Specification Test for Partially Functional Linear Spatial Autoregressive Model

January 3, 2024

Abstract

A typical model for dealing with functional data is the functional linear model. However, data is frequently linked to spatial information in practical applications. The traditional functional linear model is no longer applicable in this case, whereas the functional linear spatial autoregressive model can handle both spatial and functional data at the same time. The specification test of model is one of the important aspects of data analysis. If the model is set wrong, it will result in errors and even have unquantifiable consequences. As data types get increasingly complicated in the era of big data, the chances of models going incorrect increase. Therefore, it is essential to investigate the problem of the specification test for functional linear spatial autoregressive models.

According to the generalized method of moments, the estimation of residual error is obtained, which is combined with the linear projection of covariates to generate the residual empirical process and establish the test statistic, allowing the specification test for partially functional linear spatial autoregressive models to be performed. The proposed test method may successfully prevent subjective parameter selection and has a wide range of applications. First, under certain regularity conditions, the asymptotic distributions of the test statistic under the null and the alternative hypothesis are shown to be Gaussian processes, which shows that it can detect Pitman local alternatives converging to the null hypothesis at the usual parameter rate, and further gives the consistence of the test statistic. Second, we apply the bootstrap method to approximate the distribution of the test statistic, after which we provide an approximate calculation method for the test statistic's critical value, whose reasonableness is theoretically shown. Furthermore, different data generation processes are constructed and the finite sample properties of the statistic are investigated by Monte Carlo simulations. It is shown that, under the null hypothesis, the test statistic's nominal level tends to the given nominal level. Under the alternative hypothesis, the test statistic's power grows as the sample size grows, quickly approaching 1. Finally, an example analysis of the Spanish Meteorological data is presented to illustrate the practicality of the proposed method.

Keywords: Partially functional linear spatial autoregressive; Random projection; Specification test; Residual empirical process

1 Introduction

Functional data, first proposed by [16], are widely used in many fields such as medicine, meteorology, neuroscience, etc., and thus have gained the attention of many scholars. [2] investigated the estimation of functional linear models in the case where the response variable is a scalar using the B spline estimation method and gave an upper bound on the speed of convergence of the estimation. Under functional principal component analysis method and spectral decomposition method, [8] showed that the convergence speed of the estimates obtained based on the FPCA is optimal. [1] studied the parameter estimation problem in the case where the response variable is a functional variable based on the regularization method of L2 penalties, taking into account the case where the covariates are endogenous. [13] studied of sparse functional linear regression models by adding penalties to the traditional least-squares criterion for estimating functional-type parameters using variable selection methods.

With the continuous improvement and development of the theory, scholars focus on the study of various types of functional models to solve complex practical problems. [15] considered the generalized functional-type linear model and prove the asymptotic normality of the estimation, while [14] considered the application of the coefficients of variation under this model, focusing on the variable selection problem and broadening the perspective of the study of the functional-type model.

The spatial autoregressive model, which aims to deal with the spatial correlation between data, was first proposed by [4], which can fully take into account the spatial properties between variables, and the research on its estimation method has more mature results. [10] discussed the parameter estimation problem of the model when the perturbation term of the model has also autoregressive form. The problem of parameter estimation of the model when the perturbation term of the model also has an autoregressive form was discussed, and the Generalized Two-Stage Least Squares Estimation was proposed and its theoretical properties were investigated. Later on [11] further proposed the GMM method and proved the large-sample nature of the estimation, which is computationally simple and effective and does not need to consider the sample size. [12] proposed the Quasi-Maximum Likelihood Estimation, which is a method for the estimation of a large number of samples. [20]considered the problem of variable selection for spatial autoregressive models. [18] extended spatial autoregressive models to semiparametric spatial dynamical models and applied them to the analysis of real data. [5] studied the problem of parameter estimation for partially linear additive spatial autoregressive models using the GMM method, and deduced the asymptotic distribution.

This article focus on the problem of setting tests for partial functional linear spatial autoregressive models under the generalized moment estimates of the parameters. Firstly, for the special functional data, they are simply processed using the FPCA method, which transforms the infinite-dimensional data into a linear combination of finite number of terms. Further, the objective function is constructed by moment conditions and the estimates of the parameters are obtained with the help of generalized moment estimation methods. Second, the test statistics are constructed based on the projection method and the empirical process of residuals to investigate the properties of the test proposed in this paper under different hypothesis tests. Finally, using the bootstrap method to approximate the critical value of the distribution for test statistic.

The rest of this article is structured as follows. Section 2 formulates the problem, defines the estimators and test statistics. Section 3 provides the results relating to the asymptotic validity and consistency of the bootstrap tests. Section 4 describes a simulation study and discusses a empirical illustration. Section 5 concludes the article. The proofs are relegated to the Appendix.

2 Estimation and test construction

2.1 Partially functional linear spatial autoregressive model

We consider the partially functional linear spatial autoregressive model

$$\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{Z}_n^T \boldsymbol{\beta} + \int_0^1 \boldsymbol{X}_n(t) \boldsymbol{\gamma}(t) dt + \boldsymbol{\varepsilon}_n, \tag{1}$$

where $\boldsymbol{Y}_n$ is a response variable; ρ is a scalar parameter denoting the lag coefferent between -1 and 1, the closer ρ is to 0, the weaker the corelation is; $\boldsymbol{W}_n$ is a spatial weight matrix with zero diagonal; $\boldsymbol{Z}_n = (\boldsymbol{Z}_1, \boldsymbol{Z}_2, \cdots, \boldsymbol{Z}_p)$ is a $n \times p$-dimensional linear regression; $\boldsymbol{\beta}$ is a unknown parametric; $\{\boldsymbol{X}_n(t) : t \in [0,1]\}$ is a zero mean, second stochastic process defined on $(\Omega, \mathcal{B}, P)$ with sample paths in $L^2([0,1])$, the Hilbert space containing square integrable functions defined on $[0,1]$ with inner product $\langle x, y \rangle = \int_0^1 x(t)y(t)dt, \forall x, y \in L^2(0,1)$ and norm $||x|| = \langle x, x \rangle^{\frac{1}{2}}$; $\boldsymbol{\gamma}(t)$ is a square inter integrable function on $[0,1]$; $\boldsymbol{\varepsilon}_n$ is an i.i.d. random variable with mean 0 and finite variance σ^2 independent of $\boldsymbol{X}_n$ and $\boldsymbol{Z}_n$.

2.2 Estimation procedures

we refer to the estimation method of [9], and use the generalized method of moments to estimate the parameters after dealing with the functional data part.

Suppose that $(X_1, Y_1, Z_1), (X_2, Y_2, Z_2), \cdots, (X_n, Y_n, Z_n)$ are n independent and identically distributed samples which are generated from model (1). To facilitate the study of functional data, the covariance function and empirical covariance function are defined as

$$K(s,t) = \mathrm{Cov(X(s), X(t))},$$

and

$$\hat{K}(s,t)\frac{1}{n}\sum_{i=1}^{n} X_i(s)X_i(t).$$

According to the Mercer theorem, the spectral decomposition of matrices $K(s,t)$ and $\hat{K}(s,t)$ can be performed as

$$K(s,t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(t)\phi_j(s),$$

and

$$\hat{K}(s,t) = \sum_{j=1}^{\infty} \hat{\lambda}_j \hat{\phi}_j(t)\hat{\phi}_j(s).$$

where λ_j and $\hat{\lambda}_j$ are the eigenvalues of $K(s,t)$ and $\hat{K}(s,t)$ respectively, satisfying $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, λ_j and $\hat{\lambda}_j$ are the corresponding eigenvectors.

Then, from the Karhunen-Loève respresentation,

$$X(t) = \sum_{j=1}^{\infty} \xi_j \phi_j(t),$$

and

$$\beta(t) = \sum_{j=1}^{\infty} \gamma_j \phi_j(t).$$

where $\xi_j = \langle X, \phi_j \rangle_{j=1}^{\infty}$, $\gamma_j = \langle \beta, \phi_j \rangle$. Thus, model (1) can be re-pressed as

$$\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{Z}_n^T \boldsymbol{\beta} + \sum_{j=1}^{\infty} \gamma_j \langle \boldsymbol{X}_n, \phi_j \rangle + \boldsymbol{\varepsilon}_n. \tag{2}$$

Then, model (2) can be approximated as

$$\boldsymbol{Y}_n \approx \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{Z}_n^T \boldsymbol{\beta} + \sum_{j=1}^{m} \gamma_j \langle \boldsymbol{X}_n, \phi_j \rangle + \boldsymbol{\varepsilon}_n, \tag{3}$$

where the exact value of m can be determined by the cumulative variance contribution or by methods such as validation. Replace ϕ_j by $\hat{\phi}_j$, model (3) is

$$\boldsymbol{Y}_n \approx \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{Z}_n^T \boldsymbol{\beta} + \boldsymbol{\Pi} \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_n, \tag{4}$$

where $\boldsymbol{\Pi} = (\langle \boldsymbol{X}_n, \hat{\phi}_1 \rangle, \langle \boldsymbol{X}_n, \hat{\phi}_2 \rangle, \cdots, \langle \boldsymbol{X}_n, \hat{\phi}_m \rangle)$, $\boldsymbol{\alpha} = (\gamma_1, \gamma_2, \cdots, \gamma_m)^T$.

Denote $\boldsymbol{Q}_n = (\boldsymbol{W}_n \boldsymbol{Y}_n, \boldsymbol{Z}_n)$, $\boldsymbol{\theta} = (\rho, \boldsymbol{\beta}^T)^T$, $\boldsymbol{S}_n = (\boldsymbol{I}_n - \rho \boldsymbol{W}_n)$, it is obvious that $\mathrm{E}((\boldsymbol{W}_{\mathrm{n}} \boldsymbol{Y}_{\mathrm{n}})^{\mathrm{T}} \boldsymbol{\varepsilon}_{\mathrm{n}}) \neq 0$. Let $\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{\Pi}(\boldsymbol{\Pi}^T \boldsymbol{\Pi})^T \boldsymbol{\Pi}^T$, thus

$$\boldsymbol{P} \boldsymbol{Y}_n \approx \boldsymbol{P} \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{P} \boldsymbol{Z}_n^T \boldsymbol{\beta} + \boldsymbol{P} \boldsymbol{\varepsilon}_n. \tag{5}$$

Here, we apply the two stage least squares procedure in [10], the proposed estimator is as follows:

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{Q}_n^T \boldsymbol{P} \boldsymbol{M} \boldsymbol{P} \boldsymbol{Q}_n)^{-1} \boldsymbol{Q}_n^T \boldsymbol{P} \boldsymbol{M} \boldsymbol{P} \boldsymbol{Y}_n,$$

where $\boldsymbol{M} = \boldsymbol{H}(\boldsymbol{H}^T \boldsymbol{H})^{-1} \boldsymbol{H}^T$, $\boldsymbol{H}$ is the instrumental variables, and

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{\Pi}^T \boldsymbol{\Pi})^{-1} \boldsymbol{\Pi} (\boldsymbol{Y}_n - \boldsymbol{Q}_n \hat{\boldsymbol{\theta}}).$$

Similar with [21], instrument variables are constructed as follows, step a) we directly use a simple linear regression of the dependent variable on the independent variable $\tilde{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\alpha}}$, step b) obtain $\bar{\boldsymbol{\theta}}$ and $\bar{\boldsymbol{\alpha}}$ by the instrument variable

$$\tilde{\boldsymbol{H}} = (\boldsymbol{W}_n (\boldsymbol{I}_n - \tilde{\boldsymbol{\theta}} \boldsymbol{W}_n)^{-1} (\boldsymbol{\Pi} \tilde{\boldsymbol{\alpha}}, \boldsymbol{Z}_n), \boldsymbol{Z}_n),$$

step c) obtain $\bar{\boldsymbol{\theta}}$ and $\bar{\boldsymbol{\alpha}}$ by the instrument variable

$$\boldsymbol{H}=(\boldsymbol{W}_n(\boldsymbol{I}_n-\bar{\boldsymbol{\theta}}\boldsymbol{W}_n)^{-1}(\boldsymbol{\Pi}\bar{\boldsymbol{\alpha}},\boldsymbol{Z}_n),\boldsymbol{Z}_n).$$

