On the Bootstrap for Spatial Econometric Models

Fei Jin[a], Lung-fei Lee[a]

[a] *Department of Economics, Ohio State University, Columbus, OH 43210 USA*

Abstract

This paper is concerned about the use of the bootstrap for spatial econometric models. We show that the bootstrap for spatial econometric models can be studied based on linear-quadratic (LQ) forms of disturbances. By proving the uniform convergence of the cumulative distribution function for LQ forms to that of normal distributions, we show that the bootstrap is generally consistent for test statistics that can be approximated by LQ forms, which include Moran's I, Cox-type and spatial J-type test statistics. Possible asymptotic refinements of the bootstrap for spatial econometric models require the existence of some asymptotic expansions for LQ forms. We discuss two cases: when the disturbances are normal, we directly show the existence of Edgeworth expansions and apply the result to Moran's I to show the second-order correctness of the bootstrap; when the disturbances are not normal, we show the existence of a one-term asymptotic expansion based on martingales.

Keywords: Bootstrap, spatial, consistency, asymptotic refinement, linear-quadratic form

JEL classification: C21, C83, R15

1. Introduction

The bootstrap is a statistical procedure that estimates the distributions of estimators or test statistics by resampling the data. Its approximations can be at least as good as those from the first-order asymptotic theory under mild conditions. Thus it can be used as an alternative when evaluating the asymptotic distributions is difficult. A more appealing feature of the bootstrap is that it is often more accurate in finite samples than the asymptotic theory, i.e., it can provide asymptotic refinements. The bootstrap is frequently used to correct the bias of estimators, estimate the critical values for hypothesis tests and construct confidence intervals. Useful survey papers on the bootstrap include, among others, DiCiccio and Efron (1996), MacKinnon (2002), Davison et al. (2003), and Horowitz (2001, 2003).

The bootstrap has been discussed and implemented by many researchers for models in spatial econometrics. Anselin (1988, 1990) discusses the bootstrap estimation in spatial autoregressive (SAR) models, which is implemented by Can (1992). Fingleton (2008) and Fingleton and Le Gallo (2008) use the bootstrap

Email addresses: `jin.120@osu.edu` (Fei Jin), `lee.1777@osu.edu` (Lung-fei Lee)

to test the significance of the moving average parameter in models with spatial moving average disturbances. Lin et al. (2011) investigate the properties of bootstrapped Moran's I under heterogeneous and non-normal disturbances by Monte Carlo study. Fingleton and Burridge (2010) and Burridge (2012) find that the bootstrap can essentially remove the size distortion of the spatial J test in Kelejian (2008) in Monte Carlo studies. Yang (2011) proposes the residual bootstrap for LM tests of spatial dependence. Jin and Lee (2012) employ the bootstrap to remove the size distortion of Cox-type tests for SAR models with SAR disturbances (SARAR models for short). Su and Yang (2008) suggest a bootstrap procedure that leads to a robust estimate of a variance-covariance matrix. Yang (2012) proposes a bootstrap procedure to correct the bias and variance of quasi-maximum likelihood estimators (QMLE) for SAR models. Monchuk et al. (2011) compare several bootstrap methods in Monte Carlo studies for constructing confidence intervals in a spatial error model.

Although there have been many applications of the bootstrap in spatial econometric models including Monte Carlo studies in the preceding papers, its validity for these models has not been formally justified. The objective of this paper is to establish the consistency of the bootstrap for several test statistics in spatial econometric models and provide a preliminary discussion of possible asymptotic refinements. We shall show that many estimators in spatial econometric models can be approximated by linear-quadratic (LQ) forms of the disturbances, and test statistics are either approximated by or closely related to LQ forms, because of the presence of spatial dependence. The bootstrap in spatial econometric models thus can be studied based on LQ forms in general. Kelejian and Prucha (2001) prove a central limit theorem for LQ forms using a central limit theorem for martingale difference arrays. We shall show that the convergence of the cumulative distribution function (CDF) for a LQ form is uniform under the same conditions. Using this uniform convergence, the bootstrap can generally be shown to be consistent for statistics that can be approximated by a LQ form. We apply the result to show the consistency of the bootstrap for Moran's I and the spatial J-type tests (Kelejian and Piras, 2011).

We shall also discuss the possible asymptotic refinements of the bootstrap for spatial econometric models based on the LQ forms. For non-spatial econometric models, the bootstrap is often considered for the statistics that are smooth functions of sample averages of independent random vectors, see, e.g., Hall (1997), or stationary dependent random vectors, see, e.g., Götze and Hipp (1983, 1994). The existence of an Edgeworth expansion for the random vectors can be used to prove the consistency and asymptotic refinements of the bootstrap. The framework does not apply to LQ forms, which cannot be written as simple averages of disturbances or their cross-products. We may investigate whether the bootstrap can provide asymptotic refinements by considering Edgeworth-type expansions of LQ forms. Such expansions, however, have not be proved to exist in the literature. For normal disturbances, we shall show the existence of Edgeworth expansions and apply the result to show that the bootstrap can provide asymptotic refinements for Moran's I; for non-normal disturbances, we verify an asymptotic expansion of LQ forms based on martingales

(Mykland, 1993). The asymptotic expansion based on martingales is not in a pointwise topology but sheds light on the bootstrap. It implies the second-order correctness of the bootstrap for LQ forms in the sense of the convergence in Mykland (1993).

The rest of the paper is organized as follows. Section 2 demonstrates a close relationship between LQ forms and estimators and test statistics in spatial econometric models; Section 3 first shows the uniform convergence of the CDF for LQ forms and then applies the result to show the bootstrap is consistent for Moran's I and spatial J-type tests; Section 4 establishes the Edgeworth expansion of LQ forms with normal disturbances, which is applied to show the second order correctness of the bootstrap for Moran's I, and establishes the asymptotic expansion in Mykland (1993) for LQ forms with non-normal disturbances; Section 5 concludes. Lemmas and proofs are collected in the appendices.

2. Statistics in Spatial Econometrics and LQ Forms

In this section, we show that several estimators for spatial econometric models can be approximated by LQ forms of disturbances, and many test statistics can be approximated by or relate closely to LQ forms. As a result, we may study the bootstrap in spatial econometric models based on LQ forms. As the SARAR model is a popular and general spatial model, which contains both the spatial lag (SAR) model and spatial error (SE) model as special cases, our discussion will mainly focus on this model. A SARAR model is specified as

$$y_n = \lambda W_n y_n + X_n \beta + u_n, \quad u_n = \rho M_n u_n + \epsilon_n, \quad \epsilon_n = (\epsilon_{n1}, \dots, \epsilon_{nn})', \tag{1}$$

where n is the sample size, y_n is an n-dimensional vector of observations on the dependent variable, X_n is an $n \times k_x$ matrix of exogenous variables, W_n and M_n are $n \times n$ spatial weights matrices with zero diagonals, ϵ_{ni}'s are i.i.d. with mean zero and variance σ^2, and $\theta = (\lambda, \rho, \beta', \sigma^2)' = (\gamma', \sigma^2)'$ is a vector of parameters. Let θ_0 be the true parameter vector, $S_n(\lambda) = I_n - \lambda W_n$ and $R_n(\rho) = I_n - \rho M_n$ with I_n being the n-dimensional identity matrix. Denote $S_n = S_n(\lambda_0)$ and $R_n = R_n(\rho_0)$ for short. The SARAR model nests the SAR and SE models. The SAR model is (1) with i.i.d. disturbances, i.e., $\rho = 0$, and the SE model is (1) without the spatially lagged term of the dependent variable, i.e., $\lambda = 0$. The spatial Durbin model adds an additional term $(+W_n X_n \zeta)$ to the r.h.s. of the equation for y_n of the SAR model. As $W_n X_n$ can be taken as an exogenous variable matrix, with some additional identification consideration in some cases, the analysis for a spatial Durbin model is similar to that for a SAR model.

For estimators of the SARAR model, the derivative of the corresponding criterion function evaluated at the true parameter vector is often a LQ form of the disturbances, rather than just a linear form, because of spatial dependence. As a result, these estimators can be approximated by a LQ form. Lee (2004) has proved the consistency and asymptotic normality of the QMLE for a SAR model without SAR disturbances. The analysis can be extended to the SARAR model (1) as in Jin and Lee (2012), from which we have

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -\Big(\frac{1}{n}\,\mathrm{E}\,\frac{\partial^2 L_n(\theta_0)}{\partial\theta\partial\theta'}\Big)^{-1}\frac{1}{\sqrt{n}}\frac{\partial L_n(\theta_0)}{\partial\theta} + o_P(1), \tag{2}$$

where $\hat{\theta}_n$ is the QMLE and $L_n(\theta)$ denotes the log likelihood function of the model. Every element of the vector $\frac{1}{\sqrt{n}}\frac{\partial L_n(\theta_0)}{\partial\theta}$ is linear in the disturbances or of the LQ form

$$(\epsilon_n' A_n \epsilon_n - \sigma_0^2\,\mathrm{tr}(A_n) + b_n'\epsilon_n)/\sqrt{n}, \tag{3}$$

where A_n is an n-dimensional square matrix and b_n is an n-dimensional vector. Thus every element of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ can be approximated asymptotically by a linear combination of LQ forms, which is still a LQ form with the same ϵ_n. For the generalized method of moments (GMM) estimator, from Lee (2001, 2007),

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) = -\Big(\big(\mathrm{E}\,\frac{\partial g_n'(\gamma_0)}{\partial\gamma}\big)a_n a_n'\big(\mathrm{E}\,\frac{\partial g_n(\gamma_0)}{\partial\gamma'}\big)\Big)^{-1}\big(\mathrm{E}\,\frac{\partial g_n'(\gamma_0)}{\partial\gamma}\big)a_n a_n'\sqrt{n}g_n(\gamma_0) + o_P(1),$$

where $\hat{\gamma}_n$ is the GMM estimator of γ, a_n is a matrix with full column rank greater than or equal to (k_x+2), and $g_n(\gamma) = \big(\epsilon_n'(\gamma)D_{1n}\epsilon_n(\gamma), \ldots, \epsilon_n'(\gamma)D_{mn}\epsilon_n(\gamma), \epsilon_n'(\gamma)Q_n\big)/n$ with $\epsilon_n(\gamma) = R_n(\rho)[S_n(\lambda)y_n - X_n\beta]$, D_{in}'s being n-dimensional square matrices with zero traces and Q_n being a matrix of instrumental variables constructed as functions of X_n, W_n and M_n in a two-stage least squares (2SLS) approach. Every element of $\sqrt{n}g_n(\gamma_0)$ is a quadratic or linear form of the disturbances, then every element of $\sqrt{n}(\hat{\gamma}_n - \gamma_0)$ can be approximated by a LQ form of the disturbances. The spatial generalized 2SLS (G2SLS) approach in Kelejian and Prucha (1998) first estimates $(\lambda, \beta')'$ using only linear moments, then derives estimates of ρ and σ^2 based on quadratic moments using the residuals from the first step, and finally updates the estimate of $(\lambda, \beta')'$ by a G2SLS taking into account the covariance structure. As $(\lambda, \beta')'$ is estimated using only linear moments, its estimator can be approximated by a linear form of the disturbances, but the estimator of ρ is approximated by a LQ form because quadratic moments are used.

The estimators discussed above can be used to implement hypothesis tests such as the classical Wald, likelihood ratio and Lagrangian multiplier (LM) tests in the likelihood framework, or by the Wald test, the distance test, and the gradient test in the GMM framework. These asymptotically equivalent tests are based on the asymptotic normality of the estimators. As a result, these classical test statistics can be studied based on LQ forms. In addition to the classical hypothesis tests, Moran's I (Moran, 1950; Cliff and Ord, 1973, 1981) is a popular test for spatial dependence, and tests for non-nested hypotheses, such as the spatial J-type tests (Kelejian, 2008; Kelejian and Piras, 2011) and Cox-type tests (Jin and Lee, 2012), have been proposed for testing the selection of various spatial weights matrices in spatial models.

The Moran I statistic is

$$\frac{n}{l_n' M_n l_n}\cdot\frac{\hat{\epsilon}_n' M_n \hat{\epsilon}_n}{\hat{\epsilon}_n'\hat{\epsilon}_n},$$

where l_n is an n-dimensional vector of ones and $\hat{\epsilon}_n$ is the residual vector from the least squares estimation. The test is based on the asymptotic normality of the standardized test statistic by deducting the mean and dividing by the standard deviation. Burridge (1980) shows that for the SE model with normal disturbances or the spatial moving average model

$$y_n = X_n\beta + u_n, \quad u_n = \rho M_n\epsilon_n + \epsilon_n, \quad \epsilon_n \sim N(0, \sigma^2 I_n),$$

the LM test statistic is proportional to the Moran I statistic, which is

$$\mathbb{I}_n = \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \cdot \frac{\hat{\epsilon}_n' M_n \hat{\epsilon}_n}{\hat{\epsilon}_n' \hat{\epsilon}_n}. \tag{4}$$

Let $H_n = I_n - X_n(X_n'X_n)^{-1}X_n'$. Under the null hypothesis of no spatial dependence, Eq. (4) becomes

$$\begin{aligned}
\mathbb{I}_n &= \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\epsilon_n' H_n M_n H_n \epsilon_n}{\epsilon_n' H_n \epsilon_n} \\
&= \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)}{(n - k_x)\sigma_0^2} + \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\operatorname{tr}(M_n H_n)}{n - k_x} \\
&\quad - \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\epsilon_n' H_n M_n H_n \epsilon_n \big(\epsilon_n' H_n \epsilon_n - (n - k_x)\sigma_0^2\big)}{(n - k_x)\sigma_0^2 \epsilon_n' H_n \epsilon_n}.
\end{aligned} \tag{5}$$

Under some regularity assumptions, the last two terms on the r.h.s. of Eq. (5) have the order $O(n^{-1/2})$, thus the LM or Moran I statistic can be approximated by a quadratic form of the disturbances. Kelejian and Prucha (2001) propose a generalized Moran's I test for which the test statistic equals a quadratic form of some regression residuals divided by a normalizing factor. Their regularity conditions guarantee that the test statistic can be approximated by a LQ form.

The spatial J-type tests for testing one spatial econometric model against another one are based on augmenting the null model by using a predictor from the alternative model. The augmented model is estimated by the 2SLS in Kelejian (2008) or Kelejian and Piras (2011) and then they test whether the coefficient of the predictor is statistically different from zero or not. Hence, the test statistic is only a linear form of the disturbances plus a term that converges to zero in probability. But a linear form is just a special case of the more general LQ form, so the test statistic can also be studied using LQ forms. The more efficient GMM estimation with both linear and quadratic moments for the augmented model may significantly improve the power of the spatial J-type tests (Jin and Lee, 2012). The test statistics with the GMM estimation are approximated by LQ forms.

Jin and Lee (2012) derive the Cox-type specification tests for SARAR models. The Cox-type tests are based on the log likelihood ratios for the null and alternative models with proper adjustment for the nonzero mean. While the first order asymptotic expansion of estimators can be approximated by LQ forms, the adjusted log likelihood ratio itself is a LQ form at given parameters. As a result, the Cox test statistic, equal to the adjusted log likelihood ratio divided by its standard error, is the sum of a LQ form and a remainder term where the remainder converges to zero in probability.

Our study focuses on the bootstrap for test statistics which can be approximated by LQ forms, including Moran's I and spatial J-type test statistics.

3. Consistency of the Bootstrap

In this section, we first present a general result on the consistency of the bootstrap for statistics that may be approximated by LQ forms. Then we apply the result to the Moran I statistic in Eq. (5) and J-type test statistics for SARAR models.

Consider a statistic t_n for a spatial econometric model which is asymptotically normal with mean zero. The t_n would involve spatial weights matrices, exogenous variables and dependent variables. The dependent variables in t_n can be replaced by their reduced forms as functions of disturbances $\epsilon_n = (\epsilon_{n1}, \dots, \epsilon_{nn})'$, exogenous variables and the true parameter vector θ_0. The t_n may also involve the estimator $\hat{\theta}_n$ of θ_0 and the estimator $\hat{\varsigma}_n$ of other moment parameter vector ς_0 for ϵ_{ni}. To compute a bootstrapped version of t_n, a proper bootstrap procedure needs to be considered. The spatially dependent variable usually cannot be resampled directly, because doing so would destroy the inherent dependence structure. Instead, the residual bootstrap can be used as we usually assume that the disturbances ϵ_{ni}'s are i.i.d.. We may first derive a consistent estimator of parameters in a spatial econometric model and compute the residual vector $\hat{\epsilon}_n$. The $\hat{\epsilon}_n$ may not have a zero mean, so we deduct its empirical mean from the vector to obtain $\tilde{\epsilon}_n = (I_n - \frac{1}{n} l_n l_n')\hat{\epsilon}_n$.[1] Next, sample with replacement n times from the elements of $\tilde{\epsilon}_n$ to obtain a vector ϵ_n^*.[2] Then a pseudo data vector y_n^* on the dependent variable can be computed by using the reduced form with the parameter $\hat{\theta}_n$ and disturbances ϵ_n^*. For example, for the SARAR model (1), we have $y_n^* = S_n^{-1}(\hat{\lambda}_n)\big(X_n\hat{\beta}_n + R_n^{-1}(\hat{\rho}_n)\epsilon_n^*\big)$. Estimating θ using y_n^* yields $\hat{\theta}_n^*$ and a residual vector $\hat{\epsilon}_n^*$. The bootstrapped version of t_n, t_n^*, is the statistic obtained from replacing ϵ_n, θ_0, $\hat{\theta}_n$ and $\hat{\varsigma}_n$ in t_n by, respectively, ϵ_n^*, $\hat{\theta}_n$, $\hat{\theta}_n^*$ and $\hat{\varsigma}_n^*$, where $\hat{\varsigma}_n^*$ is a vector of sample moments of $\hat{\epsilon}_n^*$ that correspond to the moment parameters in ς_0.

Let the second, third and four moments of the zero-mean i.i.d. disturbances ϵ_{ni}'s be σ_0^2, μ_3 and μ_4 respectively, $A_n = [a_{n,ij}]$ be an n-dimensional nonstochastic symmetric matrix, $b_n = (b_{n1}, \dots, b_{nn})'$ be an n-dimensional nonstochastic vector, and $c_n = n^{-1/2}\big(\epsilon_n' A_n \epsilon_n - \sigma_0^2 \operatorname{tr}(A_n) + b_n'\epsilon_n\big)$ be a LQ form with mean zero and variance $\sigma_{c_n}^2 = n^{-1}\big[2\sigma_0^4 \operatorname{tr}(A_n^2) + \sigma_0^2 b_n' b_n + \sum_{i=1}^n \big((\mu_4 - 3\sigma_0^4)a_{n,ii}^2 + 2\mu_3 a_{n,ii} b_{ni}\big)\big]$. We assume that t_n can be approximated by c_n/σ_{c_n} such that $d_n = t_n - c_n/\sigma_{c_n}$ converges to zero in probability. Let $c_n^* = n^{-1/2}\big(\epsilon_n^{*\prime} A_n \epsilon_n^* - \sigma_n^{*2} \operatorname{tr}(A_n) + b_n'\epsilon_n^*\big)$ with variance $\sigma_{c_n}^{*2} = n^{-1}\big[2\sigma_n^{*4} \operatorname{tr}(A_n^2) + \sigma_n^{*2} b_n' b_n + \sum_{i=1}^n \big((\mu_{4n}^* - 3\sigma_n^{*4})a_{n,ii}^2 + 2\mu_{3n}^* a_{n,ii} b_{ni}\big)\big]$ conditional on the bootstrap sampling process, where $\sigma_n^{*2} = n^{-1}\tilde{\epsilon}_n'\tilde{\epsilon}_n$, $\mu_{3n}^* = n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^3$ and $\mu_{4n}^* = n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^4$. Define $d_n^* = t_n^* - c_n^*/\sigma_{c_n}^*$. We assume the following conditions about c_n.

Assumption 1. *The ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \dots, \epsilon_{nn})'$ are i.i.d. $(0, \sigma_0^2)$ and $\mathrm{E}\,|\epsilon_{ni}|^{4(1+\delta)} < \infty$ for some $\delta > 0$.*

---

[1] Freedman (1981) shows the necessity of recentering for regression models. For the SARAR model (1), if X_n contains l_n corresponding to an intercept term in the model, then the residuals from the quasi-maximum likelihood estimation have mean zero and there is no need to recenter.

[2] That is, generate the bootstrap error terms from the empirical distribution function of the recentered residuals.

Assumption 2. *The sequence of symmetric matrices* $\{A_n\}$ *are bounded in both row and column sum norms,*[3] *and elements of the vectors* $\{b_n\}$ *satisfy* $\sup_n n^{-1}\sum_{i=1}^n |b_{ni}|^{2(1+\delta)} < \infty$.

Assumption 3. *The* $\sigma_{c_n}^2$ *is bounded away from zero.*

The A_n and b_n are functions of spatial weights matrices and exogenous variables. As spatial weights matrices are often assumed to be bounded in both row and column sum norms and the elements of exogenous variables are assumed to be bounded constants (Kelejian and Prucha, 1998; Lee, 2004), it is reasonable to impose Assumption 2. Kelejian and Prucha (2001) have proved the asymptotic normality of c_n/σ_{c_n} under Assumptions 1–3. Under the same conditions, we can have the uniform convergence of the CDF for c_n/σ_{c_n} to that for a standard normal variable as subsequently shown. As in Kelejian and Prucha (2001), we write c_n a sum of martingale differences, then theorems in Heyde and Brown (1970) and Haeusler (1988) on the departure of c_n/σ_{c_n} from the standard normal distribution are applicable. Let $\Phi(x)$ be the CDF for a standard normal random variable, P^* and E^* be, respectively, the probability distribution and expectation induced by the bootstrap sampling process, and let K_a and K_b be constants such that for any n,

$$\sup_{1\le j\le n}\sum_{i=1}^n |a_{n,ij}| \le K_a,\ \sup_{1\le i\le n}\sum_{j=1}^n |a_{n,ij}| \le K_a,\ \text{and}\ \frac{1}{n}\sum_{i=1}^n |b_{ni}|^{2(1+\eta)} \le K_b \text{ for } -1<\eta\le\delta.$$

Theorem 1. *Under Assumptions 1–3,*

$$\sup_{x\in\mathbb{R}} |\,\mathrm{P}(c_n/\sigma_{c_n} \le x) - \Phi(x)| \le r_n, \tag{6}$$

$$\sup_{x\in\mathbb{R}} |\,\mathrm{P}^*(c_n^*/\sigma_{c_n}^* \le x) - \Phi(x)| \le r_n^*, \tag{7}$$

$$\sup_{x\in\mathbb{R}} |\,\mathrm{P}^*\big(c_n^*/\sigma_{c_n}^* + d_n^* \le x\big) - \mathrm{P}\big(c_n/\sigma_{c_n} + d_n \le x\big)| \le r_n + \mathrm{P}(|d_n|>\tau_n) + r_n^* + \mathrm{P}^*(|d_n^*|>\tau_n) + 2^{1/2}\pi^{-1/2}\tau_n, \tag{8}$$

$$\begin{aligned}\sup_{x\in\mathbb{R}}\big|\mathrm{P}^*\big((c_n^*/\sigma_{c_n}^* + d_n^*)e_n^* \le x\big) - \mathrm{P}\big((c_n/\sigma_{c_n} + d_n)e_n \le x\big)\big| \le r_n + \mathrm{P}(|d_n|>\tau_n) + r_n^* + \mathrm{P}^*(|d_n^*|>\tau_n)\\ + 2^{1/2}\pi^{-1/2}\tau_n + \sup_{x\in\mathbb{R}}\big|\Phi(x/e_n) - \Phi(x/e_n^*)\big|,\end{aligned} \tag{9}$$

where τ_n *is any positive term depending only on* n*,* e_n *is a nonstochastic term depending on* n*,* θ_0 *and moment parameters of* ϵ_{ni}*,* $r_n = K\sigma_{c_n}^{-2(1+\delta)/(3+2\delta)} n^{-\delta/(3+2\delta)}\Big((K_a+1)^{1+2\delta}\big(K_a\,\mathrm{E}\,|\epsilon_{ni}^2-\sigma_0^2|^{2+2\delta} + 2^{2+2\delta}K_a(\mathrm{E}\,|\epsilon_{ni}|^{2+2\delta})^2 + K_b\,\mathrm{E}\,|\epsilon_{ni}|^{2+2\delta}\big) + 4^{1+\delta}\big(\sigma_0^4K_a^4(\mu_4-\sigma_0^4) + 4\sigma_0^8K_a^4 + \sigma_0^2K_a^2(\mu_3^2K_a + \sigma_0^4K_b)(K_a+1) + 2|\mu_3|\sigma_0^2K_a^3(|\mu_3|K_a + \sigma_0^2K_b)\big)^{(1+\delta)/2}\Big)^{1/(3+2\delta)}$ *with* K *being a constant depending only on* δ*,* r_n^* *is a term obtained from replacing the population moment parameters of* ϵ_{ni} *in* r_n *with the corresponding sample moments of* ϵ_{ni}^**, and* e_n^* *is a term obtained from replacing* θ_0 *and population moment parameters of* ϵ_{ni} *in* e_n *by, respectively,* $\hat{\theta}_n$ *and corresponding sample moments of* ϵ_{ni}^*.

[3] As $\epsilon_n' A_n \epsilon_n = \epsilon_n'(A_n + A_n')\epsilon_n/2$, it is w.l.o.g. to assume the symmetry of A_n.

