	91	91	90.8	91.6	92.6	92.8
	5	1.8	1.8	1.6	6.2	6.4	5.6	95.4	95.8	95	97.2	97.8	97.8
	10	4	3.2	2.4	11.6	11.4	11.8	96.6	96.8	97.2	98.2	99.2	99.6
	20	8.8	6.6	5	22.2	23.4	22.6	98.2	98.2	98.2	99.8	99.8	100
400	1	0.8	0.8	0.8	0.8	0.8	1.4	100	100	100	99.8	99.8	99.8
	5	2	2.2	2	5.2	5	5.4	100	100	100	100	100	100
	10	3.6	3	3.4	12	11.4	11.4	100	100	100	100	100	100
	20	8.4	7.8	6.2	23.4	24	24	100	100	100	100	100	100

functional form of spatial weight function in semiparametric model setup and to test the functional form of regressors without pre-defining the spatial weights.

Table 4: Estimated rejection rates when H_0^2 is the null hypothesis

		DGP$_1$			DGP$_3$		
n	$\alpha\backslash r$	0.5	0.55	0.6	0.5	0.55	0.6
100	1	1.2	1.2	1.4	18.6	18.2	16
	5	3	3	4.4	28.2	26.4	24.6
	10	5.6	7	8	32	31.6	30.4
	20	14.2	14.2	16	41.2	40.2	38.4
200	1	1.4	1.2	1.2	41	40	38.8
	5	4.6	4.6	5.8	48.2	45.8	44.2
	10	7.6	8.4	8.4	54.6	52.4	50.4
	20	14.4	14	15	60.8	59.2	58
400	1	2	1.8	1.6	62.6	59.8	56.2
	5	4.6	4.4	4.4	68	66.2	64.2
	10	6.6	6.6	7.4	71.2	70.6	68.2
	20	12.2	13.6	15.8	75.4	74.4	74.8

6 Appendix: Brief Mathematical Proofs

For a matrix A_n, the spectral norm is denoted by $\|A_n\|_{sp} = \lambda_{\max}^{1/2}(A_n A_n')$, the Euclidean norm is denoted by $\|A_n\| = \sqrt{\operatorname{tr}(A_n A_n')}$. Evidently, $\|a_n\|_{sp} = \|a_n\|$ for any vector a_n. Also, as $\rho(A_n) \leq \|A_n\|_1$, we see $\|A_n\|_{sp} \leq \|A_n A_n'\|_1^{1/2} \leq \sqrt{\|A_n\|_1 \|A_n'\|_1}$. Also, throughout this Appendix, we denote $\tilde{a}_n = (I_n - G_n)^{-1} a_n$.

Proof of Theorem 2.1: It is straightforward to obtain

$$U_n = \frac{1}{n(n-1)|H|} \sum_{i=1}^{n} \sum_{j \neq i} b_{ii} b_{jj} v_j v_i K_{H,ij} + O_p\left(n^{-1}\right)$$

from (2.1) under Assumptions 1-2. Then, applying de Jong's (1987, Th. 3.2) CLT gives

$$\frac{1}{n\sqrt{|H|}} \sum_{i=1}^{n} \sum_{j \neq i} b_{ii} b_{jj} v_j v_i K_{H,ij} \xrightarrow{d} N\left(0, 2c_b^2 \kappa_{2,0} E\left[\left(m_2^2 f\right)(D_1)\right]\right)$$

if $\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} b_{ii}^2 = c_b$, a finite positive constant, $\|H\| \to 0$ and $n|H| \to \infty$ as $n \to \infty$. This completes the proof of this lemma.

Proof of Theorem 2.2: For any j, we have $E\left[\hat{U}_{nj}\left(U_{nj} - E\left(U_{nj}\right)\right)\right] = E\left(\hat{U}_{nj}^2\right)$ because applying the law of iterated expectation gives

$$\begin{aligned}
& E\left[\hat{U}_{nj}\left(U_{nj} - E\left(U_{nj}\right)\right)\right] \\
= & \sum_{i=1}^{n} E\left[E\left(U_{nj} - E\left(U_{nj}\right)|\chi_i\right)\left(U_{nj} - E\left(U_{nj}\right)\right)\right] \\
= & \sum_{i=1}^{n} E\left[E\left(U_{nj} - E\left(U_{nj}\right)|\chi_i\right)\right]^2 = Var\left(\hat{U}_{nj}\right) = E\left(\hat{U}_{nj}^2\right)
\end{aligned}$$

where the last line results from the fact that $E\left(U_{nj}-E\left(U_{nj}\right)|\chi_i\right)$ is independent across i. Therefore, we obtain

$$Var\left(U_{nj}-E\left(U_{nj}\right)-\hat{U}_{nj}\right)=Var\left(U_{nj}\right)-E\left(\hat{U}_{nj}^2\right)=O\left(n^{-2}\left|H\right|^{-1}\right)$$

by Lemmas A.1 and A.2. Hence, we obtain $U_n-E\left(U_n\right)-\hat{U}_n=o_p\left(n^{-1/2}\right)$ if $n\left|H\right|\to\infty$ as $n\to\infty$. As $\sqrt{n}\hat{U}_n\xrightarrow{d}N\left(0,\sigma_\psi^2\right)$ by Lemma A.3, we obtain $U_n=E\left(U_n\right)+O_P\left(n^{-1/2}\right)$. This completes the proof of this theorem.

Lemma A.1 *Under Assumptions 1-2 and* $m_1\left(d\right)\neq 0$ *over at least one non-empty subset of* $\mathcal{S}_D$, $Var\left(\hat{U}_{nj}\right)$ *are given in (A.1), (A.2), and (A.3) for j=1,2,3, respectively.*

Proof: (i) We consider $j=1$. For $s=1,\ldots,n$, we have

$$\begin{aligned}&E\left(U_{n1}-E\left(U_{n1}\right)|\chi_s\right)\\ =\;&\frac{2}{n\left(n-1\right)}\sum_{i\neq s}\left\{E\left[p_1\left(\chi_i,\chi_s\right)|\chi_s\right]-E\left[p_1\left(\chi_i,\chi_s\right)\right]\right\}\end{aligned}$$

where

$$\begin{aligned}&E\left[p_1\left(\chi_i,\chi_s\right)|\chi_s\right]-E\left[p_1\left(\chi_i,\chi_s\right)\right]\\ =\;&\left|H\right|^{-1}b_{ii}b_{si}\left[E\left(v_i^2K_{H,is}|\chi_s\right)-E\left(v_i^2K_{H,is}\right)\right]\\ &+\left|H\right|^{-1}b_{is}b_{ss}\left[E\left(v_s^2K_{H,is}|\chi_s\right)-E\left(v_s^2K_{H,is}\right)\right]\\ &+\left|H\right|^{-1}\left(b_{ii}b_{ss}+b_{is}b_{si}\right)\left[v_sE\left(v_iK_{H,is}|\chi_s\right)-E\left(v_iv_sK_{H,is}\right)\right]\\ =\;&\left[b_{ii}b_{si}\varphi_{1s}+b_{is}b_{ss}\varphi_{2s}+\left(b_{ii}b_{ss}+b_{is}b_{si}\right)\varphi_{3s}\right]\left(1+O\left(\left\|H\right\|^2\right)\right),\end{aligned}$$

denoting $\varphi_{1s}=\left(m_2f\right)\left(D_s\right)-E\left[\left(m_2f\right)\left(D_1\right)\right]$ and $\varphi_{2s}=v_s^2f\left(D_s\right)-E\left[\left(m_2f\right)\left(D_1\right)\right]$ and $\varphi_{3s}=v_s\left(m_1f\right)\left(D_s\right)-E\left[\left(m_1^2f\right)\left(D_1\right)\right]$. It then follows

$$\begin{aligned}Var\left(\hat{U}_{n1}\right)&=\sum_{s=1}^{n}E\left[E\left(U_{n1}-E\left(U_{n1}\right)|\chi_s\right)\right]^2\\ &\approx\frac{4}{n^2\left(n-1\right)^2}\sum_{s=1}^{n}E\left\{\sum_{i\neq s}\left[b_{ii}b_{si}\varphi_{1s}+b_{is}b_{ss}\varphi_{2s}+\left(b_{ii}b_{ss}+b_{is}b_{si}\right)\varphi_{3s}\right]\right\}^2\\ &=\frac{4Var\left[\left(m_1f\right)\left(D_1\right)\right]}{n^2\left(n-1\right)^2}\left(\sum_{s=1}^{n}b_{ss}^2\right)\left(\sum_{i=1}^{n}b_{ii}\right)^2+O\left(n^{-2}\right).\qquad\text{(A.1)}\end{aligned}$$

(ii) We consider $j=2$. For $s=1,\ldots,n$, we have

$$\begin{aligned}&E\left(U_{n2}-E\left(U_{n2}\right)|\chi_s\right)\\ =\;&\frac{1}{n\left(n-1\right)}\sum_{j\neq s}\sum_{l=j+1,l\neq s}^{n}\left\{E\left[p_2^*\left(\chi_s,\chi_j,\chi_l\right)|\chi_s\right]-E\left[p_2^*\left(\chi_s,\chi_j,\chi_l\right)\right]\right\}\\ \approx\;&\frac{2}{n\left(n-1\right)}\left(\mu_v\omega_{1s}\eta_{2s}+\mu_v\omega_{2s}\eta_{3s}+\omega_{3s}\left(v_s-\mu_v\right)E\left[\left(m_1f\right)\left(D_1\right)\right]\right)\end{aligned}$$

where we denote $\mu_v = E(v_i)$, $\eta_{1s} = f(D_s) - E[f(D_1)]$, $\eta_{2s} = v_s f(D_s) - E[(m_1 f)(D_1)]$, $\eta_{3s} = (m_1 f)(D_s) - E[(m_1 f)(D_1)]$, and

$$\omega_{1s} = b_{ss} \sum_{j \neq s} \sum_{l=j+1, l \neq s}^{n} (b_{jl} + b_{lj}), \; \omega_{2s} = \sum_{j \neq s} \sum_{l=j+1, l \neq s}^{n} (b_{sj} b_{ll} + b_{sl} b_{jj}), \; \omega_{3s} = \sum_{j \neq s} \sum_{l=j+1, l \neq s}^{n} (b_{js} b_{ll} + b_{ls} b_{jj})$$

because we have

$$\begin{aligned} & E\left[p_2^*\left(\chi_s, \chi_j, \chi_l\right) | \chi_s\right] - E\left[p_2^*\left(\chi_s, \chi_j, \chi_l\right)\right] \\ \approx \; & 2\mu_2 (b_{sl} b_{jl} + b_{sj} b_{lj}) \eta_{1s} + 2\mu_v ((b_{jl} + b_{lj}) b_{ss} + b_{js} b_{sl} + b_{ls} b_{sj}) \eta_{2s} \\ & + 2\mu_v (b_{sj} b_{jl} + b_{sl} b_{lj} + b_{sj} b_{ll} + b_{sl} b_{jj}) \eta_{3s} + 2 b_{ls} b_{js} \left(v_s^2 - \sigma_v^2\right) E[f(D_1)] \\ & + 2 (b_{js} b_{ll} + b_{ls} b_{jj} + b_{jl} b_{ls} + b_{lj} b_{js}) (v_s - \mu_v) E[(m_1 f)(D_1)] \end{aligned}$$

denoting $\mu_2 = E(v_i^2)$. Then, we have

$$\begin{aligned} & Var\left(\hat{U}_{n2}\right) \\ = \; & \sum_{s=1}^{n} E\left[E\left(U_{n2} - E(U_{n2}) | \chi_s\right)\right]^2 \\ \approx \; & \frac{4}{n^2 (n-1)^2} \sum_{s=1}^{n} Var\left\{\mu_v \omega_{1s} v_1 f(D_1) + \mu_v \omega_{2s} (m_1 f)(D_1) + \omega_{3s} v_1 E[(m_1 f)(D_1)]\right\} \end{aligned} \quad (A.2)$$