2.3 Test construction

To focuses on the problem of setting tests for partially functional linear spatial autoregressive models, the null and alternative hypotheses considered are

$$\boldsymbol{H}_0:\boldsymbol{P}\left(m\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right)=\boldsymbol{Z}_n\boldsymbol{\beta}_0+\int_0^1\boldsymbol{X}_n(t)\boldsymbol{\gamma}_0(t)dt\right)=1,\text{for some }\boldsymbol{\beta}_0\in\boldsymbol{R}^q\text{ and }\boldsymbol{\gamma}_0(t)\in L^2([0,1]),$$

$$\boldsymbol{H}_1:\boldsymbol{P}\left(m\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right)=\boldsymbol{Z}_n\boldsymbol{\beta}+\int_0^1\boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt\right)<1,\text{for all }\boldsymbol{\beta}\in\boldsymbol{R}^q\text{ and }\boldsymbol{\gamma}(t)\in L^2([0,1]).$$

When $\boldsymbol{H}_0$ holds, the model specific expression is

$$\boldsymbol{Y}_n=\rho\boldsymbol{W}_n\boldsymbol{Y}_n+\boldsymbol{Z}_n\boldsymbol{\beta}_0+\int_0^1\boldsymbol{X}_n(t)\boldsymbol{\gamma}_0(t)dt+\boldsymbol{\varepsilon}_n.$$

According to the parameter estimation method proposed above, the estimation of $\boldsymbol{\varepsilon}_n$ can be obtained as

$$\hat{\boldsymbol{\varepsilon}}_n=\boldsymbol{Y}_n-\hat{\rho}\boldsymbol{W}_n\boldsymbol{Y}_n-\boldsymbol{Z}_n\hat{\boldsymbol{\beta}}-\boldsymbol{\Pi}\hat{\boldsymbol{\alpha}}.$$

Remember that $\boldsymbol{U}_n=(\boldsymbol{Z}_n,\boldsymbol{X}_n)$ is the observation matrix, and let $\boldsymbol{U}_n=\left(\boldsymbol{U}_{n1},\boldsymbol{U}_{n2},\ldots,\boldsymbol{U}_{nn}\right)^T$, $\boldsymbol{U}=(\boldsymbol{Z},\boldsymbol{X})$. When the null hypothesis $\boldsymbol{H}_0$ holds,

$$\boldsymbol{Y}_n=\rho\boldsymbol{W}_n\boldsymbol{Y}_n+\boldsymbol{Z}_n\boldsymbol{\beta}_0+\int_0^1\boldsymbol{X}_n(t)\boldsymbol{\gamma}_0(t)dt+m(\boldsymbol{Z}_n,\boldsymbol{X}_n)-\boldsymbol{Z}_n\boldsymbol{\beta}_0-\int_0^1\boldsymbol{X}_n(t)\boldsymbol{\gamma}_0(t)dt+\boldsymbol{\varepsilon}_n.$$

holds. Let $\boldsymbol{\varepsilon}=m\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right)-\boldsymbol{Z}_n\boldsymbol{\beta}_0-\int_0^1\boldsymbol{X}_n(t)\boldsymbol{\gamma}_0(t)dt+\boldsymbol{\varepsilon}_n$, then $E\left(\boldsymbol{\varepsilon}|\boldsymbol{U}\right)=0$. According to [6], $E\left(\boldsymbol{\varepsilon}H\left(\boldsymbol{U},u\right)\right)=0$ holds for all $u\in R$, where $H\left(\boldsymbol{U},u\right)$ is some kind of weighting function. If indicator function is weighting function, the expression $\boldsymbol{h}$ holds for any projection direction $E\left(\boldsymbol{\varepsilon}I\left(<\boldsymbol{U},\boldsymbol{h}>\leqslant u\right)\right)=0$, where $\boldsymbol{h}\in L^2[0,1]$ and there is $\|\boldsymbol{h}\|=1$. For the projection direction $\boldsymbol{h}$, if $\boldsymbol{h}=(\boldsymbol{h}_1,\boldsymbol{h}_2)^T$, where $\boldsymbol{h}_1$ is a vector, and $\boldsymbol{h}_2$ is a function or process, then $<\boldsymbol{U},\boldsymbol{h}>=<\boldsymbol{Z},\boldsymbol{h}_1>+<\boldsymbol{X},\boldsymbol{h}_2>=\boldsymbol{Z}^T\boldsymbol{h}_1+\int_0^1\boldsymbol{X}(t)\boldsymbol{h}_2(t)dt$.

Construct residual empirical process:

$$C\bar{R}_n(u)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\boldsymbol{\varepsilon}_iI\left(<\boldsymbol{U}_{ni},\boldsymbol{h}>\leqslant u\right),$$

the expression for the Cramér-von Mises test statistic is as follows

$$C\bar{T}_n = \int \left(C\bar{R}_n(u)\right)^2 dF_n(u),$$

where $F_n(u)$ is the empirical distribution function of $\{<\boldsymbol{U}_{n1},\boldsymbol{h}>,<\boldsymbol{U}_{n2},\boldsymbol{h}>,\ldots,<\boldsymbol{U}_{nn},\boldsymbol{h}>\}$. Since the parameters are unknown, $\boldsymbol{\varepsilon}_i$ cannot be obtained directly, and its estimate $\hat{\boldsymbol{\varepsilon}}_i$ is used instead of $\boldsymbol{\varepsilon}_i$, in which case the residual empirical process $C\hat{R}_n(u)$ has an expression

$$C\hat{R}_n(u) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\boldsymbol{\varepsilon}}_i I\left(<\boldsymbol{U}_{ni},\boldsymbol{h}>\leqslant u\right),$$

The expression for the test statistic CT_n is

$$CT_n = \int \left(C\hat{R}_n(u)\right)^2 dF_n(u).$$

Subsequently, based on CT_n, the p-value of the test is calculated to make a judgment on this hypothesis testing problem.

3 Assumptions and asymptotic distributions

In this section, we first introduce some basic assumptions and then derive the asymptotic distributions of our test under the null hypotheses and study the power properties under the alternative hypotheses. Then, we propose bootstrap tests and justify their consistency.

3.1 Basic assumption

Before giving the assumptions, the following notation is given

$$\boldsymbol{I}_n(u) = \left(I\left(<\boldsymbol{U}_{n1},\boldsymbol{h}>\leqslant u\right), I\left(<\boldsymbol{U}_{n2},\boldsymbol{h}>\leqslant u\right),\ldots, I\left(<\boldsymbol{U}_{nn},\boldsymbol{h}>\leqslant u\right)\right)^T,$$

$$\boldsymbol{V}(u) = \lim_{n\to\infty}\frac{1}{n}\begin{pmatrix} \left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\left(\boldsymbol{Z}_n\boldsymbol{\beta}+\boldsymbol{\eta}\right)\right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u).$$

Assumption 1. *For Matrices $\boldsymbol{W}_n$ and $\left(\boldsymbol{I}_n-\rho\boldsymbol{W}_n\right)^{-1}$, the row sums and column sums are consistently bounded in absolute values.*

Assumption 2. *$\boldsymbol{I}_n - \rho \boldsymbol{W}_n$ is non-singular, elements in matrix $\boldsymbol{Z}_n$ are uniformly bounded and columns are full-ranked.*

Assumption 3. *The instrumental variable $\boldsymbol{H}$ is full rank and its elements are uniformly bounded.*

Assumption 4. *(1) Let $\boldsymbol{A} = \lim\limits_{n\to\infty} \frac{1}{n}\mathbf{E}\big(\boldsymbol{Q}_n^T \boldsymbol{P}\boldsymbol{H}\big)$, $\boldsymbol{A}$ exists;*

(2) Let $\boldsymbol{B} = \lim\limits_{n\to\infty} \frac{1}{n}\mathbf{E}\big(\boldsymbol{H}^T \boldsymbol{H}\big)$, $\boldsymbol{B}$ exists and is full rank;

(3) The eigenvalues of $\boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^T\boldsymbol{P}$ are bounded in probability.

Assumption 5. *There exist constants $a > 1$ and $b > a/2 + 1$ such that $\lambda_j - \lambda_{j+1} \geq Cj^{-a-1}$ and $\gamma_j| \leq Cj^{-b}$ hold for any $j \geq 1$.*

Assumption 6. *For the truncation parameter m, there are $m \sim O(n^{\frac{1}{a+2b}})$ and $nm^{-2b} \to 0$ as $m \to \infty$.*

Assumption 7. *$\boldsymbol{V}(u)$ is the bounded vector.*

Assumption 8. *Suppose $\boldsymbol{\Upsilon} = \lim\limits_{n\to\infty} \frac{1}{n} E\left(\boldsymbol{H}^T\boldsymbol{P}D(\boldsymbol{Z}_n, \boldsymbol{X}_n)\right) \neq 0$, $D(\cdot,\cdot)$ is a measurable function.*

Assumption 9. *Suppose $\boldsymbol{\varsigma} = \lim\limits_{n\to\infty} \frac{1}{n} E\left(D^T(\boldsymbol{Z}_n, \boldsymbol{X}_n)\boldsymbol{P}\boldsymbol{I}_n(u)\right) \neq 0$, $D(\cdot,\cdot)$ is a measurable function.*

Remark 1. *Assumption 1 is a fundamental condition for the subsequent study of the properties of the theory, the consistently bounded condition on the absolute value of the matrix $\boldsymbol{W}_n$ places certain constraints on the degree of spatial autocorrelation. Assumption 2 yields an explicit expression for the response variable $\boldsymbol{Y}_n$. Assumption 3 is a basic requirement for instrumental variables in the study of statistical inference, and its detailed construction can be found in [5]. Assumption 5 is the assumption underlying the functional linear model. Assumption 6 gives the rate of convergence required for the test and is important for subsequent proofs of theoretical properties. Assumption 8 and Assumption 9 are the additional conditions needed to prove the asymptotic properties of the parameter estimates $\hat{\boldsymbol{\theta}}$ and asymptotic distributions of the statistic under the alternative hypothesis.*

3.2 Asymptotic properties

This subsection focuses on the asymptotic distribution of the residual empirical process $C\hat{R}_n(u)$ and the test statistic CT_n under the null hypothesis and different alternative hypotheses.

Theorem 1. *Suppose that Assumptions 3.1-3.5 and $\boldsymbol{H}_0$ are satisfied, then*

$$C\hat{R}_n(u) \xrightarrow{L} CR(u),$$

where $CR(u)$ is a centred Gaussian process with a covariance function of

$$\lim_{n\to\infty} E\left(J(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u_1)J^T(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u_2)\right),$$

and

$$J(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u) = \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{\varepsilon}_n - \boldsymbol{V}^T(u)(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n,$$

$F(u)$ is the conditional distribution of $\{< \boldsymbol{U}_{n1}, \boldsymbol{h} >, < \boldsymbol{U}_{n2}, \boldsymbol{h} >, \cdots, < \boldsymbol{U}_{nn}, \boldsymbol{h} >\}$ when given $\boldsymbol{h}$.

Theorem 1 which gives the limiting distribution of the test statistic under $\boldsymbol{H}_0$ shows that the probability of making a Type I error can be well controlled.