The l.h.s. of (6) is the Kolmogorov-Smirnov distance between the CDFs of two random variables. The inequality gives a rate of convergence, $O(n^{-\delta/(3+2\delta)})$, of the CDF of c_n/σ_{c_n} to that of a standard normal random variable. The larger is δ, i.e., the higher moments of ϵ_{ni} assumed to exist, the faster is the convergence. The convergence rate approaches $O(n^{-1/2})$, the rate for a sample average of i.i.d. random variables, as δ becomes larger. A similar result for the bootstrapped version of c_n/σ_{c_n} is given in (7). The result in (8) is shown by using (6) and (7). To prove the consistency of the bootstrapped t_n, we may show that the r.h.s. of (8) converges to zero in probability. This type of convergence with respect to the Kolmogorov-Smirnov distance implies the asymptotic consistency of confidence intervals. If we can show that the sample moments of ϵ^*_{ni} converge in probability to the relevant population moments of ϵ_{ni}, then the continuous mapping theorem implies that r^*_n is of order $O_P(n^{-\delta/(3+2\delta)})$. The remainder term d_n is often of order $O_P(n^{-1/2})$, thus we may let $\tau_n = o(n^{-1/2})$. It only remains to show that $\mathrm{P}^*(|d^*_n| > \tau_n)$ converges to zero in probability. For asymptotically normal statistics with non-unit variances, e.g., various estimators, we may rescale terms in (8) to obtain (9), which can be more convenient for the proof of consistency. Now we apply the results in Theorem 1 to show the consistency of bootstrapped Moran's I and spatial J-type test statistics for SARAR models.

3.1. Moran's I

To show the consistency of the bootstrap for Moran's I in Eq. (5), we write $\mathbb{I}_n$ in the form on the l.h.s. of (9). Note that the variance of $\epsilon'_n H_n M_n H_n \epsilon_n$ is $\sigma_0^4 \operatorname{tr}\big(H_n M_n H_n (M_n + M'_n)\big)$ when $\epsilon_n \sim N(0, \sigma_0^2 I_n)$, we may let

$$c_n = n^{-1/2}\big(\epsilon'_n H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)\big), \tag{10}$$

$$\sigma^2_{c_n} = n^{-1}\sigma_0^4 \operatorname{tr}[H_n M_n H_n (M_n + M'_n)], \tag{11}$$

$$e_n = \frac{n}{n-k_x} \frac{\sqrt{\operatorname{tr}[H_n M_n H_n (M_n + M'_n)]}}{\sqrt{\operatorname{tr}(M_n^2 + M'_n M_n)}}, \tag{12}$$

$$d_n = \mathbb{I}_n / e_n - c_n/\sigma_{c_n}. \tag{13}$$

Let $\mathbb{I}^*_n$ be the bootstrapped $\mathbb{I}_n$. The $\mathbb{I}^*_n$, and the corresponding c^*_n, $\sigma^*_{c_n}$, e^*_n and d^*_n are derived as described earlier.

Proposition 1. *Under H_0 and Assumptions I1–I4 in Appendix A.1, the Moran I statistic in Eq. (5) satisfies* $\sup_{x\in\mathbb{R}} |\mathrm{P}^*(\mathbb{I}^*_n \le x) - \mathrm{P}(\mathbb{I}_n \le x)| = o_P(1)$.

The above proposition is the case where $\epsilon_n \sim N(0, \sigma_0^2 I_n)$, which guarantees that $\mathbb{I}_n$ in (5) is asymptotically standard normal. When the i.i.d. disturbances are not normal, $\mathbb{I}_n$ is still asymptotically normal but with a non-unit variance in general, since the variance of $\epsilon'_n H_n M_n H_n \epsilon_n$ is $(\mu_4 - 3\sigma_0^4)\sum_{i=1}^n (H_n M_n H_n)^2_{ii} +$

$\sigma_0^4 \operatorname{tr}[H_n M_n H_n (M_n + M_n')]$. To make the test statistic robust to the distribution of the disturbances, we consider the following statistic

$$\mathbb{I}_n' = \frac{\epsilon_n' H_n M_n H_n \epsilon_n}{\sqrt{n}\hat{\sigma}_{c_n}}, \tag{14}$$

where $\hat{\sigma}_{c_n}^2 = n^{-1}(\hat{\mu}_{4n} - 3\hat{\sigma}_n^4)\sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + n^{-1}\hat{\sigma}_n^4 \operatorname{tr}[H_n M_n H_n (M_n + M_n')]$ with $\hat{\mu}_{3n} = n^{-1}\sum_{i=1}^n \hat{\epsilon}_{ni}^3$ and $\hat{\mu}_{4n} = n^{-1}\sum_{i=1}^n \hat{\epsilon}_{ni}^4$. The $\mathbb{I}_n'$ is asymptotically standard normal. We use (8) to show the consistency of the bootstrap for $\mathbb{I}_n'$. Now let

$$c_n = n^{-1/2}\big(\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)\big), \tag{15}$$

$$\sigma_{c_n}^2 = n^{-1}(\mu_4 - 3\sigma_0^4)\sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + n^{-1}\sigma_0^4 \operatorname{tr}[H_n M_n H_n (M_n + M_n')], \tag{16}$$

$$d_n = \mathbb{I}_n' - c_n/\sigma_{c_n}. \tag{17}$$

Denote the bootstrapped $\mathbb{I}_n'$ by $\mathbb{I}_n'^*$. Correspondingly, we have c_n^*, $\sigma_{c_n}^{*2}$ and d_n^*.

Proposition 2. *Under H_0 and Assumptions I1–I3 and I4' in Appendix A.1,* $\sup_{x\in\mathbb{R}} |\,\mathrm{P}^*(\mathbb{I}_n'^* \le x) - \mathrm{P}(\mathbb{I}_n' \le x)| = o_P(1)$.

3.2. Spatial J-type Tests

In this subsection, we show the consistency of the bootstrapped spatial J-type tests for SARAR models (Kelejian and Piras, 2011). Consider the problem of testing one SARAR model against another one:

$$H_0: \quad y_n = \lambda_1 W_{1n} y_n + X_{1n}\beta_1 + u_{1n}, \quad u_{1n} = \rho_1 M_{1n} u_{1n} + \epsilon_{1n}, \quad \epsilon_{1n} = (\epsilon_{1n,1}, \ldots, \epsilon_{1n,n})', \tag{18}$$

$$H_1: \quad y_n = \lambda_2 W_{2n} y_n + X_{2n}\beta_2 + u_{2n}, \quad u_{2n} = \rho_2 M_{2n} u_{2n} + \epsilon_{2n}, \quad \epsilon_{2n} = (\epsilon_{2n,1}, \ldots, \epsilon_{2n,n})', \tag{19}$$

where $\epsilon_{1n,i}$'s are i.i.d. $(0, \sigma_1^2)$ and $\epsilon_{2n,i}$'s are i.i.d. $(0, \sigma_2^2)$. Other terms in the above models, with subscripts indicating different models, have similar meanings as those for the model (1). For $i = 1, 2$, let $\theta_i = (\lambda_i, \rho_i, \beta_i', \sigma_i^2)'$, $S_{in}(\lambda_i) = I_n - \lambda_i W_{in}$, $R_{in}(\rho_i) = I_n - \rho_i M_{in}$. The true parameter vector for the model (18) is θ_{10}. The idea of the J-type tests is to augment the null model using a predictor $\hat{y}_n$ for the dependent variable from the alternative model and test whether the coefficient of the predictor is significantly different from zero. In specific, the augmented model is

$$R_{1n}(\rho_1) y_n = \lambda_1 R_{1n}(\rho_1) W_{1n} y_n + R_{1n}(\rho_1) X_{1n}\beta_1 + \alpha R_{1n}(\rho_1)\hat{y}_n + \epsilon_n, \tag{20}$$

Note that a spatial Cochrane-Orcutt transformation has been used for the efficiency of the predictor $\hat{y}_n$. Given an estimator $\hat{\theta}_{2n}$ for the alternative model, a predictor of y_n can be $\hat{\lambda}_{2n} W_{2n} y_n + X_{2n}\hat{\beta}_{2n}$ from the r.h.s. of the equation for y_n in (19) or $S_{2n}^{-1}(\hat{\lambda}_{2n}) X_{2n}\hat{\beta}_{2n}$ from the reduced form.[4] In Kelejian and Piras

[4]The analyses for the two predictors are similar. In the following part, we only focus on the predictor $\hat{\lambda}_{2n} W_{2n} y_n + X_{2n}\hat{\beta}_{2n}$ for simplicity.

(2011), a spatial 2SLS estimator ρ_1 is plugged in (20) and $\hat{y}_n$ is also computed using the spatial 2SLS estimator, then (20) is estimated by the 2SLS. Alternatively, we can use the QMLE to compute $\hat{y}_n$ and then estimate ρ_1 jointly with λ_1, β_1 and α in (20) by the GMM. Under the null hypothesis, each estimator of α is asymptotically normal and the test is based on such a distribution. We first investigate the case with the estimation method in Kelejian and Piras (2011), and then study the case with the alternative estimation method.

The spatial 2SLS estimation of a SARAR model (Kelejian and Prucha, 1998), (18) or (19), involves several steps: $\gamma_i = (\lambda_i, \beta_i')'$ is first estimated by the 2SLS, then the residuals are used to estimate $\xi_i = (\rho_i, \sigma_i^2)'$ by a GMM with quadratic moments of the form $\mathrm{E}(\epsilon_{in}' D_{ij,n} \epsilon_{in}) = \sigma_{i0}^2 \operatorname{tr}(D_{ij,n})$, where $D_{ij,n}$ is an n-dimensional square matrix and σ_{i0}^2 is the true second moment when the ith SARAR model generates the data, and finally the estimates of λ_i and β_i are updated by the 2SLS estimation of the Cochrane-Orcutt transformed spatial model, for $i = 1, 2$. Kelejian and Prucha (1998) use the matrices I_n, M_{in} and $M_{in}' M_{in}$ for their quadratic moments in the second step. Let $Z_{in} = (W_{in} y_n, X_{in})$, $P_{A_n} = A_n (A_n' A_n)^{-1} A_n'$ for any full rank matrix A_n with row dimension n, Υ_{in} be the instruments for the first step estimation, $\check{\gamma}_{in}$ be the first step 2SLS estimator of γ_i, $\hat{\xi}_{in}$ be the estimator of ξ_i in the second step, Ξ_{in} be the instruments for the final step and $\hat{\gamma}_{in}$ be the estimator of γ_i from the final step.[5] With these notations, we have $\check{\gamma}_{in} = (Z_{in}' P_{\Upsilon_{in}} Z_{in})^{-1} Z_{in}' P_{\Upsilon_{in}} y_n$, the objective function of the second step in the spatial 2SLS is $g_n'(\xi_i; \check{\gamma}_{in}) g_n(\xi_i; \check{\gamma}_{in})$, where $g_n(\xi_i; \check{\gamma}_{in}) = n^{-1}[\epsilon_{in}'(\rho_i; \check{\gamma}_{in}) \epsilon_{in}(\rho_i; \check{\gamma}_{in}) - n\sigma_i^2, \epsilon_{in}'(\rho_i; \check{\gamma}_{in}) M_{in}' M_{in} \epsilon_{in}(\rho_i; \check{\gamma}_{in}) - \sigma_i^2 \operatorname{tr}(M_{in}' M_{in}), \epsilon_{in}'(\rho_i; \check{\gamma}_{in}) M_{in} \epsilon_{in}(\rho_i; \check{\gamma}_{in})]$ with $\epsilon_{in}(\rho_i; \check{\gamma}_{in}) = R_{in}(\rho_i)[S_{in}(\check{\lambda}_{in}) y_n - X_{in} \check{\beta}_{in}]$, and $\hat{\gamma}_{in} = [Z_{in}' R_{in}'(\hat{\rho}_{in}) P_{\Xi_{in}} R_{in}(\hat{\rho}_{in}) Z_{in}]^{-1} Z_{in}' R_{in}'(\hat{\rho}_{in}) P_{\Xi_{in}} R_{in}(\hat{\rho}_{in}) y_n$. For the estimation of (20), the instruments Δ_n can be from both models, so they can be generated from X_{1n}, X_{2n}, W_{1n}, W_{2n}, M_{1n} and M_{2n}. By the Frisch-Waugh-Lovell theorem on partitioned regressions,

$$\begin{aligned}
\hat{\alpha}_n &= \left[\left(P_{\Delta_n} R_{1n}(\hat{\rho}_{1n}) \hat{y}_n\right)' (I_n - P_{V_n(\hat{\rho}_{1n})}) P_{\Delta_n} R_{1n}(\hat{\rho}_{1n}) \hat{y}_n\right]^{-1} \left(P_{\Delta_n} R_{1n}(\hat{\rho}_{1n}) \hat{y}_n\right)' (I_n - P_{V_n(\hat{\rho}_{1n})}) R_{1n}(\hat{\rho}_{1n}) y_n \\
&= [\hat{y}_n' R_{1n}'(\hat{\rho}_{1n}) P_{\Delta_n} (I_n - P_{V_n(\hat{\rho}_{1n})}) P_{\Delta_n} R_{1n}(\hat{\rho}_{1n}) \hat{y}_n]^{-1} \hat{y}_n' R_{1n}'(\hat{\rho}_{1n}) P_{\Delta_n} (I_n - P_{V_n(\hat{\rho}_{1n})}) R_{1n}(\hat{\rho}_{1n}) R_{1n}^{-1} \epsilon_{1n},
\end{aligned} \tag{21}$$

where $V_n(\hat{\rho}_{1n}) = P_{\Delta_n} R_{1n}(\hat{\rho}_{1n}) Z_{1n}$. As $R_{1n}(\hat{\rho}_{1n}) R_{1n}^{-1} = I_n + (\rho_{10} - \hat{\rho}_{1n}) M_{1n} R_{1n}^{-1}$, the spatial J test statistic

$$J_{1n} = \hat{\alpha}_n / \hat{\sigma}_{\hat{\alpha}_n} = \hat{\alpha}_n [\hat{y}_n' R_{1n}'(\hat{\rho}_{1n}) P_{\Delta_n} (I_n - P_{V_n(\hat{\rho}_{1n})}) P_{\Delta_n} R_{1n}(\hat{\rho}_{1n}) \hat{y}_n]^{1/2} / \hat{\sigma}_{1n}, \tag{22}$$

where $\hat{\sigma}_{1n}^2 = n^{-1} \hat{\epsilon}_{1n}' \hat{\epsilon}_{1n}$ with $\hat{\epsilon}_{1n} = R_{1n}(\hat{\rho}_{1n})[S_{1n}(\hat{\lambda}_{1n}) y_n - X_{1n} \hat{\beta}_{1n}]$, is asymptotically standard normal under the null hypothesis and the assumption that $n^{-1} \Delta_n R_{1n}(\hat{\rho}_{1n}) \hat{y}_n$ converges to a non-zero limit in probability along with other regularity conditions. The assumption on $\hat{y}_n$ is on the whole term $n^{-1} \Delta_n R_{1n}(\hat{\rho}_{1n}) \hat{y}_n$, but remains implicit on the specific behavior of $\hat{\gamma}_{2n}$ under the null hypothesis. As we would like to study the

[5]The Υ_{in} can be generated from W_{in} and X_{in}, say the linear independent columns of X_{in}, $W_{in} X_{in}$ and $W_{in}^2 X_{in}$, and Ξ_{in} can be generated from W_{in}, M_{in} and X_{in}, say the linear independent columns of X_{in}, $W_{in} X_{in}$, $W_{in}^2 X_{in}$, $M_{in} X_{in}$ and $M_{in}^2 X_{in}$.

consistency of the bootstrapped spatial J tests, there is a need to investigate the remainder term of the spatial J test statistic after being approximated by a linear form of the disturbances. This can be done by using the pseudo-true values. The alternative model may have different functional forms or variables from those for the null model, thus the estimator for the alternative model generally would not converge to the true parameter value of the null model. But we can often find a sequence of non-stochastic vectors, i.e., pseudo-true values, such that the difference between the estimator and the pseudo-true value converges to zero in probability. As the spatial 2SLS involves three steps, we have a pseudo-true value in each step. In the first step, as $\check{\gamma}_{in} = (Z'_{in}P_{\Upsilon_{in}}Z_{in})^{-1}Z'_{in}P_{\Upsilon_{in}}y_n$, the pseudo-true value $\tilde{\gamma}_{in,1}$ can be $\tilde{\gamma}_{in,1} = (\mathrm{E}\, Z'_{in}P_{\Upsilon_{in}}\,\mathrm{E}\, Z_{in})^{-1}\,\mathrm{E}\, Z'_{in}P_{\Upsilon_{in}}\,\mathrm{E}\, y_n$.[6] As shown in Lemma 9, $n^{1/2}(\check{\gamma}_{in} - \tilde{\gamma}_{in,1}) = O_P(1)$. Then in step 2, the pseudo-true value $\bar{\xi}_{in,1}$ can be $\bar{\xi}_{in,1} = \arg\min_{\xi_i} n^{-1}\,\mathrm{E}\, g'_n(\xi_i;\tilde{\gamma}_{in,1})\,\mathrm{E}\, g_n(\xi_i;\tilde{\gamma}_{in,1})$. In the last step, the pseudo-true value $\bar{\gamma}_{in,1}$ is $\bar{\gamma}_{in,1} = [(R_{in}\,\mathrm{E}\, Z_{in})'P_{\Xi_{in}}R_{in}\,\mathrm{E}\, Z_{in}]^{-1}(R_{in}\,\mathrm{E}\, Z_{in})'P_{\Xi_{in}}R_{in}\,\mathrm{E}\, y_n$, where R_{in} denotes $R_{in}(\bar{\rho}_{in,1})$ for short. Let $\bar{\sigma}^2_{\hat{\alpha}_n} = \sigma^2_{10}[\bar{\gamma}'_{2n,1}\,\mathrm{E}(Z'_{2n})R'_{1n}P_{\Delta_n}(I_n - P_{V_n})P_{\Delta_n}R_{1n}\,\mathrm{E}(Z_{2n})\bar{\gamma}_{2n,1}]^{-1}$ and $\bar{\alpha}_n = \sigma_{10}^{-2}\bar{\sigma}^2_{\hat{\alpha}_n}\bar{\gamma}'_{2n,1}\,\mathrm{E}(Z'_{2n})R'_{1n}P_{\Delta_n}(I_n - P_{V_n})\epsilon_{1n}$ with $V_n = P_{\Delta_n}R_{1n}\,\mathrm{E}\, Z_{1n}$. Then as shown in the proof of Proposition 3, $J_{1n} = \bar{\alpha}_n/\bar{\sigma}_{\hat{\alpha}_n} + o_P(1)$. Although J_{1n} is approximated by a linear form of the disturbances, the bootstrap for J_{1n} can be proved to be consistent using a LQ form.[7] Corresponding to the bootstrapped data vector y^*_n,[8] let $\check{\gamma}^*_{in} = (Z^{*'}_{in}P_{\Upsilon_{in}}Z^*_{in})^{-1}Z^{*'}_{in}P_{\Upsilon_{in}}y^*_n$ with $Z^*_{in} = (W_{in}y^*_n, X_{in})$, $\hat{\xi}^*_{in} = \arg\min_{\xi_i} g^{*'}_n(\xi_i;\check{\gamma}^*_{in})g^*_n(\xi_i;\check{\gamma}^*_{in})$, where $g^*_n(\xi_i;\check{\gamma}^*_{in}) = n^{-1}[\epsilon^{*'}_{in}(\rho_i;\check{\gamma}^*_{in})\epsilon^*_{in}(\rho_i;\check{\gamma}^*_{in}) - n\sigma^2_i, \epsilon^{*'}_{in}(\rho_i;\check{\gamma}^*_{in})M'_{in}M_{in}\epsilon^*_{in}(\rho_i;\check{\gamma}^*_{in}) - \sigma^2_i\,\mathrm{tr}(M'_{in}M_{in}),$ $\epsilon^{*'}_{in}(\rho_i;\check{\gamma}^*_{in})M_{in}\epsilon^*_{in}(\rho_i;\check{\gamma}^*_{in})]$ with $\epsilon^*_{in}(\rho_i;\check{\gamma}^*_{in}) = R_{in}(\rho_i)[S_{in}(\check{\lambda}^*_{in})y^*_n - X_{in}\check{\beta}^*_{in}]$, and $\hat{\gamma}^*_{in} = [(R_{in}(\hat{\rho}^*_{in})Z^*_{in})'P_{\Xi_n}R_{in}(\hat{\rho}^*_{in})Z^*_{in}]^{-1}(R_{in}(\hat{\rho}^*_{in})Z^*_{in})'P_{\Xi_n}R_{in}(\hat{\rho}^*_{in})y^*_n$. Then the bootstrapped spatial J test statistic is

$$J^*_{1n} = \hat{\alpha}^*_n/\hat{\sigma}^*_{\hat{\alpha}^*_n} = \hat{\alpha}^*_n[\hat{y}^{*'}_n R'_{1n}(\hat{\rho}^*_{1n})P_{\Delta_n}(I_n - P_{V^*_n(\hat{\rho}^*_{1n})})P_{\Delta_n}R_{1n}(\hat{\rho}^*_{1n})\hat{y}^*_n]^{1/2}/\hat{\sigma}^*_{1n}, \tag{23}$$

where $\hat{y}^*_n = Z^*_{2n}\hat{\gamma}^*_{2n}$, $\hat{\sigma}^{*2}_{1n} = n^{-1}\epsilon^{*'}_{1n}(\hat{\rho}^*_{1n};\hat{\gamma}^*_{1n})\epsilon^*_{1n}(\hat{\rho}^*_{1n};\hat{\gamma}^*_{1n})$, $V^*_n(\hat{\rho}^*_{1n}) = P_{\Delta_n}R_{1n}(\hat{\rho}^*_{1n})Z^*_{1n}$, and

$$\hat{\alpha}^*_n = [\hat{y}^{*'}_n R'_{1n}(\hat{\rho}^*_{1n})P_{\Delta_n}(I_n - P_{V^*_n(\hat{\rho}^*_{1n})})P_{\Delta_n}R_{1n}(\hat{\rho}^*_{1n})\hat{y}^*_n]^{-1}\hat{y}^{*'}_n R'_{1n}(\hat{\rho}^*_{1n})P_{\Delta_n}(I_n - P_{V^*_n(\hat{\rho}^*_{1n})})R_{1n}(\hat{\rho}^*_{1n})R^{-1}_{1n}(\hat{\rho}_{1n})\epsilon^*_{1n}.$$

Proposition 3. *Under H_0 and the assumptions in Appendix A.2,* $\sup_{x\in\mathbb{R}}|\mathrm{P}^*(J^*_{1n} \le x) - \mathrm{P}(J_{1n} \le x)| = o_P(1)$.

Consider now the alternative estimation method of the augmented model (20). Let $\ddot{\theta}_{2n} = (\ddot{\gamma}'_{2n}, \ddot{\sigma}^2_{2n})'$ be the QMLE of the model (19) with $\bar{\theta}_{2n,1}$ being the pseudo-true value under H_0. For the estimation of (20), we can use both linear moments and quadratic moments for the GMM. Let $D_{1n}, \ldots, D_{mn}$ be n-dimensional square matrices with zero traces for the quadratic moments and Δ_n be the instrumental matrix used in the

[6]For generality, we use the term pseudo-true value for both $i = 1$ and $i = 2$. Note that $\tilde{\gamma}_{1n,1} = \gamma_{10}$.

[7]An alternative is to use the Mallows metric as in regression models. See Freedman (1981).

[8]Here $y^*_n = S^{-1}_n(\hat{\lambda}_{1n})[X_{1n}\hat{\beta}_{1n} + R^{-1}_{1n}(\hat{\rho}_{1n})\epsilon^*_{1n}]$, where ϵ^*_{1n} is an n-dimensional vector of random samples from the elements of $(I_n - l_n l'_n/n)\hat{\epsilon}_{1n}$, with $\hat{\epsilon}_{1n}$ being the residual vector from the spatial 2SLS estimation of the model (18).