(ii) We consider $j = 3$. For $s = 1, \ldots, n$, we have

$$\begin{aligned} & E\left(U_{n3} - E(U_{n3}) | \chi_s\right) \\ = \; & \frac{1}{n(n-1)} \sum_{i=1}^{n-3} \sum_{j=i+1}^{n-2} \sum_{l=j+1}^{n-1} \sum_{l'=l+1}^{n} \left\{E\left[p_3^*\left(\chi_i, \chi_j, \chi_l, \chi_{l'}\right) | \chi_s\right] - E\left[p_3^*\left(\chi_i, \chi_j, \chi_l, \chi_{l'}\right)\right]\right\} \\ = \; & \frac{1}{n(n-1)} \sum_{j \neq s}^{n-2} \sum_{l=j+1}^{n-1} \sum_{l'=l+1}^{n} \left\{E\left[p_3^*\left(\chi_s, \chi_j, \chi_l, \chi_{l'}\right) | \chi_s\right] - E\left[p_3^*\left(\chi_s, \chi_j, \chi_l, \chi_{l'}\right)\right]\right\} \\ \approx \; & \frac{2\mu_v^2}{n(n-1)} \varsigma_{1s} \eta_{1s} + \frac{2\mu_v E[f(D_1)]}{n(n-1)} \varsigma_{2s} (v_s - \mu_v) \end{aligned}$$

denoting

$$\begin{aligned} \varsigma_{1s} &= \sum_{j \neq s}^{n-2} \sum_{l=j+1}^{n-1} \sum_{l'=l+1}^{n} \left[b_{sl} (b_{jl'} + b_{l'j}) + b_{sl'} (b_{lj} + b_{jl}) + b_{sj} (b_{l'l} + b_{ll'})\right] \\ \varsigma_{2s} &= \sum_{j \neq s}^{n-2} \sum_{l=j+1}^{n-1} \sum_{l'=l+1}^{n} \left[b_{ls} (b_{jl'} + b_{l'j}) + b_{l's} (b_{lj} + b_{jl}) + b_{js} (b_{l'l} + b_{ll'})\right] \end{aligned}$$

where we have

$$
\begin{aligned}
& E\left[p_3^*\left(\chi_s, \chi_j, \chi_l, \chi_{l'}\right)|\chi_s\right] - E\left[p_3^*\left(\chi_s, \chi_j, \chi_l, \chi_{l'}\right)\right] \\
= \; & 2\mu_v^2\left[b_{sl}\left(b_{jl'}+b_{l'j}\right)+b_{sl'}\left(b_{jl}+b_{lj}\right)+b_{sj}\left(b_{ll'}+b_{l'l}\right)\right]\eta_{1s} \\
& +2\mu_v\left[b_{ls}\left(b_{jl'}+b_{l'j}\right)+b_{l's}\left(b_{lj}+b_{jl}\right)+b_{js}\left(b_{l'l}+b_{ll'}\right)\right]\left(v_s-\mu_v\right)E\left[f\left(D_1\right)\right]
\end{aligned}
$$

Hence, we obtain

$$
Var\left(\hat{U}_{n3}\right) \approx \frac{4\mu_v^2}{n^2\left(n-1\right)^2}\sum_{s=1}^{n} Var\left(\varsigma_{1s}\mu_v f\left(D_1\right)+\varsigma_{2s}E\left[f\left(D_1\right)\right]v_1\right). \qquad \text{(A.3)}
$$

This completes the proof of this lemma.

Lemma A.2 *Under Assumptions 1-2 and $m_1(d) \neq 0$ over at least one non-empty subset of $\mathcal{S}_D$, we have $Var\left(U_{nj}\right) = Var\left(\hat{U}_{nj}\right) + O\left(n^{-2}\left|H\right|^{-1}\right)$ for j=1,2,3.*

Proof: Firstly, we consider $Var\left(U_{n1}\right)$. By definition we have

$$
U_{n1} - E\left(U_{n1}\right) = \hat{U}_{n1} + \frac{2}{n\left(n-1\right)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} e_{ij}
$$

where $e_{ij} = p_1\left(\chi_i,\chi_j\right) - E\left(p_1\left(\chi_i,\chi_j\right)|\chi_i\right) - E\left[p_1\left(\chi_i,\chi_j\right)|\chi_j\right]$ for $i \neq j$. As $e_{ij} = e_{ji}$, $Cov\left(e_{ij}, e_{i'j'}\right) = 0$ if $(i,j) \neq (i',j')$, and $Cov\left(\hat{U}_{n1}, U_{n1}-\hat{U}_{n1}\right) = 0$, we have

$$
\begin{aligned}
Var\left(U_{n1}\right) &= Var\left(\hat{U}_{n1}\right) + Var\left(U_{n1}-\hat{U}_{n1}\right) \\
&= Var\left(\hat{U}_{n1}\right) + \frac{4}{n^2\left(n-1\right)^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Var\left(e_{ij}\right) \\
&= Var\left(\hat{U}_{n1}\right) + \; O\left(n^{-2}\left|H\right|^{-1}\right).
\end{aligned}
$$

Secondly, we consider $Var\left(U_{n2}\right)$, where we decompose U_{n2} as

$$
U_{n2} - E\left(U_{n2}\right) = \hat{U}_{n2} + \frac{1}{n\left(n-1\right)}\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{l=j+1}^{n} e_{ijl}
$$

where $e_{ijl} = p_2^*\left(\chi_i,\chi_j,\chi_l\right) - E\left[p_2^*\left(\chi_i,\chi_j,\chi_l\right)|\chi_i\right] - E\left[p_2^*\left(\chi_i,\chi_j,\chi_l\right)|\chi_j\right] - E\left[p_2^*\left(\chi_i,\chi_j,\chi_l\right)|\chi_l\right]$ is symmetric in (i,j,l). As $Cov\left(\hat{U}_{n2}, U_{n2}-\hat{U}_{n2}\right) = 0$ and $Cov\left(e_{ijl}, e_{i'j'l'}\right) = 0$ if (i,j,l) and

(i', j', l') have two or more indices different, we have

$$\begin{aligned} & Var(U_{n2}) = Var\left(\hat{U}_{n2}\right) + Var\left(U_{n2} - \hat{U}_{n2}\right) \\ = \; & Var\left(\hat{U}_{n2}\right) + \frac{1}{n^2(n-1)^2} \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{l=j+1}^{n} Var(e_{ijl}) \\ & + \frac{2}{n^2(n-1)^2} \sum_{i<j<j} \sum_{(i',j',l') \neq (i,j,l)} Cov(e_{ijl}, e_{i'j'l'}) \\ = \; & Var\left(\hat{U}_{n2}\right) + O\left(n^{-2}|H|^{-1}\right). \end{aligned}$$

Thirdly, we consider U_{n3}, where we decompose U_{n3} as

$$U_{n3} - E(U_{n3}) = \hat{U}_{n3} + \frac{1}{n(n-1)} \sum_{i=1}^{n-3} \sum_{j=i+1}^{n-2} \sum_{l=j+1}^{n-1} \sum_{l'=l+1}^{n} e_{ijll'}$$

where

$$\begin{aligned} e_{ijll'} = \; & p_3^*(\chi_i, \chi_j, \chi_l, \chi_{l'}) - E\left[p_3^*(\chi_i, \chi_j, \chi_l, \chi_{l'})|\chi_i\right] - E\left[p_3^*(\chi_i, \chi_j, \chi_l, \chi_{l'})|\chi_j\right] \\ & - E\left[p_3^*(\chi_i, \chi_j, \chi_l, \chi_{l'})|\chi_l\right] - E\left[p_3^*(\chi_i, \chi_j, \chi_l, \chi_{l'})|\chi_{l'}\right] \end{aligned}$$

is symmetric in (i, j, l, l'). As $Cov\left(\hat{U}_{n3}, U_{n3} - \hat{U}_{n3}\right) = 0$ and $Cov\left(e_{ijll'}, e_{i_1j_1l_1l_1'}\right) = 0$ if (i, j, l, l') and (i_1, j_1, l_1, l_1') have three or more indices different, we obtain

$$\begin{aligned} & Var(U_{n3}) = Var\left(\hat{U}_{n3}\right) + Var\left(U_{n3} - \hat{U}_{n3}\right) \\ = \; & Var\left(\hat{U}_{n3}\right) + \frac{1}{n^2(n-1)^2} Var\left(\sum_{i=1}^{n-3} \sum_{j=i+1}^{n-2} \sum_{l=j+1}^{n-1} \sum_{l'=l+1}^{n} e_{ijll'}\right) \\ & + \frac{1}{n^2(n-1)^2} \sum_{i<j<k<l'} \sum_{(i,j,k,l') \neq \left(i_1,j_1,k_1,l_1'\right)} Cov\left(e_{ijll'}, e_{i_1j_1l_1l_1'}\right) \\ = \; & Var\left(\hat{U}_{n3}\right) + O\left(n^{-2}|H|^{-1}\right). \end{aligned}$$

This completes the proof of this lemma.

Lemma A.3 *Under Assumptions 1-2, we have* $\sqrt{n}\hat{U}_n \xrightarrow{d} N\left(0, \sigma_\psi^2\right)$, *where* σ_ψ^2 *is defined in (A.4).*

Proof: By Lemma A.1 we have

$$\hat{U}_n = \hat{U}_{n1} + \hat{U}_{n2} + \hat{U}_{n3} \approx \frac{1}{n} \sum_{i=1}^{n} \psi_i$$

where

$$\psi_i = \frac{2\mu_v^2\varsigma_{1i}}{n-1}\eta_{1i} + \frac{2\mu_v\omega_{1i}}{n-1}\eta_{2i} + \frac{2\mu_v\omega_{2i}}{n-1}\eta_{3i} + \left(\frac{b_{ii}}{n-1}\sum_{j\neq i} b_{jj}\right)\varphi_{3i} + \frac{2\mu_v E\left[f\left(D_1\right)\right]}{n-1}\varsigma_{2i}\left(v_i-\mu_v\right)$$
$$= \gamma_{1i}\eta_{1i} + \gamma_{2i}\eta_{2i} + \gamma_{3i}\eta_{3i} + \gamma_{4i}\varphi_{3i} + \gamma_{5i}\left(v_i-\mu_v\right)$$

is independent but not identically distributed across index i and has a zero mean and finite variance $\max_{1\leq i\leq n} E\left(\psi_i^2\right) \leq M$ under Assumptions 1-2, and the definition of γ_{ji} for j=1,2,3,4,5 should be apparent from the context.