To further illustrate the goodness-of-fit nature of the test statistic, consider Pitman local alternative hypothesis model

$$\boldsymbol{H}_1' : m(\boldsymbol{Z}_n, \boldsymbol{X}_n) = \boldsymbol{Z}_n\boldsymbol{\beta} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt + n^{-\frac{1}{2}}D(\boldsymbol{Z}_n, \boldsymbol{X}_n),$$

where $D(\cdot, \cdot)$ is the measurable function. If we make $\mathcal{F} = \{\boldsymbol{Z}_n\boldsymbol{\beta} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt | \boldsymbol{\beta} \in \boldsymbol{R}^q, \boldsymbol{\gamma}(t) \in L^2([0, 1])\}$ be a family of functions, for any natural number n, for any function G in $\mathcal{F}$, the measurable function $D(\cdot, \cdot)$ satisfies $\inf_{G\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n} ||D(\boldsymbol{Z}_n, \boldsymbol{X}_n) - G(\boldsymbol{Z}_n, \boldsymbol{X}_n)||^2 > 0$. That is, Pitman local alternative hypothes is $\boldsymbol{H}_1'$ shows that it is not possible to find $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}(t)$ such that $D(\boldsymbol{Z}_n, \boldsymbol{X}_n)$ can be expressed as $\boldsymbol{Z}_n\boldsymbol{\beta} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt$ which is of the form $E(\boldsymbol{\varepsilon}_n|\boldsymbol{Z}_n, \boldsymbol{X}_n) \neq 0$.

Theorem 2. *Suppose that Assumptions 3.1-3.7 and Pitman local alternative hypotheses $\boldsymbol{H}_1'$ are satisfied, then*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = (\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{\Upsilon} + (\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + o_p(1).$$

Based on the estimates of the parameters $\hat{\boldsymbol{\theta}}$ under $\boldsymbol{H}_1'$, we next give the limiting distributions of $C\hat{R}_n(u)$ and CT_n under $\boldsymbol{H}_1'$.

Theorem 3. *Suppose that Assumptions 3.1-3.7 and Pitman local alternative hypotheses $\boldsymbol{H}_1'$ are satisfied, then*

$$C\hat{R}_n(u) \xrightarrow{L} CR(u) + \boldsymbol{\Omega},$$

where offset term $\boldsymbol{\Omega} = -\boldsymbol{V}^T(u)(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{\Upsilon} + \boldsymbol{\varsigma}$, $CR(u)$ is described in Theorem 1.

Theorem 3 which gives the limiting distribution of the test statistic under $\boldsymbol{H}_1'$, shows that the this method is capable of distinguishing between the null hypothesis and the Pitman local alternative hypothesis at the usual rate of parameter convergence.

Consider the global alternative hypothesis model,

$$\boldsymbol{H}_1 : m(\boldsymbol{Z}_n, \boldsymbol{X}_n) = \boldsymbol{Z}_n\boldsymbol{\beta} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt + a_n D(\boldsymbol{Z}_n, \boldsymbol{X}_n),$$

where a_n satisfies $\sqrt{n}a_n \to \infty$.

Next we give the power of test under the global alternative hypothesis.

Theorem 4. *Suppose that Assumptions 3.1-3.7 and Pitman local alternative hypotheses $\boldsymbol{H}_1'$ are satisfied, then*

$$C\hat{R}_n(u) \to \infty,$$

$$CT_n \to \infty.$$

Combined with Theorem 1, Theorem 4 shows that the proposed test statistic is consistency.

3.3 Bootstrapped test statiatics

According to the theorem in subsection 3.2, it is easy to see that the distribution structure of the test statistic is more complex, and its critical value is not easy to calculate directly. Moreover, when the sample size is small, if we use the limiting distribution of the statistic to approximate the true distribution of the statistic, the error is large and distortion will occur. Therefore, we develop bootstrap method to calculate the critical value by bootstrap sampling method. Similar to [19], the detailed steps are as follows

Step 1: Compute the estimated residual $\hat{\boldsymbol{\varepsilon}}_n = \boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}} - \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}}$ and test statistic CT_n based on original sample $(\boldsymbol{Y}_n, \boldsymbol{Z}_n, \boldsymbol{X}_n)$;

Step 2: Generate a series of sequences of independent identically distributed random variables $\xi_1, \xi_2, \ldots, \xi_n$ with mean 0 and variance 1. Let $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)^T$, $\boldsymbol{\varepsilon}_n^* = (\hat{\varepsilon}_{n1}\xi_1, \hat{\varepsilon}_{n2}\xi_2, \ldots, \hat{\varepsilon}_{nn}\xi_n)^T$, $\boldsymbol{\varepsilon}_n^* = \hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi}$, Then the bootstrap sample of $\boldsymbol{Y}_n$ is expressed as $\boldsymbol{Y}_n^* = (\boldsymbol{I}_n - \hat{\rho}\boldsymbol{W}_n)^{-1}\left(\boldsymbol{Z}_n\hat{\boldsymbol{\beta}} + \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} + \boldsymbol{\varepsilon}_n^*\right)$.

Step 3: Using $(\boldsymbol{Y}_n^*, \boldsymbol{Z}_n, \boldsymbol{X}_n)$, compute $(\hat{\boldsymbol{\beta}}^*, \hat{\rho}^*, \hat{\boldsymbol{\alpha}}^*)$, the bootstrap analogue $(\hat{\boldsymbol{\beta}}, \hat{\rho}, \hat{\boldsymbol{\alpha}})$ and the test statistic CT_n^*;

Step 4: Repeat steps 2-3 Br times to get the specific value of the test statistic $CT_i^*, i = 1, 2, \cdots, Br$, the bootstrap value is defined

$$p = Br^{-1}\sum_{i=1}^{Br} I(CT_i^* \geqslant CT_n).$$

The next theorem establishes the consistency of the bootstrap method.

Theorem 5. *Suppose that Assumptions 3.1-3.7 are satisfied, based on sample* $(\boldsymbol{Y}_n, \boldsymbol{Z}_n, \boldsymbol{X}_n)$ *and bootstrap sample* $(\boldsymbol{Y}_n^*, \boldsymbol{Z}_n, \boldsymbol{X}_n)$*, then*

$$\sup_{\iota \in R} \left|P\left(CT_n^* \leqslant \iota\right) - P\left(CT_n \leqslant \iota\right)\right| = o_p(1).$$

Theorem 5 shows that it is reasonable to approximate the true distribution of the statistic CT_n by the bootstrap method, whether or not the null hypothesis holds.

4 Numerical and Empirical Studies

In this section, we examine the finite-sample performances of test through Monte Carlo simulations and a real data example.

4.1 Numerical Study

In this subsection, the Monte Carlo is used to investigate the finite sample performance of our method.

In the selection of the spatial weight matrix, the spatial weight matrix is selected as $\boldsymbol{W}_n = \boldsymbol{I}_R \bigotimes \boldsymbol{B}_m$ as in[3], where $\boldsymbol{B}_m = (\boldsymbol{\ell}_m \boldsymbol{\ell}_m^T - \boldsymbol{I}_m)/(m-1)$ and R is the number of spatial regions and m is the number of individuals in each region. Different spatial regions are uncorrelated, but different individuals within the same spatial region interact with each other and have the same degree of influence. The sample size n is $n = R \times m$.

In the constructed data generation models DGP1 to DGP6, two cases of homoskedasticity and heteroskedasticity are considered separately, and different degrees of spatial correlation $\rho_0 = 0.25, 0.5, 0.75$ are discussed in each case. Drift terms d and significance levels $\alpha = 0.01, 0.05, 0.1$, as a means of validating the performance of the proposed test statistic CT_n. Among them, DGP1 and DGP2 are the null hypothesis model; DGP3 and DGP4 are the alternative hypothesis model; DGP5 and DGP6, both hypothesis models are included, where d denotes the degree of deviation from the original hypothesis. All the simulation results are based on 1000 data repetitions with 500 bootstrap replications in each repetition. Each result of indicates the probability of making a Type I error or the power of the test statistic. The details of DGPs as follows:

DGP1: $\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{z}_{n1} + \boldsymbol{z}_{n2} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt + \boldsymbol{\varepsilon}_n.$

DGP2: $\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{z}_{n1} + \boldsymbol{z}_{n2} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt + \boldsymbol{\varepsilon}_n.$

DGP3: $\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + 2\left(\boldsymbol{z}_{n1} + \boldsymbol{z}_{n2} + \boldsymbol{z}_{n3} + \boldsymbol{z}_{n4} + \boldsymbol{z}_{n5}\right) + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt.$

DGP4: $\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + 2\left(\boldsymbol{z}_{n1} + \boldsymbol{z}_{n2} + \boldsymbol{z}_{n3} + \boldsymbol{z}_{n4} + \boldsymbol{z}_{n5}\right) + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt$
$+ 0.8\left(\cos\left(0.6\pi \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt\right) + \sin\left(\sum_{j=1}^{5} \boldsymbol{z}_{nj}\right)\right) + \boldsymbol{\varepsilon}_n.$

DGP5: $\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + \boldsymbol{z}_{n1} + \boldsymbol{z}_{n2} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt + d\exp\left(\frac{1}{5}\sum_{j=1}^{2} \boldsymbol{z}_{nj}^2 + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt\right)$
$+ \boldsymbol{\varepsilon}_n.$

DGP6: $\boldsymbol{Y}_n = \rho \boldsymbol{W}_n \boldsymbol{Y}_n + 2\boldsymbol{z}_{n1} + \boldsymbol{z}_{n2} + \boldsymbol{z}_{n3} + \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt + d\exp\left(2 * \sum_{j=1}^{3} \boldsymbol{z}_{nj}^2\right.$
$\left.+ \int_0^1 \boldsymbol{X}_n(t)\boldsymbol{\gamma}(t)dt\right) + \boldsymbol{\varepsilon}_n.$

For the explanatory variable $\boldsymbol{Z}_n$, there are two ways of generating it,

(1) In DGP1, DGP3, DGP4, DGP5, $\boldsymbol{Z}_n$ obeys a multivariate normal distribution;

(2) In DGP2, DGP6, $\boldsymbol{Z}_n$ obeys a uniform distribution.

Functional data part is similar to [17], In DGP1-DGP5, the value of the functional parameter is set to be $\gamma(t) = \sqrt{2}\sin(\pi t/2) + 3\sqrt{2}\sin(3\pi t/2)$, $X_i(t) = \sum_{j=1}^{100} \gamma_j \phi_j(t)$, where γ_j is an independent identically distributed normal random variable with mean 0 and variance $\lambda_j = ((j-0.5)\pi)^{-2}$, and $\phi_j(t) = \sqrt{2}\sin((j-0.5)\pi)$. In DGP6, the functional parameterr becomes $\gamma(t) = \sin(\pi t/2) + \frac{1}{2}\sin(3\pi t/2) + \frac{1}{4}\sin(5\pi t/2)$, and the values of λ_j and $\phi_j(t)$ are kept constant.

For the error term, there are two ways of generating homoskedasticity and heteroskedasticity respectivel,

(1) ε_{ni} obeys a normal distribution;

(2) $\varepsilon_{ni} = \sqrt{2}(1 + x_{n1})\bar{\varepsilon}_{ni}$, where $\bar{\varepsilon}_{ni}$ obeys a normal distribution.

From Table 1 and Table 2, it can be seen that under the null hypothesis, the empirical level of the test statistic is close to α, as is also the case with heteroscedasticity, suggesting that the probability of committing a Type I error can be well controlled by the proposed test statistic. In Table 3 and Table 4, the power of the test under different alternative hypotheses converges to 1 as the sample size increases, and this result holds for the heteroskedastic case. And in Table 5 and Table 6, as the offset d increases, the power of the test statistic becomes larger and tends to 1. The above simulation results show the validity of the method proposed in this article.