2SLS estimation approach. The D_{in}'s can be constructed from W_{1n}, M_{1n}, W_{2n} and M_{2n}. The moment vector is $g_n(\psi;\ddot{\gamma}_{2n}) = n^{-1}(\epsilon_n'(\psi;\ddot{\gamma}_{2n})D_{1n}\epsilon_n(\psi;\ddot{\gamma}_{2n}),\ldots,\epsilon_n'(\psi;\ddot{\gamma}_{2n})D_{mn}\epsilon_n(\psi;\ddot{\gamma}_{2n}),\epsilon_n'(\psi;\ddot{\gamma}_{2n})\Delta_n)'$, where $\psi = (\lambda_1,\rho_1,\beta_1',\alpha)'$ and $\epsilon_n(\psi;\ddot{\gamma}_{2n}) = R_{1n}(\rho_1)[S_{1n}(\lambda_1)y_n - X_{1n}\beta_1 - \alpha(\ddot{\lambda}_{2n}W_{2n}y_n + X_{2n}\ddot{\beta}_{2n})]$. The true parameter vector of ψ is $\psi_0 = (\lambda_{10},\rho_{10},\beta_{10}',0)'$. A general objective function of the GMM is $g_n'(\psi;\ddot{\gamma}_{2n})a_na_n'g_n(\psi;\ddot{\gamma}_{2n})$, where $\{a_n\}$ is a sequence of full rank matrices that converges to a constant matrix a_0. By the generalized Cauchy-Schwarz inequality, the optimal weighting matrix is the variance-covariance (VC) matrix Ω_n of $n^{1/2}g_n(\psi_0;\gamma_2)$. For the feasible optimal GMM, a first step consistent estimator $\check{\psi}_n$ can be derived from minimizing $g_n'(\psi;\ddot{\gamma}_{2n})g_n(\psi;\ddot{\gamma}_{2n})$, then an estimator $\hat{\psi}_n$ can be the minimizer of $g_n'(\psi;\ddot{\gamma}_{2n})\hat{\Omega}_n^{-1}g_n(\psi;\ddot{\gamma}_{2n})$, where $\hat{\Omega}_n$ is the matrix obtained by replacing the ψ_0 and other moment parameters of $\epsilon_{1n,i}$ in Ω_n by, respectively, $\check{\psi}_n$ and the corresponding sample moments of the first-step residuals. Under some regularity conditions, $\hat{\psi}_n$ is consistent for ψ_0 and $n^{1/2}(\hat{\psi}_n - \psi_0)$ is asymptotically normal with limiting VC matrix $\lim_{n\to\infty}[\mathrm{E}\,G_n'(\psi_0;\gamma_2)\Omega_n^{-1}\,\mathrm{E}\,G_n(\psi_0;\gamma_2)]^{-1}$, where $G_n(\psi;\gamma_2) = \frac{\partial g_n(\psi;\gamma_2)}{\partial\psi'}$. Then we may let the spatial J test statistic be

$$J_{2n} = n^{1/2}e_\psi'\hat{\psi}_n/[e_\psi'(G_n'(\hat{\psi}_n;\ddot{\gamma}_{2n})\hat{\Omega}_n^{-1}G_n(\hat{\psi}_n;\ddot{\gamma}_{2n}))^{-1}e_\psi]^{1/2}, \tag{24}$$

where e_ψ is a vector with length equal to that of ψ, whose last element is 1 and other elements are zero. As shown in Section 2, J_{2n} can be approximated by a LQ form as every element of $g_n(\psi_0;\gamma_2)$ is a linear or quadratic form of ϵ_{1n}. With the bootstrapped data vector y_n^*,[9] let $\ddot{\theta}_{2n}^*$ be the QMLE of the model (19), the moment vector for the GMM estimation of (20) be $g_n^*(\psi;\ddot{\gamma}_{2n}^*) = n^{-1}(\epsilon_n^{*'}(\psi;\ddot{\gamma}_{2n}^*)D_{1n}\epsilon_n^*(\psi;\ddot{\gamma}_{2n}^*),\ldots,$ $\epsilon_n^{*'}(\psi;\ddot{\gamma}_{2n}^*)D_{mn}\epsilon_n^*(\psi;\ddot{\gamma}_{2n}^*),\epsilon_n^{*'}(\psi;\ddot{\gamma}_{2n}^*)\Delta_n)'$ with $\epsilon_n^*(\psi;\ddot{\gamma}_{2n}^*) = R_{1n}(\rho_1)[S_{1n}(\lambda_1)y_n^* - X_{1n}\beta_1 - \alpha(\ddot{\lambda}_{2n}^*W_{2n}y_n^* + X_{2n}\ddot{\beta}_{2n}^*)]$, $\check{\psi}_n^*$ and $\hat{\psi}_n^*$ be the first-step and second-step estimators in the feasible optimal GMM approach respectively, $G_n^*(\psi;\gamma_2) = \frac{\partial g_n^*(\psi;\gamma_2)}{\partial\psi'}$, and $\hat{\Omega}_n^*$ be the matrix obtained by replacing the estimators in $\hat{\Omega}_n$ by the corresponding ones with y_n^*. Then the bootstrapped J_{2n} is

$$J_{2n}^* = n^{1/2}e_\psi'\hat{\psi}_n^*/[e_\psi'(G_n^{*'}(\hat{\psi}_n^*;\ddot{\gamma}_{2n}^*)\hat{\Omega}_n^{*-1}G_n^*(\hat{\psi}_n^*;\ddot{\gamma}_{2n}^*))^{-1}e_\psi]^{1/2}. \tag{25}$$

Proposition 4. *Under H_0 and the assumptions in Appendix A.3,* $\sup_{x\in\mathbb{R}}|\mathrm{P}^*(J_{2n}^* \le x) - \mathrm{P}(J_{2n} \le x)| = o_P(1)$.

4. Asymptotic Refinements

The Edgeworth expansion has been well established for a smooth function of sample averages of independent random vectors and/or stationary dependent random vectors. It provides a useful tool to prove that the bootstrap may provide asymptotic refinements. The LQ forms for spatial econometric models

[9] Here we may let $y_n^* = S_n^{-1}(\ddot{\lambda}_{1n})[X_{1n}\ddot{\beta}_{1n} + R_{1n}^{-1}(\ddot{\rho}_{1n})\epsilon_{1n}^*]$, where $\ddot{\theta}_{1n} = (\ddot{\lambda}_{1n},\ddot{\rho}_{1n},\ddot{\beta}_{1n}',\ddot{\sigma}_{1n}^2)'$ is the QMLE of the model (18), and ϵ_{1n}^* is an n-dimensional vector of random samples from the elements of $(I_n - l_nl_n'/n)\hat{\epsilon}_{1n}$, with $\hat{\epsilon}_{1n}$ being the residual vector from the QML estimation of the model (18).

involve spatial weights matrices and cannot be written as simple sample averages of disturbances or their cross-products. If we would like to investigate possible asymptotic refinements of the bootstrap in spatial econometric models using some expansions, we need to justify the validity for such expansions first. When the disturbances in a LQ form are normally distributed, Edgeworth expansions can be established without much difficulty. But when the disturbances are not normal, directly investigating possible expansions can be hard.[10] An alternative approach is to decompose a LQ form into the sum of martingale differences and then study the expansions for martingales. Mykland (1993) establishes an asymptotic expansion for martingales, but the expansions are not in a pointwise topology. Below we discuss the cases of normal and non-normal disturbances separately.

4.1. Normal Disturbances

When the disturbances in a LQ form $c_n/\sigma_{c_n} = n^{-1/2}\big(\epsilon_n' A_n \epsilon_n - \sigma_0^2 \operatorname{tr}(A_n) + b_n'\epsilon_n\big)/\sigma_{c_n}$ are i.i.d. normal, we can easily derive its characteristic function. By a smoothing inequality in Feller (1970), the difference between two functions has an upper bound generated from the Fourier transforms relating to these two functions. The inequality is used to establish the Berry-Esseen bound for the error of the approximation of the normal distribution to the true distribution for a sample mean of i.i.d. disturbances or the Edgeworth expansion of the sample mean. It can also be used to establish the Edgeworth expansion of a LQ form. Let $f^{(k)}(x)$ be the kth order derivative of a function $f(x)$.

Theorem 2. *Under Assumptions 2 and 3, when $\epsilon_n \sim N(0, \sigma_0^2 I_n)$,*

$$\sup_{x\in\mathbb{R}} |\,\mathrm{P}(c_n/\sigma_{c_n} \le x) - [\Phi(x) + \kappa_n(1-x^2)\Phi^{(1)}(x)]| = O(n^{-1}), \tag{26}$$

$$\sup_{x\in\mathbb{R}} |\,\mathrm{P}^*(c_n^*/\sigma_{c_n}^* \le x) - [\Phi(x) + \kappa_n^*(1-x^2)\Phi^{(1)}(x)]| = O_P(n^{-1}), \tag{27}$$

where $\kappa_n = n^{-3/2}\sigma_{c_n}^{-3}[4\sigma_0^6 \operatorname{tr}(A_n^3)/3 + \sigma_0^4 b_n' A_n b_n] = O(n^{-1/2})$ with $\sigma_{c_n}^2 = n^{-1}[2\sigma_0^4 \operatorname{tr}(A_n^2) + \sigma_0^2 b_n' b_n]$ and $\kappa_n^* = n^{-3/2}\sigma_{c_n}^{*-3}[4\sigma_n^{*6} \operatorname{tr}(A_n^3)/3 + \sigma_n^{*4} b_n' A_n b_n] = O_P(n^{-1/2})$ with $\sigma_{c_n}^{*2} = n^{-1}[2\sigma_n^{*4} \operatorname{tr}(A_n^2) + \sigma_n^{*2} b_n' b_n]$, and for $r \ge 3$, there exist real polynomials $P_{n3}(x), \ldots, P_{nr}(x)$ with bounded coefficients such that

$$\sup_{x\in\mathbb{R}} \Big|\mathrm{P}(c_n/\sigma_{c_n} \le x) - \Phi(x) - \Phi^{(1)}(x)\sum_{i=3}^{r} n^{-(i-2)/2} P_{ni}(x)\Big| = O(n^{-(r-1)/2}). \tag{28}$$

Eqs. (26) and (27) can be used to show that the bootstrap can provide asymptotic refinements for some statistics that can be approximated by a LQ form. Eq. (28) presents a general high order expansion for

[10] Götze et al. (2007) establishes a one term Edgeworth expansion for a quadratic form. Their proof is based on a symmetrization inequality and the differential inequality method. The quadratic matrix in Götze et al. (2007) has some special feature not shared by a spatial weights matrix. With a spatial weights matrix in the quadratic form, the expansion established using similar methods may not generate a remainder term of a desirable order. In addition, the generalization to a LQ form is not straightforward.

the CDF of a LQ form. Note that κ_n has a relatively simple form. Instead of bootstrapping tests, we may correct the bias distortion for test statistics that can be approximated by a LQ form.[11] The above theorem can be applied to show that the bootstrap for Moran's I is more accurate than the first-order asymptotic theory.

Proposition 5. *Under H_0 and Assumptions I1–I4 in Appendix A.1, the Moran I statistic in Eq. (5) satisfies* $\mathrm{P}^*(\mathbb{I}_n^* \leq x) - \mathrm{P}(\mathbb{I}_n \leq x) = O_P(n^{-1})$.

4.2. Non-normal Disturbances

For LQ forms with non-normal disturbances, a theorem on asymptotic expansions for martingales in Mykland (1993) can be applied to establish an expansion, which the author calls the Edgeworth expansion for martingales. The conditions needed are mainly imposed on the variation measures associated with martingales, e.g., the optional kth-order variation, which is defined as the sum of the kth powers of the martingale differences. One condition is the central limit theorem which relates to the optional second-order variations. The c_n/σ_{c_n} can be decomposed as the sum of martingale differences that are quadratic in the disturbances. We need the existence of $\mathrm{E}\,|\epsilon_{ni}|^{4(1+\delta)}$ for some $\delta > 0$ to show the asymptotic normality of c_n/σ_{c_n}, which is based a central limit theorem for martingales. For the central limit theorem relating to the optional second-order variations, higher moments for ϵ_{ni} are required to exist. Correspondingly, a stronger condition on b_n is also assumed.

Assumption 1'. *The ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})'$ are i.i.d. $(0, \sigma_0^2)$ and $\mathrm{E}\,|\epsilon_{ni}|^{8(1+\delta)} < \infty$ for some $\delta > 0$.*

Assumption 2'. *The sequence of symmetric matrices $\{A_n\}$ are bounded in both row and column sum norms and the elements of the vectors $\{b_n\}$ satisfy $\sup_n n^{-1} \sum_{i=1}^n |b_{ni}|^{4(1+\delta)} < \infty$.*

Theorem 3. *Under Assumptions 1', 2' and 3, we have*

$$\int_{-\infty}^{+\infty} h(x)\, dF_n(x) = \int_{-\infty}^{+\infty} h(x)\, d\Phi(x) + \frac{1}{6} n^{-1/2}\, \mathrm{E}\big[\big(\psi_o(Y) + 2\psi_p(Y)\big) h^{(2)}(Y)\big] + o(n^{-1/2}), \tag{29}$$

where $F_n(x) = \mathrm{P}(c_n/\sigma_{c_n} \leq x)$, Y is the normal random variable that c_n/σ_{c_n} converges to, and expressions for $\psi_o(Y)$ and $\psi_p(Y)$ are given in (C.17)–(C.20), uniformly on a set ℓ of functions h which are twice differentiable, with h, $h^{(1)}$ and $h^{(2)}$ uniformly bounded, and with $\{h^{(2)}, h \in \ell\}$ being equicontinuous a.e. Lebesgue. Denote the convergence in (29) by $o_2(n^{-1/2})$ (Mykland, 1993), then

$$F_n(x) = \Phi(x) + \frac{1}{6} n^{-1/2} \big(\psi_o^{(1)}(x) + 2\psi_p^{(1)}(x) - [\psi_o(x) + 2\psi_p(x)]x\big)\Phi^{(1)}(x) + o_2(n^{-1/2}). \tag{30}$$

[11] Robinson and Rossi (2010) have considered a finite sample correction of Moran's I test for a pure SAR model. They have not shown the validity of their expansion for the CDF of Moran's I test statistic, which is in terms of the CDF for a chi-square distribution.

As pointed out by Mykland (1993), the expansion generally does not hold when h is an indicator function of an interval, so it is a "smoothed" expansion. Note that $\psi_o(x)$ and $\psi_p(x)$ are linear in x, then $\psi_o^{(1)}(x) + 2\psi_p^{(1)}(x) - [\psi_o(x) + 2\psi_p(x)]x = (1-x^2)[\psi_o^{(1)}(x) + 2\psi_p^{(1)}(x)]$. In the special case that ϵ_{ni}'s are i.i.d. normal, we can verify that $\frac{1}{6}n^{-1/2}[\psi_o^{(1)}(x) + 2\psi_p^{(1)}(x) - [\psi_o(x) + 2\psi_p(x)]x] = (1-x^2)\lim_{n\to\infty}\kappa_n$, thus (30) has similar terms as the usual one-term Edgeworth expansion (26).

5. Conclusion

In this paper, we consider the use of the bootstrap in spatial econometric models. We show that the bootstrap for estimators and test statistics in spatial econometric models can be studied based on LQ forms. We have established the uniform convergence of the CDF for a LQ form to that of the standard normal random variable. Based on this result, we show that the bootstrap is consistent for Moran's I and spatial J-type test statistics. As possible asymptotic refinements for the bootstrap are usually shown by using some asymptotic expansions, we discuss expansions for LQ forms: for normal disturbances, we have established the Edgeworth expansions for LQ forms and applied the result to show the second-order correctness of the bootstrap for Moran's I; for non-normal disturbances, we have established an asymptotic expansion based on martingales.

There are some extensions which can be of interest for future research. Some asymptotic chi-square tests in spatial econometrics, e.g., hypothesis tests with multiple constraints, are constructed from vectors of LQ forms. The current uniform convergence result, which is only about a single LQ form, does not cover vectors of LQ forms. It is of interest to establish the uniform convergence result for vectors of LQ forms so that the bootstrap can be shown to be consistent for asymptotic chi-square tests. It also remains to show high order expansions of a vector of LQ forms for asymptotic refinements of the bootstrap.

Appendix A. Assumptions

Appendix A.1. Assumptions for Moran's I

Assumption I1. *The sequence of matrices $\{M_n\}$ have zero diagonals and are bounded in both row and column sum norms.*

Assumption I2. *The sequence of full rank matrices $\{X_n\}$ have uniformly bounded constant elements, and $\lim_{n\to\infty}\frac{1}{n}X_n'X_n$ exists and is nonsingular.*

Assumption I3. *The sequence $\{n^{-1}[(\mu_4 - 3\sigma_0^4)\sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + \sigma_0^4 \operatorname{tr}(M_n^2 + M_n' M_n)]\}$ is bounded away from zero.*

Assumption I4. *The disturbance vector $\epsilon_n \sim N(0, \sigma_0^2 I_n)$.*

Assumption I4'. *The ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})'$ are i.i.d. and* $\mathrm{E}\,\epsilon_{ni}^8 < \infty$.

The variance of $n^{1/2}\epsilon_n' H_n M_n H_n \epsilon_n$ is guaranteed to be bounded away from zero in Assumption I3, as $n^{-1}\operatorname{tr}[H_n M_n H_n (M_n + M_n')] = n^{-1}\operatorname{tr}(M_n^2 + M_n' M_n) + o(1)$ by Lemma 1. When the disturbances are not assumed to be normal, $\mathbb{I}_n'$ generally involves the estimated fourth moment of ϵ_{ni}. To prove the consistency of the bootstrapped $\mathbb{I}_n'$ using Theorem 1, we need to know the rate of convergence of the estimated fourth moment to the true one, thus a strong condition on ϵ_{ni} is imposed in Assumption I4'.

Appendix A.2. Assumptions for the Spatial J Tests: J_{1n}

Assumption J1. *The $\epsilon_{1n,i}$'s are i.i.d. $(0, \sigma_{10}^2)$ and the moment $\mathrm{E}(\epsilon_{1n,i}^4)$ exists.*

Assumption J2. *The matrices X_{1n} and X_{2n} have full ranks and uniformly bounded constants. The limits $\lim_{n\to\infty} \frac{1}{n} X_{1n}' X_{1n}$ and $\lim_{n\to\infty} \frac{1}{n} X_{2n}' X_{2n}$ exist and are nonsingular.*

Assumption J3. *Matrices S_{1n} and R_{1n} are nonsingular.*

Assumption J4. *The sequences of matrices $\{W_{1n}\}$, $\{M_{1n}\}$, $\{R_{1n}^{-1}\}$ and $\{S_{1n}^{-1}\}$ are bounded in both row and column sum norms. The $\{W_{1n}\}$ and $\{M_{1n}\}$ have zero diagonals.*

Assumption J5. *The $n^{-1}\Upsilon_{1n}'\Upsilon_{1n}$, $n^{-1}\Xi_{1n}'\Xi_{1n}$, $n^{-1}\Upsilon_{1n}'(W_{1n}S_{1n}^{-1}X_{1n}\beta_{10}, X_{1n})$ and $n^{-1}\Xi_{1n}'R_{1n}(W_{1n}S_{1n}^{-1}X_{1n}\beta_{10}, X_{1n})$ converge to full rank matrices.*

Assumption J6. *The minimum eigenvalue of the matrix*

$$\frac{1}{n}\begin{pmatrix} n & 2\sigma_{10}^2 \operatorname{tr}(M_{1n}R_{1n}^{-1}) & \sigma_{10}^2 \operatorname{tr}(R_{1n}'^{-1}M_{1n}'M_{1n}R_{1n}^{-1}) \\ \operatorname{tr}(M_{1n}'M_{1n}) & 2\sigma_{10}^2 \operatorname{tr}(M_{1n}'M_{1n}^2R_{1n}^{-1}) & \sigma_{10}^2 \operatorname{tr}(R_{1n}'^{-1}M_{1n}'^2M_{1n}^2R_{1n}^{-1}) \\ 0 & \sigma_{10}^2 \operatorname{tr}[(M_{1n}+M_{1n}')M_{1n}R_{1n}^{-1}] & \sigma_{10}^2 \operatorname{tr}(R_{1n}'^{-1}M_{1n}'M_{1n}^2R_{1n}^{-1}) \end{pmatrix}$$

is bounded away from zero, $|\lambda_1| < 1$, $|\rho_1| < 1$ and $0 < \sigma_1^2 < c$ for some $c > 0$.

Assumption J7. *The $n^{-1}\Upsilon_{2n}'\Upsilon_{2n}$, $n^{-1}\Xi_{2n}'\Xi_{2n}$, $n^{-1}\Upsilon_{2n}'(W_{2n}S_{1n}^{-1}X_{1n}\beta_{10}, X_{2n})$ and $n^{-1}\Xi_{2n}'R_{2n}(W_{2n}S_{1n}^{-1}X_{1n}\beta_{10}, X_{2n})$ converge to full rank matrices.*

Assumption J8. *For any $\eta > 0$, there exists $\kappa > 0$ such that, when $||\xi_2 - \bar{\xi}_{2n,1}|| > \eta$, $n^{-1}[\mathrm{E}\, g_n'(\xi_2; \tilde{\gamma}_{2n,1})\,\mathrm{E}\, g_n(\xi_2; \tilde{\gamma}_{2n,1}) - \mathrm{E}\, g_n'(\bar{\xi}_{2n,1}; \tilde{\gamma}_{2n,1})\,\mathrm{E}\, g_n(\bar{\xi}_{2n,1}; \tilde{\gamma}_{2n,1})] > \kappa$ for all large enough n. The $\bar{\xi}_{2n,1}$ is in the interior of the compact parameter space of ξ_2.*

Assumption J9. *The $n^{-1}\Delta_n'\Delta_n$ and $n^{-1}\Delta_n'R_{1n}(W_{2n}S_{1n}^{-1}X_{1n}\beta_{10}, X_{2n})\bar{\gamma}_{2n,1}$ converge to full rank matrices.*

Assumptions J1–J6 are similar to those in Kelejian and Prucha (1998). Assumption J7 is for the estimators $\tilde{\gamma}_{2n}$ and $\hat{\gamma}_{2n}$, similar to Assumption J5 for $\tilde{\gamma}_{1n}$ and $\hat{\gamma}_{1n}$. Assumption J8 states the identification uniqueness condition for $\bar{\xi}_{2n,1}$. The condition for the estimation of the augmented model (20), Assumption J9, is stated in terms of the pseudo-true value $\bar{\gamma}_{2n,1}$.

Appendix A.3. Assumptions for the Spatial J Tests: J_{2n}

Let $L_{1n}(\theta_1)$ be the log likelihood function of the model (18), $L_{2n}(\theta_2)$ be the log likelihood function of the model (19), and $\bar{\theta}_{in,1} = \arg\max \bar{L}_{in}(\theta_i;\theta_{10})$ with $\bar{L}_{in}(\theta_i;\theta_{10}) = \mathrm{E}\, L_{in}(\theta_i)$ under H_0, for $i = 1, 2$. Maximizing $L_{in}(\theta_i)$ and $\bar{L}_{in}(\theta_i;\theta_1)$ for given β_i and σ_i^2 yields functions $L_{in}(\phi_i)$ and $\bar{L}_{in}(\phi_i;\theta_1)$ respectively, where $\phi_i = (\lambda_i, \rho_i)'$.