For some $\delta > 0$ we have

$$E\left(\left|\psi_i\right|^{2+\delta}\right) \leq 2^{\delta+1}\left[\sum_{j=1}^{3}\left|\gamma_{ji}\right|^{2+\delta} E\left(\left|\eta_{ji}\right|^{2+\delta}\right) + \left|\gamma_{4i}\right|^{2+\delta} E\left(\left|\varphi_{3i}\right|^{2+\delta}\right) + \left|\gamma_{5i}\right|^{2+\delta} E\left(\left|v_i-\mu_v\right|^{2+\delta}\right)\right]$$

holds under Assumption 1, where we apply the c_r inequality (White, 2004, p.35) to obtain the above inequality. Also, it is readily seen that

$$\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} E\left(\psi_i^2\right) = \frac{1}{n}\sum_{i=1}^{n} E\left[\gamma_{1i}\eta_{1i} + \gamma_{2i}\eta_{2i} + \gamma_{3i}\eta_{3i} + \gamma_{4i}\varphi_{3i} + \gamma_{5i}\left(v_i-\mu_v\right)\right]^2 \overset{n\to\infty}{\to} \sigma_\psi^2 > 0, \tag{A.4}$$

where σ_ψ^2 is a finite positive value under Assumption 1. Hence, applying the Liapounov's CLT we obtain

$$\sqrt{n}\hat{U}_n \overset{d}{\to} N\left(0, \sigma_\psi^2\right).$$

Note that we have

$$\psi_i = \left(\frac{b_{ii}}{n-1}\sum_{j\neq i} b_{jj}\right)\varphi_{3i}$$

if B_n is a diagonal matrix, under which case U_n becomes a standard weighted second-order U statistic for i.i.d. data and

$$\bar{\sigma}_n^2 = \frac{c_b^2}{n}\sum_{i=1}^{n} b_{ii}^2 Var\left(v_1\left(m_1 f\right)\left(D_1\right)\right) \overset{n\to\infty}{\to} c_b^2 c_{b,2} Var\left(v_1\left(m_1 f\right)\left(D_1\right)\right)$$

where $c_{b,2} = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} b_{ii}^2$. This completes the proof of this lemma.

Proof of Theorem 3.1: Let $\hat{u}$ be an $n\times 1$ vector that stacks up $\{\hat{u}_i\}$ and $G_{n,0}$ be the $n\times n$ spatial weight matrix with its (i,j)th element being $g_0\left(\mathcal{Z}_{ij}\right)$. Then, we have $\hat{u} = \left(I_n - \hat{\rho}G_{n,0}\right)Y - X\hat{\theta} = \left(I_n - \hat{\rho}G_{n,0}\right)\left(I_n - G_n\right)^{-1}\left(\text{mtk}\left\{X, \theta\left(D\right)\right\} + u\right) - X\hat{\theta} = S_n u + \chi_n$, where

$$S_n = \left(I_n - \hat{\rho}G_{n,0}\right)\left(I_n - G_n\right)^{-1} \text{ and } \chi_n = S_n \text{mtk}\left\{X, \theta\left(D\right)\right\} - X\hat{\theta}.$$

Then, we have

$$T_n = \frac{1}{n^2\left|H\right|}\hat{u}'\mathbf{K}_H\hat{u} = T_{n1} + 2T_{n2} + T_{n3} \tag{A.5}$$

where

$$T_{n1}=\frac{1}{n^{2}\left|H\right|}u'S_n'\mathbf{K}_HS_nu,\ T_{n2}=\frac{1}{n^{2}\left|H\right|}\chi_n'\mathbf{K}_HS_nu \text{ and } T_{n3}=\frac{1}{n^{2}\left|H\right|}\chi_n'\mathbf{K}_H\chi_n. \tag{A.6}$$

By Lemmas A.4-A.6 derived below, we have

$$n\sqrt{|H|}T_n=n\sqrt{|H|}T_{n1}+O_p\left(\sqrt{|H|}\right)\xrightarrow{d}N\left(0,2\sigma_u^2\pi_G^2\kappa_{2,0}E\left[f\left(D_1\right)\right]\right)$$

where π_G is given in (A.10) and takes different values when $H_0^1\cap H_0^2$ hold and when $H_0^1\cap H_0^2$ fails to hold. Also, when $H_0^1\cap H_0^2$ fails to hold, $n\sqrt{|H|}T_n=n\sqrt{|H|}T_{n3}+O_p\left(\sqrt{n\left|H\right|}\right)$ $=O_p\left(n\sqrt{|H|}\right)$ and $T_{n3}=E\left(T_{n3,1}\right)+O_p\left(n^{-1/2}\right)$ with $E\left(T_{n3,1}\right)$ converging to a positive value. Combining this result with Lemma A.7 completes the proof of this theorem.

Lemma A.4 *Under Assumptions 2-5 and $\pi_G\neq 0$, we have $n\sqrt{|H|}T_{n1}\xrightarrow{d}N\left(0,2\sigma_u^2\pi_G^2\kappa_{2,0}E\left[f\left(D_1\right)\right]\right)$, where π_G is defined in (A.10), and $\pi_G=1$ under $H_0^1\cap H_0^2$.*

Proof: Under $H_0^1\cap H_0^2$, $G_n=\rho_0G_{n,0}$, so we have

$$S_n=I_n-\left(\hat\rho-\rho_0\right)G_{n,0}\left(I_n-\rho_0G_{n,0}\right)^{-1}. \tag{A.7}$$

Note that $\hat\rho-\rho_0=O_p\left(n^{-1/2}\right)$ under Assumption 5(i). Then, we obtain

$$\begin{aligned} n\sqrt{|H|}T_{n1} &= \frac{1}{n\sqrt{|H|}}u'\mathbf{K}_Hu-\frac{2\left(\hat\rho-\rho_0\right)}{n\sqrt{|H|}}u'\mathbf{K}_H\left[G_{n,0}\left(I_n-\rho_0G_{n,0}\right)^{-1}\right]u \\ &+\frac{\left(\hat\rho-\rho_0\right)^2}{n\sqrt{|H|}}u'\left[G_{n,0}\left(I_n-\rho_0G_{n,0}\right)^{-1}\right]'\mathbf{K}_H\left[G_{n,0}\left(I_n-\rho_0G_{n,0}\right)^{-1}\right]u \\ &= \frac{1}{n\sqrt{|H|}}u'\mathbf{K}_Hu+O_p\left(n^{-1/2}\right)\end{aligned}$$

where Theorem 2.1 is used to derive the above result. Then, applying Hall's (1984) CLT gives

$$\frac{1}{n\sqrt{|H|}}u'\mathbf{K}_Hu\xrightarrow{d}N\left(0,2\sigma_u^2\kappa_{2,0}E\left[f\left(D_1\right)\right]\right) \tag{A.8}$$

under Assumptions 1-2.

Second, if $H_0^1\cap H_0^2$ fails to hold, we have

$$S_n=I_n-\left(\hat\rho-\rho\right)G_{n,0}\left(I_n-G_n\right)^{-1}-\left(\rho G_{n,0}-G_n\right)\left(I_n-G_n\right)^{-1}. \tag{A.9}$$

Then, under Assumptions 2-5, applying Theorem 2.1 gives

$$\begin{aligned}
& n\sqrt{|H|}T_{n1} \\
= & \frac{1}{n\sqrt{|H|}}u'\left[I_n-(\rho G_{n,0}-G_n)(I_n-G_n)^{-1}\right]\mathbf{K}_H\left[I_n-(\rho G_{n,0}-G_n)(I_n-G_n)^{-1}\right]u \\
& +\frac{2(\hat{\rho}-\rho)}{n\sqrt{|H|}}u'\left[G_{n,0}(I_n-G_n)^{-1}\right]\mathbf{K}_H\left[I_n-(\rho G_{n,0}-G_n)(I_n-G_n)^{-1}\right]u \\
& +\frac{(\hat{\rho}-\rho)^2}{n\sqrt{|H|}}u'\left[G_{n,0}(I_n-G_n)^{-1}\right]\mathbf{K}_H\left[G_{n,0}(I_n-G_n)^{-1}\right]u \\
= & \frac{1}{n\sqrt{|H|}}u'\left[I_n-(\rho G_{n,0}-G_n)(I_n-G_n)^{-1}\right]\mathbf{K}_H\left[I_n-(\rho G_{n,0}-G_n)(I_n-G_n)^{-1}\right]u+O_p\left(n^{-1/2}\right) \\
& \xrightarrow{d} N\left(0,2\sigma_u^2\pi_G^2\kappa_{2,0}E\left[f(D_1)\right]\right)
\end{aligned}$$

where

$$\pi_G=1-n^{-1}\lim_{n\to\infty}\operatorname{tr}\left((\rho G_{n,0}-G_n)(I_n-G_n)^{-1}\right). \tag{A.10}$$

This completes the proof of this lemma.