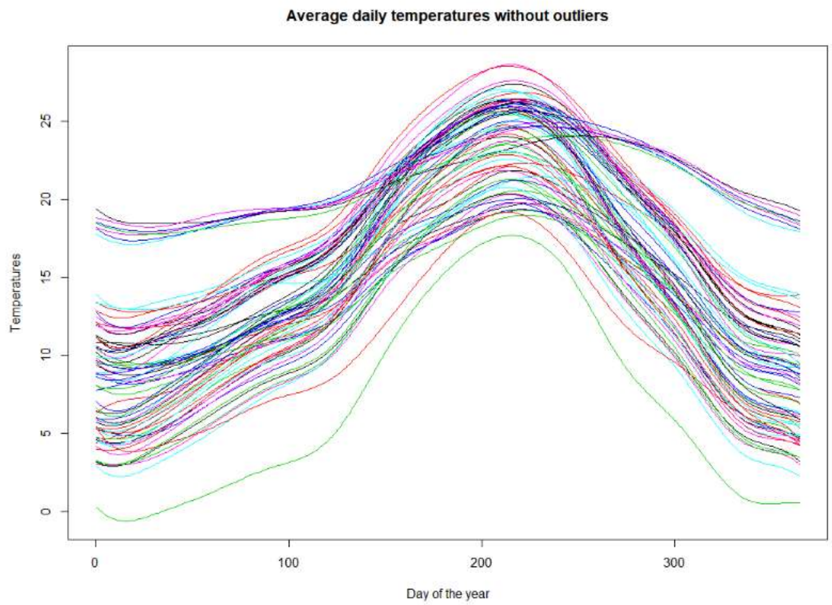

Figure 1:

4.2 Illustrative Example

In order to better illustrate the usefulness of our method in real data analysis, this subsection considers the application of Spanish meteorological data. The dataset contains meteorological monitoring data from 1980 to 2009, 73 of Spanish meteorological stations, and the main variables are the geographic location information of each station, as well as the daily data of temperature, precipitation, and wind speed, where the geographic location information includes the variables of the province to which it belongs to, the longitude, the latitude and others. Similar to [7], the average temperature curve for the period from 1980 to 2009 as an explanatory variable, the functional variable temperature $X_n(t)$ contains the daily temperature values of all the meteorological stations, as shown in the Figure 1.

From Figure 1, it can be seen that there are 7 curves clearly clustered together and have a large deviation from the majority of the curves, and at the same time there is 1 curve that also have a large deviation. In order to make more accurate judgments, it is considered as an outlier and deleted the 8 curve data. Considering the response variable Y as the number of sunny days per year from 1980 to 2009, it is also deleted because the response variable contains 14 missing values. Therefore actual data analysis contains 51 observations. The variables longitude Z_1 and latitude Z_2 are selected and the following model is constructed,

$$Y = \rho \boldsymbol{W}_n Y + Z_1\beta_1 + Z_2\beta_2 + \int_0^{365} X(t)\gamma(t)dt + V.$$

For the above model, the spatial weight matrix is constructed similarly to [18] with the following expression

$$W_{ij} = \exp(-||s_i - s_j|| \Big/ \sum_{l\neq i}(-||s_i - s_l||),$$

where $s_i = (Log_i, Lat_i)$ denotes the geographic location of the i weather station, which consists of latitude and longitude, with the specific expression and $||\cdot||$ denotes the Euclidean distance. Based on our method, the p-value of the test is calculated to be 0.96, so the setting of the partial functional linear spatial autoregressive model cannot be rejected.

5 Conclusion

This article consider the combination of functional data and spatial data to study the setting test problem of partially functional linear spatial autoregressive models. Based on the generalized moment estimation form of the parameters, the residuals are estimated, which are combined with the linear projections of the covariates to generate the empirical process and establish the test statistic to realize the specification test of the partially functional linear spatial autoregressive model. The proposed test avoids the problem of subjective selection of unnecessary parameters and has wide applicability. First, under some assumptions, the limiting distribution of the proposed test is proved and shown to be able to distinguish between the null hypothesis and the Pitman local alternative hypothesis at the usual rate of parameter convergence, which further gives the compatibility of the statistics. In finite size, we approximate the distribution of the statistic by the bootstrap method, and the simulation results show that this method has well empirical level or power, both in null hypothesis and alternative hypotheses. Meanwhile, the simulation results also show that our method performs well in the heteroskedasticity case.

The test method proposed in this articl can be applied to other functional-type data and spatial data models with strong flexibility. Future research can consider the performance of our method in semiparametric models such as functional semiparametric spatial autoregressive models or consider the performance of this method under different weighting functions.

References

[1] Benatia D, Carrasco M, Florens J P. Functional linear regression with functional response[J]. Journal of Econometrics, 2017, 201(2): 269-291.

[2] Cardot H, Ferraty F, Sarda P. Splines estimators for the functional linear model[J]. Statistica Sinica, 2003, 13(3): 571-591.

[3] Case A C. Spatial patterns in household demand[J]. Econometrica, 1991, 59(4): 953-965.

[4] Cliff A D, Ord J K. Spatial processes: models and applications[M]. London: Pion, 1981.

[5] Du J, Sun X Q, Cao R Y, Zhang Z Z. Statistical inference for partially linear additive spatial autoregressive models[J]. Spatial Statistics, 2018, 25: 52-67.

[6] Escanciano J C. A consistent diagnostic test for regression models using projections[J]. Econometric Theory, 2006, 22: 1030-1051.

[7] Febrero-Bande M, Galeano P, Gonzalez-Manteiga W. Estimaiton, imputation and prediction for the functional linear model with scalar response with responses missing at random[J]. Computational Statistics and Data Analysis, 2019, 131: 91-103.

[8] Hall P, Horowitz J L. Methodology and convergence rates for functional linear regression[J]. The Annals of Statistics, 2007, 35(1): 70-91.

[9] Hu Y P, Wu S Y, Feng S Y, Jin J L. Estimation in partial functional linear spatial autoregressive model[J]. Mathematics, 2020, 8(10): 1680.

[10] Kelejian H H, Prucha I R. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances[J]. The Journal of Real Estate Finance and Economics, 1998, 17: 99-121.

[11] Kelejian H H, Prucha I R. A generalized moments estimator for the autoregressive parameter in a spatial model[J]. International Economic Review, 1999, 40(2): 509-533.

[12] Lee L F. Asymptotic distribution of quasi-maximum likelihood estimators for spatial autoregressive models[J]. Econometrica, 2004, 72: 1899-1925.

[13] Lee E R, Park B U. Sparse estimation in functional linear regression[J]. Jornal of Multivariate Analysis, 2012, 105: 1-17.

[14] Matsui H. Sparse varying-coefficient functional linear model[J]. arXiv preprint arXiv: 2110. 12599, 2021.

[15] Müller H G, Stadtmuller U. Generalized functional linear models[J]. The Annals of Statistics, 2005, 33(2): 774-805.

[16] Ramsay J O. When the data are functions[J]. Psychometrika, 1982, 47(4): 379-396.

[17] Shin H. Partial functional linear regression[J]. Journal of Statistical Planning and Inference, 2009, 139(10): 3405-3418.

[18] Sun Y, Yan H J, Zhang W Y, Lu Z D. A semiparametric spatial dynamic model[J]. The Annals of Statistics, 2014, 42(2): 700-727.

[19] Sun Z, Ye X, Sun L. Consistent test for parametric models with right-censored data using projections[J]. Computational Statistics and Data Analysis, 2018, 118: 112-125.

[20] Xie T F, Cao R Y, Du J. Variable selection for spatial autoregressive models with a diverging number of parameters[J]. Statistical Papers, 2020, 61: 1125-1145.

[21] Zhang Y Q, Shen D M. Estimaiton of semi-parametric varying-coefficient spatial panel data models with random-effects[J]. Journal of Statistical Planning and Inference, 2015, 159: 64-80.

Appendix: Proofs of the Theorems

Proof of Theorem 1. When H_0 is satisfied, the parameters are estimated to be

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Y}_n,$$

and

$$\hat{\boldsymbol{\alpha}} = \left(\boldsymbol{\Pi}^T\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}^T\left(\boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\right).$$

According to the specific form of $\hat{\boldsymbol{\theta}}$

$$\begin{aligned}
\hat{\boldsymbol{\theta}} &= \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Y}_n \\
&= \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P} \\
&\quad \left(\boldsymbol{Q}_n\boldsymbol{\theta} + \boldsymbol{\Pi}\boldsymbol{\alpha} + \boldsymbol{e}_n + \boldsymbol{\varepsilon}_n\right) \\
&= \boldsymbol{\theta} + \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{e}_n \\
&\quad + \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n \\
&= \boldsymbol{\theta} + L_1 + L_2,
\end{aligned}$$

where

$$L_1 = \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{e}_n,$$

and

$$L_2 = \left(\boldsymbol{Q}_n^T \boldsymbol{P}\boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n.$$

From Assumption 4,

$$\begin{aligned}
&\lim_{n\to\infty}\Big(\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\big(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\big)^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\Big)^{-1}\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\big(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\big)^{-1}\\
= & \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1} + o_p(1)
\end{aligned}$$

exists and bounded.

Because

$$\begin{aligned}
\frac{1}{n}\boldsymbol{e}_n^T\boldsymbol{P}\boldsymbol{H}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{e}_n &= \frac{1}{n^2}\boldsymbol{e}_n^T\boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{e}_n\\
&\leq \frac{1}{n^2}\lambda_{\max}\left(\boldsymbol{P}\boldsymbol{H}\boldsymbol{H}^T\boldsymbol{P}\right)\boldsymbol{e}_n^T\boldsymbol{e}_n\\
&= O\Big(\frac{1}{n^2}nm^{-2b+1}\Big)\\
&= O\Big(\frac{m^{-2b+1}}{n}\Big)\\
&= O\Big(n^{-\frac{a+4b-1}{a+2b}}\Big).
\end{aligned}$$

Then $L_1 = O_p\Big(n^{-\frac{a+4b-1}{2(a+2b)}}\Big)$. Similarly to L_1,

$$L_2 = \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + o_p(1).$$

Thus

$$\begin{aligned}
\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) &= \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + \sqrt{n}O_p\Big(n^{-\frac{a+4b-1}{2(a+2b)}}\Big)\\
&= \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + O_p\Big(n^{-\frac{2b-1}{2(a+2b)}}\Big)\\
&= \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + o_p(1),
\end{aligned}$$

and

$$\begin{aligned}
\hat{\boldsymbol{\varepsilon}}_n &= \boldsymbol{Y}_n - \hat{\rho}\boldsymbol{W}_n\boldsymbol{Y}_n - \boldsymbol{Z}_n\hat{\boldsymbol{\beta}} - \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}}\\
&= \boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}} - \boldsymbol{\Pi}(\boldsymbol{\Pi}^T\boldsymbol{\Pi})^{-1}\boldsymbol{\Pi}^T\big(\boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\big)\\
&= \big(\boldsymbol{I}_n - \boldsymbol{\Pi}(\boldsymbol{\Pi}^T\boldsymbol{\Pi})^{-1}\boldsymbol{\Pi}^T\big)\big(\boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\big)
\end{aligned}$$

$$\begin{aligned}
&= \big(\boldsymbol{I}_n - \boldsymbol{\Pi}(\boldsymbol{\Pi}^T\boldsymbol{\Pi})^{-1}\boldsymbol{\Pi}^T\big)\big(\boldsymbol{Q}_n\boldsymbol{\theta} + \boldsymbol{\Pi}\boldsymbol{\alpha} + \boldsymbol{e}_n + \boldsymbol{\varepsilon}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\big) \\
&= \boldsymbol{P}\left(\boldsymbol{Q}_n\boldsymbol{\theta} + \boldsymbol{\Pi}\boldsymbol{\alpha} + \boldsymbol{e}_n + \boldsymbol{\varepsilon}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\right) \\
&= \boldsymbol{P}\boldsymbol{Q}_n(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) + \boldsymbol{P}\boldsymbol{e}_n + \boldsymbol{P}\boldsymbol{\varepsilon}_n.
\end{aligned}$$

Due to $\sqrt{n}\big(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\big) = O_p(1)$,

$$\begin{aligned}
C\hat{R}_n(u) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\varepsilon}_i I\big(\langle \boldsymbol{U}_{ni}, \boldsymbol{h}\rangle \leqslant u\big) \\
&= \frac{1}{\sqrt{n}}\hat{\boldsymbol{\varepsilon}}_n^T\boldsymbol{I}_n(u) \\
&= \frac{1}{\sqrt{n}}\big(\boldsymbol{P}\boldsymbol{Q}_n(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) + \boldsymbol{P}\boldsymbol{e}_n + \boldsymbol{P}\boldsymbol{\varepsilon}_n\big)^T\boldsymbol{I}_n(u) \\
&= \frac{1}{\sqrt{n}}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{I}_n(u) + \frac{1}{\sqrt{n}}\boldsymbol{e}_n^T\boldsymbol{P}\boldsymbol{I}_n(u) + \frac{1}{\sqrt{n}}\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \Pi_1 + \Pi_2 + \Pi_3,
\end{aligned}$$

where $\Pi_1 = \frac{1}{\sqrt{n}}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{I}_n(u)$, $\Pi_2 = \frac{1}{\sqrt{n}}\boldsymbol{e}_n^T\boldsymbol{P}\boldsymbol{I}_n(u)$, $\Pi_3 = \frac{1}{\sqrt{n}}\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u)$.