Assumption J10. *The $\epsilon_{1n,i}$'s are i.i.d. $(0, \sigma_{10}^2)$ and the moment $\mathrm{E}(\epsilon_{1n,i}^8)$ exists.*

Assumption J11. *The matrices X_{1n} and X_{2n} have full ranks and uniformly bounded constants. The limits $\lim_{n\to\infty} \frac{1}{n} X_{1n}' X_{1n}$ and $\lim_{n\to\infty} \frac{1}{n} X_{2n}' X_{2n}$ exist and are nonsingular.*

Assumption J12. *Matrices S_{1n} and R_{1n} are nonsingular.*

Assumption J13. *The sequences of matrices $\{W_{1n}\}$, $\{M_{1n}\}$, $\{R_{1n}^{-1}\}$, $\{S_{1n}^{-1}\}$, $\{W_{2n}\}$ and $\{M_{2n}\}$ are bounded in both row and column sum norms. The $\{W_{1n}\}$, $\{M_{1n}\}$, $\{W_{2n}\}$ and $\{M_{2n}\}$ have zero diagonals.*

Assumption J14. *Each sequence of matrices $\{S_{1n}^{-1}(\lambda_1)\}$, $\{R_{1n}^{-1}(\rho_1)\}$, $\{S_{2n}^{-1}(\lambda_2)\}$ and $\{R_{2n}^{-1}(\rho_2)\}$ is bounded in either row or column sum norm uniformly in the compact parameter space. The λ_{10}, ρ_{10}, $\bar{\lambda}_{2n,1}$ and $\bar{\rho}_{2n,1}$ are in the interiors of their parameter spaces.*

Assumption J15. *The limits $\lim_{n\to\infty} \frac{1}{n} X_{1n}' R_{1n}'(\rho_1) R_{1n}(\rho_1) X_{1n}$ and $\lim_{n\to\infty} \frac{1}{n} X_{2n}' R_{2n}'(\rho_2) R_{2n}(\rho_2) X_{2n}$ exist and are nonsingular for any ρ_1 and ρ_2 in their respective parameter spaces. The smallest eigenvalues of $R_{1n}'(\rho_1) R_{1n}(\rho_1)$ and $R_{2n}'(\rho_2) R_{2n}(\rho_2)$ are bounded away from zero uniformly on their respective parameter spaces.*

Assumption J16. *For the identification of the model (18), either (i) $\lim_{n\to\infty} \frac{1}{n}[\ln|\sigma_{10}^2 S_{1n}^{-1} R_{1n}^{-1} R_{1n}'^{-1} S_{1n}'^{-1}| - \ln|\bar{\sigma}_{1n,a}^2(\phi_1) S_{1n}^{-1}(\lambda_1) R_{1n}^{-1}(\rho_1) R_{1n}'^{-1}(\rho_1) S_{1n}'^{-1}(\lambda_1)|]$ exists and is nonzero for any $\phi_1 \neq \phi_{10}$, where $\bar{\sigma}_{1n,a}^2(\phi_1) = \frac{\sigma_{10}^2}{n} \mathrm{tr}[R_{1n}'^{-1} S_{1n}'^{-1} S_{1n}'(\lambda_1) R_{1n}'(\rho_1) R_{1n}(\rho_1) S_{1n}(\lambda_1) S_{1n}^{-1} R_{1n}^{-1}]$, or (ii) $\lim_{n\to\infty} \frac{1}{n}(Q_{1n} X_{1n}\beta_{10}, X_{1n})'(Q_{1n} X_{1n}\beta_{10}, X_{1n})$ exists and is nonsingular, and $\lim_{n\to\infty} \frac{1}{n}[\ln|\sigma_{10}^2 S_{1n}^{-1} R_{1n}^{-1} R_{1n}'^{-1} S_{1n}'^{-1}| - \ln|\bar{\sigma}_{1n,a}^2(\lambda_{10}, \rho_1) S_{1n}^{-1} R_{1n}^{-1}(\rho_1) R_{1n}'^{-1}(\rho_1) S_{1n}'^{-1}|]$ exists and is nonzero for any $\rho_1 \neq \rho_{10}$, where $Q_{1n} = W_{1n} S_{1n}^{-1}$. For the model (19), for $\eta > 0$, there exists $\kappa > 0$ such that, when $||\phi_2 - \bar{\phi}_{2n,1}|| > \eta$, $n^{-1}\big(\bar{L}_{2n}(\bar{\phi}_{2n,1};\theta_{10}) - \bar{L}_{2n,1}(\phi_2;\theta_{10})\big) > \kappa$ for any large enough n.*

Assumption J17. *The limits $\lim_{n\to\infty} \frac{1}{n} \frac{\partial^2 \bar{L}_{1n}(\phi_{10};\theta_{10})}{\partial\phi_1 \partial\phi_1'}$ and $\lim_{n\to\infty} \frac{1}{n} \frac{\partial^2 \bar{L}_{2n}(\bar{\phi}_{2n,1};\theta_{10})}{\partial\phi_2 \partial\phi_2'}$ exist and are nonsingular.*

Assumption J18. *The $\{\bar{\sigma}_{2n,1}^2\}$ is bounded away from zero.*

Assumption J19. *Either (i) $\lim_{n\to\infty} n^{-1} \Delta_n' R_{1n}(\rho_1) \Gamma_n$, where $\Gamma_n = (W_{1n} S_{1n}^{-1} X_{1n}\beta_{10}, X_{1n}, \bar{\lambda}_{2n,1} W_{2n} S_{1n}^{-1} X_{1n}\beta_{10} + X_{2n}\bar{\beta}_{2n,1})$, has full rank $k_{x_1} + 2$ for each possible ρ_1 in its parameter space, and the moment equations*

$\mathrm{tr}[R_{1n}^{\prime -1} R_{1n}^{\prime}(\rho_1) P_{im} R_{1n}(\rho_1) R_{1n}^{-1}] = 0$, *for* $i = 1, \ldots, m$, *have the unique solution at* ρ_{10}, *or (ii)* $\lim_{n\to\infty} n^{-1} \Delta_n^{\prime} R_{1n}(\rho_1) X_{1n}$ *has full rank* k_{x_1} *for each possible* ρ_1 *in its parameter space, and the moment equations* $\mathrm{tr}[R_{1n}^{\prime -1} S_{1n}^{\prime -1}(S_{1n}^{\prime}(\lambda_1) - \alpha\bar{\lambda}_{2n,1} W_{2n}^{\prime}) R_{1n}^{\prime}(\rho_1) P_{im} R_{1n}(\rho_1)(S_{1n}(\lambda_1) - \alpha\bar{\lambda}_{2n,1} W_{2n}) S_{1n}^{-1} R_{1n}^{-1}] = 0$, *for* $i = 1, \ldots, m$, *have the unique solution at the true parameter values.*

Assumptions J10–J18 are directly from Jin and Lee (2012) with a few modifications. A strong condition is needed in Assumption J10 as explained in Appendix A.1 for Assumption I4'. Assumption J16 strengthens the original identification condition for θ_2 so that we have the result on $\hat{\theta}_{2n}^{*}$ in Lemma 17. Assumption J19 is the identification uniqueness condition of the GMM estimation for the augmented model (20), which resembles a condition for the GMM estimation of high order SARAR models in Lee and Liu (2010).

Appendix B. Lemmas

Appendix B.1. Elementary Lemmas

Lemmas 1–4 can be found in, e.g., Lin and Lee (2010).

Lemma 1. *Suppose that* $n \times n$ *matrices* $\{A_n\}$ *are bounded in both row and column sum norms. Elements of* $n \times k$ *matrices* $\{X_n\}$ *are uniformly bounded and* $\lim_{n\to\infty} n^{-1} X_n^{\prime} X_n$ *exists and is nonsingular. Let* $H_n = I_n - X_n(X_n^{\prime} X_n)^{-1} X_n^{\prime}$. *Then* $\{H_n\}$ *are bounded in both row and column sum norms and* $\mathrm{tr}(H_n A_n) = \mathrm{tr}(A_n) + O(1)$.

Lemma 2. *Suppose that* $A_n = [a_{n,ij}]$ *and* $B_n = [b_{n,ij}]$ *are* $n \times n$ *matrices and* ϵ_{ni}*'s in* $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})^{\prime}$ *are i.i.d. with mean zero and variance* σ_0^2. *Then,*

(1) $\mathrm{E}(\epsilon_n \cdot \epsilon_n^{\prime} A_n \epsilon_n) = \mathrm{E}(\epsilon_{ni}^3)(a_{n,11}, \ldots, a_{n,nn})^{\prime}$, *and*

(2) $\mathrm{E}(\epsilon_n^{\prime} A_n \epsilon_n \cdot \epsilon_n^{\prime} B_n \epsilon_n) = [\mathrm{E}(\epsilon_{ni}^4) - 3\sigma_0^4] \sum_{i=1}^{n} a_{n,ii} b_{n,ii} + \sigma_0^4 \,\mathrm{tr}(A_n)\,\mathrm{tr}(B_n) + \sigma_0^4 \,\mathrm{tr}[A_n(B_n + B_n^{\prime})]$.

Lemma 3. *Suppose that* $n \times n$ *matrices* $\{A_n\}$ *are bounded in both row and column sum norms, elements of the* $n \times k$ *matrices* $\{C_n\}$ *are uniformly bounded, and* ϵ_{ni}*'s in* $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})^{\prime}$ *are independent* $(0, \sigma_{ni}^2)$. *The sequences* $\{\sigma_{ni}^2\}$ *and* $\{\mathrm{E}(\epsilon_{ni}^4)\}$ *are bounded. Then* $\epsilon_n^{\prime} A_n \epsilon_n = O_P(n)$, $\mathrm{E}(\epsilon_n^{\prime} A_n \epsilon_n) = O(n)$, $n^{-1}[\epsilon_n^{\prime} A_n \epsilon_n - \mathrm{E}(\epsilon_n^{\prime} A_n \epsilon_n)] = o_P(1)$ *and* $n^{1/2} C_n^{\prime} A_n \epsilon_n = O_P(1)$.

Lemma 4. *Suppose that* $\{A_n\}$ *is a sequence of symmetric* $n \times n$ *matrices with row and column sum norms bounded and* $b_n = (b_{n1}, \ldots, b_{nn})^{\prime}$ *is an* n*-dimensional column vector such that* $\sup_n n^{-1} \sum_{i=1}^{n} |b_{ni}|^{2+\eta_1} < \infty$ *for some* $\eta_1 > 0$. *Furthermore, suppose that* $\epsilon_{n1}, \cdots, \epsilon_{nn}$ *are mutually independent with zero means and the moments* $\mathrm{E}(|\epsilon_{ni}|^{4+\eta_2})$ *for some* $\eta_2 > 0$ *exist and are uniformly bounded for all* n *and* i. *Let* $\sigma_{Q_n}^2$ *be the variance of* Q_n *where* $Q_n = \epsilon_n^{\prime} A_n \epsilon_n + b_n^{\prime} \epsilon_n - tr(A_n \Sigma_n)$ *with* Σ_n *being a diagonal matrix with* $\mathrm{E}\,\epsilon_{ni}^2$*'s on its diagonal. Assume that* $n^{-1}\sigma_{Q_n}^2$ *is bounded away from zero. Then* $\frac{Q_n}{\sigma_{Q_n}} \xrightarrow{d} N(0,1)$.

Lemmas 5–8 are for the SARAR model (1), where ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \dots, \epsilon_{nn})'$ are i.i.d. with mean zero, variance σ_0^2, third moment μ_3 and finite fourth moment μ_4, and $\hat{\epsilon}_n = R_n(\hat{\rho}_n)[S_n(\hat{\lambda}_n)y_n - X_n\hat{\beta}_n]$ with $\hat{\theta}_n$ being $n^{1/2}$-consistent, i.e., $n^{1/2}(\hat{\theta}_n - \theta_0) = O_P(1)$. The ϵ_n^*, y_n^* and $\hat{\theta}_n^*$ are derived as described in Section 3. Let $||\cdot||$ be the Euclidean norm of a vector.

Lemma 5. *Let $P_{ln} = [p_{ln,ij}]$ be $n \times n$ matrices which are bounded in row sum norms, for $l = 1, \dots, s$. If $\sup_{n,j} \mathrm{E}\,|\epsilon_{nj}|^s < \infty$, then $n^{-1}\sum_{i=1}^n \prod_{l=1}^s \sum_{j=1}^n |p_{ln,ij}\epsilon_{nj}| = O_P(1)$.*

Proof. For $s = 1$, the result is immediate. So consider $s > 1$. For $s > 1$, there exists a finite r such that $\frac{1}{r} + \frac{1}{s} = 1$. Hölder's inequality implies that

$$\begin{aligned}\sum_{j=1}^n |p_{ln,ij}||\epsilon_{nj}| &\le \sum_{j=1}^n |p_{ln,ij}|^{\frac{1}{r}}|p_{ln,ij}|^{\frac{1}{s}}|\epsilon_{nj}| \le [\sum_{j=1}^n (|p_{ln,ij}|^{\frac{1}{r}})^r]^{\frac{1}{r}}[\sum_{j=1}^n (|p_{ln,ij}|^{\frac{1}{s}}|\epsilon_{nj}|)^s]^{\frac{1}{s}} \\ &\le c^{\frac{1}{r}}[\sum_{j=1}^n |p_{ln,ij}||\epsilon_{nj}|^s]^{\frac{1}{s}} \le c^{\frac{1}{r}}[\sum_{j=1}^n(\sum_{l=1}^s |p_{ln,ij}|)|\epsilon_{nj}|^s]^{\frac{1}{s}},\end{aligned}$$

where $c = \sup_{l=1,\cdots,s}||P_{ln}||_\infty$. It follows that

$$\prod_{l=1}^s\sum_{j=1}^n |p_{ln,ij}||\epsilon_{nj}| \le c^{\frac{s}{r}}[\sum_{j=1}^n(\sum_{l=1}^s |p_{ln,ij}|)|\epsilon_{nj}|^s].$$

Hence,

$$\mathrm{E}(\prod_{l=1}^s\sum_{j=1}^n |p_{ln,ij}||\epsilon_{nj}|) \le c^{\frac{s}{r}}(\sum_{l=1}^s\sum_{j=1}^n |p_{ln,ij}|)\sup_{n,j}\mathrm{E}\,|\epsilon_{nj}|^s \le sc^{1+\frac{s}{r}}\sup_{n,j}\mathrm{E}\,|\epsilon_{nj}|^s = sc^s\sup_{n,j}\mathrm{E}\,|\epsilon_{nj}|^s = O(1).$$

The result of stochastic boundedness follows from Markov's inequality. □

Lemma 6. *For any integer r, if $\mathrm{E}\,|\epsilon_{ni}|^r < \infty$, $\mathrm{E}^*\,\epsilon_{ni}^{*r} = \mathrm{E}\,\epsilon_{ni}^r + o_P(1)$, $n^{-1}\sum_{i=1}^n \hat{\epsilon}_{ni}^r = \mathrm{E}\,\epsilon_{ni}^r + o_P(1)$, $\mathrm{E}^*\,|\epsilon_{ni}^*|^r = \mathrm{E}\,|\epsilon_{ni}|^r + o_P(1)$ and $n^{-1}\sum_{i=1}^n |\hat{\epsilon}_{ni}|^r = \mathrm{E}\,|\epsilon_{ni}|^r + o_P(1)$. If $\mathrm{E}\,\epsilon_{ni}^{2r} < \infty$, $n^{1/2}[\mathrm{E}^*\,\epsilon_{ni}^{*r} - \mathrm{E}\,\epsilon_{ni}^r] = O_P(1)$ and $n^{1/2}[n^{-1}\sum_{i=1}^n \hat{\epsilon}_{ni}^r - \mathrm{E}\,\epsilon_{ni}^r] = O_P(1)$.*

Proof. Let $J_n = I_n - \frac{1}{n}l_n l_n'$. As $y_n = S_n^{-1}(X_n\beta_0 + R_n^{-1}\epsilon_n)$,

$$\begin{aligned}\epsilon_n^* &= J_n\hat{\epsilon}_n \\ &= J_n\big(R_n + (\rho_0 - \hat{\rho}_n)M_n\big)\big(S_n y_n - X_n\beta_0 + (\lambda_0 - \hat{\lambda}_n)W_n y_n + X_n(\beta_0 - \hat{\beta}_n)\big) \\ &= \epsilon_n - \frac{l_n'\epsilon_n}{n}l_n + (\lambda_0 - \hat{\lambda}_n)\big(J_n R_n + (\rho_0 - \hat{\rho}_n)J_n M_n\big)W_n S_n^{-1}X_n\beta_0 \\ &\quad + \big(J_n R_n X_n + (\rho_0 - \hat{\rho}_n)J_n M_n X_n\big)(\beta_0 - \hat{\beta}_n) \\ &\quad + (\lambda_0 - \hat{\lambda}_n)\big(J_n R_n + (\rho_0 - \hat{\rho}_n)J_n M_n\big)W_n S_n^{-1}R_n^{-1}\epsilon_n + (\rho_0 - \hat{\rho}_n)J_n M_n R_n^{-1}\epsilon_n.\end{aligned} \tag{B.1}$$

Write $\epsilon_n^* = \epsilon_n + \sum_{j=1}^r \zeta_{1n,j}p_{nj} + \sum_{j=1}^s \zeta_{2n,j}Q_{nj}\epsilon_n$, where $p_{nj} = [p_{nj,i}]$ is an n-dimensional vector with bounded constant elements, $Q_{nj} = [q_{nj,il}]$ is an $n \times n$ matrix with bounded row and column sum norms, and $\zeta_{1n,j}$ and $\zeta_{2n,j}$'s are equal to $l_n'\epsilon_n/n$, $\lambda_0 - \hat{\lambda}_n$, $\rho_0 - \hat{\rho}_n$, elements of $\beta_0 - \hat{\beta}_n$ or their products. Then $\zeta_{1n,j} = O_P(n^{-1/2})$ and $\zeta_{2n,j} = O_P(n^{-1/2})$. The ϵ_{ni}^{*r} can be expanded by the multinomial theorem, which states that $(x_1+\cdots+x_m)^r = \sum_{k_1,\ldots,k_m} \binom{r}{k_1,\ldots,k_m} x_1^{k_1}\ldots x_m^{k_m}$, where $\binom{r}{k_1,\ldots,k_m}$ is a multinomial coefficient and the summation is taken over all sequences of nonnegative integer indices k_1 through k_m such that their sum is r. Then we have an expansion form for $n^{-1}\sum_{i=1}^n \epsilon_{ni}^{*r} - n^{-1}\sum_{i=1}^n \epsilon_{ni}^r$, where each term in the expansion has the product form $T_{1n}T_{2n}$ with T_{1n} being products of $\zeta_{1n,j}$ and $\zeta_{2n,j}$'s and T_{2n} not involving $\zeta_{1n,j}$ and $\zeta_{2n,j}$'s. The T_{2n} is either bounded or stochastically bounded by Lemma 5. It follows that $\mathrm{E}^*\epsilon_{ni}^{*r} = \mathrm{E}\,\epsilon_{ni}^r + o_P(1)$ by the law of large numbers and $n^{1/2}[\mathrm{E}^*\epsilon_{ni}^{*r} - \mathrm{E}\,\epsilon_{ni}^r] = O_P(1)$ by Chebyshev's inequality. Other results are similarly derived. $\square$

Lemma 7. *Let $P_{ln} = [p_{ln,ij}]$ be $n\times n$ matrices with bounded row sum norms for $l = 1,\ldots,s$, then $\mathrm{P}^*(n^{-1}\sum_{i=1}^n \prod_{k=1}^s \sum_{j=1}^n |p_{kn,ij}\epsilon_{nj}^*| > \eta) = O_P(1)$ for $\eta > 0$, if $\mathrm{E}|\epsilon_{ni}|^s < \infty$.*

Proof. The proof is similar to that for Lemma 5 except for the application of Lemma 6. $\square$

Lemma 8. *For $\eta > 0$ and an integer r, $\mathrm{P}^*(|n^{-1}\sum_{i=1}^n \hat{\epsilon}_{ni}^{*r} - \mathrm{E}^*\epsilon_{ni}^{*r}| > \eta) = o_P(1)$ if $\mathrm{E}|\epsilon_{ni}|^r < \infty$ and $\mathrm{P}^*(||\hat{\theta}_n^* - \hat{\theta}_n|| > \kappa) = o_P(1)$ for $\kappa > 0$, and $\mathrm{P}^*(n^a|n^{-1}\sum_{i=1}^n \hat{\epsilon}_{ni}^{*r} - \mathrm{E}^*\epsilon_{ni}^{*r}| > \eta) = o_P(1)$ for $0 \le a < 1/2$ if $\mathrm{E}|\epsilon_{ni}|^{2r} < \infty$ and $\mathrm{P}^*(n^a||\hat{\theta}_n^* - \hat{\theta}_n|| > \kappa) = o_P(1)$ for $\kappa > 0$.*

Proof. As $y_n^* = S_n^{-1}(\hat{\lambda}_n)(X_n\hat{\beta}_n + R_n^{-1}(\hat{\rho}_n)\epsilon_n^*)$,

$$\begin{aligned}
\hat{\epsilon}_n^* &= \big(R_n(\hat{\rho}_n) + (\hat{\rho}_n - \hat{\rho}_n^*)M_n\big)\big(S_n(\hat{\lambda}_n)y_n^* - X_n\hat{\beta}_n + (\hat{\lambda}_n - \hat{\lambda}_n^*)W_n y_n^* + X_n(\hat{\beta}_n - \hat{\beta}_n^*)\big) \\
&= \epsilon_n^* + (\hat{\lambda}_n - \hat{\lambda}_n^*)\big(R_n(\hat{\rho}_n) + (\hat{\rho}_n - \hat{\rho}_n^*)M_n\big)W_n S_n^{-1}(\hat{\lambda}_n)X_n\hat{\beta}_n \\
&\quad + \big(R_n(\hat{\rho}_n)X_n + (\hat{\rho}_n - \hat{\rho}_n^*)M_n X_n\big)(\hat{\beta}_n - \hat{\beta}_n^*) \\
&\quad + (\hat{\lambda}_n - \hat{\lambda}_n^*)\big(R_n(\hat{\rho}_n) + (\hat{\rho}_n - \hat{\rho}_n^*)M_n\big)W_n S_n^{-1}(\hat{\lambda}_n)R_n^{-1}(\hat{\rho}_n)\epsilon_n^* + (\hat{\rho}_n - \hat{\rho}_n^*)M_n R_n^{-1}(\hat{\rho}_n)\epsilon_n^*.
\end{aligned}$$

Write $\hat{\epsilon}_n^* = \epsilon_n^* + \sum_{j=1}^r \zeta_{1n,j}p_{nj} + \sum_{j=1}^s \zeta_{2n,j}Q_{nj}\epsilon_n^*$, where $p_{nj} = [p_{nj,i}]$ is an n-dimensional vector with bounded constant elements, $Q_{nj} = [q_{nj,il}]$ is an $n\times n$ matrix with bounded row and column sum norms, and $\zeta_{1n,j}$ and $\zeta_{2n,j}$'s are equal to $\hat{\lambda}_n - \hat{\lambda}_n^*$, $\hat{\rho}_n - \hat{\rho}_n^*$, elements of $\hat{\beta}_n - \hat{\beta}_n^*$ or their products. Now the argument is similar to that for Lemma 6 except for the application of Lemma 7. $\square$

Appendix B.2. Lemmas for the Spatial J Tests: J_{1n}

Lemma 9. $n^{1/2}(\check{\gamma}_{1n} - \gamma_{10}) = O_P(1)$ *and* $n^{1/2}(\check{\gamma}_{2n} - \tilde{\gamma}_{2n,1}) = O_P(1)$.

Proof. Noting that $\check{\gamma}_{in} = [\frac{1}{n}Z_{in}'\Upsilon_{in}(\frac{1}{n}\Upsilon_{in}'\Upsilon_{in})^{-1}\frac{1}{n}\Upsilon_{in}'Z_{in}]^{-1}\frac{1}{n}Z_{in}'\Upsilon_{in}(\frac{1}{n}\Upsilon_{in}'\Upsilon_{in})^{-1}\frac{1}{n}\Upsilon_{in}'y_n$, where $n^{-1}\Upsilon_{in}'Z_{in} = n^{-1}\Upsilon_{in}'\,\mathrm{E}\,Z_{in} + o_P(1) = O_P(1)$ and $n^{-1/2}\Upsilon_{in}'(y_n - \mathrm{E}\,y_n) = O_P(1)$ for $i = 1, 2$, the result follows. $\square$

Lemma 10. $n^{1/2}(\hat{\xi}_{1n} - \xi_{10}) = O_P(1)$ *and* $n^{1/2}(\hat{\xi}_{2n} - \bar{\xi}_{2n,1}) = O_P(1)$.

Proof. It has been shown in Kelejian and Prucha (1998) that $\hat{\xi}_{1n} - \xi_{10} = o_P(1)$. By the mean value theorem,

$$0 = \frac{\partial g_n'(\hat{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1} g_n(\hat{\xi}_{1n};\check{\gamma}_{1n}) = \frac{\partial g_n'(\hat{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1}\big(g_n(\xi_{10};\check{\gamma}_{1n}) + \frac{\partial g_n(\tilde{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1'}(\hat{\xi}_{1n} - \xi_{10})\big),$$

where $\tilde{\xi}_{1n}$ is between $\hat{\xi}_{1n}$ and ξ_{10}. Then $n^{1/2}(\hat{\xi}_{1n}-\xi_{10}) = -\big(\frac{\partial g_n'(\hat{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1}\frac{\partial g_n(\tilde{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1'}\big)^{-1}\frac{\partial g_n'(\hat{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1} n^{1/2} g_n(\xi_{10};\check{\gamma}_{1n})$. Noting that $g_n(\xi_1;\gamma_1)$ is linear in σ_1^2 and quadratic in ρ_1 and γ_1, we can write $\tilde{\rho}_{1n} = (\tilde{\rho}_{1n} - \rho_{10}) + \rho_{10}$ and $\check{\gamma}_{1n} = (\check{\gamma}_{1n} - \gamma_{10}) + \gamma_{10}$, and expand relevant terms in $\frac{\partial g_n(\tilde{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1'}$. Then it is easy to see that $\frac{\partial g_n(\tilde{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1'} = \frac{\partial g_n(\xi_{10};\gamma_{10})}{\partial \xi_1'} + o_P(1)$. In addition, $\frac{\partial g_n(\xi_{10};\gamma_{10})}{\partial \xi_1'} = \mathrm{E}\,\frac{\partial g_n(\xi_{10};\gamma_{10})}{\partial \xi_1'} + o_P(1)$ by Chebyshev's inequality. Thus, $\frac{\partial g_n(\tilde{\xi}_{1n};\check{\gamma}_{1n})}{\partial \xi_1'} = \mathrm{E}\,\frac{\partial g_n(\xi_{10};\gamma_{10})}{\partial \xi_1'} + o_P(1) = O_P(1)$. Furthermore, $n^{1/2} g_n(\xi_{10};\check{\gamma}_{1n}) = n^{1/2} g_n(\xi_{10};\gamma_{10}) + \frac{\partial g_n(\xi_{10};\gamma_{10})}{\partial \gamma_1'} n^{1/2}(\check{\gamma}_{1n} - \gamma_{10}) + \frac{1}{2}(\check{\gamma}_{1n} - \gamma_{10})'\frac{\partial^2 g_n(\xi_{10};\gamma_{10})}{\partial \gamma_1 \partial \gamma_1'} n^{1/2}(\check{\gamma}_{1n} - \gamma_{10}) = O_P(1)$, as $n^{1/2} g_n(\xi_{10};\gamma_{10}) = O_P(1)$ by Chebyshev's inequality and $n^{1/2}(\check{\gamma}_{1n} - \gamma_{10}) = O_P(1)$. Thus, $n^{1/2}(\hat{\xi}_{1n} - \xi_{10}) = O_P(1)$.