Lemma A.5 *Under Assumptions 2-5, we have* $n\sqrt{|H|}T_{n2}=O_p\left(\sqrt{|H|}\right)$ *if* $H_0^1\cap H_0^2$ *holds, and* $n\sqrt{|H|}T_{n2}=O_p\left(\sqrt{n|H|}\right)$ *otherwise.*

Proof: Under $H_0^1\cap H_0^2$ and by (A.7), we have

$$\chi_n=S_nX\theta_0-X\hat{\theta}=X\left(\theta_0-\hat{\theta}\right)-(\hat{\rho}-\rho_0)G_{n,0}(I_n-\rho_0G_{n,0})^{-1}X\theta_0. \tag{A.11}$$

It follows that

$$\begin{aligned}
n\sqrt{|H|}T_{n2}= & \frac{\left(\theta_0-\hat{\theta}\right)'}{n\sqrt{|H|}}X'\mathbf{K}_Hu-\frac{\hat{\rho}-\rho_0}{n\sqrt{|H|}}\left(\theta_0-\hat{\theta}\right)'X'\mathbf{K}_HA_{n,0}u \\
& -\frac{\hat{\rho}-\rho_0}{n\sqrt{|H|}}(A_{n,0}X\theta_0)'\mathbf{K}_Hu+\frac{(\hat{\rho}-\rho_0)^2}{n\sqrt{|H|}}(A_{n,0}X\theta_0)'\mathbf{K}_HA_{n,0}u
\end{aligned}$$

where we denote $A_{n,0}=G_{n,0}(I_n-\rho_0G_{n,0})^{-1}$. It is straightforward to show that

$$X'\mathbf{K}_Hu=O_p\left(n^{3/2}|H|\right),\ X'\mathbf{K}_HA_{n,0}u=O_p\left(n^{3/2}|H|\right),\text{ and }(A_{n,0}X)'\mathbf{K}_Hu=O_p\left(n^{3/2}|H|\right) \tag{A.12}$$

under Assumptions 2-4. Also, we have $E\left[(A_{n,0}X\theta_0)'\mathbf{K}_HA_{n,0}u\right]=0$ and

$$\begin{aligned}
& E\left|(A_{n,0}X\theta_0)'\mathbf{K}_HA_{n,0}u\right|^2=\sigma_u^2E\left[(A_{n,0}X\theta_0)'\mathbf{K}_HA_{n,0}A_{n,0}'\mathbf{K}_HA_{n,0}X\theta_0\right] \\
\leq\ & \sigma_u^2\|A_{n,0}\|_{sp}^2\theta_0'E\left(X'A_{n,0}'\mathbf{K}_H^2A_{n,0}X\right)\theta_0 \\
=\ & \sigma_u^2\left\|G_{n,0}(I_n-\rho_0G_{n,0})^{-1}\right\|_{sp}^2\sum_{i=1}^nE\left(\sum_{j\neq i}\sum_{l\neq j}g_{0,jl}\theta_0'\tilde{X}_lK_{H,ij}\right)^2=O\left(n^3|H|^2\right),
\end{aligned}$$

which imply that

$$\left(A_{n,0}X\theta_0\right)'\mathbf{K}_H A_{n,0}u = O_p\left(n^{3/2}\left|H\right|\right). \tag{A.13}$$

Therefore, we obtain

$$n\sqrt{|H|}T_{n2} = O_p\left(\frac{n^{-1/2}}{n\sqrt{|H|}}\right)O_p\left(n^{3/2}\left|H\right|\right) = O_p\left(\sqrt{|H|}\right).$$

Next, if $H_0^1 \cap H_0^2$ fails to hold, by (A.9) we have

$$\begin{aligned}
\chi_n &= S_n \mathrm{mtk}\left(X, \theta\left(D\right)\right) - X\hat{\theta} \\
&= \mathrm{mtk}\left(X, \theta\left(D\right) - \hat{\theta}\right) - \left(\hat{\rho} - \rho\right)G_{n,0}\left(I_n - G_n\right)^{-1}\mathrm{mtk}\left(X, \theta\left(D\right)\right) \\
&\quad - \left(\rho G_{n,0} - G_n\right)\left(I_n - G_n\right)^{-1}\mathrm{mtk}\left(X, \theta\left(D\right)\right)
\end{aligned} \tag{A.14}$$

and it follows that

$$\begin{aligned}
& n\sqrt{|H|}T_{n2} \\
=\; & \frac{1}{n\sqrt{|H|}}\left[\mathrm{mtk}\left(X, \theta\left(D\right) - \theta\right)\right]'\mathbf{K}_H S_n u \\
& + \frac{1}{n\sqrt{|H|}}\left(\theta - \hat{\theta}\right)' X'\mathbf{K}_H S_n u - \frac{\hat{\rho} - \rho}{n\sqrt{|H|}}\left[G_{n,0}\left(I_n - G_n\right)^{-1}\mathrm{mtk}\left(X, \theta\left(D\right)\right)\right]'\mathbf{K}_H S_n u \\
& - \frac{1}{n\sqrt{|H|}}\left[\left(\rho G_{n,0} - G_n\right)\left(I_n - G_n\right)^{-1}\mathrm{mtk}\left(X, \theta\left(D\right)\right)\right]'\mathbf{K}_H S_n u \\
=\; & \frac{1}{n\sqrt{|H|}}O_p\left(n^{3/2}\left|H\right|\right) = O_p\left(\sqrt{n\left|H\right|}\right)
\end{aligned}$$

where $\hat{\rho} - \rho = O_p\left(n^{-1/2}\right)$ and $\hat{\theta} - \theta = O_p\left(n^{-1/2}\right)$ under Assumption 5(ii), and the similar method used to prove (A.12) and (A.13) is applied to obtain our results here. This completes the proof of this lemma.

Lemma A.6 *Under Assumptions 2-5, we have* $n\sqrt{|H|}T_{n3} = O_p\left(\sqrt{|H|}\right)$ *if* $H_0^1 \cap H_0^2$ *holds, and* $n\sqrt{|H|}T_{n3} = O_p\left(n\sqrt{|H|}\right)$ *otherwise.*

Proof: Under $H_0^1 \cap H_0^2$, by (A.11) and Theorem 2.2, we have

$$\begin{aligned}
T_{n3} &= \frac{\left(\theta_0 - \hat{\theta}\right)'}{n^2\left|H\right|}X'\mathbf{K}_H X\left(\theta_0 - \hat{\theta}\right) \\
&\quad - \frac{2\left(\hat{\rho} - \rho_0\right)}{n^2\left|H\right|}\left(\theta_0 - \hat{\theta}\right)' X'\mathbf{K}_H G_{n,0}\left(I_n - \rho_0 G_{n,0}\right)^{-1}X\theta_0 \\
&\quad + \frac{\left(\hat{\rho} - \rho_0\right)^2}{n^2\left|H\right|}\left[G_{n,0}\left(I_n - G_{n,0}\right)^{-1}X\theta_0\right]'\mathbf{K}_H G_{n,0}\left(I_n - \rho_0 G_{n,0}\right)^{-1}X\theta_0 \\
&= O_p\left(n^{-1}\right).
\end{aligned}$$

If $H_0^1 \cap H_0^2$ fails to hold, by (A.14) we have $T_{n3} = T_{n3,1} + O_p\left(n^{-1/2}\right)$, where

$$\begin{aligned} T_{n3,1} &= \frac{1}{n^2 |H|} \left[\text{mtk}\left(X, \theta\left(D\right) - \theta\right) - \left(\rho G_{n,0} - G_n\right)\left(I_n - G_n\right)^{-1} \text{mtk}\left(X, \theta\left(D\right)\right)\right]' \mathbf{K}_H \\ &\quad \times \left[\text{mtk}\left(X, \theta\left(D\right) - \theta\right) - \left(\rho G_{n,0} - G_n\right)\left(I_n - G_n\right)^{-1} \text{mtk}\left(X, \theta\left(D\right)\right)\right] \end{aligned} \tag{A.15}$$

and applying Theorem 2.2 gives $T_{n3,1} = E\left(T_{n3,1}\right) + O_p\left(n^{-1/2}\right)$ and $E\left(T_{n3,1}\right)$ evidently converges to a positive value. This completes the proof of this lemma.

Lemma A.7 *Under Assumptions 2-5, we have* $\hat{\sigma}_n^2 \xrightarrow{p} \sigma_0^2$*, where* $\sigma_0^2 = 2\sigma_u^4 \kappa_{2,0} E\left[f\left(D_1\right)\right]$*, under* $H_0^1 \cap H_0^2$*, and* $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$ *otherwise, where* σ^2 *is defined in eq. (A.16) below.*

Proof: Under $H_0^1 \cap H_0^2$, by (A.7) and (A.11) we have $\hat{u} = X\left(\theta_0 - \hat{\theta}\right) - \left(\hat{\rho} - \rho_0\right) G_{n,0}\left(I_n - \rho_0 G_{n,0}\right)^{-1}\left(X\theta_0\right.$ u, where $\hat{\theta} - \theta_0 = O_p\left(n^{-1/2}\right)$ and $\hat{\rho} - \rho_0 = O_p\left(n^{-1/2}\right)$. Then, under Assumptions 2-5, we can easily show that

$$\begin{aligned} \hat{\sigma}_n^2 &= \frac{2}{n^2 |H|} \sum_{i=1}^n \sum_{j \neq i} \hat{u}_i^2 \hat{u}_j^2 K^2\left(H^{-1}\left(D_i - D_j\right)\right) \\ &= \frac{2}{n^2 |H|} \sum_{i=1}^n \sum_{j \neq i} K_{H,ij}^2 u_i^2 u_j^2 + O_p\left(\sqrt{|H|}\right) \\ &= \sigma_0^2 + o_p(1) \end{aligned}$$

where the last line is obtained by following the proof of Lemma 3.3e in Zheng (1996) and applying Lemma 3.1 in Powell et al. (1989).

If $H_0^1 \cap H_0^2$ fails to hold, by (A.9) and (A.14) and denoting $V = \left[I_n - \left(\rho G_{n,0} - G_n\right)\left(I_n - G_n\right)^{-1}\right] u$ $+\text{mtk}\{X, \theta\left(D\right) - \theta\} - \left(\rho G_{n,0} - G_n\right)\left(I_n - G_n\right)^{-1} \text{mtk}\{X, \theta\left(D\right)\}$, we have

$$\begin{aligned} \hat{u} &= \left(I_n - \hat{G}_{n,0}\right)\left(I_n - G_n\right)^{-1}\left(\text{mtk}\left\{X, \theta\left(D\right)\right\} + u\right) - X\hat{\theta} \\ &= V - X\left(\hat{\theta} - \theta\right) - \left(\hat{\rho} - \rho\right) G_{n,0}\left(I_n - G_n\right)^{-1}\left(\text{mtk}\left\{X, \theta\left(D\right)\right\} + u\right) = V + O_p\left(n^{-1/2}\right) \end{aligned}$$

under Assumption 5. We then can show that

$$\hat{\sigma}_n^2 = \frac{2}{n^2 |H|} \sum_{i=1}^n \sum_{j \neq i} K_{H,ij}^2 v_i^2 v_j^2 + o_p(1)$$

where v_i is the ith element of V. Applying the similar method used to prove Theorem 2.2 we can show that $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$ with

$$\sigma^2 = 2\kappa_{2,0} \lim_{n \to \infty} \frac{1}{n^2} \sum_{i=1}^n \sum_{j \neq i} E\left[c_{ij}\left(D_1, D_1\right) f\left(D_1\right)\right] \tag{A.16}$$

being a finite positive constant, where $c_{ij}(d_1, d_2) = E\left(v_i^2 v_j^2 | D_i = d_1, D_j = d_2\right)$. This completes the proof of this lemma.

Proof of Theorem 3.2: The residual calculated under H_0^1 becomes $\hat{u} = \left(I_n - \hat{G}_n\right) Y - X\hat{\theta} = S_n u + \chi_n$, where $S_n = \left(I_n - \hat{G}_n\right)(I_n - G_n)^{-1} = I_n - \left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1}$, $\chi_n = S_n \text{mtk}\{X, \theta(D)\} - X\hat{\theta}$, and $\hat{G}_n$ is an $n \times n$ matrix with zero diagonal elements and its (i,j)th element equal to $\hat{g}_{ij} = \hat{g}(\mathcal{Z}_{ij})$ for any $i \neq j$. The decomposition of our test statistic given in (A.5)-(A.6) continues to hold. By Lemmas A.8-A.11, we complete the proof of this theorem.