For Π_1,

$$\begin{aligned}
\Pi_1 &= \frac{1}{\sqrt{n}}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \frac{1}{n}\sqrt{n}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \frac{1}{n}\sqrt{n}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\left(\big(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\boldsymbol{Z}_n\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_n)\big)^T, \boldsymbol{Z}_n^T\right)^T\boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \frac{1}{n}\sqrt{n}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\left(\big(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\boldsymbol{Z}_n\boldsymbol{\beta} + \boldsymbol{\eta})\big)^T, \boldsymbol{Z}_n^T\right)^T\boldsymbol{P}\boldsymbol{I}_n(u) \\
&\quad + \frac{1}{n}\sqrt{n}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\left(\big(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\boldsymbol{\varepsilon}_n\big)^T, \boldsymbol{0}\right)^T\boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \mathrm{III}_1 + \mathrm{III}_2,
\end{aligned}$$

where

$$\mathrm{III}_1 = \frac{1}{\mathrm{n}}\sqrt{\mathrm{n}}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^{\mathrm{T}}\left(\big(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{S}_{\mathrm{n}}^{-1}(\boldsymbol{Z}_{\mathrm{n}}\boldsymbol{\beta} + \boldsymbol{\eta})\big)^{\mathrm{T}}, \boldsymbol{Z}_{\mathrm{n}}^{\mathrm{T}}\right)^{\mathrm{T}}\boldsymbol{P}\boldsymbol{I}_{\mathrm{n}}(\mathrm{u}),$$

$$\mathrm{III}_2 = \frac{1}{\mathrm{n}}\sqrt{\mathrm{n}}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^{\mathrm{T}}\left(\big(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{S}_{\mathrm{n}}^{-1}\boldsymbol{\varepsilon}_{\mathrm{n}}\big)^{\mathrm{T}}, \boldsymbol{0}\right)^{\mathrm{T}}\boldsymbol{P}\boldsymbol{I}_{\mathrm{n}}(\mathrm{u}).$$

For III_1,

$$\mathrm{III}_1 = \sqrt{n}\big(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\big)^T\frac{1}{n}\left(\big(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\boldsymbol{Z}_n\boldsymbol{\beta} + \boldsymbol{\eta})\big)^T, \boldsymbol{Z}_n^T\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)$$

$$
\begin{aligned}
&= \sqrt{n}\big(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\big)^T\big(\boldsymbol{V}(u)+o(1)\big) \\
&= -\boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n+o_p\left(1\right).
\end{aligned}
$$

For III_2, because

$$
\mathbf{E}\Big(\frac{1}{n}\boldsymbol{\varepsilon}_n^T\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)\Big)=\mathbf{E}\Big(\frac{1}{n}\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\boldsymbol{\varepsilon}_n\Big)=0,
$$

and

$$
\begin{aligned}
&\frac{1}{n^2}\mathbf{E}\left(||\boldsymbol{\varepsilon}_n^T\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)||^2\right) \\
=\;& \frac{1}{n^2}\mathbf{E}\Big(\mathrm{trace}\Big(\boldsymbol{I}_\mathrm{n}^\mathrm{T}(\mathrm{u})\boldsymbol{P}\boldsymbol{W}_\mathrm{n}\boldsymbol{S}_\mathrm{n}^{-1}\boldsymbol{\varepsilon}_\mathrm{n}\boldsymbol{\varepsilon}_\mathrm{n}^\mathrm{T}\left(\boldsymbol{W}_\mathrm{n}\boldsymbol{S}_\mathrm{n}^{-1}\right)^\mathrm{T}\boldsymbol{P}\boldsymbol{I}_\mathrm{n}(\mathrm{u})\Big)\Big) \\
\leq\;& \sigma^2\frac{1}{n^2}\mathbf{E}\Big(\mathrm{trace}\Big(\boldsymbol{I}_\mathrm{n}^\mathrm{T}\left(\mathrm{u}\right)\boldsymbol{P}\boldsymbol{W}_\mathrm{n}\boldsymbol{S}_\mathrm{n}^{-1}\left(\boldsymbol{W}_\mathrm{n}\boldsymbol{S}_\mathrm{n}^{-1}\right)^\mathrm{T}\boldsymbol{P}\boldsymbol{I}_\mathrm{n}(\mathrm{u})\Big)\Big) \\
=\;& \sigma^2\frac{1}{n^2}O\left(n\right) \\
=\;& O\left(\frac{1}{n}\right),
\end{aligned}
$$

we have

$$
\frac{1}{n}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\boldsymbol{\varepsilon}_n\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)=O_p\left(\frac{1}{\sqrt{n}}\right)=o_p\left(1\right).
$$

So

$$
\begin{aligned}
\mathrm{III}_2 &= \sqrt{n}\big(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\big)^T\frac{1}{n}\Big(\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\boldsymbol{\varepsilon}_n\right)^T,\mathbf{0}\Big)^T\boldsymbol{P}\boldsymbol{I}_n(u) \\
&= O_p\left(1\right)o_p\left(1\right) \\
&= o_p\left(1\right),
\end{aligned}
$$

organize to get

$$
\Pi_1=-\boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n+o_p(1).
$$

For Π_2,

$$
||\boldsymbol{e}_n||^2=nm^{-2b+1},
$$

$$
||\boldsymbol{P}\boldsymbol{I}_n(u)||\leqslant||\boldsymbol{I}_n(u)||=\sqrt{n},
$$

then

$$
\frac{1}{\sqrt{n}}\boldsymbol{e}_n^T\boldsymbol{P}\boldsymbol{I}_n(u)\;\;\leq\;\;\frac{1}{\sqrt{n}}||\boldsymbol{e}_n||\;||\boldsymbol{P}\boldsymbol{I}_n(u)||
$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{n}}\sqrt{n}m^{\frac{-2b+1}{2}}\sqrt{n} \\
&= \left(nm^{-2b+1}\right)^{\frac{1}{2}} \\
&= o_p(1).
\end{aligned}
$$

We conclude

$$\Pi_2 = o_p\left(1\right).$$

Combining the above results,

$$
\begin{aligned}
\hat{\mathrm{CR}}_{\mathrm{n}}\left(\mathrm{u}\right) &= \frac{1}{\sqrt{n}}\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u) - \boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + o_p\left(1\right) \\
&= \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{\varepsilon}_n - \boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + o_p\left(1\right) \\
&= J\left(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u\right) + o_p\left(1\right).
\end{aligned}
$$

Thus

$$\hat{\mathrm{CR}}_{\mathrm{n}}(\mathrm{u}) \xrightarrow{\mathrm{L}} \mathrm{CR}(\mathrm{u}),$$

where CR(u) is a centered Gaussian process with a covariance function of

$$
\begin{aligned}
&\lim_{n\to\infty}\mathbf{E}\left(J(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u_1)J^T(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u_2)\right) \\
= &\lim_{n\to\infty}\mathbf{E}\Bigg(\left(\frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u_1)\boldsymbol{P}\boldsymbol{\varepsilon}_n - \boldsymbol{V}^T(u_1)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n\right) \\
&\left(\frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u_2)\boldsymbol{P}\boldsymbol{\varepsilon}_n - \boldsymbol{V}^T(u_2)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n\right)^T\Bigg) \\
= &\lim_{n\to\infty}\Big(\frac{1}{n}\mathbf{E}\left(\boldsymbol{I}_n^T(u_1)\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u_2)\right) \\
&-\frac{1}{n}\mathbf{E}\big(\boldsymbol{I}_n^T(u_1)\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{H}\boldsymbol{B}^{-1}\boldsymbol{A}^T\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{V}(u_2)\big) \\
&-\frac{1}{n}\mathbf{E}\big(\boldsymbol{V}^T(u_1)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u_2)\big) \\
&+\frac{1}{n}\mathbf{E}\big(\boldsymbol{V}^T(u_1)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{H}\boldsymbol{B}^{-1}\boldsymbol{A}^T\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{V}(u_2)\big)\Big) \\
= &\frac{\sigma^2}{n}\Big(\boldsymbol{I}_n^T(u_1)\boldsymbol{P}\boldsymbol{I}_n(u_2) - \boldsymbol{I}_n^T(u_1)\boldsymbol{P}\boldsymbol{H}\boldsymbol{B}^{-1}\boldsymbol{A}^T\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{V}(u_2) \\
&-\boldsymbol{V}^T(u_1)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{I}_n(u_2) + \boldsymbol{V}^T(u_1)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{V}(u_2)\Big).
\end{aligned}
$$

Because the integral operator is continuous, according to the continuous mapping theorem on the sampling space,

$$CT_n \xrightarrow{L} \int \left(CR(u)\right)^2 dF(u).$$

Proof of Theorem 2. When local alternative hypotheses $\boldsymbol{H}_1'$ is satisfied,

$$\begin{aligned}
\hat{\boldsymbol{\theta}} &= \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Y}_n \\
&= \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P} \\
&\quad \left(\boldsymbol{Q}_n\boldsymbol{\theta}+\boldsymbol{\Pi}\boldsymbol{\alpha}+\boldsymbol{e}_n+n^{-\frac{1}{2}}D(\boldsymbol{Z}_n,\boldsymbol{X}_n)+\boldsymbol{\varepsilon}_n\right) \\
&= \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P} \\
&\quad \left(\boldsymbol{Q}_n\boldsymbol{\theta}+\boldsymbol{e}_n+n^{-\frac{1}{2}}D(\boldsymbol{Z}_n,\boldsymbol{X}_n)+\boldsymbol{\varepsilon}_n\right) \\
&= \boldsymbol{\theta}+M_1+M_2+M_3,
\end{aligned}$$

where

$$M_1 = \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{e}_n,$$

$$M_2 = \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right),$$

$$M_3 = \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n.$$

From Assumption 8 and Theorem 1,

$$\begin{aligned}
M_1 &= \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{e}_n \\
&= \left(\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\Big(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\Big)^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\Big(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\Big)^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{e}_n \\
&= O_p\left(\frac{m^{\frac{-2b+1}{2}}}{\sqrt{n}}\right) \\
&= O_p\left(n^{-\frac{a+4b-1}{2(a+2b)}}\right),
\end{aligned}$$

$$\begin{aligned}
M_2 &= \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right) \\
&= \left(\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right) \\
&= \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}n^{-\frac{1}{2}}\boldsymbol{\Upsilon}+o_p\left(1\right),
\end{aligned}$$

and

$$M_3 = \left(\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n$$

$$
\begin{aligned}
&= \left(\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{Q}_n\right)^{-1}\frac{1}{n}\boldsymbol{Q}_n^T\boldsymbol{P}\boldsymbol{H}\left(\frac{1}{n}\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n \\
&= \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{n}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + o_p(1).
\end{aligned}
$$

Combining the above results,

$$
\begin{aligned}
&\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right) \\
= {}& \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\sqrt{n}n^{-\frac{1}{2}}\boldsymbol{\Upsilon} + \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + \sqrt{n}O_p\left(n^{-\frac{a+4b-1}{2(a+2b)}}\right) \\
= {}& \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\sqrt{n}n^{-\frac{1}{2}}\boldsymbol{\Upsilon} + \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + O_p\left(n^{-\frac{2b-1}{2(a+2b)}}\right) \\
= {}& \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{\Upsilon} + \left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n + o_p(1).
\end{aligned}
$$