As $g_{2n}(\xi_2;\gamma_2)$ is quadratic in ξ_2, it is easy to see that $g_{2n}(\xi_2;\check{\gamma}_{2n}) - \mathrm{E}\, g_{2n}(\xi_2;\bar{\gamma}_{2n,1}) = o_P(1)$ uniformly in the parameter space of ξ_2, by showing that $g_{2n}(\xi_2;\check{\gamma}_{2n}) - g_{2n}(\xi_2;\bar{\gamma}_{2n,1}) = o_P(1)$ and $g_{2n}(\xi_2;\bar{\gamma}_{2n,1}) - \mathrm{E}\, g_{2n}(\xi_2;\bar{\gamma}_{2n,1}) = o_P(1)$ uniformly in the parameter space of ξ_2. In addition, $\mathrm{E}\, g_{2n}'(\xi_2;\bar{\gamma}_{2n,1})\,\mathrm{E}\, g_{2n}(\xi_2;\bar{\gamma}_{2n,1})$ is uniformly equicontinuous. Then the identification uniqueness assumption implies that $\hat{\xi}_{2n} - \bar{\xi}_{2n} = o_P(1)$. With this result, it can be proved that $n^{1/2}(\hat{\xi}_{2n} - \bar{\xi}_{2n,1}) = O_P(1)$ in a way similar to the proof for the result on $\hat{\xi}_{1n}$. □

Lemma 11. *$n^{1/2}(\hat{\gamma}_{1n} - \gamma_{10}) = O_P(1)$ and $n^{1/2}(\hat{\gamma}_{2n} - \bar{\gamma}_{2n,1}) = O_P(1)$.*

Proof. Write $\hat{\gamma}_{in} = [\frac{1}{n} Z_{in}' R_{in}'(\hat{\rho}_{in})\Xi_{in}(\frac{1}{n}\Xi_{in}'\Xi_{in})^{-1}\frac{1}{n}\Xi_{in}' R_{in}(\hat{\rho}_{in}) Z_{in}]^{-1}\frac{1}{n} Z_{in}' R_{in}'(\hat{\rho}_{in})\Xi_{in}(\frac{1}{n}\Xi_{in}'\Xi_{in})^{-1}\frac{1}{n}\Xi_{in}' R_{in}(\hat{\rho}_{in}) y_n$. Since $n^{-1}\Xi_{in}' R_{in}(\hat{\rho}_{in}) Z_{in} = n^{-1}\Xi_{in}' R_{in} Z_{in} + o_P(1) = n^{-1}\Xi_{in}' R_{in}\,\mathrm{E}\, Z_{in} + o_P(1) = O_P(1)$ and $n^{-1/2}\Xi_{in}'(R_{in}(\hat{\rho}_{in}) y_n - R_{in}\,\mathrm{E}\, y_n) = n^{-1/2}\Xi_{in}' R_{in}(y_n - \mathrm{E}\, y_n) + n^{-1}\Xi_{in} M_{in} y_n n^{1/2}(\bar{\rho}_{in,1} - \hat{\rho}_{in}) = O_P(1)$ for $i = 1, 2$, the result follows.
□

Lemma 12. *For $i = 1, 2$, $\eta > 0$ and $0 \le a < \frac{1}{2}$, $\mathrm{P}^*(n^a ||\check{\gamma}_{in}^* - \check{\gamma}_{in}|| > \eta) = o_P(1)$.*

Proof. Since $\check{\gamma}_{in}^* = [\frac{1}{n} Z_{in}^{*\prime}\Upsilon_{in}(\frac{1}{n}\Upsilon_{in}'\Upsilon_{in})^{-1}\frac{1}{n}\Upsilon_{in}' Z_{in}^*]^{-1}\frac{1}{n} Z_{in}^{*\prime}\Upsilon_{in}(\frac{1}{n}\Upsilon_{in}'\Upsilon_{in})^{-1}\frac{1}{n}\Upsilon_{in}' y_n^*$, where $\mathrm{P}^*(||n^{-1}\Upsilon_{in}'(Z_{in}^* - \mathrm{E}^*\, Z_{in}^*)|| > \eta) = o_P(1)$ and $\mathrm{P}^*(||n^{a-1/2}\Upsilon_{in}'(y_n^* - \mathrm{E}^*\, y_n^*)|| > \eta) = o_P(1)$ by Chebyshev's inequality and Lemma 6, the result follows. □

Lemma 13. *For $i = 1, 2$ and $\eta > 0$, $\mathrm{P}^*(||\hat{\xi}_{in}^* - \hat{\xi}_{in}|| > \eta) = o_P(1)$.*

Proof. Note that $\epsilon_{in}^*(\rho_i;\check{\gamma}_{in}^*) = R_{in}(\rho_i)[S_{in}(\check{\lambda}_{in}^*) y_n^* - X_{in}\check{\beta}_{in}^*] = \epsilon_{in}^*(\rho_i;\check{\gamma}_{in}) + R_{in}(\rho_i)[(\check{\lambda}_{in} - \check{\lambda}_{in}^*) W_n y_n^* + X_{in}(\check{\beta}_{in} - \check{\beta}_{in}^*)]$, then $\mathrm{P}^*(\sup_{\xi_i \in \varpi_i}\big|||g_n^*(\xi_i;\check{\gamma}_{in}^*)||^2 - ||g_n^*(\xi_i;\check{\gamma}_{in})||^2\big| > \eta) = o_P(1)$ for $\eta > 0$, where ϖ_i denotes the parameter space of ξ_i, as $R_{in}(\rho_i)$ is linear in ρ_i. Since $\epsilon_{in}(\rho_i;\bar{\gamma}_{in,1}) = R_{in}(\rho_i)[S_{in}(\bar{\lambda}_{in,1}) y_n - X_{in}\bar{\beta}_{in,1}]$ and $\epsilon_{in}^*(\rho_i;\check{\gamma}_{in}) = R_{in}(\rho_i)[S_{in}(\check{\lambda}_{in}) y_n^* - X_{in}\check{\beta}_{in}]$, $\sup_{\xi_i \in \varpi_i} ||\mathrm{E}^*\, g_n^*(\xi_i;\check{\gamma}_{in}) - \mathrm{E}\, g_n(\xi_i;\bar{\gamma}_{in,1})|| = o_P(1)$ by the mean value theorem. As a result, $\sup_{\xi_i\in\varpi_i}\big|\|\mathrm{E}^*\, g_n^*(\xi_i;\check{\gamma}_{in})\|^2 - \|\mathrm{E}\, g_n(\xi_i;\tilde{\gamma}_{in,1})\|^2\big| = o_P(1)$, since $\sup_{\xi_i\in\varpi_i}\|\mathrm{E}\, g_n(\xi_i;\tilde{\gamma}_{in,1})\| = O_P(1)$.

If $\|\xi_i - \hat{\xi}_{in}\| > \eta$, $\|\xi_i - \bar{\xi}_{in,1}\| \geq \|\xi_i - \hat{\xi}_{in}\| - \|\hat{\xi}_{in} - \bar{\xi}_{in,1}\| > \eta/2$ with probability $1-o(1)$. Then given $\eta > 0$, there exists a $\kappa > 0$, such that $\|\xi_i - \hat{\xi}_{in}\| > \eta$ implies that $\|\mathrm{E}^*\, g_n^*(\xi_i;\check{\gamma}_{in})\|^2 - \|\mathrm{E}^*\, g_n^*(\bar{\xi}_{in,1};\check{\gamma}_{in})\|^2 \geq \kappa$ with probability $1-o(1)$. Then for $\eta_1 > 0$,

$$
\begin{aligned}
&\mathrm{P}\big(\mathrm{P}^*(\|\hat{\xi}_{in}^* - \hat{\xi}_{in}\| > \eta) > \eta_1\big) \\
&\leq \mathrm{P}\big(\mathrm{P}^*(\|\mathrm{E}^*\, g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in})\|^2 - \|\mathrm{E}^*\, g_n^*(\bar{\xi}_{in,1};\check{\gamma}_{in})\|^2 \geq \kappa) > \eta_1\big) + o(1) \\
&\leq \mathrm{P}\big(\mathrm{P}^*(\|\mathrm{E}^*\, g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in})\|^2 - \|g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in})\|^2 + \|g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in}^*)\|^2 - \|\mathrm{E}^*\, g_n^*(\bar{\xi}_{in,1};\check{\gamma}_{in})\|^2 \geq \kappa) > \eta_1\big) + o(1) \\
&\leq \mathrm{P}\big(\mathrm{P}^*(\|\mathrm{E}^*\, g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in})\|^2 - \|g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in})\|^2 + \|g_n^*(\bar{\xi}_{in,1};\check{\gamma}_{in}^*)\|^2 - \|\mathrm{E}^*\, g_n^*(\bar{\xi}_{in,1};\check{\gamma}_{in})\|^2 \geq \kappa) > \eta_1\big) + o(1) \\
&\leq \mathrm{P}\big(\mathrm{P}^*(\|\mathrm{E}^*\, g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in})\|^2 - \|g_n^*(\hat{\xi}_{in}^*;\check{\gamma}_{in})\|^2 + \|g_n^*(\bar{\xi}_{in,1};\check{\gamma}_{in})\|^2 - \|\mathrm{E}^*\, g_n^*(\bar{\xi}_{in,1};\check{\gamma}_{in})\|^2 \geq \kappa) > \eta_1\big) + o(1) \\
&\leq \mathrm{P}\big(\mathrm{P}^*\big(\sup_{\xi_i\in\varpi_i} 2\big(\big|\|g_n^*(\xi_i;\check{\gamma}_{in})\|^2 - \|\mathrm{E}^*\, g_n^*(\xi_i;\check{\gamma}_{in})\|^2\big| \geq \kappa\big) > \eta_1\big)\big) + o(1).
\end{aligned}
$$

The $g_n^*(\xi_i;\check{\gamma}_{in})$ is a 3×1 vector with each element being of the form $g_{nj}^*(\xi_i;\check{\gamma}_{in}) = n^{-1}[\epsilon_{in}^{*\prime}(\rho_i;\check{\gamma}_{in})D_n\epsilon_{in}^*(\rho_i;\check{\gamma}_{in}) - \sigma_i^2\,\mathrm{tr}(D_n)]$, where D_n is an $n\times n$ matrix. By Chebyshev's inequality, $\mathrm{P}(\mathrm{P}^*(\sup_{\xi_i\in\varpi_i}|g_{nj}^*(\xi_i;\check{\gamma}_{in}) - \mathrm{E}^*\, g_{nj}^*(\xi_i;\check{\gamma}_{in})| \geq \kappa) > \eta_1) = \mathrm{P}(\kappa^{-2}\,\mathrm{E}^*(\sup_{\xi_i\in\varpi_i}|g_{nj}^*(\xi_i;\check{\gamma}_{in}) - \mathrm{E}^*\, g_{nj}^*(\xi_i;\check{\gamma}_{in})|)^2 > \eta_1) = o(1)$ by Lemma 6. Then $\mathrm{P}(\mathrm{P}^*\big(\sup_{\xi_i\in\varpi_i}\big(\big|\|g_n^*(\xi_i;\check{\gamma}_{in}) - \mathrm{E}^*\, g_n^*(\xi_i;\check{\gamma}_{in})\|^2\big| \geq \kappa\big) > \eta_1) = o(1)$ and $\mathrm{P}(\mathrm{P}^*\big(\sup_{\xi_i\in\varpi_i} 2\big(\big|\|g_n^*(\xi_i;\check{\gamma}_{in})\|^2 - \|\mathrm{E}^*\, g_n^*(\xi_i;\check{\gamma}_{in})\|^2\big| \geq \kappa\big) > \eta_1)) = o(1)$. Thus $\mathrm{P}(\mathrm{P}^*(\|\hat{\xi}_{in}^* - \hat{\xi}_{in}\| > \eta) > \eta_1) = o(1)$. □

Lemma 14. *For $i = 1, 2$, $\eta > 0$ and $0 \leq a < \frac{1}{2}$, $\mathrm{P}^*(n^a\|\hat{\gamma}_{in}^* - \hat{\gamma}_{in}\| > \eta) = o_P(1)$.*

Proof. Since $\hat{\gamma}_{in} = [Z_{in}'R_{in}'(\hat{\rho}_{in})P_{\Xi_{in}}R_{in}(\hat{\rho}_{in})Z_{in}]^{-1}Z_{in}'R_{in}'(\hat{\rho}_{in})P_{\Xi_{in}}R_{in}(\hat{\rho}_{in})y_n$ and $\hat{\gamma}_{in}^* = [Z_{in}^{*\prime}R_{in}'(\hat{\rho}_{in}^*)P_{\Xi_{in}}R_{in}(\hat{\rho}_{in}^*)Z_{in}^*]^{-1}Z_{in}^{*\prime}R_{in}'(\hat{\rho}_{in}^*)P_{\Xi_{in}}R_{in}(\hat{\rho}_{in}^*)y_n^*$, we only need to show that $\mathrm{P}^*(n^{-1}\|\Xi_{in}R_{in}(\hat{\rho}_{in}^*)Z_{in}^* - \Xi_{in}R_{in}(\hat{\rho}_{in})Z_{in}\| > \eta) = o_P(1)$ and $\mathrm{P}^*(n^{a-1}\|\Xi_{in}R_{in}(\hat{\rho}_{in}^*)y_{in}^* - \Xi_{in}R_{in}(\hat{\rho}_{in})y_{in}\| > \eta) = o_P(1)$. Write $n^{-1}[\Xi_{in}R_{in}(\hat{\rho}_{in}^*)Z_{in}^* - \Xi_{in}R_{in}(\hat{\rho}_{in})Z_{in}] = n^{-1}\Xi_{in}R_{in}(\hat{\rho}_{in})(Z_{in}^* - Z_{in}) + n^{-1}\Xi_{in}M_{in}Z_{in}^*(\hat{\rho}_{in} - \hat{\rho}_{in}^*)]$, where $Z_{in} = [W_{in}S_{1n}^{-1}(X_{1n}\beta_{10} + R_{1n}^{-1}\epsilon_{1n}), X_{in}]$ and $Z_{in}^* = [W_{in}S_{1n}^{-1}(\hat{\lambda}_{1n})(X_{1n}\hat{\beta}_{1n} + R_{1n}^{-1}(\hat{\rho}_{1n})\epsilon_{1n}^*), X_{in}]$. Since $n^{-1}\Xi_{in}R_{in}(\hat{\rho}_{in})W_{in}S_{1n}^{-1}R_{1n}^{-1}\epsilon_{1n} = n^{-1}\Xi_{in}R_{in}W_{in}S_{1n}^{-1}R_{1n}^{-1}\epsilon_{1n} + n^{-1}\Xi_{in}M_{in}W_{in}S_{1n}^{-1}R_{1n}^{-1}\epsilon_{1n}(\hat{\rho}_{in} - \bar{\rho}_{in,1}) = o_P(1)$, $\mathrm{P}^*\big(\|n^{-1}\Xi_{in}R_{in}(\hat{\rho}_{in})W_{in}S_{1n}^{-1}(\hat{\lambda}_{in})R_{1n}^{-1}(\hat{\rho}_{in})\epsilon_{1n}^*\| > \eta\big) = o_P(1)$ by first using Chebyshev's inequality and then using the mean value theorem, and $\mathrm{P}^*\big(\|n^{-1}\Xi_{in}R_{in}(\hat{\rho}_{in})W_{in}[S_{1n}^{-1}(\hat{\lambda}_{in})X_{1n}\hat{\beta}_{1n} - S_{1n}^{-1}X_{1n}\beta_{10}]\| > \eta\big) = o_P(1)$ by the mean value theorem, we have $\mathrm{P}^*\big(\|n^{-1}\Xi_{in}R_{in}(\hat{\rho}_{in})(Z_{in}^* - Z_{in})\| > \eta\big) = o_P(1)$. Similarly, $\mathrm{P}^*\big(\|n^{-1}\Xi_{in}M_{in}(Z_{in}^* - Z_{in})\| > \eta\big) = o_P(1)$, and $\mathrm{P}^*(n^{a-1}\|\Xi_{in}R_{in}(\hat{\rho}_{in}^*)y_{in}^* - \Xi_{in}R_{in}(\hat{\rho}_{in})y_{in}\| > \eta) = o_P(1)$ by a similar argument with adjustments of orders. Thus the result in the lemma holds. □

Appendix B.3. Lemmas for the Spatial J Tests: J_{2n}

Lemma 15. *$n^{1/2}(\ddot{\theta}_{1n} - \theta_{10}) = O_P(1)$ and $n^{1/2}(\ddot{\theta}_{2n} - \bar{\theta}_{2n,1}) = O_P(1)$.*

Proof. See Jin and Lee (2012). □

Lemma 16. $\frac{1}{\sqrt{n}}\|\frac{\partial^2 L_{2n}(\tilde{\theta}_{2n})}{\partial\theta_2\partial\theta_2'} - \frac{\partial^2 \bar{L}_{2n}(\bar{\theta}_{2n,1};\theta_{10})}{\partial\theta_2\partial\theta_2'}\| = O_P(1)$*, where* $\tilde{\theta}_{2n}$ *is between* $\ddot{\theta}_{2n}$ *and* $\bar{\theta}_{2n,1}$*.*

Proof. We prove the result by showing that (i) $n^{-1/2}\|\frac{\partial^2 L_{2n}(\tilde{\theta}_{2n})}{\partial\theta_2\partial\theta_2'} - \frac{\partial^2 L_{2n}(\bar{\theta}_{2n,1})}{\partial\theta_2\partial\theta_2'}\| = O_P(1)$ and (ii) $n^{-1/2}\|\frac{\partial^2 L_{2n}(\bar{\theta}_{2n,1})}{\partial\theta_2\partial\theta_2'} - \mathrm{E}\,\frac{\partial^2 L_{2n}(\bar{\theta}_{2n,1})}{\partial\theta_2\partial\theta_2'}\| = O_P(1)$. To prove (i), apply the mean value theorem to each term in the second order derivative. Specifically, we investigate $n^{-1/2}|\frac{\partial^2 L_{2n}(\tilde{\theta}_{2n})}{\partial\lambda_2^2} - \mathrm{E}\,\frac{\partial^2 L_{2n}(\bar{\theta}_{2n,1})}{\partial\lambda_2^2}|$. Results for other terms can be derived similarly. The $L_{2n}(\theta_2)$ is equal to

$$L_{2n}(\theta_2) = -\frac{n}{2}\ln(2\pi\sigma_2^2)+\ln|S_{2n}(\lambda_2)|+\ln|R_{2n}(\rho_2)|-\frac{1}{2\sigma_2^2}[S_{2n}(\lambda_2)y_n-X_{2n}\beta_2]'R_{2n}'(\rho_2)R_{2n}(\rho_2)[S_{2n}(\lambda_2)y_n-X_{2n}\beta_2].$$

By the mean value theorem,

$$\frac{1}{\sqrt{n}}\Big(\frac{\partial^2 L_{2n}(\tilde{\theta}_{2n})}{\partial\lambda_2^2} - \frac{\partial^2 L_{2n}(\bar{\theta}_{2n,1})}{\partial\lambda_2^2}\Big) = B_{1n} + \frac{2}{\dot{\sigma}_{2n}^2}B_{2n}\sqrt{n}(\tilde{\rho}_{2n}-\dot{\rho}_{2n}) + B_{3n},$$

where $B_{1n} = -2n^{-1}\,\mathrm{tr}\big[\big(W_{2n}S_{2n}^{-1}(\dot{\lambda}_{2n})\big)^3\big]n^{1/2}(\tilde{\lambda}_{2n}-\hat{\lambda}_{2n})$, $B_{2n} = n^{-1}y_n'W_{2n}'M_{2n}'R_{2n}(\dot{\rho}_{2n})W_{2n}y_n$ and $B_{3n} = (2n\dot{\sigma}_{2n}^4)^{-1}y_n'W_{2n}'R_{2n}'(\dot{\rho}_{2n})R_{2n}(\dot{\rho}_{2n})W_{2n}y_n n^{1/2}(\tilde{\sigma}_{2n}^2-\hat{\sigma}_{2n}^2)$ with $\dot{\theta}_{2n}$ being between $\tilde{\theta}_{2n}$ and $\bar{\theta}_{2n,1}$. By the uniform boundedness of $S_{2n}^{-1}(\lambda_2)$ in the parameter space, $B_{1n} = O_P(1)$. Note that $B_{2n} = B_{2n,1} + B_{2n,2}(\bar{\rho}_{2n,1} - \dot{\rho}_{2n})$, where $B_{2n,1} = n^{-1}y_n'W_{2n}'M_{2n}'R_{2n}W_{2n}y_n = O_P(1)$ and $B_{2n,2} = n^{-1}y_n'W_{2n}'M_{2n}'M_{2n}W_{2n}y_n = O_P(1)$, then $2\dot{\sigma}_{2n}^{-2}B_{2n}n^{1/2}(\tilde{\rho}_{2n}-\dot{\rho}_{2n}) = O_P(1)$. Similarly, $B_{3n} = O_P(1)$. Hence (i) holds. (ii) follows from Chebychev's inequality. □

Lemma 17. *For* $\eta > 0$, $\mathrm{P}^*(\|\ddot{\theta}_{2n}^* - \ddot{\theta}_{2n}\| > \eta) = o_P(1)$*.*

Proof. Let $\bar{L}_{2n}(\phi_2;\theta_{10}) = \max_{\beta_2,\sigma_2^2}\bar{L}_{2n}(\theta_2;\theta_{10})$ and $\bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a}) = \max_{\beta_2,\sigma_2^2}\mathrm{E}^*\,L_{2n}^*(\theta_2)$, where $\ddot{\theta}_{1n,a} = (\ddot{\lambda}_{1n},\ddot{\rho}_{1n},\ddot{\beta}_{1n}',\mathrm{E}^*\,\epsilon_{1n,i}^{*2})'$, then

$$\bar{L}_{2n}(\phi_2;\theta_{10}) = -\frac{n}{2}[\ln(2\pi)+1] - \frac{n}{2}\ln\bar{\sigma}_{2n}^2(\phi_2) + \ln|S_{2n}(\lambda_2)| + \ln|R_{2n}(\rho_2)|,$$
$$\bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a}) = -\frac{n}{2}[\ln(2\pi)+1] - \frac{n}{2}\ln\bar{\sigma}_{2n}^{*2}(\phi_2) + \ln|S_{2n}(\lambda_2)| + \ln|R_{2n}(\rho_2)|,$$

where

$$\begin{aligned}\bar{\sigma}_{2n}^2(\phi_2) &= \frac{1}{n}\sigma_{10}^2\,\mathrm{tr}\big(R_{1n}'^{-1}S_{1n}'^{-1}S_{2n}'(\lambda_2)R_{2n}'(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}R_{1n}^{-1}\big)\\ &\quad + \frac{1}{n}(X_{1n}\beta_{10})'S_{1n}'^{-1}S_{2n}'(\lambda_2)R_{2n}'(\rho_2)H_{2n}(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}X_{1n}\beta_{10},\\ \bar{\sigma}_{2n}^{*2}(\phi_2) &= \frac{1}{n}(\mathrm{E}^*\,\epsilon_{1n,i}^{*2})\,\mathrm{tr}\big(R_{1n}'^{-1}(\ddot{\rho}_{1n})S_{1n}'^{-1}(\ddot{\rho}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\ddot{\lambda}_{1n})R_{1n}^{-1}(\ddot{\rho}_{1n})\big)\\ &\quad + \frac{1}{n}(X_{1n}\ddot{\beta}_{1n})'S_{1n}'^{-1}(\ddot{\lambda}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)H_{2n}(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\ddot{\lambda}_{1n})X_{1n}\ddot{\beta}_{1n},\end{aligned}$$

with $\bar{\sigma}_{2n}^2(\phi_2)$ being bounded away from zero uniformly in the parameter space φ_2 and $H_{2n}(\rho_2) = I_n - R_{2n}(\rho_2)X_{2n}[X_{2n}'R_{2n}'(\rho_2)R_{2n}(\rho_2)X_{2n}]^{-1}X_{2n}'R_{2n}'(\rho_2)$ being bounded in both row and column sum norms uniformly in ρ_2 (see the proof of Proposition 3 in Jin and Lee (2012)). By the mean value theorem,

$$\frac{1}{n}[\bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a}) - \bar{L}_{2n}(\phi_2;\theta_{10})] = -\frac{1}{2}\frac{\bar{\sigma}_{2n}^{*2}(\phi_2) - \bar{\sigma}_{2n}^2(\phi_2)}{\tilde{\sigma}_{2n}^2},$$

where $\tilde{\sigma}_{2n}^2$ is between $\bar{\sigma}_{2n}^2(\phi_2)$ and $\bar{\sigma}_{2n}^{*2}(\phi_2)$, and

$$
\begin{aligned}
&\bar{\sigma}_{2n}^{*2}(\phi_2) - \bar{\sigma}_{2n}^2(\phi_2) \\
&= \frac{1}{n}(\mathrm{E}^*\, \epsilon_{1n,i}^{*2} - \sigma_0^2)\operatorname{tr}\big(R_{1n}'^{-1}(\dot{\rho}_{1n})S_{1n}'^{-1}(\dot{\rho}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\dot{\lambda}_{1n})R_{1n}^{-1}(\dot{\rho}_{1n})\big) \\
&\quad + \frac{2}{n}(X_{1n}\dot{\beta}_{1n})'S_{1n}'^{-1}(\dot{\lambda}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)H_{2n}(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\dot{\lambda}_{1n})X_{1n}(\ddot{\beta}_{1n} - \beta_{10}) \\
&\quad + \frac{2\dot{\sigma}_{1n}^2}{n}\operatorname{tr}\big(R_{1n}'^{-1}(\dot{\rho}_{1n})S_{1n}'^{-1}(\dot{\rho}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\dot{\lambda}_{1n})R_{1n}^{-1}(\dot{\rho}_{1n})M_{1n}R_{1n}^{-1}(\dot{\rho}_{1n})\big)(\ddot{\rho}_{1n} - \rho_{10}) \\
&\quad + \frac{2\dot{\sigma}_{1n}^2}{n}\operatorname{tr}\big(R_{1n}'^{-1}(\dot{\rho}_{1n})S_{1n}'^{-1}(\dot{\rho}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\dot{\lambda}_{1n})W_{1n}S_{1n}^{-1}(\dot{\lambda}_{1n})R_{1n}^{-1}(\dot{\rho}_{1n})\big)(\ddot{\lambda}_{1n} - \lambda_{10}) \\
&\quad + \frac{2}{n}(X_{1n}\dot{\beta}_{1n})'S_{1n}'^{-1}(\dot{\lambda}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)H_{2n}(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\dot{\lambda}_{1n})W_{1n}S_{1n}^{-1}(\dot{\lambda}_{1n})X_{1n}\dot{\beta}_{1n}(\ddot{\lambda}_{1n} - \lambda_{10}),
\end{aligned}
$$

with $\dot{\gamma}_{1n} = (\dot{\lambda}_{1n}, \dot{\rho}_{1n}, \dot{\beta}_{1n}, \dot{\sigma}_{1n}^2)'$ being between θ_{10} and $\ddot{\theta}_{1n,a}$. By Lemma 6, $\sup_{\phi_2\in\varphi_2}|\bar{\sigma}_{2n}^{*2}(\phi_2) - \bar{\sigma}_{2n}^2(\phi_2)| = o_P(1)$. Then $\sup_{\phi_2\in\varphi_2}|n^{-1}[\bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a}) - \bar{L}_{2n}(\phi_2;\theta_{10})]| = o_P(1)$.