Lemma A.8 *Under Assumptions 2-4 and 6-8, we have* $n\sqrt{|H|}T_{n1} \xrightarrow{d} N\left(0, 2\sigma_u^2 \pi_G^2 \kappa_{2,0} E\left[f(D_1)\right]\right)$ *where* π_G *is defined in (A.18) and* $\pi_G = 1$ *under* H_0^1.

Proof: Under H_0^1, we have $\hat{G}_n - G_n = \hat{G}_n - G_{0,n}^* + G_{0,n}^* - G_n$, where the typical element of $\hat{G}_n - G_{0,n}^*$ is $\hat{g}_{ij} - g_{0,ij}^* = (\hat{\alpha} - \alpha_0)' \boldsymbol{\phi}_{L_n}(\mathcal{Z}_{ij})$ and the typical element of $G_{0,n}^* - G_n$ is $g_{0,ij}^* - g_{ij} = g_0^*(\mathcal{Z}_{ij}) - g(\mathcal{Z}_{ij})$, and $\|\hat{\alpha} - \alpha_0\| = O_p(\vartheta_n)$ under Assumption 7(i). Now, we have

$$n\sqrt{|H|}T_{n1} = A_{n1} - 2A_{n2} + A_{n3}$$

where the limiting distribution of $A_{n1} = \left(n\sqrt{|H|}\right)^{-1} u' \mathbf{K}_H u$ is given in (A.8).

Next, we consider

$$\begin{aligned} A_{n2} &= \frac{1}{n\sqrt{|H|}} u' \mathbf{K}_H \left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1} u \\ &= \frac{1}{n\sqrt{|H|}} u' \mathbf{K}_H \left(\hat{G}_n - G_{0,n}^*\right)(I_n - G_n)^{-1} u + \frac{1}{n\sqrt{|H|}} u' \mathbf{K}_H \left(G_{0,n}^* - G_n\right)(I_n - G_n)^{-1} u \\ &= (\hat{\alpha} - \alpha_0)' A_{n2,1} + A_{n2,2} = O_p\left(\vartheta_n \sqrt{L_n |H|}\right) + O_p\left(L_n^{-\xi}\right) \qquad \text{(A.17)} \end{aligned}$$

where under Assumptions 2-4 and 6 we obtain

$$\begin{aligned} A_{n2,2} &= \frac{1}{n\sqrt{|H|}} \sum_{i=1}^n \sum_{j \neq i} K_{H,ij} u_i \sum_{l \neq j} \left(g_{0,jl}^* - g_{jl}\right) \tilde{u}_l \\ &= \frac{1}{n\sqrt{|H|}} \sum_{i=1}^n u_i^2 \sum_{j \neq i} K_{H,ij} \sum_{l \neq j} a_{li} \left(g_{0,jl}^* - g_{jl}\right) \\ &\quad + \frac{1}{n\sqrt{|H|}} \sum_{i=1}^n \sum_{l' \neq i} u_i u_{l'} \sum_{j \neq i} K_{H,ij} \sum_{l \neq j} a_{ll'} \left(g_{0,jl}^* - g_{jl}\right) \\ &= O_p\left(\sqrt{|H|} L_n^{-\xi}\right) + O_p\left(L_n^{-\xi}\right) = O_p\left(L_n^{-\xi}\right) \end{aligned}$$

and

$$
\begin{aligned}
A_{n2,1} &= \frac{1}{n\sqrt{|H|}}\sum_{i=1}^{n}\sum_{j\neq i}K_{H,ij}u_i\sum_{l\neq j}\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{jl}\right)\tilde{u}_l \\
&= \frac{1}{n\sqrt{|H|}}\sum_{i=1}^{n}u_i^2\sum_{j\neq i}K_{H,ij}\sum_{l\neq j}a_{li}\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{jl}\right) \\
&\quad +\frac{1}{n\sqrt{|H|}}\sum_{i=1}^{n}\sum_{l'\neq i}u_iu_{l'}\sum_{j\neq i}K_{H,ij}\sum_{l\neq j}a_{ll'}\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{jl}\right) \\
&= A_{n2,1,1}+A_{n2,1,2}=O_p\left(\sqrt{L_n\left|H\right|}\right)
\end{aligned}
$$

because we have

$$
\begin{aligned}
&E\left(\left\|A_{n2,1,1}\right\|\right)\leq\frac{\sigma_u^2}{n\sqrt{|H|}}\sum_{i=1}^{n}\sum_{j\neq i}E\left(K_{H,ij}\right)\left\|\sum_{l\neq j}a_{li}\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{jl}\right)\right\| \\
\leq\ & M\frac{\sqrt{|H|}}{n}\left[n^2\sum_{i=1}^{n}\sum_{j\neq i}\left\|\sum_{l\neq j}a_{li}\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{jl}\right)\right\|^2\right]^{1/2} \\
=\ & M\frac{\sqrt{|H|}}{n}\left[n^2\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l\neq j}\sum_{l'\neq j}a_{li}a_{l'i}\boldsymbol{\phi}_{L_n}'\left(\mathcal{Z}_{jl}\right)\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{jl'}\right)\right]^{1/2}=O\left(\sqrt{L_n\left|H\right|}\right)
\end{aligned}
$$

by the fact that $\left(\sum_{i=1}^n z_i\right)^2\leq n\sum_{i=1}^n z_i^2$, $\max_{1\leq i\leq n,1\leq s\leq L_n}\sum_{j\neq i}\left|\phi_s\left(\mathcal{Z}_{ij}\right)\right|\leq M$ by Assumptions 4(i) and 6(ii) and $\left\|\left(I_n-G_n\right)^{-1}\right\|_1\leq M$, and $E\left(A_{n2,1,2}\right)=0$ and

$$
\begin{aligned}
&E\left\|A_{n2,1,2}\right\|^2 \\
=\ & \frac{\sigma_u^4}{n^2\left|H\right|}\sum_{i=1}^{n}\sum_{l'\neq i}\sum_{j\neq i}\sum_{j_1\neq i\neq j}E\left(K_{H,ij}K_{H,ij_1}\right)\sum_{l\neq j}a_{ll'}\sum_{l_1\neq j_1}a_{l_1l'}\boldsymbol{\phi}_{L_n}'\left(\mathcal{Z}_{jl}\right)\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{j_1l_1}\right) \\
&+\frac{\sigma_u^4}{n^2\left|H\right|}\sum_{i=1}^{n}\sum_{l'\neq i}\sum_{j\neq i}E\left(K_{H,ij}^2\right)\sum_{l\neq j}a_{ll'}\sum_{l_1\neq j}a_{l_1l'}\boldsymbol{\phi}_{L_n}'\left(\mathcal{Z}_{jl}\right)\boldsymbol{\phi}_{L_n}\left(\mathcal{Z}_{jl_1}\right) \\
=\ & O\left(\frac{n^2\left|H\right|^2L_n}{n^2\left|H\right|}\right)+O\left(\frac{n\left|H\right|L_n}{n^2\left|H\right|}\right)=O\left(L_n\left(\left|H\right|+n^{-1}\right)\right).
\end{aligned}
$$

Lastly, we consider

$$
\begin{aligned}
A_{n3} &= \frac{1}{n\sqrt{|H|}}u'\left[\left(\hat{G}_n-G_n\right)\left(I_n-G_n\right)^{-1}\right]'\mathbf{K}_H\left(\hat{G}_n-G_n\right)\left(I_n-G_n\right)^{-1}u \\
&= A_{n3,1}+2A_{n3,2}+A_{n3,3}
\end{aligned}
$$

where

$$\begin{aligned}
A_{n3,1} &= \frac{1}{n\sqrt{|H|}}u'\left[\left(\hat{G}_n - G_n^*\right)(I_n - G_n)^{-1}\right]' \mathbf{K}_H \left(\hat{G}_n - G_n^*\right)(I_n - G_n)^{-1} u \\
&= \frac{(\hat{\alpha}-\alpha_0)'}{n\sqrt{|H|}} \sum_{i=1}^{n}\sum_{j\neq i} K_{H,ij} \sum_{l\neq i}\sum_{l'\neq j\neq l} \boldsymbol{\phi}_{L_n}(\mathcal{Z}_{il})\,\boldsymbol{\phi}'_{L_n}(\mathcal{Z}_{jl'})\,\tilde{u}_l\tilde{u}_{l'}(\hat{\alpha}-\alpha_0) \\
&\quad + \frac{(\hat{\alpha}-\alpha_0)'}{n\sqrt{|H|}} \sum_{i=1}^{n}\sum_{j\neq i} K_{H,ij} \sum_{l\neq i\neq j} \boldsymbol{\phi}_{L_n}(\mathcal{Z}_{il})\,\boldsymbol{\phi}'_{L_n}(\mathcal{Z}_{jl})\,\tilde{u}_l^2(\hat{\alpha}-\alpha_0) \\
&= O_p\left(\vartheta_n^2 L_n \sqrt{|H|}\right),
\end{aligned}$$

$$\begin{aligned}
A_{n3,2} &= \frac{1}{n\sqrt{|H|}}u'\left[\left(\hat{G}_n - G_n^*\right)(I_n - G_n)^{-1}\right]' \mathbf{K}_H (G_n^* - G_n)(I_n - G_n)^{-1} u \\
&\leq \sqrt{A_{n3,1}A_{n3,2}} = O_p\left(\left(\vartheta_n^2 L_n\sqrt{|H|}\right)^{1/2}\right)\ O_p\left(\left(L_n^{-2\xi}\sqrt{|H|}\right)^{1/2}\right) \\
&= O_p\left(\vartheta_n L_n^{1/2-\xi}\sqrt{|H|}\right)
\end{aligned}$$

by Cauchy-Schwarz inequality, and

$$\begin{aligned}
A_{n3,3} &= \frac{1}{n\sqrt{|H|}}u'\left[(G_n^* - G_n)(I_n - G_n)^{-1}\right]' \mathbf{K}_H (G_n^* - G_n)(I_n - G_n)^{-1} u \\
&= O\left(L_n^{-2\xi}\sqrt{|H|}\right)
\end{aligned}$$

because we have

$$\begin{aligned}
E(A_{n3,3}) &= \frac{1}{n\sqrt{|H|}}\mathrm{tr}\left\{E\left\{u'\left[(G_n^* - G_n)(I_n - G_n)^{-1}\right]' \mathbf{K}_H (G_n^* - G_n)(I_n - G_n)^{-1} u\right\}\right\} \\
&= \frac{\sigma_u^2}{n\sqrt{|H|}}\mathrm{tr}\left\{E(\mathbf{K}_H)(G_n^* - G_n)(I_n - G_n)^{-1}\left[(G_n^* - G_n)(I_n - G_n)^{-1}\right]'\right\} \\
&\leq \frac{\sigma_u^2}{n\sqrt{|H|}}\mathrm{tr}\left\{(G_n^* - G_n)' E(\mathbf{K}_H)(G_n^* - G_n)\right\}\lambda_{\max}\left\{\left[(I_n - G_n)^{-1}\right]'(I_n - G_n)^{-1}\right\} \\
&\leq \frac{M}{n\sqrt{|H|}}\sum_{l=1}^{n}\sum_{i=1}^{n}\sum_{j\neq i} E(K_{H,ji})\left(g_{jl}^* - g_{jl}\right)(g_{il}^* - g_{il}) \\
&= O\left(\frac{\sqrt{|H|}}{n} n L_n^{-2\xi}\right) = O\left(L_n^{-2\xi}\sqrt{|H|}\right).
\end{aligned}$$

Hence, $A_{n3} = O_p\left(\vartheta_n^2 L_n\sqrt{|H|}\right) + O_p\left(L_n^{-2\xi}\sqrt{|H|}\right) = o_p(1)$ under Assumption 8.