Proof of Theorem 3. From local alternative hypotheses $\boldsymbol{H}_1'$ and Theorem 2,

$$
\begin{aligned}
\hat{\boldsymbol{\varepsilon}}_n &= \boldsymbol{Y}_n - \hat{\rho}\boldsymbol{W}_n\boldsymbol{Y}_n - \boldsymbol{Z}_n\hat{\boldsymbol{\beta}} - \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} \\
&= \boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}} - \boldsymbol{\Pi}\left(\boldsymbol{\Pi}^T\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}^T\left(\boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\right) \\
&= \left(\boldsymbol{I}_n - \boldsymbol{\Pi}\left(\boldsymbol{\Pi}^T\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}^T\right)\left(\boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\right) \\
&= \boldsymbol{P}\left(\boldsymbol{Y}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\right) \\
&= \boldsymbol{P}\left(\boldsymbol{Q}_n\boldsymbol{\theta} + \boldsymbol{\Pi}\boldsymbol{\alpha} + \boldsymbol{e}_n + n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n, \boldsymbol{X}_n\right) + \boldsymbol{\varepsilon}_n - \boldsymbol{Q}_n\hat{\boldsymbol{\theta}}\right) \\
&= \boldsymbol{P}\boldsymbol{Q}_n(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}) + \boldsymbol{P}\boldsymbol{e}_n + \boldsymbol{P}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n, \boldsymbol{X}_n\right) + \boldsymbol{P}\boldsymbol{\varepsilon}_n,
\end{aligned}
$$

then

$$
\begin{aligned}
C\hat{R}_n(u) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\varepsilon}_i I\left(<\boldsymbol{U}_{ni}, \boldsymbol{h}>\leqslant u\right) \\
&= \frac{1}{\sqrt{n}}\hat{\boldsymbol{\varepsilon}}_n^T\boldsymbol{I}_n(u) \\
&= \frac{1}{\sqrt{n}}\left(\boldsymbol{P}\boldsymbol{Q}_n(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}) + \boldsymbol{P}\boldsymbol{e}_n + \boldsymbol{P}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n, \boldsymbol{X}_n\right) + \boldsymbol{P}\boldsymbol{\varepsilon}_n\right)^T\boldsymbol{I}_n(u) \\
&= \frac{1}{\sqrt{n}}\left(\boldsymbol{P}\boldsymbol{Q}_n(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\right)^T\boldsymbol{I}_n(u) + \frac{1}{\sqrt{n}}\left(\boldsymbol{P}\boldsymbol{e}_n\right)^T\boldsymbol{I}_n(u) \\
&\quad + \frac{1}{\sqrt{n}}\left(\boldsymbol{P}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n, \boldsymbol{X}_n\right)\right)^T\boldsymbol{I}_n(u) + \frac{1}{\sqrt{n}}\left(\boldsymbol{P}\boldsymbol{\varepsilon}_n\right)^T\boldsymbol{I}_n(u) \\
&= \psi_1 + \psi_2 + \psi_3 + \psi_4,
\end{aligned}
$$

where

$$\psi_1 = \frac{1}{\sqrt{n}} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \boldsymbol{Q}_n^T \boldsymbol{P}\boldsymbol{I}_n(u),$$

$$\psi_2 = \frac{1}{\sqrt{n}} \boldsymbol{e}_n^T \boldsymbol{P}\boldsymbol{I}_n(u),$$

$$\psi_3 = \frac{1}{\sqrt{n}} n^{-\frac{1}{2}} D^T \left(\boldsymbol{Z}_n, \boldsymbol{X}_n\right) \boldsymbol{P}\boldsymbol{I}_n(u),$$

and

$$\psi_4 = \frac{1}{\sqrt{n}} \boldsymbol{\varepsilon}_n^T \boldsymbol{P}\boldsymbol{I}_n(u).$$

For ψ_1,

$$\begin{aligned}
\psi_1 &= \frac{1}{\sqrt{n}} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \boldsymbol{Q}_n^T \boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \frac{1}{n} \boldsymbol{Q}_n^T \boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \frac{1}{n} \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1} \left(\boldsymbol{Z}_n \boldsymbol{\beta} + \boldsymbol{\eta} + n^{-\frac{1}{2}} D\left(\boldsymbol{Z}_n, \boldsymbol{X}_n\right) + \boldsymbol{\varepsilon}_n\right)\right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix} \boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \tau_1 + \tau_2 + \tau_3,
\end{aligned}$$

where

$$\tau_1 = \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \frac{1}{n} \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1} \left(\boldsymbol{Z}_n \boldsymbol{\beta} + \boldsymbol{\eta}\right)\right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix} \boldsymbol{P}\boldsymbol{I}_n(u),$$

$$\tau_2 = \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \frac{1}{n} \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1} n^{-\frac{1}{2}} D(\boldsymbol{Z}_n, \boldsymbol{X}_n)\right)^T \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{P}\boldsymbol{I}_n(u)$$

and

$$\tau_3 = \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \frac{1}{n} \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1} \boldsymbol{\varepsilon}_n\right)^T \boldsymbol{0} \end{pmatrix} \boldsymbol{P}\boldsymbol{I}_n(u).$$

From the proof of Theorem 1 and Theorem 2,

$$\begin{aligned}
\tau_1 &= \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \frac{1}{n} \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1} \left(\boldsymbol{Z}_n \boldsymbol{\beta} + \boldsymbol{\eta}\right)\right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix} \boldsymbol{P}\boldsymbol{I}_n(u) \\
&= \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \left(\boldsymbol{V}(u) + o(1)\right),
\end{aligned}$$

$$\tau_3 = \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \frac{1}{n} \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1} \boldsymbol{\varepsilon}_n\right)^T \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{P}\boldsymbol{I}_n(u) = o_p(1).$$

For τ_2

$$\begin{aligned}
&\frac{1}{n^2}\|\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right)\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)\|^2\\
= \;& O\left(\frac{1}{n^3}\right)O\left(n\right)\\
= \;& o(1),
\end{aligned}$$

thus

$$\frac{1}{n}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}n^{-\frac{1}{2}}D\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right)\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)=o_p(1),$$

then

$$\begin{aligned}
\tau_2 &= \sqrt{n}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T\frac{1}{n}\begin{pmatrix}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}n^{-\frac{1}{2}}D(\boldsymbol{Z}_n,\boldsymbol{X}_n)\right)^T\\ \boldsymbol{0}\end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u)\\
&= o_p(1).
\end{aligned}$$

We conclude

$$\psi_1=-\boldsymbol{V}^T(u)\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)+o_p(1).$$

According to the proof of Theorem 1, we have $\psi_2=o_p(1)$. From Assumption 7,

$$\begin{aligned}
\psi_3 &= \frac{1}{\sqrt{n}}n^{-\frac{1}{2}}D^T(\boldsymbol{Z}_n,\boldsymbol{X}_n)\boldsymbol{P}\boldsymbol{I}_n(u)\\
&= \frac{1}{n}D^T(\boldsymbol{Z}_n,\boldsymbol{X}_n)\boldsymbol{P}\boldsymbol{I}_n(u)\\
&= \boldsymbol{\varsigma}+o(1).
\end{aligned}$$

So

$$\begin{aligned}
C\hat{R}_n(u) &= \frac{1}{\sqrt{n}}\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u)-\boldsymbol{V}^T(u)\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)+\boldsymbol{\varsigma}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{\varepsilon}_n-\boldsymbol{V}^T(u)\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)+\boldsymbol{\varsigma}+o_p(1).
\end{aligned}$$

Besides,

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)=\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{\Upsilon}+\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n+o_p\left(1\right),$$

then

$$C\hat{R}_n(u) = \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{\varepsilon}_n-\boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n$$

$$-\boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{\Upsilon}+\boldsymbol{\varsigma}+o_p(1)$$
$$= J(\boldsymbol{Y}_n,\rho,\boldsymbol{Z}_n,\boldsymbol{\beta},\boldsymbol{X}_n,u)+\boldsymbol{\Omega}+o_p(1).$$

Then

$$C\hat{R}_n(u)\xrightarrow{L} CR(u)+\boldsymbol{\Omega},$$

where $\boldsymbol{\Omega}=-\boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{\Upsilon}+\boldsymbol{\varsigma}$ and $CR(u)$ is described in Theorem 1.

Because the integral operator is continuous, according to the continuous mapping theorem on the sampling space,

$$CT_n\xrightarrow{L}\int\left(CR(u)+\boldsymbol{\Omega}\right)^2 dF(u).$$

Proof of Theorem 4. Let

$$\boldsymbol{\omega}(a,u)=\lim_{n\to\infty}\frac{1}{n}\begin{pmatrix}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}a_nD\left(\boldsymbol{Z}_n,\boldsymbol{X}_n\right)\right)^T\\ \mathbf{0}\end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n\left(u\right).$$

According to the proof of Theorem 3,

$$\begin{aligned}C\hat{R}_n(u) &= \sqrt{n}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T\frac{1}{n}\begin{pmatrix}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}a_nD(\boldsymbol{Z}_n,\boldsymbol{X}_n)\right)^T\\ \mathbf{0}\end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u)\\
&\quad+\sqrt{n}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T\frac{1}{n}\begin{pmatrix}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\boldsymbol{\varepsilon}_n\right)^T\\ \mathbf{0}\end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u)\\
&\quad+\sqrt{n}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T\left(\boldsymbol{V}(u)+o(1)\right)+\sqrt{n}a_n\boldsymbol{\varsigma}+\frac{1}{\sqrt{n}}\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u)+o_p(1)\\
&= \frac{1}{\sqrt{n}}\boldsymbol{\varepsilon}_n^T\boldsymbol{P}\boldsymbol{I}_n(u)-\left(\boldsymbol{V}(u)+\boldsymbol{\omega}(a,u)\right)^T\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\\
&\quad-\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\frac{1}{n}\begin{pmatrix}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\boldsymbol{\varepsilon}_n\right)^T\\ \mathbf{0}\end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u)+\sqrt{n}a_n\boldsymbol{\varsigma}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{\varepsilon}_n-\left(\boldsymbol{V}(u)+\boldsymbol{\omega}(a,u)\right)^T\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\\
&\quad-\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\frac{1}{n}\begin{pmatrix}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}\boldsymbol{\varepsilon}_n\right)^T\\ \mathbf{0}\end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u)+\sqrt{n}a_n\boldsymbol{\varsigma}+o_p(1),\end{aligned}$$

$\sqrt{n}a_n\boldsymbol{\varsigma}\to\infty$ as $\sqrt{n}a_n\boldsymbol{\varsigma}\to\infty$, so

$$C\hat{R}_n(u)\to\infty,$$

and

$$CT_n \to \infty.$$

Proof of Theorem 5. Because

$$\boldsymbol{Y}_n^* = \hat{\rho}\boldsymbol{W}_n\boldsymbol{Y}_n^* + \boldsymbol{Z}_n\hat{\boldsymbol{\beta}} + \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} + \boldsymbol{\varepsilon}_n^*,$$

Let $\hat{\boldsymbol{\theta}}^* = \left(\hat{\rho}^*, \hat{\boldsymbol{\beta}}^{*T}\right)^T$, $\boldsymbol{Q}_n^* = (\boldsymbol{W}_n\boldsymbol{Y}_n^*, \boldsymbol{Z}_n)$, $\boldsymbol{H}^*$ is an instrumental variable, then the estimate of $\hat{\boldsymbol{\theta}}$ based on the bootstrap sample is