If $||\phi_2 - \ddot{\phi}_{2n}|| > \eta$, $||\phi_2 - \bar{\phi}_{2n,1}|| \geq ||\phi_2 - \ddot{\phi}_{2n}|| - ||\ddot{\phi}_{2n} - \bar{\phi}_{2n,1}|| > \eta/2$ with probability $1 - o(1)$. Note that

$$
\begin{aligned}
\frac{1}{n}\big(\bar{L}_{2n}(\ddot{\phi}_{2n};\ddot{\theta}_{1n,a}) - \bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a})\big) &= \frac{1}{n}\big(\bar{L}_{2n}(\ddot{\phi}_{2n};\ddot{\theta}_{1n,a}) - \bar{L}_{2n}(\ddot{\phi}_{2n};\theta_{10})\big) - \frac{1}{n}\big(\bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a}) - \bar{L}_{2n}(\phi_2;\theta_{10})\big) \\
&\quad + \frac{1}{n}\big(\bar{L}_{2n}(\bar{\phi}_{2n,1};\theta_{10}) - \bar{L}_{2n}(\phi_2;\theta_{10})\big) - \frac{1}{n}\big(\bar{L}_{2n}(\bar{\phi}_{2n,1};\theta_{10}) - \bar{L}_{2n}(\ddot{\phi}_{2n};\theta_{10})\big),
\end{aligned}
$$

given $\eta > 0$, there exists a $\kappa > 0$, such that $||\phi_2 - \ddot{\phi}_{2n}|| > \eta$ implies that $n^{-1}\big(\bar{L}_{2n}(\ddot{\phi}_{2n};\ddot{\theta}_{1n,a}) - \bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a})\big) \geq \kappa$ with probability $1 - o(1)$. Then for $\eta_1 > 0$,

$$
\begin{aligned}
&\mathrm{P}(\mathrm{P}^*(||\ddot{\phi}_{2n}^* - \ddot{\phi}_{2n}|| > \eta) > \eta_1) \\
&\leq \mathrm{P}\big(\mathrm{P}^*\big(n^{-1}\big(\bar{L}_{2n}(\ddot{\phi}_{2n};\ddot{\theta}_{1n,a}) - \bar{L}_{2n}(\ddot{\phi}_{2n}^*;\ddot{\theta}_{1n,a})\big) \geq \kappa\big) > \eta_1\big) + o(1) \\
&\leq \mathrm{P}\big(\mathrm{P}^*\big(n^{-1}\big(\bar{L}_{2n}(\ddot{\phi}_{2n};\ddot{\theta}_{1n,a}) - L_{2n}^*(\ddot{\phi}_{2n}) + L_{2n}^*(\ddot{\phi}_{2n}^*) - \bar{L}_{2n}(\ddot{\phi}_{2n}^*;\ddot{\theta}_{1n,a})\big) \geq \kappa\big) > \eta_1\big) + o(1) \\
&\leq \mathrm{P}\big(\mathrm{P}^*\big(2n^{-1}\sup_{\phi_2\in\varphi_2}\big|L_{2n}^*(\phi_2) - \bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a})\big| \geq \kappa\big) > \eta_1\big) + o(1),
\end{aligned}
$$

where

$$
\frac{1}{n}\big(L_{2n}^*(\phi_2) - \bar{L}_{2n}(\phi_2;\ddot{\theta}_{1n,a})\big) = -\frac{\ddot{\sigma}_{2n}^{*2}(\phi_2) - \bar{\sigma}_{2n}^{*2}(\phi_2)}{2\dot{\sigma}_{2n}^{*2}(\phi_2)},
$$

with $\dot{\sigma}_{2n}^{*2}(\phi_2)$ being between $\ddot{\sigma}_{2n}^{*2}(\phi_2)$ and $\bar{\sigma}_{2n}^{*2}(\phi_2)$, and

$$
\begin{aligned}
\ddot{\sigma}_{2n}^{*2}(\phi_2) - \bar{\sigma}_{2n}^{*2}(\phi_2) &= \frac{1}{n}\epsilon_{1n}^{*\prime}R_{1n}'^{-1}(\ddot{\rho}_{1n})S_{1n}'^{-1}(\ddot{\lambda}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)H_{2n}(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\ddot{\lambda}_{1n})R_{1n}^{-1}(\ddot{\rho}_{1n})\epsilon_{1n}^* \\
&\quad - \frac{\mathrm{E}^*\,\epsilon_{1n,i}^{*2}}{n}\operatorname{tr}[R_{1n}'^{-1}(\ddot{\rho}_{1n})S_{1n}'^{-1}(\ddot{\lambda}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\ddot{\lambda}_{1n})R_{1n}^{-1}(\ddot{\rho}_{1n})] \\
&\quad + \frac{2}{n}(X_{1n}\ddot{\beta}_{1n})'S_{1n}'^{-1}(\ddot{\lambda}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)H_{2n}(\rho_2)R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\ddot{\lambda}_{1n})R_{1n}^{-1}(\ddot{\rho}_{1n})\epsilon_{1n}^*.
\end{aligned}
$$

The $\ddot{\sigma}_{2n}^{*2}(\phi_2) - \bar{\sigma}_{2n}^{*2}(\phi_2)$ is equal to a LQ form plus $n^{-1}(\mathrm{E}^*\,\epsilon_{1n,i}^{*2})\operatorname{tr}\big(R_{1n}'^{-1}(\ddot{\rho}_{1n})S_{1n}'^{-1}(\ddot{\lambda}_{1n})S_{2n}'(\lambda_2)R_{2n}'(\rho_2)[H_{2n}(\rho_2) - I_n]R_{2n}(\rho_2)S_{2n}(\lambda_2)S_{1n}^{-1}(\ddot{\lambda}_{1n})R_{1n}^{-1}(\ddot{\rho}_{1n})\big)$. Since $R_{2n}(\rho_2)$ is linear in ρ_2, $S_{2n}(\lambda_2)$ is linear in λ_2 and $H_{2n}(\rho_2)$ is bounded in both row and column sum norms uniformly in the parameter space of ρ_2, Chebyshev's inequality implies that $n\,\mathrm{P}^*(\sup_{\phi_2\in\varphi_2}|\ddot{\sigma}^{*2}_{2n}(\phi_2)-\bar{\sigma}^{*2}_{2n}(\phi_2)|>\eta)$ for $\eta>0$ is bounded by a term depending only on $\ddot{\theta}_{1n}$, $\mathrm{E}\,\epsilon^{*2}_{1n,i}$, $\mathrm{E}\,\epsilon^{*3}_{1n,i}$ and $\mathrm{E}\,\epsilon^{*4}_{1n,i}$, which has the order $O_P(1)$ by Lemma 6. Then $\mathrm{P}^*(\sup_{\phi_2\in\varphi_2}|\ddot{\sigma}^{*2}_{2n}(\phi_2)-\bar{\sigma}^{*2}_{2n}(\phi_2)|>\eta)=o_P(1)$. It has been shown that $\sup_{\phi_2\in\varphi_2}|\bar{\sigma}^{*2}_{2n}(\phi_2)-\bar{\sigma}^{2}_{2n}(\phi_2)|=o_P(1)$ with $\bar{\sigma}^{2}_{2n}(\phi_2)$ being bounded away from zero uniformly in ϕ_2, then $\mathrm{P}^*(||\ddot{\phi}^*_{2n}-\ddot{\phi}_{2n}||>\eta)=o_P(1)$. Now the mean value theorem and the formulas of $\ddot{\beta}^*_{2n}$ and $\ddot{\sigma}^{*2}_{2n}$ as functions of $\ddot{\phi}^*_{2n}$ can be used to show that we also have $\mathrm{P}^*(||\ddot{\beta}^*_{2n}-\ddot{\beta}_{2n}||>\eta)=o_P(1)$ and $\mathrm{P}^*(||\ddot{\sigma}^{*2}_{2n}-\ddot{\sigma}^{2}_{2n}||>\eta)=o_P(1)$. □

Lemma 18. *For $\eta>0$, $\mathrm{P}^*\big(n^{-1}\big\|\frac{\partial^2 L^*_{2n}(\tilde{\theta}^*_{2n})}{\partial\theta_2\partial\theta_2'}-\mathrm{E}^*\,\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2\partial\theta_2'}\big\|>\eta\big)=o_P(1)$, where $\tilde{\theta}^*_{2n}$ is between $\ddot{\theta}^*_{2n}$ and $\ddot{\theta}_{2n}$.*

Proof. The result is proved by showing that (i) $\mathrm{P}^*\big(n^{-1}\big\|\frac{\partial^2 L^*_{2n}(\tilde{\theta}^*_{2n})}{\partial\theta_2\partial\theta_2'}-\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2\partial\theta_2'}\big\|>\eta\big)=o_P(1)$ and (ii) $\mathrm{P}^*\big(n^{-1}\big\|\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2\partial\theta_2'}-\frac{\partial^2\,\mathrm{E}^*\,L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2\partial\theta_2'}\big\|>\eta\big)=o_P(1)$. As in the proof of Lemma 16, use the mean value theorem for each term in the second order derivative to prove (i). Here we only investigate $n^{-1}\big|\frac{\partial^2 L^*_{2n}(\tilde{\theta}_{2n})}{\partial\lambda_2^2}-\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\lambda_2^2}\big|$. Results for other terms are similarly derived. By the mean value theorem,

$$\frac{1}{n}\Big(\frac{\partial^2 L^*_{2n}(\tilde{\theta}^*_{2n})}{\partial\lambda_2^2}-\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\lambda_2^2}\Big)=B^*_{1n}+\frac{2}{\dot{\sigma}^{*2}_{2n}}B^*_{2n}(\tilde{\rho}^*_{2n}-\ddot{\rho}_{2n})+B^*_{3n},$$

where $B^*_{1n}=-2n^{-1}\,\mathrm{tr}\big((W_{2n}S^{-1}_{2n}(\dot{\lambda}^*_{2n}))^3\big)(\tilde{\lambda}^*_{2n}-\ddot{\lambda}_{2n})$, $B^*_{2n}=n^{-1}y^{*\prime}_n W'_{2n}M'_{2n}R_{2n}(\dot{\rho}^*_{2n})W_{2n}y^*_n$ and $B^*_{3n}=(n\dot{\sigma}^{*4}_{2n})^{-1}y^{*\prime}_n W'_{2n}R'_{2n}(\dot{\rho}^*_{2n})R_{2n}(\dot{\rho}^*_{2n})W_{2n}y^*_n(\tilde{\sigma}^{*2}_{2n}-\ddot{\sigma}^2_{2n})$ with $\dot{\theta}^*_{2n}$ being between $\tilde{\theta}^*_{2n}$ and $\ddot{\theta}_{2n}$. By Lemma 17 and the uniform boundedness of $S^{-1}_{2n}(\lambda_2)$, $\mathrm{P}\big(\mathrm{P}^*(|B^*_{1n}|>\eta_1)>\eta_2\big)=O(n^{-1})$ for $\eta_1>0$ and $\eta_2>0$. Let $B^*_{2n,1}=n^{-1}y^{*\prime}_n W'_{2n}M'_{2n}R_{2n}(\ddot{\rho}_{2n})W_{2n}y^*_n$ and $B^*_{2n,2}=n^{-1}y^{*\prime}_n W'_{2n}M'_{2n}M_{2n}W_{2n}y^*_n$. Then $\mathrm{P}^*(|B^*_{2n,1}-\mathrm{E}^*\,B^*_{2n,1}|>\eta)=o_P(1)$ and $\mathrm{P}^*(|B^*_{2n,2}-\mathrm{E}^*\,B^*_{2n,2}|>\eta)=o_P(1)$. Since $B^*_{2n}=B^*_{2n,1}+B^*_{2n,2}(\ddot{\rho}_{2n}-\dot{\rho}^*_{2n})$, $\mathrm{P}^*\big(\big|2\dot{\sigma}^{*-2}_{2n}B^*_{2n}(\tilde{\rho}^*_{2n}-\ddot{\rho}_{2n})\big|>\eta\big)=o_P(1)$. Similarly, $\mathrm{P}^*(|B^*_{3n}|>\eta)=o_P(1)$. Therefore, $\mathrm{P}^*\big(\frac{1}{n}\big|\frac{\partial^2 L^*_{2n}(\tilde{\theta}^*_{2n})}{\partial\lambda_2^2}-\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\lambda_2^2}\big|>\eta\big)=o_P(1)$. (ii) is proved by using Chebyshev's inequality and Lemma 6. □

Lemma 19. $n^{-1}\big\|\mathrm{E}^*\,\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2\partial\theta_2'}-\mathrm{E}\,\frac{\partial^2 L_{2n}(\bar{\theta}_{2n,1})}{\partial\theta_2\partial\theta_2'}\big\|=o_P(1)$.

Proof. The lemma is proved by using the mean value theorem and Lemma 6. □

Lemma 20. *For $\eta>0$ and $0\le a<\frac{1}{2}$, $\mathrm{P}^*(n^a||\ddot{\theta}^*_{2n}-\ddot{\theta}_{2n}||>\eta)=o_P(1)$.*

Proof. By the mean value theorem,

$$n^a(\ddot{\theta}^*_{2n}-\ddot{\theta}_{2n})=\Big(-\frac{1}{n}\frac{\partial^2 L^*_{2n}(\tilde{\theta}^*_{2n})}{\partial\theta_2\partial\theta_2'}\Big)^{-1}n^{a-1}\frac{\partial L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2},\tag{B.2}$$

where $\tilde{\theta}^*_{2n}$ is between $\ddot{\theta}^*_{2n}$ and $\ddot{\theta}_{2n}$. Then

$$\begin{aligned}\mathrm{P}^*(n^a||\ddot{\theta}^*_{2n}-\ddot{\theta}_{2n}||>\eta)\le{}&\mathrm{P}^*\Big(\Big\|\frac{1}{n}\frac{\partial^2 L^*_{2n}(\tilde{\theta}^*_{2n})}{\partial\theta_2\partial\theta_2'}-\frac{1}{n}\mathrm{E}^*\,\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2\partial\theta_2'}\Big\|>\eta\Big)\\&+\mathrm{P}^*\Big(n^a||\ddot{\theta}^*_{2n}-\ddot{\theta}_{2n}||>\eta,\Big\|\frac{1}{n}\frac{\partial^2 L^*_{2n}(\tilde{\theta}^*_{2n})}{\partial\theta_2\partial\theta_2'}-\frac{1}{n}\mathrm{E}^*\,\frac{\partial^2 L^*_{2n}(\ddot{\theta}_{2n})}{\partial\theta_2\partial\theta_2'}\Big\|\le\eta\Big).\end{aligned}$$

Using (B.2), the result follows from Lemmas 6, 18–19 and Chebyshev's inequality. □

Lemma 21. *$n^{1/2}(\check{\psi}_n-\psi_0) \xrightarrow{d} N(0,\Sigma)$ and $n^{1/2}(\hat{\psi}_n-\psi_0) \xrightarrow{d} N(0, \lim_{n\to\infty}[\mathrm{E}\, G_n'(\psi_0;\gamma_2)\Omega_n^{-1}\,\mathrm{E}\, G_n(\psi_0;\gamma_2)]^{-1})$, where $\Sigma = \lim_{n\to\infty}[\mathrm{E}\, G_n'(\psi_0;\gamma_2)\,\mathrm{E}\, G_n(\psi_0;\gamma_2)]^{-1}\,\mathrm{E}\, G_n'(\psi_0;\gamma_2)\Omega_n\,\mathrm{E}\, G_n(\psi_0;\gamma_2)[\mathrm{E}\, G_n'(\psi_0;\gamma_2)\,\mathrm{E}\, G_n(\psi_0;\gamma_2)]^{-1}$.*

Proof. As $y_n = S_{1n}^{-1}(X_{1n}\beta_{10}+R_{1n}^{-1}\epsilon_{1n})$, $\mathrm{E}\, g_n(\psi;\bar{\gamma}_{2n,1}) = n^{-1}(\Gamma_{1n}'(\psi)D_{1n}\Gamma_{1n}(\psi)+\sigma_{10}^2\,\mathrm{tr}[\Gamma_{2n}'(\psi)D_{1n}\Gamma_{2n}(\psi)], \ldots, \Gamma_{1n}'(\psi)D_{mn}\Gamma_{1n}(\psi)+\sigma_{10}^2\,\mathrm{tr}[\Gamma_{2n}'(\psi)D_{mn}\Gamma_{2n}(\psi)], \Gamma_{1n}'(\psi)\Delta_n)'$ with $\Gamma_{1n}(\psi) = R_{1n}(\rho_1)[W_{1n}S_{1n}^{-1}X_{1n}\beta_{10}(\lambda_{10}-\lambda_1)+X_{1n}(\beta_1-\beta_{10})-(\bar{\lambda}_{2n,1}W_{2n}S_{1n}^{-1}X_{1n}\beta_{10}+X_{2n}\bar{\beta}_{2n,1})(\alpha-\alpha_0)]$ and $\Gamma_{2n}(\psi) = R_{1n}(\rho_1)[S_{1n}(\lambda_1)-\alpha\bar{\lambda}_{2n,1}W_{2n}]S_{1n}^{-1}R_{1n}^{-1}$. Since $\epsilon_n(\psi;\gamma_2)$ is linear in ψ, it is straightforward to verify that $||g_n(\psi;\ddot{\gamma}_{2n}) - \mathrm{E}\, g_n(\psi;\bar{\gamma}_{2n,1})||$ converges to zero uniformly in the parameter space of ψ by the mean value theorem. The $\mathrm{E}\, g_n(\psi;\bar{\gamma}_{2n,1})$ has a similar form to that of $\mathrm{E}\, g_n(\lambda_1,\rho_1,\beta_1,0;\gamma_2)$, i.e., the expected value of the moment conditions for the SARAR model (18), then Assumption J19 ensures that $\mathrm{E}\, g_n(\psi;\bar{\gamma}_{2n,1}) = 0$ has a unique solution at ψ_0, according to Lee and Liu (2010). As $\mathrm{E}\, g_n(\psi;\bar{\gamma}_{2n,1})$ is quadratic in ψ, it is equicontinuous in ψ. It follows from the uniform convergence and identification uniqueness condition that $\check{\psi}_n - \psi_0 = o_P(1)$ (White, 1994). For the distribution of $\check{\psi}_n$, applying the the mean value theorem to the first order condition yields

$$G_n'(\check{\psi}_n;\ddot{\gamma}_{2n})g_n(\check{\psi}_n;\ddot{\gamma}_{2n}) = 0 = G_n'(\check{\psi}_n;\ddot{\gamma}_{2n})[g_n(\psi_0;\ddot{\gamma}_{2n}) + G_n(\tilde{\psi}_n;\ddot{\gamma}_{2n})(\check{\psi}_n-\psi_0)],$$

where $\tilde{\psi}_n$ is between ψ_0 and $\check{\psi}_n$. Then

$$\sqrt{n}(\check{\psi}_n-\psi_0) = -[G_n'(\check{\psi}_n;\ddot{\gamma}_{2n})G_n(\tilde{\psi}_n;\ddot{\gamma}_{2n})]^{-1}G_n(\check{\psi}_n;\ddot{\gamma}_{2n})\sqrt{n}g_n(\psi_0;\ddot{\gamma}_{2n}).$$

As $\tilde{\psi}_n - \psi_0 = O_P(n^{-1/2})$, by writing $\tilde{\psi}_n$ in $G_n(\tilde{\psi}_n;\ddot{\gamma}_{2n})$ as $(\tilde{\psi}_n-\psi_0)+\psi_0$ and expanding $G_n(\tilde{\psi}_n;\ddot{\gamma}_{2n})$, we have $G_n(\tilde{\psi}_n;\ddot{\gamma}_{2n}) = G_n(\psi_0;\ddot{\gamma}_{2n}) + O_P(n^{-1/2})$, where $G_n(\psi_0;\ddot{\gamma}_{2n}) = G_n(\psi_0;\gamma_2) = \mathrm{E}\, G_n(\psi_0;\gamma_2) + o_P(1)$ by Chebyshev's inequality. Noting that elements of $\sqrt{n}g_n(\psi_0;\ddot{\gamma}_{2n})$ are either linear or quadratic in ϵ_{1n}, the distribution of $\check{\gamma}_n$ follows from Lemma 4.

By Lemmas 2 and 6, $\hat{\Omega}_n - \Omega_n = O_P(n^{-1/2})$. Following the argument for Proposition 1 in Lee (2007), we have $\hat{\psi}_n - \psi_0 = o_P(1)$. The distribution of $\hat{\psi}_n$ again follows from the expansion of the first order condition.
□

Lemma 22. *For $\eta > 0$ and $0 \le a < 1/2$, $\mathrm{P}^*(n^a||\check{\psi}_n^*-\hat{\psi}_n|| > \eta) = o_P(1)$ and $\mathrm{P}^*(n^a||\hat{\psi}_n^*-\hat{\psi}_n|| > \eta) = o_P(1)$.*

Proof. By an argument similar to that for Lemma 13, we have $\mathrm{P}^*(||\check{\psi}_n^*-\hat{\psi}_n|| > \eta) = o_P(1)$ and $\mathrm{P}^*(||\hat{\psi}_n^*-\hat{\psi}_n|| > \eta) = o_P(1)$. Then by an argument similar to that for Lemma 20, we have $\mathrm{P}^*(n^a||\check{\psi}_n^*-\hat{\psi}_n|| > \eta) = o_P(1)$ and $\mathrm{P}^*(n^a||\hat{\psi}_n^*-\hat{\psi}_n|| > \eta) = o_P(1)$. □

Appendix C. Proofs

Proof of Theorem 1. As in Kelejian and Prucha (2001), write c_n as $c_n = \sum_{i=1}^n c_{ni}$ with

$$c_{ni} = n^{-1/2}\Big(a_{n,ii}(\epsilon_{ni}^2-\sigma_0^2) + 2\epsilon_{ni}\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{nj} + b_{ni}\epsilon_{ni}\Big).$$

Obviously, $\mathrm{E}\,|c_{ni}| < \infty$. Consider the σ-fields $\mathscr{F}_{n0} = \{\varnothing, \Omega\}$, $\mathscr{F}_{ni} = \sigma(\epsilon_{n1}, \ldots, \epsilon_{ni})$, $1 \le i \le n$, where Ω is the sample space. Then $\{c_{ni}, \mathscr{F}_{ni}, 1 \le i \le n, n \ge 1\}$ forms a martingale difference array and $\sigma_{c_n}^2 = \sum_{i=1}^n \mathrm{E}(c_{ni}^2)$, where

$$\mathrm{E}(c_{ni}^2) = n^{-1}\Big(a_{n,ii}^2(\mu_4 - \sigma_0^4) + 4\sigma_0^4 \sum_{j=1}^{i-1} a_{n,ij}^2 + b_{ni}^2\sigma_0^2 + 2\mu_3 a_{n,ii} b_{ni}\Big).$$

By a theorem in Heyde and Brown (1970), if there is a constant δ with $0 < \delta \le 1$ such that

$$\mathrm{E}\,|c_{ni}|^{2+2\delta} < \infty, \tag{C.1}$$

then there exists a finite constant K depending only on δ, such that[12]

$$\sup_x |\,\mathrm{P}(c_n \le \sigma_{c_n} x) - \Phi(x)| \le K\Big\{\sigma_{c_n}^{-2-2\delta}\Big(\sum_{i=1}^n \mathrm{E}\,|c_{ni}|^{2+2\delta} + \mathrm{E}\Big|\Big(\sum_{i=1}^n \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\Big) - \sigma_{c_n}^2\Big|^{1+\delta}\Big)\Big\}^{1/(3+2\delta)}. \tag{C.2}$$

Thus if

$$\lim_{n\to\infty} \sigma_{c_n}^{-2-2\delta} \sum_{i=1}^n \mathrm{E}\,|c_{ni}|^{2+2\delta} = 0, \tag{C.3}$$

$$\lim_{n\to\infty} \mathrm{E}\Big|\Big(\sigma_{c_n}^{-2} \sum_{i=1}^n \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\Big) - 1\Big|^{1+\delta} = 0, \tag{C.4}$$