Under H_1^1, we have $\hat{G}_n - G_n = \hat{G}_n - G_n^* + G_n^* - G_n$, where the typical element of $\hat{G}_n - G_n^*$ is $\hat{g}_{ij} - g_{ij}^* = (\hat{\alpha}-\alpha)'\boldsymbol{\phi}_{L_n}(\mathcal{Z}_{ij})$ and the typical element of $G_n^* - G_n$ is

$$g_{ij}^* - g_{ij} = g^*(\mathcal{Z}_{ij}) - g_0^*(\mathcal{Z}_{ij}) + g_0^*(\mathcal{Z}_{ij}) - g(\mathcal{Z}_{ij}) = (\alpha-\alpha_0)'\boldsymbol{\phi}_{L_n}(\mathcal{Z}_{ij}) + g_0^*(\mathcal{Z}_{ij}) - g(\mathcal{Z}_{ij}),$$

and $\|\hat{\alpha}-\alpha\| = O_p(\vartheta_n)$ under Assumption 7(ii). Applying the proof method used above gives

$$\begin{aligned}
& n\sqrt{|H|}T_{n1} \\
\approx\ & \frac{1}{n\sqrt{|H|}}u'\left[I_n - \left(G_n^* - G_{0,n}^*\right)(I_n - G_n)^{-1}\right]'\mathbf{K}_H\left[I_n - \left(G_n^* - G_{0,n}^*\right)(I_n - G_n)^{-1}\right]u \\
& \xrightarrow{d} N\left(0, 2\pi_G^2\kappa_{2,0}E\left[\left(m_2^2 f\right)(D_1)\right]\right)
\end{aligned}$$

where the limit result is obtained by Theorem 2.1 and

$$\begin{aligned}
\pi_G &= \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}\left\{I_n - \left(G_n^* - G_{0,n}^*\right)(I_n - G_n)^{-1}\right\} \\
&= 1 - \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}\left\{\left(G_n^* - G_{0,n}^*\right)(I_n - G_n)^{-1}\right\}. \qquad \text{(A.18)}
\end{aligned}$$

This completes the proof of this lemma.

Lemma A.9 *Under Assumptions 2-4 and 6-8, we have* $n\sqrt{|H|}T_{n2} = o_p(1)$ *under* H_0^1 *and* $n\sqrt{|H|}T_{n2} = O_p\left(\sqrt{n|H|}\right)$ *under* H_1^1.

Proof: Under H_0^1, we have

$$\chi_n = S_n X\theta_0 - X\hat{\theta} = X\left(\theta_0 - \hat{\theta}\right) - \left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1}X\theta_0 \qquad \text{(A.19)}$$

so that

$$\begin{aligned}
T_{n2} &= \frac{\left(\theta_0 - \hat{\theta}\right)'}{n^2|H|}X'\mathbf{K}_H u - \frac{\left(\theta_0 - \hat{\theta}\right)'}{n^2|H|}X'\mathbf{K}_H\left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1}u \\
&\quad - \frac{1}{n^2|H|}\left[\left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1}X\theta_0\right]'\mathbf{K}_H u \\
&\quad + \frac{1}{n^2|H|}\left[\left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1}X\theta_0\right]'\mathbf{K}_H\left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1}u \\
&= \left(\theta_0 - \hat{\theta}\right)'T_{n2,1} - \left(\theta_0 - \hat{\theta}\right)'T_{n2,2} - T_{n2,3} + T_{n2,4},
\end{aligned}$$

where $T_{n2,1} = O_p\left(n^{-1/2}\right)$ by (A.12), and we have

$$\begin{aligned}
T_{n2,2} &= \frac{1}{n^2|H|}X'\mathbf{K}_H\left(\hat{G}_n - G_{0,n}^*\right)(I_n - G_n)^{-1}u - \frac{1}{n^2|H|}X'\mathbf{K}_H\left(G_{0,n}^* - G_n\right)(I_n - G_n)^{-1}u \\
&= \frac{(\hat{\alpha} - \alpha_0)'}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}X_i K_{H,ij}\sum_{l\neq j}\boldsymbol{\phi}_{L_n}(Z_{jl})\tilde{u}_l - \frac{1}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}X_i K_{H,ij}\sum_{l\neq j}\left(g_{0,jl}^* - g_{jl}\right)\tilde{u}_l \\
&= O_p\left(n^{-1/2}\vartheta_n\sqrt{L_n} + n^{-1/2}L_n^{-\xi}\right),
\end{aligned}$$

$$
\begin{aligned}
T_{n2,3} &= \frac{1}{n^2|H|}\left[\left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1} X\theta_0\right]' \mathbf{K}_H u \\
&= \frac{1}{n^2|H|}\left[\left(\hat{G}_n - G^*_{0,n}\right)(I_n - G_n)^{-1} X\theta_0\right]' \mathbf{K}_H u \\
&\quad + \frac{1}{n^2|H|}\left[\left(G^*_{0,n} - G_n\right)(I_n - G_n)^{-1}\right]' \mathbf{K}_H u \\
&= O_p\left(\vartheta_n\sqrt{L_n/n} + L_n^{-\xi} n^{-1/2}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
T_{n2,4} &= \frac{1}{n^2|H|}\left[\left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1} X\theta_0\right]' \mathbf{K}_H \left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1} u \\
&\leq \sqrt{T_{n2,2}T_{n2,3}} = O_p\left(\vartheta_n\sqrt{L_n/n} + L_n^{-\xi} n^{-1/2}\right)
\end{aligned}
$$

because we have

$$
\begin{aligned}
& E\left\|\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l\neq i}\boldsymbol{\phi}_{L_n}(Z_{il})\tilde{X}_i'\theta_0 K_{H,ij}u_j\right\|^2 \\
&= \sigma_u^2 \sum_{i=1}^{n}\sum_{l\neq i}\sum_{i_1\neq i}^{n}\sum_{j\neq i, j\neq i_1}\sum_{l_1\neq i_1}\boldsymbol{\phi}'_{L_n}(Z_{il})\boldsymbol{\phi}_{L_n}(Z_{i_1 l_1})\theta_0' E\left(K_{H,ij}K_{H,i_1 j}\tilde{X}_i\tilde{X}_{i_1}'\right)\theta_0 \\
&\quad + \sigma_u^2 \sum_{i=1}^{n}\sum_{l\neq i}\sum_{j\neq i}\sum_{l_1\neq i}\boldsymbol{\phi}'_{L_n}(Z_{il})\boldsymbol{\phi}_{L_n}(Z_{il_1})\theta_0' E\left(K^2_{H,ij}\tilde{X}_i\tilde{X}_i'\right)\theta_0 \\
&= O\left(n^3 L_n|H|^2 + n^2 L_n|H|\right) = O\left(n^3 L_n|H|^2\right)
\end{aligned}
$$

and

$$
E\left\|\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l\neq i}\boldsymbol{\phi}_{L_n}(Z_{il})\tilde{X}_i'\theta_0 K_{H,ij}\sum_{l'\neq j}\boldsymbol{\phi}'_{L_n}(Z_{jl'})\tilde{u}_{l'}\right\|^2 = O\left(n^3 L_n^2|H|^2\right).
$$

Therefore, we obtain $n\sqrt{|H|}T_{n2} = n\sqrt{|H|}O_p\left(n^{-1/2}\vartheta_n L_n^{1/2}\right) = O_p\left(\vartheta_n\sqrt{nL_n|H|}\right) = o_p(1)$ under Assumption 8.

Next, under H_1^1, we have

$$
\begin{aligned}
\chi_n &= S_n \mathrm{mtk}(X, \theta(D)) - X\hat{\theta} \\
&= \mathrm{mtk}(X, \theta(D) - \theta) + X\left(\theta - \hat{\theta}\right) - \left(\hat{G}_n - G_n\right)(I_n - G_n)^{-1}\mathrm{mtk}(X, \theta(D)) \quad \text{(A.20)}
\end{aligned}
$$

and applying the similar methods used above gives

$$\begin{aligned}
T_{n2} &= \frac{1}{n^2|H|}\left[\mathrm{mtk}\left(X,\theta(D)-\theta\right)\right]'\mathbf{K}_H S_n u + \frac{\left(\theta-\hat{\theta}\right)'}{n^2|H|}X'\mathbf{K}_H S_n u \\
&\quad -\frac{1}{n^2|H|}\left[\left(\hat{G}_n-G_n\right)\left(I_n-G_n\right)^{-1}\mathrm{mtk}\left(X,\theta(D)\right)\right]'\mathbf{K}_H S_n u \\
&\approx \frac{1}{n^2|H|}\left[\mathrm{mtk}\left(X,\theta(D)-\theta\right)\right]'\mathbf{K}_H\left[I_n-\left(G_n^*-G_{0,n}^*\right)\left(I_n-G_n\right)^{-1}\right]u \\
&\quad -\frac{1}{n^2|H|}\left[\left(G_n^*-G_{0,n}^*\right)\left(I_n-G_n\right)^{-1}\mathrm{mtk}\left(X,\theta(D)\right)\right]' \\
&\quad \times\mathbf{K}_H\left[I_n-\left(G_n^*-G_{0,n}^*\right)\left(I_n-G_n\right)^{-1}\right]u \\
&= \frac{1}{n^2|H|}O_p\left(n^{3/2}|H|\right)=O_p\left(n^{-1/2}\right)
\end{aligned}$$

so that $n\sqrt{|H|}T_{n2}=O_p\left(\sqrt{n|H|}\right)$. This completes the proof of this lemma.

Lemma A.10 *Under Assumptions 2-4 and 6-8, we have* $n\sqrt{|H|}T_{n3}=O_p\left(\sqrt{|H|}\right)$ *under* H_0^1 *and* $n\sqrt{|H|}T_{n3}=O_p\left(n\sqrt{|H|}\right)$ *under* H_1^1.