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^* &= \left(\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*\left(\boldsymbol{H}^{*T}\boldsymbol{H}^*\right)^{-1}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{Q}_n^*\right)^{-1}\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*\left(\boldsymbol{H}^{*T}\boldsymbol{H}^*\right)^{-1}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{Y}_n^* \\
&= \left(\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*\left(\boldsymbol{H}^{*T}\boldsymbol{H}^*\right)^{-1}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{Q}_n^*\right)^{-1}\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*\left(\boldsymbol{H}^{*T}\boldsymbol{H}^*\right)^{-1}\boldsymbol{H}^{*T}\boldsymbol{P} \\
&\quad \left(\boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}} + \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} + \boldsymbol{\varepsilon}_n^*\right) \\
&= \hat{\boldsymbol{\theta}} + \left(\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*(\boldsymbol{H}^{*T}\boldsymbol{H}^*)^{-1}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{Q}_n^*\right)^{-1}\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*\left(\boldsymbol{H}^{*T}\boldsymbol{H}^*\right)^{-1}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{\varepsilon}_n^*,
\end{aligned}$$

the estimate of $\hat{\alpha}$ based on the bootstrap sample is

$$\hat{\boldsymbol{\alpha}}^* = \left(\boldsymbol{\Pi}^T\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}^T\left(\boldsymbol{Y}_n^* - \boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}}^*\right).$$

then

$$\begin{aligned}
\sqrt{n}\left(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}\right) &= \left(\frac{1}{n}\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*(\frac{1}{n}\boldsymbol{H}^{*T}\boldsymbol{H}^*)^{-1}\frac{1}{n}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{Q}_n^*\right)^{-1} \\
&\quad \frac{1}{n}\boldsymbol{Q}_n^{*T}\boldsymbol{P}\boldsymbol{H}^*\left(\frac{1}{n}\boldsymbol{H}^{*T}\boldsymbol{H}^*\right)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{\varepsilon}_n^* \\
&= \left(\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\boldsymbol{A}^{*T}\right)^{-1}\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{\varepsilon}_n^* + o_p(1),
\end{aligned}$$

the estimate of $\hat{\varepsilon_n}$ based on the bootstrap sample is

$$\begin{aligned}
\hat{\boldsymbol{\varepsilon}}_n^* &= \boldsymbol{Y}_n^* - \boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}}^* - \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}}^* \\
&= \boldsymbol{Y}_n^* - \boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}}^* - \boldsymbol{\Pi}\left(\boldsymbol{\Pi}^T\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}^T\left(\boldsymbol{Y}_n^* - \boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}}^*\right) \\
&= \left(\boldsymbol{I}_n - \boldsymbol{\Pi}\left(\boldsymbol{\Pi}^T\boldsymbol{\Pi}\right)^{-1}\boldsymbol{\Pi}^T\right)\left(\boldsymbol{Y}_n^* - \boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}}^*\right) \\
&= \boldsymbol{P}\left(\boldsymbol{Y}_n^* - \boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}}^*\right) \\
&= \boldsymbol{P}\left(\boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}} + \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} + \boldsymbol{\varepsilon}_n^* - \boldsymbol{Q}_n^*\hat{\boldsymbol{\theta}}^*\right) \\
&= \boldsymbol{P}\boldsymbol{Q}_n^*(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*) + \boldsymbol{P}\boldsymbol{\varepsilon}_n^*.
\end{aligned}$$

Let $\boldsymbol{S}_n(\hat\rho) = \boldsymbol{I}_n - \hat\rho \boldsymbol{W}_n$, the estimate of $C\hat{R}_n(u)$ based on the bootstrap sample is

$$\begin{aligned}
C\hat{R}_n^*(u) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat\varepsilon_i^* \boldsymbol{I}(<\boldsymbol{U}_{ni}, \boldsymbol{h}> \leqslant u) \\
&= \frac{1}{\sqrt{n}} \hat{\boldsymbol{\varepsilon}}_n^{*T} \boldsymbol{I}_n(u) \\
&= \frac{1}{\sqrt{n}} \left(\boldsymbol{P}\boldsymbol{Q}_n^*(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*) + \boldsymbol{P}\boldsymbol{\varepsilon}_n^* \right)^T \boldsymbol{I}_n(u) \\
&= \frac{1}{\sqrt{n}} \left(\boldsymbol{P}\boldsymbol{Q}_n^*(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*) \right)^T \boldsymbol{I}_n(u) + \frac{1}{\sqrt{n}} \left(\boldsymbol{P}\boldsymbol{\varepsilon}_n^* \right)^T \boldsymbol{I}_n(u) \\
&= VII_1 + VII_2,
\end{aligned}$$

where

$$VII_1 = \frac{1}{\sqrt{n}} \left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^* \right)^T \boldsymbol{Q}_n^{*T} \boldsymbol{P} \boldsymbol{I}_n(u),$$

and

$$VII_2 = \frac{1}{\sqrt{n}} \boldsymbol{\varepsilon}_n^{*T} \boldsymbol{P} \boldsymbol{I}_n(u).$$

For VII_1,

$$\begin{aligned}
VII_1 &= \frac{1}{\sqrt{n}} \left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^* \right)^T \boldsymbol{Q}_n^{*T} \boldsymbol{P} \boldsymbol{I}_n(u) \\
&= \frac{1}{n}\sqrt{n} \left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^* \right)^T \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat\rho) \left(\boldsymbol{Z}_n \hat{\boldsymbol{\beta}} + \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} + \boldsymbol{\varepsilon}_n^* \right) \right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix} \boldsymbol{P} \boldsymbol{I}_n(u) \\
&= \frac{1}{n}\sqrt{n} \left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^* \right)^T \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat\rho) \left(\boldsymbol{Z}_n \hat{\boldsymbol{\beta}} + \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} \right) \right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix} \boldsymbol{P} \boldsymbol{I}_n(u) \\
&\quad + \frac{1}{n}\sqrt{n} \left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^* \right)^T \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat\rho) \boldsymbol{\varepsilon}_n^* \right)^T \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{P} \boldsymbol{I}_n(u) \\
&= VIII_1 + VIII_2,
\end{aligned}$$

where

$$VIII_1 = \frac{1}{n}\sqrt{n} \left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^* \right)^T \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat\rho) \left(\boldsymbol{Z}_n \hat{\boldsymbol{\beta}} + \boldsymbol{\Pi}\hat{\boldsymbol{\alpha}} \right) \right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix} \boldsymbol{P} \boldsymbol{I}_n(u),$$

and

$$VIII_2 = \frac{1}{n}\sqrt{n} \left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^* \right)^T \begin{pmatrix} \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat\rho) \boldsymbol{\varepsilon}_n^* \right)^T \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{P} \boldsymbol{I}_n(u).$$

Let

$$\hat{\boldsymbol{V}}(u) = \lim_{n\to\infty}\frac{1}{n}\begin{pmatrix} \left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\left(\boldsymbol{Z}_n\hat{\boldsymbol{\beta}}+\boldsymbol{\Pi}\hat{\boldsymbol{\alpha}}\right)\right)^T \\ \boldsymbol{Z}_n^T \end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u).$$

Because $\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}=o_p(1)$ and the order of the element in $\boldsymbol{\Pi}\hat{\boldsymbol{\alpha}}-\boldsymbol{\eta}$ is $o_p(1)$, we have

$$\hat{\boldsymbol{V}}(u)\xrightarrow{p}\boldsymbol{V}(u).$$

Due to

$$\hat{\boldsymbol{\varepsilon}}_n=\boldsymbol{Y}_n-\hat{\rho}\boldsymbol{W}_n\boldsymbol{Y}_n-\boldsymbol{Z}_n\hat{\boldsymbol{\beta}}-\boldsymbol{\Pi}\hat{\boldsymbol{\alpha}},$$

and

$$\boldsymbol{\varepsilon}_n=\boldsymbol{Y}_n-\rho\boldsymbol{W}_n\boldsymbol{Y}_n-\boldsymbol{Z}_n\boldsymbol{\beta}-\boldsymbol{\eta},$$

then

$$\hat{\boldsymbol{\varepsilon}}_n-\boldsymbol{\varepsilon}_n=o_p(1),$$

$$\hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi}-\boldsymbol{\varepsilon}_n\boldsymbol{\xi}=o_p(1).$$

Thus $\boldsymbol{\varepsilon}_n^*=\hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi}$ is a n dimensional random vector and the elements are independent of each other, with a mean of 0 and a variance of σ^2. Beside

$$\mathbf{E}\left(\frac{1}{n}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\boldsymbol{\varepsilon}_n^*\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)\right)=\mathbf{E}\left(\frac{1}{n}\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\boldsymbol{\varepsilon}_n^*\right)=0,$$

and

$$\begin{aligned}
&\frac{1}{n^2}E\left(||\boldsymbol{\varepsilon}_n^{*T}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)||^2\right)\\
=\;&\frac{1}{n^2}E\left(trace\left(\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\boldsymbol{\varepsilon}_n^*\boldsymbol{\varepsilon}_n^{*T}\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)\right)\right)\\
\le\;&\sigma^2\frac{1}{n^2}E\left(trace\left(\boldsymbol{I}_n^T(u)\boldsymbol{P}\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\right)^T\boldsymbol{P}\boldsymbol{I}_n(u)\right)\right)\\
=\;&\sigma^2\,\frac{1}{n^2}\,O(n)\\
=\;&O\left(\frac{1}{n}\right).
\end{aligned}$$

So

$$\frac{1}{n}\begin{pmatrix} \left(\boldsymbol{W}_n\boldsymbol{S}_n^{-1}(\hat{\rho})\boldsymbol{\varepsilon}_n^*\right)^T \\ \mathbf{0} \end{pmatrix}\boldsymbol{P}\boldsymbol{I}_n(u)=O_p\left(\frac{1}{\sqrt{n}}\right)=o_p(1).$$

Because

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*) = O_p(1),$$

we have

$$\begin{aligned} VIII_2 &= \frac{1}{n}\sqrt{n}\left(\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^*\right)^T \begin{pmatrix} (\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat{\rho})\boldsymbol{\varepsilon}_n^*)^T \\ \mathbf{0} \end{pmatrix} \boldsymbol{P}\boldsymbol{I}_n(u) \\ &= O_p(1)o_p(1) \\ &= o_p(1). \end{aligned}$$

Then

$$\begin{aligned} VII_1 &= -\hat{\boldsymbol{V}}^T(u)\sqrt{n}(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}) + o_p(1) \\ &= -\boldsymbol{V}^T(u)\sqrt{n}(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}) + o_p(1). \end{aligned}$$

Besides

$$\begin{aligned} \boldsymbol{Q}_n^* &= (\boldsymbol{W}_n \boldsymbol{Y}_n^*, \boldsymbol{Z}_n) \\ &= \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat{\rho})\left(\boldsymbol{Z}_n\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\eta}} + \boldsymbol{\varepsilon}_n^*\right), \boldsymbol{Z}_n\right) \\ &= \left(\boldsymbol{W}_n \boldsymbol{S}_n^{-1}(\hat{\rho})\left(\boldsymbol{Z}_n\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\eta}} + \hat{\varepsilon}_n\boldsymbol{\xi}\right), \boldsymbol{Z}_n\right), \end{aligned}$$

then $\boldsymbol{Q}_n^* \xrightarrow{p} \boldsymbol{Q}_n$. Similarly, $\boldsymbol{H}^* \xrightarrow{p} \boldsymbol{H}$. Let

$$\begin{aligned} J^*\left(\boldsymbol{Y}_n^*, \hat{\rho}, \boldsymbol{Z}_n, \hat{\boldsymbol{\beta}}, \boldsymbol{X}_n, u\right) &= -\hat{\boldsymbol{V}}^T(u)\left(\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\boldsymbol{A}^{*T}\right)^{-1}\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^{*T}\boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi} \\ &\quad + \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_n. \end{aligned}$$

Combining the above results,

$$\begin{aligned} C\hat{R}_n^*(u) &= \frac{1}{\sqrt{n}}\boldsymbol{\varepsilon}_n^{*T}\boldsymbol{P}\boldsymbol{I}_n(u) - \hat{\boldsymbol{V}}^T(u)\sqrt{n}\left(\hat{\boldsymbol{\theta}}^* - \hat{\boldsymbol{\theta}}\right) + o_p(1) \\ &= \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi} - \hat{\boldsymbol{V}}^T(u)\left(\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\boldsymbol{A}^{*T}\right)^{-1}\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^{*T}\boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi} + o_p(1) \\ &= J^*\left(\boldsymbol{Y}_n^*, \hat{\rho}, \boldsymbol{Z}_n, \hat{\boldsymbol{\beta}}, \boldsymbol{X}_n, u\right) + o_p(1). \end{aligned}$$