$\mathrm{P}(c_n \le \sigma_{c_n} x)$ converges uniformly to $\Phi(x)$ and a bound on the rate of convergence is given by (C.2). Now we check that (C.1), (C.3) and (C.4) hold. Let $q = 2 + 2\delta$ for $0 < \delta \le 1$ and $1/p + 1/q = 1$. By the triangle and Hölder's inequalities,

$$\begin{aligned}
\sum_{i=1}^n \mathrm{E}\,|c_{ni}|^q &\le n^{-q/2}\,\mathrm{E} \sum_{i=1}^n \Big(|a_{n,ii}|^{1/p}|a_{n,ii}|^{1/q}|\epsilon_{ni}^2 - \sigma_0^2| + \sum_{j=1}^{i-1} |a_{n,ij}|^{1/p} 2|a_{n,ij}|^{1/q}|\epsilon_{ni}||\epsilon_{nj}| + |b_{ni}||\epsilon_{ni}|\Big)^q \\
&\le n^{-q/2}\,\mathrm{E} \sum_{i=1}^n \Big(\sum_{j=1}^i |a_{n,ij}| + 1\Big)^{q/p} \Big(|a_{n,ii}||\epsilon_{ni}^2 - \sigma_0^2|^q + \sum_{j=1}^{i-1} 2^q |a_{n,ij}||\epsilon_{ni}|^q|\epsilon_{nj}|^q + |b_{ni}|^q|\epsilon_{ni}|^q\Big) \\
&\le n^{(2-q)/2}(K_a + 1)^{q/p}\Big(K_a\,\mathrm{E}\,|\epsilon_{ni}^2 - \sigma_0^2|^q + 2^q K_a(\mathrm{E}\,|\epsilon_{ni}|^q)^2 + K_b\,\mathrm{E}\,|\epsilon_{ni}|^q\Big).
\end{aligned} \tag{C.5}$$

Thus (C.1) holds. As $\sigma_{c_n}^2 = \sum_{i=1}^n \mathrm{E}\, c_{ni}^2$, (C.5) implies that $\sigma_{c_n}^2$ is bounded. Then Eq. (C.3) holds by Assumption 3.

$$\begin{aligned}
&\mathrm{E}\Big|\Big(\sigma_{c_n}^{-2} \sum_{i=1}^n \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\Big) - 1\Big|^{1+\delta} \\
&= \sigma_{c_n}^{-2(1+\delta)}\,\mathrm{E}\Big|\sum_{i=1}^n \Big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}\, c_{ni}^2\Big)\Big|^{1+\delta}
\end{aligned}$$

[12]Note that the result in Heyde and Brown (1970) is on a fixed square integrable martingale difference sequence with $0 < \delta \le 1$, but the result also applies to a triangular array of martingale differences with $\delta > 1$ (Haeusler, 1988).

$$
\begin{aligned}
&\le \sigma_{c_n}^{-2(1+\delta)} \Big(\mathrm{E}\Big|\sum_{i=1}^{n}\big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}\, c_{ni}^2\big)\Big|^2\Big)^{(1+\delta)/2} \\
&= 4^{1+\delta} n^{-1-\delta} \sigma_{c_n}^{-2(1+\delta)} \bigg(\mathrm{E}\Big(\sigma_0^2 \sum_{i=1}^{n}\sum_{j=1}^{i-1} a_{n,ij}^2(\epsilon_{nj}^2 - \sigma_0^2) + 2\sigma_0^2 \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} a_{n,ij}a_{n,ik}\epsilon_{nj}\epsilon_{nk} \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{i-1}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}\epsilon_{nj}\Big)^2\bigg)^{(1+\delta)/2} \\
&= 4^{1+\delta} n^{-1-\delta} \sigma_{c_n}^{-2(1+\delta)} \bigg(\sigma_0^4(\mu_4 - \sigma_0^4)\sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^{n} a_{n,ij}^2\Big)^2 + 4\sigma_0^8 \sum_{j=1}^{n-1}\sum_{k=1}^{j-1}\Big(\sum_{i=j+1}^{n} a_{n,ij}a_{n,ik}\Big)^2 \\
&\quad + \sigma_0^2 \sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^{n}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}\Big)^2 + 2\mu_3\sigma_0^2 \sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^{n} a_{n,ij}^2\Big)\Big(\sum_{i=j+1}^{n}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}\Big)\bigg)^{(1+\delta)/2} \\
&\le 4^{1+\delta} n^{-1-\delta} \sigma_{c_n}^{-2(1+\delta)} \bigg(n\sigma_0^4 K_a^4(\mu_4 - \sigma_0^4) + 4\sigma_0^8 K_a^2 \sum_{j=1}^{n-1}\sum_{k=1}^{j-1}\sum_{i=j+1}^{n} |a_{n,ij}||a_{n,ik}| \\
&\quad + \sigma_0^2 \sum_{j=1}^{n-1}\sum_{i=j+1}^{n}(|\mu_3^2 a_{n,ii}a_{n,ij}| + \sigma_0^4 b_{ni}^2|a_{n,ij}|)\sum_{i=j+1}^{n}(|a_{n,ii}a_{n,ij}| + |a_{n,ij}|) \\
&\quad + 2|\mu_3|\sigma_0^2 K_a^2 \sum_{j=1}^{n-1}\sum_{i=j+1}^{n}|(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}|\bigg)^{(1+\delta)/2} \\
&\le 4^{1+\delta} n^{-\delta} \sigma_{c_n}^{-2(1+\delta)} \big(\sigma_0^4 K_a^4(\mu_4 - \sigma_0^4) + 4\sigma_0^8 K_a^4 + \sigma_0^2 K_a^2(\mu_3^2 K_a + \sigma_0^4 K_b)(K_a + 1) \\
&\quad + 2|\mu_3|\sigma_0^2 K_a^3(|\mu_3|K_a + \sigma_0^2 K_b)\big)^{(1+\delta)/2}.
\end{aligned} \tag{C.6}
$$

Thus Eq. (C.4) holds. Using (C.2), (C.5) and (C.6), we have

$$
\begin{aligned}
&\sup_x |\mathrm{P}(c_n \le \sigma_{c_n} x) - \Phi(x)| \\
&\le K\sigma_{c_n}^{-2(1+\delta)/(3+2\delta)} n^{-\delta/(3+2\delta)} \Big((K_a+1)^{1+2\delta}\big(K_a\,\mathrm{E}\,|\epsilon_{ni}^2 - \sigma_0^2|^{2+2\delta} + 2^{2+2\delta}K_a(\mathrm{E}\,|\epsilon_{ni}|^{2+2\delta})^2 + K_b\,\mathrm{E}\,|\epsilon_{ni}|^{2+2\delta}\big) \\
&\quad + 4^{1+\delta}\big(\sigma_0^4 K_a^4(\mu_4 - \sigma_0^4) + 4\sigma_0^8 K_a^4 + \sigma_0^2 K_a^2(\mu_3^2 K_a + \sigma_0^4 K_b)(K_a + 1) \\
&\quad + 2|\mu_3|\sigma_0^2 K_a^3(|\mu_3|K_a + \sigma_0^2 K_b)\big)^{(1+\delta)/2}\Big)^{1/(3+2\delta)} = r_n,
\end{aligned}
$$

i.e., (6) holds. Similarly, (7) holds. Since

$$
\begin{aligned}
\mathrm{P}\big(c_n/\sigma_{c_n} + d_n \le x\big) - \Phi(x) &\le \mathrm{P}(c_n/\sigma_{c_n} + d_n \le x, |d_n| \le \tau_n) - \Phi(x) + \mathrm{P}(|d_n| > \tau_n) \\
&\le [\mathrm{P}(c_n/\sigma_{c_n} \le x + \tau_n) - \Phi(x + \tau_n)] + [\Phi(x + \tau_n) - \Phi(x)] + \mathrm{P}(|d_n| > \tau_n),
\end{aligned}
$$

and similarly

$$
\mathrm{P}\,\big(c_n/\sigma_{c_n} + d_n \le x\big) - \Phi(x) \ge [\mathrm{P}(c_n/\sigma_{c_n} \le x - \tau_n) - \Phi(x - \tau_n)] - [\Phi(x) - \Phi(x - \tau_n)] - \mathrm{P}(|d_n| > \tau_n),
$$

we have

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}}\big|\mathrm{P}\big(c_n/\sigma_{c_n}+d_n\le x\big)-\Phi(x)\big|\\
&\le \max\Big\{\sup_{x\in\mathbb{R}}\big|[\mathrm{P}(c_n/\sigma_{c_n}\le x+\tau_n)-\Phi(x+\tau_n)]\big|+\sup_{x\in\mathbb{R}}[\Phi(x+\tau_n)-\Phi(x)]+\mathrm{P}(|d_n|>\tau_n),\\
&\quad \sup_{x\in\mathbb{R}}\big|[\mathrm{P}(c_n/\sigma_{c_n}\le x-\tau_n)-\Phi(x-\tau_n)]\big|+\sup_{x\in\mathbb{R}}[\Phi(x)-\Phi(x-\tau_n)]+\mathrm{P}(|d_n|>\tau_n)\Big\}\\
&\le r_n+(2\pi)^{-1/2}\tau_n+\mathrm{P}(|d_n|>\tau_n).
\end{aligned}
\tag{C.7}
$$

Similarly,

$$
\sup_{x\in\mathbb{R}}\big|\mathrm{P}^*\big(c_n^*/\sigma_{c_n}^*+d_n^*\le x\big)-\Phi(x)\big|\le r_n^*+(2\pi)^{-1/2}\tau_n+\mathrm{P}^*(|d_n^*|>\tau_n). \tag{C.8}
$$

Thus,

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}}\big|\mathrm{P}\big(c_n/\sigma_{c_n}+d_n\le x\big)-\mathrm{P}^*\big(c_n^*/\sigma_{c_n}^*+d_n^*\le x\big)\big|\\
&\le \sup_{x\in\mathbb{R}}\big|\mathrm{P}\big(c_n/\sigma_{c_n}+d_n\le x\big)-\Phi(x)\big|+\sup_{x\in\mathbb{R}}\big|\mathrm{P}^*\big(c_n^*/\sigma_{c_n}^*+d_n^*\le x\big)-\Phi(x)\big|\\
&= r_n+\mathrm{P}(|d_n|>\tau_n)+r_n^*+\mathrm{P}^*(|d_n^*|>\tau_n)+2^{1/2}\pi^{-1/2}\tau_n,
\end{aligned}
$$

i.e., (8) holds. Since

$$
\begin{aligned}
&\mathrm{P}^*\big((c_n^*/\sigma_{c_n}^*+d_n^*)e_n^*\le x\big)-\mathrm{P}\big((c_n/\sigma_{c_n}+d_n)e_n\le x\big)\\
&=\big(\mathrm{P}^*(c_n^*/\sigma_{c_n}^*+d_n^*\le x/e_n^*)-\Phi(x/e_n^*)\big)-\big(\mathrm{P}(c_n/\sigma_{c_n}+d_n\le x/e_n)-\Phi(x/e_n)\big)+\big(\Phi(x/e_n^*)-\Phi(x/e_n)\big),
\end{aligned}
$$

(9) holds by (C.7) and (C.8). □

Proof of Proposition 1. We use (9) in Theorem 1 to prove the result. From (10)–(13),

$$
d_n=\frac{\sigma_0^2\epsilon_n'H_n\epsilon_n\operatorname{tr}(M_nH_n)-\epsilon_n'H_nM_nH_n\epsilon_n[\epsilon_n'H_n\epsilon_n-(n-k_x)\sigma_0^2]}{\sigma_0^2\epsilon_n'H_n\epsilon_n\sqrt{H_nM_nH_n(M_n+M_n')}}.
$$

By Lemmas 1 and 3, $\operatorname{tr}(M_nH_n)=\operatorname{tr}(M_n)+O(1)=O(1)$, $n^{-1/2}[\epsilon_n'H_nM_nH_n\epsilon_n-\sigma_0^2\operatorname{tr}(M_nH_n)]=O_P(1)$, $n^{-1/2}[\epsilon_n'H_n\epsilon_n-(n-k_x)\sigma_0^2]=O_P(1)$ and $n^{-1}\operatorname{tr}[H_nM_nH_n(M_n+M_n')]=n^{-1}\operatorname{tr}(M_n^2+M_n'M_n)+o(1)=O(1)$ is bounded away from zero by Assumption I3, then $d_n=O_P(n^{-1/2})$. Let $\tau_n=n^{-1/4}$. As $e_n=e_n^*$, $r_n=O(n^{-\delta/(3+2\delta)})$ and $r_n^*=O_P(n^{-\delta/(3+2\delta)})$ by Lemma 6 if we let $\delta=1/2$, it remains to show that $\mathrm{P}^*(|d_n^*|>\tau_n)=o_P(1)$.

$$
\begin{aligned}
&\mathrm{P}^*(|d_n^*|>\tau_n)\\
&\le \mathrm{P}^*\Big(|d_n^*|>\tau_n,\frac{1}{n}|\epsilon_n^{*\prime}H_n\epsilon_n^*-(n-k_x)\sigma_n^{*2}|\le\kappa_n\Big)+\mathrm{P}^*\Big(\frac{1}{n}|\epsilon_n^{*\prime}H_n\epsilon_n^*-(n-k_x)\sigma_n^{*2}|>\kappa_n\Big)\\
&\le \mathrm{P}^*\Big(\frac{|\epsilon_n^{*\prime}H_nM_nH_n\epsilon_n^*|\kappa_n}{\sigma_n^{*2}\big(\frac{n-k_x}{n}\sigma_n^{*2}-\kappa_n\big)\sqrt{\operatorname{tr}\big(H_n(M_n+M_n')H_nM_n\big)}}\ge\tau_n-\frac{|\operatorname{tr}(M_nH_n)|}{\sqrt{\operatorname{tr}\big(H_n(M_n+M_n')H_nM_n\big)}}\Big)\\
&\quad+\mathrm{P}^*\Big(\frac{1}{n}|\epsilon_n^{*\prime}H_n\epsilon_n^*-(n-k_x)\sigma_n^{*2}|>\kappa_n\Big)
\end{aligned}
$$

$$\le \frac{\kappa_n^2 \,\mathrm{E}^* |\epsilon_n^{*'} H_n M_n H_n \epsilon_n^*|^2}{\sigma_n^{*4}\left(\frac{n-k_x}{n}\sigma_n^{*2}-\kappa_n\right)^2 \mathrm{tr}\big(M_n H_n (M_n+M_n')H_n\big)\Big(\tau_n - |\,\mathrm{tr}(M_n H_n)|\,\mathrm{tr}^{-1/2}\big(H_n(M_n+M_n')H_n M_n\big)\Big)^2}$$
$$+ \frac{(\mathrm{E}^*\,\epsilon_{ni}^{*4}-3\sigma_n^{*4})\sum_{i=1}^n (H_n)_{ii}^2 + 2(n-k_x)\sigma_n^{*4}}{n^2\kappa_n^2} \xrightarrow{p} 0, \text{ if } \kappa_n = n^{-3/8},$$

where $\mathrm{E}^* |\epsilon_n^{*'} H_n M_n H_n \epsilon_n^*|^2 = (\mathrm{E}^*\,\epsilon_{ni}^{*4}-3\sigma_n^{*4})\sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + \sigma_n^{*4}\,\mathrm{tr}^2(H_n M_n) + \sigma_n^{*4}\,\mathrm{tr}(H_n M_n H_n (M_n + M_n'))$. Thus the result follows. □

Proof of Proposition 2. From (15)–(17),

$$d_n = \mathbb{I}_n' - c_n/\sigma_{c_n} = \frac{(\sigma_{c_n}-\hat\sigma_{c_n})[\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2\,\mathrm{tr}(M_n H_n)]}{\sqrt{n}\sigma_{c_n}\hat\sigma_{c_n}} + \frac{\sigma_0^2\,\mathrm{tr}(M_n H_n)}{\sqrt{n}\hat\sigma_{c_n}}.$$

By Lemma 6, $n^{1/2}(\hat\sigma_{c_n}^2 - \sigma_{c_n}^2) = O_P(1)$; by Lemma 3, $n^{-1/2}[\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2\,\mathrm{tr}(M_n H_n)] = O_P(1)$; by Lemma 1, $\mathrm{tr}(M_n H_n) = \mathrm{tr}(M_n) + O(1) = O(1)$. Then $d_n = O_P(n^{-1/2})$. Let $\tau_n = \kappa n^{-1/4}$ with $\kappa > 0$. It remains to show that $\mathrm{P}^*(|d_n^*| > \tau_n) = o_P(1)$ by (8) in Theorem 1. The d_n^* is

$$d_n^* = \mathbb{I}_n'^* - c_n^*/\sigma_{c_n}^* = \frac{(\sigma_{c_n}^*-\hat\sigma_{c_n}^*)[\epsilon_n^{*'} H_n M_n H_n \epsilon_n^* - \sigma_n^{*2}\,\mathrm{tr}(M_n H_n)]}{\sqrt{n}\sigma_{c_n}^*\hat\sigma_{c_n}^*} + \frac{\sigma_n^{*2}\,\mathrm{tr}(M_n H_n)}{\sqrt{n}\hat\sigma_{c_n}^*}.$$

By Lemma 8, $\mathrm{P}^*(n^{1/4}|\hat\sigma_{c_n}^{*^2} - \sigma_{c_n}^{*2}| > \eta) = o_P(1)$ for $\eta > 0$. By Chebyshev's inequality, $\mathrm{P}^*(n^{1/4}|n^{-1}\epsilon^{*'} H_n M_n H_n \epsilon_n^* - \sigma_n^{*2}\,\mathrm{tr}(M_n H_n)| > \kappa) \le \kappa^{-2} n^{1/2}\,\mathrm{E}^* |n^{-1}\epsilon^{*'} H_n M_n H_n \epsilon_n^* - \sigma_n^{*2}\,\mathrm{tr}(M_n H_n)|^2 = o_P(1)$. Then $\mathrm{P}^*(|d_n^*| > \tau_n) = o_P(1)$. □

Proof of Proposition 3. We first show the following results: (i) $n^{-1/2}\Delta_n'[R_{1n}(\hat\rho_{1n})Z_{1n} - R_{1n}\,\mathrm{E}\,Z_{1n}] = O_P(1)$, (ii) $n^{-1/2}\Delta_n'[R_{1n}(\hat\rho_{1n})Z_{2n}\hat\gamma_{2n} - R_{1n}\,\mathrm{E}\,Z_{2n}\bar\gamma_{2n,1}] = O_P(1)$, (iii) $\mathrm{P}^*(||n^{a-1}\Delta_n'[R_{1n}(\hat\rho_{1n}^*)Z_{1n}^* - R_{1n}(\hat\rho_{1n})\,\mathrm{E}^*\,Z_{1n}^*]|| > \eta) = o_P(1)$, and (iv) $\mathrm{P}^*(||n^{a-1}\Delta_n'[R_{1n}(\hat\rho_{1n}^*)Z_{2n}^*\hat\gamma_{2n}^* - R_{1n}(\hat\rho_{1n})\,\mathrm{E}^*\,Z_{2n}^*\hat\gamma_{2n}]|| > \eta) = o_P(1)$, for $0 \le a < 1/2$ and $\eta > 0$. (i) holds since $n^{-1/2}\Delta_n'[R_{1n}(\hat\rho_{1n})Z_{1n} - R_{1n}\,\mathrm{E}\,Z_{1n}] = n^{-1/2}\Delta_n' R_{1n}[Z_{1n} - \mathrm{E}\,Z_{1n}] + n^{-1}\Delta_n' M_{1n} Z_{1n} n^{1/2}(\rho_{10} - \hat\rho_{1n}) = O_P(1)$. Similar to (i), $n^{-1/2}\Delta_n'[R_{1n}(\hat\rho_{1n})Z_{2n} - R_{1n}\,\mathrm{E}\,Z_{2n}] = O_P(1)$. Then by Lemma 11, (ii) holds. (iii) holds since $n^{a-1}\Delta_n'[R_{1n}(\hat\rho_{1n}^*)Z_{1n}^* - R_{1n}(\hat\rho_{1n})\,\mathrm{E}^*\,Z_{1n}^*] = n^{a-1}\Delta_n' R_{1n}(\hat\rho_{1n})[Z_{1n}^* - \mathrm{E}^*\,Z_{1n}^*] + n^{a-1}\Delta_n' M_{1n} Z_{1n}^*(\hat\rho_{1n} - \hat\rho_{1n}^*)$, where $\mathrm{P}^*(||n^{a-1}\Delta_n' R_{1n}(\hat\rho_{1n})[Z_{1n}^* - \mathrm{E}^*\,Z_{1n}^*]|| > \eta) = o_P(1)$ and $\mathrm{P}^*(||n^{a-1}\Delta_n' M_{1n}[Z_{1n}^* - \mathrm{E}^*\,Z_{1n}^*]|| > \eta) = o_P(1)$ by first applying Chebyshev's inequality and then the mean value theorem. (iv) holds by an argument similar to that for (iii) and Lemma 14. By (i) and (ii), $n^{-1}\hat\sigma_{\hat\alpha_n}^2 = n^{-1}\bar\sigma_{\hat\alpha_n}^2 + O_P(n^{-1/2})$ and $n^{1/2}\hat\alpha_n = n^{1/2}\bar\alpha_n + O_P(n^{-1/2})$. Thus, $J_{1n} = \bar\alpha_n/\bar\sigma_{\hat\alpha_n} + O_P(n^{-1/2})$. Let $\bar\sigma_{\hat\alpha_n}^{*2} = \hat\sigma_{1n}^2[\hat\gamma_{2n}'\,\mathrm{E}^*\,Z_{2n}^{*'} R_{1n}'(\hat\rho_{1n}) P_{\Delta_n}(I_n - P_{\bar V_n^*(\hat\rho_{1n})}) P_{\Delta_n} R_{1n}(\hat\rho_{1n})\,\mathrm{E}^*\,Z_{2n}^*\hat\gamma_{2n}]^{-1}$ and $\bar\alpha_n^* = \hat\sigma_{1n}^{-2}\bar\sigma_{\hat\alpha_n}^{*2}\hat\gamma_{2n}'\,\mathrm{E}^*\,Z_{2n}^{*'} R_{1n}'(\hat\rho_{1n}) P_{\Delta_n}(I_n - P_{\bar V_n^*(\hat\rho_{1n})})\epsilon_{1n}^*$, where $\bar V_n^*(\hat\rho_{1n}) = P_{\Delta_n} R_{1n}(\hat\rho_{1n})\,\mathrm{E}^*\,Z_{1n}^*$. Then by (iii) and (iv), $\mathrm{P}^*(n^a|J_{1n}^* - n^{1/2}\bar\alpha_n^*/\bar\sigma_{\hat\alpha_n}^*| > \eta) = o_P(1)$ for $0 \le a < 1/2$ and $\eta > 0$. The result now follows from Lemma 6 and Theorem 1. □

Proof of Proposition 4. By the mean value theorem,

$$G_n'(\hat\psi_n;\ddot\gamma_{2n})\hat\Omega_n^{-1} g_n(\hat\psi_n;\ddot\gamma_{2n}) = 0 = G_n'(\hat\psi_n;\ddot\gamma_{2n})\hat\Omega_n^{-1}[g_n(\psi_0;\ddot\gamma_{2n}) + G_n(\tilde\psi_n;\ddot\gamma_{2n})(\hat\psi_n - \psi_0)],$$

where $\tilde{\psi}_n$ is between ψ_0 and $\hat{\psi}_n$. Then

$$\sqrt{n}(\hat{\psi}_n - \psi_0) = -[G_n'(\hat{\psi}_n; \ddot{\gamma}_{2n})\hat{\Omega}_n^{-1} G_n(\tilde{\psi}_n; \ddot{\gamma}_{2n})]^{-1} G_n(\tilde{\psi}_n; \ddot{\gamma}_{2n})\hat{\Omega}_n^{-1}\sqrt{n} g_n(\psi_0; \ddot{\gamma}_{2n}).$$

Let

$$\begin{aligned}
c_n &= -e_\psi' [\mathrm{E}\, G_n'(\psi_0; \gamma_2)\Omega_n^{-1}\, \mathrm{E}\, G_n(\psi_0; \gamma_2)]^{-1}\, \mathrm{E}\, G_n(\psi_0; \gamma_2)\Omega_n^{-1}\sqrt{n} g_n(\psi_0; \gamma_2), \\
\sigma_{c_n}^2 &= e_\psi' [\mathrm{E}\, G_n'(\psi_0; \gamma_2)\Omega_n^{-1}\, \mathrm{E}\, G_n(\psi_0; \gamma_2)]^{-1} e_\psi.
\end{aligned}$$

As in the proof of Lemma 21, $\hat{\Omega}_n - \Omega_n = O_P(n^{-1/2})$ and $G_n(\tilde{\psi}_n; \ddot{\gamma}_{2n}) = \mathrm{E}\, G_n(\psi_0; \gamma_2) + O_P(n^{-1/2})$, then $J_{2n} - c_n/\sigma_{c_n} = O_P(n^{-1/2})$. In Eq. (8), let $\tau_n = \eta_1 n^{-1/4}$ for some $\eta_1 > 0$, then it remains to show that $\mathrm{P}(n^{1/4}|J_{2n}^* - c_n^*/\sigma_{c_n}^*| > \eta_2) = o_P(1)$ for $\eta_2 > 0$ by Theorem 1, where

$$\begin{aligned}
c_n^* &= -e_\psi' [\mathrm{E}^*\, G_n^{*'}(\hat{\psi}_n; \ddot{\gamma}_{2n})\hat{\Omega}_n^{-1}\, \mathrm{E}^*\, G_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n})]^{-1}\, \mathrm{E}^*\, G_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n})\hat{\Omega}_n^{-1}\sqrt{n} g_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n}), \\
\sigma_{c_n}^{*2} &= e_\psi' [\mathrm{E}^*\, G_n^{*'}(\hat{\psi}_n; \ddot{\gamma}_{2n})\hat{\Omega}_n^{-1}\, \mathrm{E}^*\, G_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n})]^{-1} e_\psi.
\end{aligned}$$