Proof: By (A.19) under H_0^1, we have

$$\begin{aligned}
T_{n3} &= \frac{\left(\theta_0-\hat{\theta}\right)'}{n^2|H|}X'\mathbf{K}_H X'\left(\theta_0-\hat{\theta}\right)-\frac{2\left(\theta_0-\hat{\theta}\right)'}{n^2|H|}X'\mathbf{K}_H\left(\hat{G}_n-G_n\right)\left(I_n-G_n\right)^{-1}X\theta_0 \\
&\quad +\frac{1}{n^2|H|}\left[\left(\hat{G}_n-G_n\right)\left(I_n-G_n\right)^{-1}X\theta_0\right]'\mathbf{K}_H\left(\hat{G}_n-G_n\right)\left(I_n-G_n\right)^{-1}X\theta_0 \\
&= O_p\left(n^{-1/2}\right)+O_p\left(\vartheta_n^2L_n+L_n^{-2\xi}\right)
\end{aligned}$$

because we have

$$\begin{aligned}
E\left[\sum_{i=1}^{n}\sum_{j\neq i}K_{H,ij}\sum_{l\neq i}\sum_{l'\neq j}\boldsymbol{\phi}_{L_n}'\left(Z_{il}\right)\boldsymbol{\phi}_{L_n}\left(Z_{jl'}\right)\theta_0'\tilde{X}_i\tilde{X}_j'\theta_0\right] &= O\left(n^2|H|L_n\right) \\
E\left[\sum_{i=1}^{n}\sum_{j\neq i}K_{H,ij}\sum_{l\neq i}\sum_{l'\neq j}\left(g_{il}^*-g_{il}\right)\left(g_{jl'}^*-g_{jl'}\right)\theta_0'\tilde{X}_i\tilde{X}_j'\theta_0\right] &= O\left(n^2|H|L_n^{-2\xi}\right).
\end{aligned}$$

By (A.20) under H_1^1, we have $T_{n3}=T_{n3,1}+O_p\left(n^{-1/2}+\vartheta_n^2L_n+L_n^{-2\xi}\right)$, where

$$\begin{aligned}
T_{n3,1} &= \frac{1}{n^2|H|}\left[\mathrm{mtk}\left(X,\theta(D)-\theta\right)-\left(G_n^*-G_{0,n}^*\right)\left(I_n-G_n\right)^{-1}\mathrm{mtk}\left(X,\theta(D)\right)\right]'\mathbf{K}_H \\
&\quad \times\left[\mathrm{mtk}\left(X,\theta(D)-\theta\right)-\left(G_n^*-G_{0,n}^*\right)\left(I_n-G_n\right)^{-1}\mathrm{mtk}\left(X,\theta(D)\right)\right] \\
&= E\left(T_{n3,1}\right)+O_p\left(n^{-1/2}\right) \qquad \text{(A.21)}
\end{aligned}$$

by Theorem 2.2 and $E\left(T_{n3,1}\right)$ converges to a positive constant. This completes the proof of this lemma.

Lemma A.11 *Under Assumptions 2-4 and 6-8, we have* $\hat{\sigma}_n^2 \xrightarrow{p} \sigma_0^2$, *where* $\sigma_0^2 = 2\sigma_u^4\kappa_{2,0}E\left[f\left(D_1\right)\right]$, *under* H_0^1, *and* $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$ *otherwise, where* σ^2 *is defined in eq. (A.22) below.*

Proof: Under H_0^1 and by (A.19), we have $\hat{u} = X\left(\theta_0 - \hat{\theta}\right) + u - \left(\hat{G}_n - G_n\right)\left(I_n - G_n\right)^{-1}\left(X\theta_0 + u\right)$, where $\hat{\theta} - \theta_0 = O_p\left(n^{-1/2}\right)$. Then, under Assumptions 2-4 and 6-8, we can show that

$$\begin{aligned}
\hat{\sigma}_n^2 &= \frac{1}{n^2\left|H\right|}\sum_{i=1}^{n}\sum_{j\neq i}\hat{u}_i^2\hat{u}_j^2K_{H,ij}^2 \\
&= \frac{1}{n^2\left|H\right|}\sum_{i=1}^{n}\sum_{j\neq i}K_{H,ij}^2u_i^2u_j^2 + O_p\left(\vartheta_n^4L_n^2\right) \\
&= \sigma_0^2 + o_p\left(1\right).
\end{aligned}$$

Under H_1^1 and by (A.20), and denoting $V = \text{mtk}\{X, \theta\left(D\right) - \theta\} + u - \left(G_n^* - G_{0,n}^*\right)\left(I_n - G_n\right)^{-1}$ $\times\left(\text{mtk}\left\{X, \theta\left(D\right)\right\} + u\right)$, we have

$$\begin{aligned}
\hat{u} &= \left(I_n - \hat{G}_n\right)\left(I_n - G_n\right)^{-1}\left(\text{mtk}\left\{X, \theta\left(D\right)\right\} + u\right) - X\hat{\theta} \\
&= \text{mtk}\left\{X, \theta\left(D\right) - \theta\right\} + u - X\left(\hat{\theta} - \theta\right) - \left(\hat{G}_n - G_n\right)\left(I_n - G_n\right)^{-1}\left(\text{mtk}\left\{X, \theta\left(D\right)\right\} + u\right) \\
&= V - X\left(\hat{\theta} - \theta\right) - \left(\hat{G}_n - G_n^* + G_{0,n}^* - G_n\right)\left(I_n - G_n\right)^{-1}\left(\text{mtk}\left\{X, \theta\left(D\right)\right\} + u\right)
\end{aligned}$$

which implies that

$$\hat{\sigma}_n^2 = \frac{1}{n^2\left|H\right|}\sum_{i=1}^{n}\sum_{j\neq i}K_{H,ij}^2v_i^2v_j^2 + o_p\left(1\right) = \sigma^2 + o_p\left(1\right)$$

where v_i is the ith element of V. Applying the similar method used to prove Theorem 2.2 we can show that

$$\sigma^2 = 2\kappa_{2,0}\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}E\left[c_{ij}\left(D_1, D_1\right)f\left(D_1\right)\right] \tag{A.22}$$

being a finite positive constant, where $c_{ij}\left(d_1, d_2\right) = E\left(v_i^2v_j^2|D_i = d_1, D_j = d_2\right)$. This completes the proof of this lemma.

Proof of Theorem 3.3: The residuals equal $\hat{u} = \left(I_n - \hat{\rho}G_{n,0}\right)Y - X\hat{\theta}\left(D\right) = S_nu + \chi_n$, where $S_n = \left(I_n - \hat{\rho}G_{n,0}\right)\left(I_n - G_n\right)^{-1}$ and $\chi_n = S_n\text{mtk}(X, \theta_0\left(D\right)) - X\hat{\theta}\left(D\right)$. The decomposition of our test statistic (A.5)-(A.6) continue to hold. By Lemmas A.12-A.15, we complete the proof of this theorem.

Lemma A.12 *Under Assumptions 2-(ii) and 9-10, and* $\pi_G \neq 0$, *we have* $n\sqrt{|H|}T_{n1} \xrightarrow{d} N\left(0, 2\sigma_u^2\pi_G^2\kappa_{2,0}E\left[f\left(D_1\right)\right]\right)$, *where* π_G *is defined in (A.10), and* $\pi_G = 1$ *under* H_0^2.

Proof: The proof is omitted as it is the same as the proof of Lemma A.4.

Lemma A.13 *Under Assumptions 2-4(ii) and 9-10, we have* $n\sqrt{|H|}T_{n2} = o_p(1)$ *under* H_0^2, *and* $n\sqrt{|H|}T_{n2} = O_p\left(\sqrt{n|H|}\right)$ *under* H_1^2.

Proof: Under H_0^2 and by (A.7), we have

$$\chi_n = \mathrm{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right) - (\hat{\rho} - \rho_0) G_{n,0}(I_n - G_{n,0})^{-1} \mathrm{mtk}(X, \theta(D)) \tag{A.23}$$

and following the proof of Lemma A.5, we obtain

$$\begin{aligned} T_{n2} &= \frac{1}{n^2|H|}\left[\mathrm{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right)\right]' \mathbf{K}_H S_n u \\ &\quad - \frac{\hat{\rho} - \rho_0}{n^2|H|}\left[G_{n,0}(I_n - G_n)^{-1} \mathrm{mtk}(X, \theta_0(D))\right]' \mathbf{K}_H S_n u \\ &= T_{n2,1} + O_p\left(\frac{n^{-1/2}}{n^2|H|}\right) O_p\left(n^{3/2}|H|\right) = T_{n2,1} + O_p\left(n^{-1}\right) \end{aligned}$$

as $\theta_0(D_i) - \hat{\theta}(D_i) = \theta_0(D_i) - E\left[\hat{\theta}(D_i)\right] + E\left[\hat{\theta}(D_i)\right] - \hat{\theta}(D_i)$ and applying tedious calculations give

$$\begin{aligned} T_{n2,1} &= \frac{1}{n^2|H|}\left[\mathrm{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right)\right]' \mathbf{K}_H S_n u \\ &= \frac{1}{n^2|H|}\sum_{i=1}^{n}\sum_{j\neq i}\left[X_i\left(\theta_0(D_i) - \hat{\theta}(D_i)\right)\right]' K_{H,ij}\left[u_j - (\hat{\rho} - \rho_0)\sum_{l\neq j} g_{0,jl}\tilde{u}_l\right] \\ &= O_p\left(n^{-1/2}\left(\|H_0\|^2 + \sqrt{\ln n/(n|H_0|)}\right)\right) \end{aligned}$$

Therefore, we obtain $n\sqrt{|H|}T_{n2} = O_p\left(\sqrt{n|H|}\left(\|H_0\|^2 + \sqrt{\ln n/(n|H_0|)}\right)\right) = o_p(1)$ under Assumption 10.

Under H_1^2, we obtain $n\sqrt{|H|}T_{n2} = O_p\left(\sqrt{n|H|}\right)$, where the proof is similar to that of Lemma A.5. This completes the proof of this lemma.

Lemma A.14 *Under Assumptions 2-4(ii) and 9-10, we have* $n\sqrt{|H|}T_{n3} = o_p(1)$ *under* H_0^2, *and* $n\sqrt{|H|}T_{n3} = O_p\left(n\sqrt{|H|}\right)$ *under* H_1^2.

Proof: Under H_0^2 and by (A.23), following the proofs of Lemmas A.6 and A.13, we have

$$\begin{aligned} T_{n3} &= \frac{1}{n^2|H|}\left[\mathrm{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right)\right]' \mathbf{K}_H \left[\mathrm{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right)\right] \\ &\quad - \frac{2(\hat{\rho} - \rho_0)}{n^2|H|}\left[\mathrm{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right)\right]' \mathbf{K}_H \left[G_{n,0}(I_n - G_{n,0})^{-1} \mathrm{mtk}(X, \theta_0(D))\right] \\ &\quad + \frac{(\hat{\rho} - \rho_0)^2}{n^2|H|}\left[G_{n,0}(I_n - G_{n,0})^{-1} \mathrm{mtk}(X, \theta_0(D))\right]' \mathbf{K}_H \left[G_{n,0}(I_n - G_{n,0})^{-1} \mathrm{mtk}(X, \theta_0(D))\right] \\ &= O_p\left(\|H_0\|^4 + \ln n/(n|H_0|)\right). \end{aligned}$$

Hence, we have $n\sqrt{|H|}T_{n3} = O_p\left(n\sqrt{|H|}\left(\|H_0\|^4 + \ln n/\left(n\,|H_0|\right)\right)\right) = o_p(1)$ under Assumption 10.