Because

$$J^*\left(\boldsymbol{Y}_n^*, \hat{\rho}, \boldsymbol{Z}_n, \hat{\boldsymbol{\beta}}, \boldsymbol{X}_n, u\right) - J\left(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u\right)\boldsymbol{\xi}$$

$$
\begin{aligned}
= \quad & \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi} - \hat{\boldsymbol{V}}^T(u)\left(\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\boldsymbol{A}^{*T}\right)^{-1}\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^{*T}\boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_n\boldsymbol{\xi} \\
& -\frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\,\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\xi} + \boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\xi} \\
= \quad & \frac{1}{\sqrt{n}}\boldsymbol{I}_n^T(u)\boldsymbol{P}(\hat{\boldsymbol{\varepsilon}}_n - \boldsymbol{\varepsilon}_n)\boldsymbol{\xi} - \hat{\boldsymbol{V}}^T(u)\left(\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\boldsymbol{A}^{*T}\right)^{-1}\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{H}^{*T}\boldsymbol{P}(\hat{\boldsymbol{\varepsilon}}_n - \boldsymbol{\varepsilon}_n)\boldsymbol{\xi} \\
& +\frac{1}{\sqrt{n}}\boldsymbol{V}^T(u)\left(\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\xi} \\
& -\frac{1}{\sqrt{n}}\hat{\boldsymbol{V}}^T(u)\left(\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\boldsymbol{A}^{*T}\right)^{-1}\boldsymbol{A}^*(\boldsymbol{B}^*)^{-1}\boldsymbol{H}^{*T}\boldsymbol{P}\boldsymbol{\varepsilon}_n\boldsymbol{\xi},
\end{aligned}
$$

and the sequence of random variables $\boldsymbol{\xi}$ which has a mean of 0 and a variance of 1 is independent of $\boldsymbol{Z}_n$ and $\boldsymbol{X}_n$. Analogous to the proof of Theorem 3, in Chen, we have

$$
J^*\left(\boldsymbol{Y}_n^*, \hat{\rho}, \boldsymbol{Z}_n, \hat{\boldsymbol{\beta}}, \boldsymbol{X}_n, u\right) - J\left(\boldsymbol{Y}_n, \rho, \boldsymbol{Z}_n, \boldsymbol{\beta}, \boldsymbol{X}_n, u\right)\boldsymbol{\xi} = o_p(1).
$$

Therefore

$$
C\hat{R}_n^*(u) \xrightarrow{L} CR(u)
$$

.

According to the Continuous Mapping Theorem, it can be seen that CT_n^* and CT_n have the same distribution, so Theorem 5 is proved.

Table 1: Simulation results for DGP1.

ρ	m	R	CvM						KS					
			homoscedasticity			heteroscedasticity			homoscedasticity			heteroscedasticity		
			1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
0.25	5	30	0.015	0.057	0.112	0.015	0.058	0.106	0.014	0.060	0.125	0.017	0.057	0.110
		40	0.011	0.061	0.107	0.024	0.063	0.112	0.021	0.065	0.133	0.020	0.067	0.128
		50	0.009	0.050	0.110	0.014	0.064	0.119						
	8	30	0.014	0.056	0.117	0.012	0.049	0.105						
		40	0.019	0.059	0.107	0.017	0.054	0.103						
		50	0.011	0.049	0.110	0.017	0.068	0.107						
0.5	5	30	0.014	0.056	0.112	0.011	0.053	0.105	0.014	0.060	0.127	0.016	0.056	0.107
		40	0.010	0.061	0.108	0.021	0.061	0.110	0.021	0.064	0.134	0.019	0.065	0.125
		50	0.009	0.050	0.110	0.012	0.062	0.119						
	8	30	0.014	0.056	0.117	0.011	0.048	0.103						
		40	0.018	0.059	0.107	0.017	0.051	0.100						
		50	0.011	0.050	0.110	0.015	0.067	0.106						
0.75	5	30	0.014	0.056	0.109	0.010	0.049	0.096	0.014	0.056	0.122	0.014	0.053	0.107
		40	0.010	0.061	0.107	0.021	0.059	0.110	0.021	0.064	0.131	0.019	0.183	0.122
		50	0.009	0.049	0.107	0.011	0.058	0.117						
	8	30	0.013	0.055	0.117	0.007	0.046	0.101						
		40	0.018	0.059	0.107	0.016	0.049	0.099						
		50	0.010	0.050	0.111	0.014	0.067	0.106						

Table 2: Simulation results for DGP2.

ρ	m	R	homoscedasticity			heteroscedasticity		
			$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0.25	5	30	0.021	0.058	0.123	0.020	0.061	0.107
		40	0.013	0.053	0.103	0.011	0.056	0.111
		50	0.017	0.051	0.106	0.106	0.051	0.101
	8	30	0.018	0.072	0.127	0.023	0.068	0.135
		40	0.023	0.072	0.125	0.031	0.075	0.130
		50	0.018	0.062	0.116	0.020	0.066	0.117
0.5	5	30	0.019	0.058	0.117	0.017	0.056	0.107
		40	0.012	0.049	0.100	0.011	0.052	0.110
		50	0.018	0.050	0.112	0.016	0.052	0.102
	8	30	0.020	0.071	0.130	0.022	0.064	0.135
		40	0.025	0.077	0.123	0.031	0.076	0.129
		50	0.017	0.064	0.117	0.019	0.064	0.115
0.75	5	30	0.017	0.056	0.121	0.010	0.049	0.100
		40	0.011	0.048	0.102	0.0 6	0.045	0.107
		50	0.015	0.048	0.107	0.014	0.046	0.099
	8	30	0.017	0.069	0.127	0.019	0.067	0.129
		40	0.026	0.071	0.124	0.028	0.076	0.127
		50	0.015	0.062	0.118	0.018	0.064	0.113

Table 3: Simulation results for DGP3.

ρ	m	R	homoscedasticity			heteroscedasticity		
			$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0.25	5	30	0.981	0.994	0.996	0.697	0.883	0.937
		40	0.999	1.000	1.000	0.810	0.912	0.948
		50	1.000	1.000	1.000	0.896	0.968	0.984
	8	30	0.998	1.000	1.000	0.870	0.957	0.977
		40	1.000	1.000	1.000	0.960	0.990	0.990
		50	1.000	1.000	1.000	0.973	0.992	0.996
0.5	5	30	0.981	0.994	0.996	0.695	0.883	0.938
		40	0.999	1.000	1.000	0.808	0.913	0.949
		50	1.000	1.000	1.000	0.893	0.968	0.984
	8	30	0.998	1.000	1.000	0.870	0.957	0.976
		40	1.000	1.000	1.000	0.960	0.990	0.990
		50	1.000	1.000	1.000	0.973	0.992	0.996
0.75	5	30	0.979	0.994	0.996	0.694	0.883	0.939
		40	0.999	1.000	1.000	0.807	0.913	0.948
		50	1.000	1.000	1.000	0.892	0.967	0.984
	8	30	0.998	1.000	1.000	0.870	0.955	0.976
		40	1.000	1.000	1.000	0.960	0.985	0.990
		50	1.000	1.000	1.000	0.973	0.992	0.996

Table 4: Simulation results for DGP4.

ρ	m	R	homoscedasticity			heteroscedasticity		
			$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0.25	5	30	0.997	1.000	1.000	0.629	0.843	0.913
		40	1.000	1.000	1.000	0.812	0.931	0.975
		50	1.000	1.000	1.000	0.915	0.965	0.985
	8	30	1.000	1.000	1.000	0.852	0.948	0.982
		40	1.000	1.000	1.000	0.945	0.982	0.990
		50	1.000	1.000	1.000	0.970	0.990	1.000
0.5	5	30	0.997	1.000	1.000	0.616	0.843	0.913
		40	1.000	1.000	1.000	0.811	0.931	0.974
		50	1.000	1.000	1.000	0.915	0.965	0.985
	8	30	1.000	1.000	1.000	0.850	0.950	0.982
		40	1.000	1.000	1.000	0.946	0.982	0.990
		50	1.000	1.000	1.000	0.970	0.990	1.000
0.75	5	30	0.997	1.000	1.000	0.616	0.840	0.913
		40	1.000	1.000	1.000	0.809	0.931	0.974
		50	1.000	1.000	1.000	0.910	0.965	0.985
	8	30	1.000	1.000	1.000	0.848	0.948	0.982
		40	1.000	1.000	1.000	0.944	0.981	0.990
		50	1.000	1.000	1.000	0.970	0.990	1.000

Table 5: Simulation results for DGP5.

d	m	R	homoscedasticity			heteroscedasticity		
			$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0	5	30	0.015	0.057	0.120	0.015	0.058	0.106
		40	0.011	0.061	0.107	0.024	0.063	0.112
		50	0.009	0.050	0.110	0.014	0.064	0.119
	8	30	0.014	0.056	0.117	0.012	0.059	0.109
		40	0.009	0.050	0.100	0.011	0.050	0.110
		50						
0.25	5	30	0.828	0.924	0.958	0.684	0.887	0.943
		40	0.869	0.945	0.970	0.788	0.935	0.971
		50	0.890	0.960	0.981	0.861	0.958	0.979
	8	30	0.866	0.937	0.971	0.831	0.932	0.962
		40	0.863	0.938	0.965	0.825	0.931	0.964
		50	0.887	0.946	0.976	0.866	0.947	0.974
0.5	5	30	0.762	0.921	0.955	0.690	0.894	0.941
		40	0.806	0.929	0.968	0.760	0.915	0.960
		50	0.832	0.947	0.975	0.814	0.939	0.975
	8	30	0.775	0.922	0.962	0.741	0.912	0.953
		40	0.785	0.923	0.961	0.762	0.916	0.952
		50	0.824	0.925	0.967	0.802	0.934	0.961
0.75	5	30	0.569	0.867	0.939	0.482	0.825	0.924
		40	0.660	0.891	0.955	0.565	0.859	0.945
		50	0.716	0.922	0.972	0.649	0.895	0.967
	8	30	0.592	0.875	0.951	0.531	0.842	0.936
		40	0.621	0.878	0.952	0.551	0.863	0.940
		50	0.664	0.901	0.947	0.618	0.888	0.947

Table 6: Simulation results for DGP6.

d	m	R	homoscedasticity			heteroscedasticity		
			$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0	5	30	0.010	0.050	0.106	0.017	0.063	0.121
		40	0.012	0.058	0.096	0.008	0.054	0.116
		50	0.008	0.044	0.084	0.012	0.040	0.088
	8	30	0.014	0.050	0.099	0.010	0.050	0.100
		40						
		50						
0.25	5	30	0.803	0.927	0.965	0.753	0.903	0.951
		40	0.875	0.959	0.979	0.848	0.952	0.978
		50	0.906	0.974	0.988	0.898	0.964	0.989
	8	30	0.843	0.952	0.980	0.805	0.943	0.979
		40	0.899	0.968	0.987	0.877	0.956	0.983
		50						
0.5	5	30	0.761	0.910	0.955	0.485	0.790	0.896
		40	0.840	0.953	0.976	0.648	0.880	0.951
		50	0.879	0.971	0.989	0.747	0.924	0.969
	8	30	0.805	0.936	0.967	0.640	0.879	0.958
		40	0.862	0.952	0.978	0.772	0.934	0.975
		50						
0.75	5	30	0.728	0.903	0.947	0.215	0.603	0.765
		40	0.823	0.946	0.971	0.364	0.711	0.867
		50	0.874	0.962	0.985	0.444	0.799	0.918
	8	30	0.775	0.918	0.967	0.327	0.699	0.865
		40	0.850	0.945	0.970	0.467	0.828	0.928
		50						