Under H_1^2, we have

$$\begin{aligned} \chi_n &= S_n \text{mtk}\left(X, \theta_0(D)\right) - X\hat{\theta}(D) \\ &= \text{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right) - \left[(\hat{\rho} - \rho) G_{n,0} + (\rho G_{n0} - G_n)\right](I_n - G_n)^{-1} \text{mtk}\left(X, \theta(D)\right) \end{aligned} \quad (A.24)$$

and it follows that $T_{n3} = T_{n3,1} + O_p\left(\|H_0\|^4 + \ln n/\left(n\,|H_0|\right)\right)$, where

$$\begin{aligned} T_{n3,1} &= \frac{1}{n^2|H|}\left[\text{mtk}\left(X, \theta_0(D) - \theta(D)\right) - (\rho G_{n0} - G_n)(I_n - G_n)^{-1}\text{mtk}\left(X, \theta(D)\right)\right]' \mathbf{K}_H \\ &\quad \times \left[\text{mtk}\left(X, \theta_0(D) - \theta(D)\right) - (\rho G_{n0} - G_n)(I_n - G_n)^{-1}\text{mtk}\left(X, \theta(D)\right)\right] \\ &= E\left(T_{n3,1}\right) + O_p\left(n^{-1/2}\right) \end{aligned} \quad (A.25)$$

and $E\left(T_{n3,1}\right) > 0$ by Theorem 2.1. This completes the proof of this lemma.

Lemma A.15 *Under Assumptions 2–4(ii) and 9-10, we have* $\hat{\sigma}_n^2 \xrightarrow{p} \sigma_0^2$, *where* $\sigma_0^2 = 2\sigma_u^4\kappa_{2,0}E\left[f(D_1)\right]$, *under* H_0^2, *and* $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$ *otherwise, where* σ^2 *is defined in eq. (A.16).*

Proof: Under H_0^2 and by (A.23), we have $\hat{u} = u + \text{mtk}\left(X, \theta_0(D) - \hat{\theta}(D)\right) - (\hat{\rho} - \rho_0)G_{n,0} \times (I_n - G_{n,0})^{-1}\left[\text{mtk}\left(X, \theta_0(D)\right) + u\right]$. Then, under Assumptions 2-4(ii) and 9-10, we can easily show that

$$\begin{aligned} \hat{\sigma}_n^2 &= \frac{2}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}\hat{u}_i^2\hat{u}_j^2K_{H,ij}^2 \\ &= \frac{2}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}K_{H,ij}^2u_i^2u_j^2 + o_p(1) = \sigma_0^2 + o_p(1). \end{aligned}$$

Under H_1^2, by (A.24) and denoting $V = \left[I_n - (\rho G_{n,0} - G_n)(I_n - G_n)^{-1}\right]u - (\rho G_{n,0} - G_n) \times (I_n - G_n)^{-1}\text{mtk}\{X, \theta(D)\}$, we have

$$\begin{aligned} \hat{u} &= \left(I_n - \hat{G}_{n,0}\right)(I_n - G_n)^{-1}\left(\text{mtk}\left\{X, \theta(D)\right\} + u\right) - X\hat{\theta}(D) \\ &= V + \text{mtk}\left\{X, \theta(D) - \hat{\theta}(D)\right\} - (\hat{\rho} - \rho)G_{n,0}(I_n - G_n)^{-1}\left(\text{mtk}\left\{X, \theta(D)\right\} + u\right) \end{aligned}$$

and obtain

$$\hat{\sigma}_n^2 = \frac{2}{n^2|H|}\sum_{i=1}^n\sum_{j\neq i}K_{H,ij}^2v_i^2v_j^2 + o_p(1)$$

where v_i is the ith element of V. Applying Lemma A.7 completes the proof of this lemma.

References

[1] Ahrens, A., Bhattacharjee, A., 2015. Two-step Lasso estimation of the spatial weights matrix. Econometrics 3, 128-155.

[2] Arcones, M.A., Gine, E., 1993. Limit theorem for U-processes. The Annals of Probability 21, 1494-1542.

[3] Baltagi, B.H., Eggerb, P., Pfaffermayr, M., 2008. Estimating regional trade agreement effects on FDI in an interdependent world. Journal of Econometrics 145, 194–208.

[4] Brett, C., Pinkse, J., 2000. The determinants of municipal tax rates in British Columbia. Canadian Journal of Economics 33: 695-714.

[5] Brueckner, J.K., 2003. Strategic interaction among governments: an overview of empirical studies. International Regional Science Review 26, 175-188.

[6] Case, A.C., Rosen, H.S., Hines, J.R., 1993. Budget spillovers and fiscal-policy interdependence: Evidence from the States. Journal of Public Economics 52, 285–307.

[7] Debarsy, N., Ertur, C., 2010. Testing for spatial autocorrelation in a fixed effects panel data model. Regional Science and Urban Economics 40, 453–470.

[8] Durlauf, S.N., Kourtellos, A., Minkin, A., 2001. The local Solow growth model. European Economic Review 45, 928–940.

[9] Ertur, C., Koch, W., 2007. Growth, technological interdependence and spatial externalities: theory and evidence. Journal of Applied Econometrics 22, 1033-1062.

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

[11] Fan, Y. and Li, Q. (1996b) Central limit theorem for degenerate

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

[13] Gao, J., Anh, V., 2000. A central limit theorem for a random quadratic form of strictly stationary processes. Statistics & Probability Letters 49, 69-79.

[14] Gao, J., Hong, Y., 2008. Central limit theorems for generalized U-statistics with applications in nonparametric specification. Journal of Nonparametric Statistics 20, 61-76.

[15] Gao, J., Lu, Z., Tjøsteim, D., 2006. Estimation in semiparametric spatial regression. The Annals of Statistics 34, 1395–1435.

[16] Han, X., Lee, L.-F., 2013. Model selection using J-test for the spatial autoregressive model vs. the matrix exponential spatial model. Regional Science and Urban Economics 43, 250–271.

[17] Hall, P., 1984. Central limit theorem for integrated square error of multivariate nonparametric density estimators. Journal of multivariate analysis 14, 1-16.

[18] Hoeffding, W., 1948. A class of statistics with aymptotically normal distribution. The Annals of Mathematical Statistics 19, 293-325.

[19] Kang, X., Li, T., 2017. Testing a linear relationship in varying coefficient spatial autoregressive models. Communications in Statistics-Simulation and Computation, 1-19.

[20] Karlin, S., Rinott, Y., 1982. Applications of ANOVA type decompositions of conditional variance statistics including Jackknife estimates. The Annals of Statistics 10, 485-501.

[21] Kashimov, SH.A., 1987. On the asymptotic distribution of the generalized U-statistics for dependent variables. Theory of Probability and its Applications 32, 373–375.

[22] Kashimov, SH.A., 1992. Limiting behavior of generalized U-statistics of weakly dependent stationary processes. Theory of Probability and its Applications 37, 148–150.

[23] Kelejian, H.H., 2008. A spatial J-test for model specification against a single or a set of non-nested alternatives. Letters in Spatial and Resource Sciences 1, 3–11.

[24] Kelejian, H.H., Piras, G., 2011. An extension of Kelejian's J-test for non-nested spatial models. Regional Science and Urban Economics 41, 281–292.

[25] Kelejian, H.H, Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics 104, 219–257.

[26] Khashimov, Sh.A., 1988. On limit theorems for the distributions of generalized von Mises functionals. Theory of Probability and its Applications 33, 566- 571.

[27] Kim, T.A., Luo, Z., Kim, C., 2011. The central limit theorem for degenerate variable U-statistics under dependence. Journal of Nonparametric Statistics 23, 683-699.

[28] Lee, A. J., 1990. U-statistics, Theory and Practice. Marcel Dekker, Inc., New York.

[29] Lee, L.F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial econometric models. Econometrica 72, 1899–1926.

[30] Lee, L.F., 2007. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. Journal of Econometrics 137, 489–514.

[31] Malikov, E., Sun, Y., 2017. Semiparametric estimation and testing of smooth coefficient spatial autoregressive model. Journal of Econometrics 199, 12-34.

[32] Mankiw, N.G., Romer, D., Weil, D.N., 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107, 407–437.

[33] Moran, P.A.P., 1950. Notes on Continuous Stochastic Phenomena. Biometrika 37, 17-23.

[34] Nasari, M.M., 2012. Strong law of large numbers for weighted U-statistics: application to incomplete U-statistics. Statistics and Probability Letters 82, 1208–1217.

[35] Pinkse, J., Slade, M.E., Brett, C., 2002. Spatial price competition: a semiparametric approach. Econometrica 70, 1111-1153.

[36] Powell,J.L., Stock, J.H., Stoker, T.M., 1989. Semiparametric estimation of index coefficients. Econometrica 57, 1403-1430.

[37] Sen, A., 1976. Large sample-size distribution of statistics used in testing for spatial correlation. Geographical Analysis 9, 175–184.

[38] Seber, G.A.F., 2008. A Matrix Handbook for Statisticians. John Wiley & Sons, Inc.

[39] Serfling, R. J., 1980. Approximation Theorems of Mathematical Statistics. Wiley, New York.

[40] Su, L., 2012. Semiparametric GMM estimation of spatial autoregressive models. Journal of Econometrics 167, 543–560.

[41] Su, L., Jin, S., 2010. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. Journal of Econometrics 157, 18-33.

[42] Su, L., and Qu, X., 2017. Specification test for spatial autoregressive models. Jorunal of Business and Economic Statistics 35, 572-584.

[43] Takahata, H., Yoshihara, K., 1987. Central limit theorems for integrated square error of nonparametric density estimators based on absolutely regular random sequences. Yokohama Mathematical Journal 35, 95–111.

[44] White, H., 2001. Asymptotic Theory for Econometricians. Revised edition, Academic Press.

[45] Yang, Z., 2015. LM tests of spatial dependence based on bootstrap critical values. Journal of Econometrics 185, 33-59.

[46] Yoshihara, K., 1976. Limiting behavior of U-statistics for stationary, absolutely regular processes. Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete 35, 237–252.

[47] Yoshihara, K., 1989. Limiting behavior of generalized quadratic forms generated by absolutely regular sequences II. Yokohama Mathematical Journal

[48] Zheng, J.X., 1996. A consistent test of functional form via nonparametric estimation techniques. Journal of Econometrics 75, 263-289.

[49] Zhang, Z., 2013. A pairwise difference estimator for partially linear spatial autoregressive models. Spatial Economic Analysis 8, 176–194.