As Ω_n is quadratic in the parameters, it can be verified that $\mathrm{P}^*(n^{1/4}||\hat{\Omega}_n^* - \hat{\Omega}_n|| > \eta) = o_P(1)$ for $\eta > 0$. Since $G_n(\psi; \gamma_2)$ is linear in ψ and quadratic in γ_2, $\mathrm{P}^*(n^{1/4}||G_n^*(\hat{\psi}_n^*; \ddot{\gamma}_{2n}^*) - \mathrm{E}^*\, G_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n})|| > \eta) = o_P(1)$ for $\eta > 0$, by showing that $\mathrm{P}^*(n^{1/4}||G_n^*(\hat{\psi}_n^*; \ddot{\gamma}_{2n}^*) - G_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n})|| > \eta) = o_P(1)$ and $\mathrm{P}^*(n^{1/4}||G_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n}) - \mathrm{E}^*\, G_n^*(\hat{\psi}_n; \ddot{\gamma}_{2n})|| > \eta) = o_P(1)$. Thus, $\mathrm{P}(n^{1/4}|J_{2n}^* - c_n^*/\sigma_{c_n}^*| > \eta_2) = o_P(1)$. □

Proof of Theorem 2. The characteristic function of c_n/σ_{c_n} is

$$\begin{aligned}
\varphi_n(t) = \mathrm{E}\exp(itc_n/\sigma_{c_n}) &= \int_{-\infty}^{+\infty} (2\pi\sigma_0^2)^{-n/2} \exp\big(-\frac{1}{2\sigma_0^2}\epsilon_n'\epsilon_n + \frac{it}{\sigma_{c_n}} c_n\big)\, d\epsilon_n \\
&= \int_{-\infty}^{+\infty} (2\pi\sigma_0^2)^{-n/2} \exp\big(-\frac{1}{2}(\epsilon_n - r_n(t))' B_n(t)(\epsilon_n - r_n(t)) + \frac{1}{2} r_n(t)' B_n(t) r_n(t) - \frac{it\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}(A_n)\big)\, d\epsilon_n \\
&= \sigma_0^{-n}|B_n(t)|^{-1/2}\exp\big(-\frac{t^2}{2n\sigma_{c_n}^2} b_n' B_n(t)^{-1} b_n - \frac{it\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}(A_n)\big) \\
&= \exp\big(g_n(t) - \frac{1}{2}t^2\big).
\end{aligned}$$

where $B_n(t) = \frac{I_n}{\sigma_0^2} - \frac{2itA_n}{\sqrt{n}\sigma_{c_n}}$, $r_n(t) = \frac{it}{\sqrt{n}\sigma_{c_n}} B_n(t)^{-1} b_n$ and $g_n(t) = -\frac{n}{2}\ln\sigma_0^2 - \frac{1}{2}\ln|B_n(t)| - \frac{t^2}{2n\sigma_{c_n}^2} b_n' B_n(t)^{-1} b_n - \frac{it\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}(A_n) + \frac{1}{2}t^2$. The derivatives of $g_n(t)$ are

$$\begin{aligned}
g_n^{(1)}(t) &= \frac{i}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}\big(A_n B_n(t)^{-1}\big) - \frac{t}{n\sigma_{c_n}^2} b_n' B_n(t)^{-1} b_n - \frac{it^2}{n^{3/2}\sigma_{c_n}^3} b_n' B_n(t)^{-1} A_n B_n(t)^{-1} b_n - \frac{i\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}(A_n) + t, \\
g_n^{(2)}(t) &= -\frac{2}{n\sigma_{c_n}^2}\operatorname{tr}\big[\big(A_n B_n(t)^{-1}\big)^2\big] - \frac{1}{n\sigma_{c_n}^2} b_n' B_n(t)^{-1} b_n - \frac{4it}{n^{3/2}\sigma_{c_n}^3} b_n' B_n(t)^{-1} A_n B_n(t)^{-1} b_n \\
&\quad + \frac{4t^2}{n^2\sigma_{c_n}^4} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^2 b_n + 1, \\
g_n^{(3)}(t) &= -\frac{8i}{n^{3/2}\sigma_{c_n}^3}\operatorname{tr}\big[\big(A_n B_n(t)^{-1}\big)^3\big] - \frac{6i}{n^{3/2}\sigma_{c_n}^3} b_n' B_n(t)^{-1} A_n B_n(t)^{-1} b_n + \frac{24t}{n^2\sigma_{c_n}^4} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^2 b_n \\
&\quad + \frac{24it^2}{n^{5/2}\sigma_{c_n}^5} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^3 b_n,
\end{aligned}$$

$$
\begin{aligned}
g_n^{(4)}(t) &= \frac{48}{n^2\sigma_{c_n}^4}\operatorname{tr}\big[\big(A_nB_n(t)^{-1}\big)^4\big] + \frac{48}{n^2\sigma_{c_n}^4}b_n'B_n(t)^{-1}\big(A_nB_n(t)^{-1}\big)^2 b_n + \frac{192it}{n^{5/2}\sigma_{c_n}^5}b_n'B_n(t)^{-1}\big(A_nB_n(t)^{-1}\big)^3 b_n \\
&\quad - \frac{192t^2}{n^3\sigma_{c_n}^6}b_n'B_n(t)^{-1}\big(A_nB_n(t)^{-1}\big)^4 b_n, \\
g_n^{(k)}(t) &= \frac{c_{k1}i^k}{n^{k/2}\sigma_{c_n}^k}\operatorname{tr}\big[\big(A_nB_n(t)^{-1}\big)^k\big] + \frac{c_{k2}i^k}{n^{k/2}\sigma_{c_n}^k}b_n'B_n(t)^{-1}\big(A_nB_n(t)^{-1}\big)^{k-2} b_n \\
&\quad + \frac{c_{k3}i^{k+1}t}{n^{(k+1)/2}\sigma_{c_n}^{k+1}}b_n'B_n(t)^{-1}\big(A_nB_n(t)^{-1}\big)^{k-1} b_n + \frac{c_{k4}i^{k+2}t^2}{n^{(k+2)/2}\sigma_{c_n}^{k+2}}b_n'B_n(t)^{-1}\big(A_nB_n(t)^{-1}\big)^{k} b_n, \\
g_n(0) &= g_n^{(1)}(0) = g_n^{(2)}(0) = 0, \\
g_n^{(3)}(0) &= -\frac{8i\sigma_0^6}{n^{3/2}\sigma_{c_n}^3}\operatorname{tr}(A_n^3) - \frac{6i\sigma_0^4}{n^{3/2}\sigma_{c_n}^3}b_n'A_nb_n, \\
g_n^{(k)}(0) &= \frac{c_{k1}\sigma_0^{2k}i^k}{n^{k/2}\sigma_{c_n}^k}\operatorname{tr}(A_n^k) + \frac{c_{k2}\sigma_0^{2(k-1)}i^k}{n^{k/2}\sigma_{c_n}^k}b_n'A_n^{k-2}b_n, \text{ for } k \geq 3,
\end{aligned}
\tag{C.9}
$$

where $c_{k1},\dots,c_{k4}$ are constants. Let $\iota_{n1},\dots,\iota_{nn}$ be A_n's eigenvalues, which are real as A_n is symmetric, and $\iota_n = \max\{|\iota_{n1}|,\dots,|\iota_{nn}|\}$. As $\sigma_{c_n}^2 = n^{-1}[2\sigma_0^4\operatorname{tr}(A_n^2) + \sigma_0^2 b_n'b_n]$, $|b_n'A_nb_n| \leq \iota_n b_n'b_n$, $|\frac{1}{\sigma_0^2} - \frac{2it\iota_{nj}}{\sqrt{n}\sigma_{c_n}}| \geq \frac{1}{\sigma_0^2}$, $|\iota_{nj}|/|\frac{1}{\sigma_0^2} - \frac{2it\iota_{nj}}{\sqrt{n}\sigma_{c_n}}| \leq \iota_n\sigma_0^2$ and $|\iota_{nj}t|/|\frac{1}{\sigma_0^2} - \frac{2it\iota_{nj}}{\sqrt{n}\sigma_{c_n}}| = (\frac{1}{\sigma_0^4\iota_{nj}^2t^2} + \frac{4}{n\sigma_{c_n}^2})^{-1/2} \leq \sqrt{n}\sigma_{c_n}/2$,

$$
\begin{aligned}
|g_n^{(4)}(t)| &\leq \frac{48\iota_n^2\sigma_0^8}{n^2\sigma_{c_n}^4}\sum_{j=1}^n \iota_{nj}^2 + \frac{48\iota_n^2\sigma_0^6}{n^2\sigma_{c_n}^4}b_n'b_n + \frac{96\iota_n^2\sigma_0^6}{n^2\sigma_{c_n}^4}b_n'b_n + \frac{48\iota_n^2\sigma_0^6}{n^2\sigma_{c_n}^4}b_n'b_n \leq \frac{192\iota_n^2\sigma_0^4}{n\sigma_{c_n}^2}, \\
|g_n^{(k)}(t)| &\leq \frac{|c_{k1}|\iota_n^{k-2}\sigma_0^{2k}}{n^{k/2}\sigma_{c_n}^k}\sum_{j=1}^n \iota_{nj}^2 + \frac{|c_{k2}|\iota_n^{k-2}\sigma_0^{2k-2}}{n^{k/2}\sigma_{c_n}^k}b_n'b_n + \frac{|c_{k3}|\iota_n^{k-2}\sigma_0^{2k-2}}{2n^{k/2}\sigma_{c_n}^k}b_n'b_n + \frac{|c_{k4}|\iota_n^{k-2}\sigma_0^{2k-2}}{4n^{k/2}\sigma_{c_n}^k}b_n'b_n \\
&\leq \frac{c_{k5}(\iota_n\sigma_0^2)^{k-2}}{n^{(k-2)/2}\sigma_{c_n}^{k-2}}. \text{ for } k \geq 3.
\end{aligned}
\tag{C.10}
$$

We first establish a one-term Edgeworth expansion for $\mathrm{P}(c_n/\sigma_{c_n} \leq x)$ separately and then consider high order expansions. Let $\gamma_n(t) = (1 - i\kappa_n t^3)\exp(-\frac{1}{2}t^2)$ be the Fourier transform of the function $\Phi^{(1)}(x) - \kappa_n\Phi^{(4)}(x)$, where $\kappa_n = \frac{4\sigma_0^6}{3n^{3/2}\sigma_{c_n}^3}\operatorname{tr}(A_n^3) + \frac{\sigma_0^4}{n^{3/2}\sigma_{c_n}^3}b_n'A_nb_n$. By a smoothing inequality in Feller (1970, p. 538), for all $T > 0$,

$$
\sup_{x\in\mathbb{R}} |\mathrm{P}(c_n/\sigma_{c_n} \leq x) - (\Phi(x) - \kappa_n\Phi^{(3)}(x))| \leq \frac{1}{\pi}\int_{-T}^{T} |\frac{\varphi_n(t) - \gamma_n(t)}{t}|\, dt + \frac{24\sup_x|\Phi^{(1)}(x) - \kappa_n\Phi^{(4)}(x)|}{\pi T},
\tag{C.11}
$$

where

$$
\begin{aligned}
|\varphi_n(t) - \gamma_n(t)| &= \exp(-t^2/2)|\exp(g_n(t)) - (1 - i\kappa_n t^3)| \\
&= \exp(-t^2/2)|\exp(g_n(t)) - \exp(-i\kappa_n t^3) + \exp(-i\kappa_n t^3) - (1 - i\kappa_n t^3)| \\
&\leq \exp(-t^2/2)[|\exp(g_n(t)) - \exp(-i\kappa_n t^3)| + |\exp(-i\kappa_n t^3) - (1 - i\kappa_n t^3)|].
\end{aligned}
$$

As

$$
|\kappa_n| = |\frac{4\sigma_0^6\sum_{i=1}^n \iota_{ni}^3 + 3\sigma_0^4 b_n'A_nb_n}{3n^{3/2}\sigma_{c_n}^3}| \leq \iota_n|\frac{4\sigma_0^6\sum_{i=1}^n \iota_{ni}^2 + 3\sigma_0^4 b_n'b_n}{3n^{3/2}\sigma_{c_n}^3}| \leq \frac{\iota_n\sigma_0^2}{n^{1/2}\sigma_{c_n}},
$$

$|\exp(-i\kappa_n t^3) - (1 - i\kappa_n t^3)| \le |i\kappa_n t^3|^2/2 = \kappa_n^2 t^6/2 \le \iota_n^2\sigma_0^4 t^6/(2n\sigma_{c_n}^2)$ (Feller 1970, p. 512). By a four-term Taylor expansion, $|g_n(t) + i\kappa_n t^3| \le \frac{8\iota_n^2\sigma_0^4 t^4}{n\sigma_{c_n}^2} \le \frac{1}{4}t^2$, when $|t| \le \frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n\sigma_0^2}$. Then $|\exp(g_n(t)) - \exp(-i\kappa_n t^3)| = |\exp(g_n(t) + i\kappa_n t^3) - 1| \le |g_n(t) + i\kappa_n t^3| \exp(|g_n(t) + i\kappa_n t^3|) \le \frac{8\iota_n^2\sigma_0^4 t^4}{n\sigma_{c_n}^2}\exp(\frac{1}{4}t^2)$ and

$$|\frac{\varphi_n(t) - \gamma_n(t)}{t}| \le \frac{\iota_n^2\sigma_0^4}{2n\sigma_{c_n}^2}(16|t|^3\exp(-t^2/4) + |t|^5\exp(-t^2/2)), \tag{C.12}$$

when $|t| \le \frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n\sigma_0^2}$.

When $|t| > \frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n\sigma_0^2}$, noting that $B_n(t)$ has eigenvalues $\frac{1}{\sigma_0^2} - \frac{2it\iota_{nj}}{\sqrt{n}\sigma_{c_n}}$, $j = 1, \dots, n$,

$$\begin{aligned}
|\varphi_n(t)| &\le \exp(-\frac{\sigma_0^2 t^2 b_n' b_n}{2n\sigma_{c_n}^2 + 8\iota_n^2\sigma_0^4 t^2})\prod_{j=1}^n (1 + \frac{4t^2\sigma_0^4\iota_{nj}^2}{n\sigma_{c_n}^2})^{-1/4} \\
&\le \exp\Big(-\frac{b_n' b_n}{72\iota_n^2\sigma_0^2} - \frac{1}{4}\sum_{j=1}^n \ln\big((1 + \frac{\iota_{nj}^2}{8\iota_n^2})\Big) \\
&\le \exp(-\frac{b_n' b_n}{72\iota_n^2\sigma_0^2} - \frac{1}{32\iota_n^2}\sum_{j=1}^n \iota_{nj}^2 + \frac{1}{512\iota_n^4}\sum_{j=1}^n \iota_{nj}^4) \quad (\ln(1+x) \ge x - \frac{x^2}{2} \text{ for } x \ge 0) \\
&\le \exp(-\frac{b_n' b_n}{72\iota_n^2\sigma_0^2} - \frac{15}{512\iota_n^2}\sum_{j=1}^n \iota_{nj}^2) \\
&\le \exp(-\frac{n\sigma_{c_n}^2}{72\iota_n^2\sigma_0^4}).
\end{aligned} \tag{C.13}$$

In (C.11), let $T = n\sigma_{c_n}^2$. By (C.12), the contribution of the integral in (C.11) when $|t| \le \frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n\sigma_0^2}$ is $O(n^{-1})$. The contribution when $\frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n\sigma_0^2} < |t| \le T$ tends to zero more rapidly than any power of n by $|\varphi_n(t) - \gamma_n(t)|/|t| \le (|\varphi_n(t)| + |\gamma_n(t)|)/|t|$ and (C.13). Therefore, $\sup_{x\in\mathbb{R}} |\mathrm{P}(c_n/\sigma_{c_n} \le x) - (\Phi(x) - \kappa_n\Phi^{(3)}(x))| = O(n^{-1})$. Eq. (27) holds by a similar argument and Lemma 6.

To establish high order expansions, use the Taylor approximation for $g_n(t)$ up to and including the term of degree r. Denote this approximation by $t^3\tau_{nr}(t) = \sum_{k=3}^r \frac{g_n^{(k)}(0)}{k!}t^k$, where $\tau_{nr}(t)$ is a polynomial of degree $r-3$. Let $p_n(t) = \sum_{k=1}^{r-2}\frac{1}{k!}[(-it)^3\tau_{nr}(-it)]^k$ be a polynomial with coefficients $p_{n1}, \dots, p_{nm}$. Note that $p_{n1}, \dots, p_{nm}$ are real by (C.9). Let $H_1(t), \dots, H_m(t)$ be hermite polynomials, then $\omega_n(t) = \Phi^{(1)}(t)(1 + \sum_{k=1}^m p_{nk}H_k(t))$ has the Fourier transform $\exp(-t^2/2)(1 + p_n(it))$. For all $T > 0$,

$$\sup_{x\in\mathbb{R}}\Big|\mathrm{P}(c_n/\sigma_{c_n} \le x) - \int_{-\infty}^x \omega_n(t)\,dt\Big| \le \frac{1}{\pi}\int_{-T}^T \Big|\frac{\varphi_n(t) - \exp(-t^2/2)(1 + p_n(it))}{t}\Big|\,dt + \frac{24\sup_x|\omega_n(x)|}{\pi T}.$$

Using the inequality that $|e^\alpha - 1 - \sum_{k=1}^{r-2}\beta^k/k!| = |(e^\alpha - e^\beta) + (e^\beta - 1 - \sum_{k=1}^{r-2}\beta^k/k!)| \le \exp(\max\{|\alpha|, |\beta|\})(|\alpha - \beta| + \frac{1}{(r-1)!}|\beta|^{r-1})$, we have

$$\begin{aligned}
&|\varphi_n(t) - \exp(-t^2/2)(1 + p_n(it))| \\
&= \big|\exp(-t^2/2)\big(\exp(g_n(t)) - 1 - p_n(it)\big)\big| \\
&\le \exp\big(-t^2/2 + \max\{|g_n(t)|, |t^3\tau_{nr}(t)|\}\big)\big(|g_n(t) - t^3\tau_{nr}(t)| + \frac{1}{(r-1)!}|t^3\tau_{nr}(t)|^{r-1}\big).
\end{aligned}$$

By (C.10), $|g_n(t) - t^3\tau_{nr}(t)| \le \frac{c_{r+1,5}(\iota_n\sigma_0^2)^{r-1}}{n^{(r-1)/2}\sigma_{c_n}^{r-1}(r+1)!}|t|^{r+1} \le t^2/8$ and $|t^3\tau_{nr}(t)| \le t^2/8$ when $t \le cn^{1/2}$ for some constant c. Then when $t \le cn^{1/2}$, $|g_n(t)| \le t^2/4$ and

$$|\varphi_n(t) - \exp(-t^2/2)(1 + p_n(it))| \le n^{-(r-1)/2}\exp(-t^2/4)\Big(\frac{c_{r+1,5}(\iota_n\sigma_0^2)^{r-1}}{\sigma_{c_n}^{r-1}(r+1)!}|t|^{r+1} + \frac{1}{(r-1)!}|t^3 n^{1/2}\tau_{nr}(t)|^{r-1}\Big).$$

Now let $T = n^{(r-1)/2}$, then $\sup_{x\in\mathbb{R}} |\operatorname{P}(c_n/\sigma_{c_n} \le x) - \int_{-\infty}^x \omega_n(t)\,dt| = O(n^{-(r-1)/2})$. Note that $\int_{-\infty}^x \omega_n(t)\,dt = \Phi(x) - p_{n1}\Phi^{(1)}(x) - \Phi^{(1)}(x)\sum_{k=2}^m p_{nk}H_{k-1}(x)$, which is a polynomial in $n^{-1/2}$ with bounded coefficients for fixed x, by (C.10). Rearranging it according to ascending powers of $n^{-1/2}$ and dropping the terms involving powers $n^{-k/2}$ with $k > r-1$ yields the expression in the proposition. □

Proof of Proposition 5. Let $A_n = [H_nM_nH_n - n^{-1}xH_n\operatorname{tr}^{1/2}(M_n^2 + M_n'M_n)]$, $\sigma_{a_n}^2 = n^{-1}\operatorname{E}[\epsilon_n'A_n\epsilon_n - \sigma_0^2\operatorname{tr}(A_n)]^2 = 2n^{-1}\sigma_0^4\operatorname{tr}(A_n^2)$, $z_n = -\sigma_0^2 n^{-1/2}\sigma_{a_n}^{-1}\operatorname{tr}(A_n) = -2^{-1/2}\operatorname{tr}(A_n)\operatorname{tr}^{-1/2}(A_n^2)$ and $\kappa_n = 4\sigma_0^6 n^{-3/2}\sigma_{a_n}^{-3}\operatorname{tr}(A_n^3)/3 = 4[2\operatorname{tr}(A_n^2)]^{-3/2}\operatorname{tr}(A_n^3)/3$. The κ_n does not involve any population parameter. Then

$$\begin{aligned}\operatorname{P}(\mathbb{I}_n \le x) &= \operatorname{P}(\epsilon_n'A_n\epsilon_n \le 0)\\ &= \operatorname{P}(n^{-1/2}\sigma_{a_n}^{-1}[\epsilon_n'A_n\epsilon_n - \sigma_0^2\operatorname{tr}(A_n)] \le z_n)\\ &= \Phi(z_n) + \kappa_n(1 - z_n^2)\Phi^{(1)}(z_n) + O(n^{-1}),\end{aligned}$$

by Theorem 2. Similarly,

$$\operatorname{P}^*(\mathbb{I}_n^* \le x) = \Phi(z_n) + \kappa_n(1 - z_n^2)\Phi^{(1)}(z_n) + O_P(n^{-1}).$$

Then $\operatorname{P}^*(\mathbb{I}_n^* \le x) - \operatorname{P}(\mathbb{I}_n \le x) = O_P(n^{-1})$. □

Proof of Theorem 3. To prove the expansions in (29) and (30), we check that the conditions of Theorem 1 in Mykland (1993) are satisfied. The same decomposition for c_n and σ-fields as in the proof of Theorem 1 are used. By (C.5), $\sum_{i=1}^n \operatorname{E}(c_{ni}^4/\sigma_{c_n}^4) = O(n^{-1})$. By (C.6), $\operatorname{E}\big|n^{1/2}\big[\big(\sigma_{c_n}^{-2}\sum_{i=1}^n \operatorname{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big) - 1\big]\big|^2 = O(1)$. Then the integrability conditions for the fourth-order and square variations are satisfied. As c_n/σ_{c_n} has constant variance 1, it remains to check the central limit condition that

$$\big(c_n/\sigma_{c_n}, n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^n c_{ni}^2 - 1\big), n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^n \operatorname{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - 1\big)\big) \tag{C.14}$$

is asymptotically trivariate normal. We may verify that $\big(c_n/\sigma_{c_n}, n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^n\big(c_{ni}^2 - \operatorname{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big),$ $n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^n \operatorname{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - 1\big)\big)$ is asymptotically trivariate normal. The c_n/σ_{c_n} is asymptotically normal by Lemma 4. Now we show that both $n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^n \operatorname{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - 1\big)$ and $n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^n\big(c_{ni}^2 - \operatorname{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)$ are asymptotically normal.

$$\begin{aligned}n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^n \operatorname{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - 1\big) = 4\sigma_{c_n}^{-2}n^{-1/2}\big[&2\sigma_0^2\sum_{j=1}^{n-1}\sum_{k=1}^{j-1}\epsilon_{nj}\epsilon_{nk}\sum_{i=j+1}^n a_{n,ij}a_{n,ik} + \sigma_0^2\sum_{j=1}^{n-1}(\epsilon_{nj}^2 - \sigma_0^2)\sum_{i=j+1}^n a_{n,ij}^2\\ &+ \sum_{j=1}^{n-1}\epsilon_{nj}\sum_{i=j+1}^n a_{n,ij}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})\big],\end{aligned}$$

which is a LQ form. The involved matrix in the LQ form is bounded in both row and column sum norms, since $\sum_{j=1}^{n-1}|\sum_{i=j+1}^{n}a_{n,ij}a_{n,ik}|+|\sum_{i=j+1}^{n}a_{n,ij}^2|\leq\sum_{i=1}^{n}|a_{n,ik}|\sum_{j=1}^{n}|a_{n,ij}|+(\sum_{i=1}^{n}|a_{n,ij}|)^2<\infty$ and $\sum_{k=1}^{n-1}|\sum_{i=j+1}^{n}a_{n,ij}a_{n,ik}|+|\sum_{i=k+1}^{n}a_{n,ik}^2|\leq\sum_{i=1}^{n}|a_{n,ij}|\sum_{k=1}^{n}|a_{n,ik}|+(\sum_{i=1}^{n}|a_{n,ik}|)^2<\infty$. In addition, for $q=1+\delta>1$ and $1/p+1/q=1$,

$$
\begin{aligned}
\frac{1}{n-1}\sum_{j=1}^{n-1}\Big|\sum_{i=j+1}^{n}a_{n,ij}(\mu_3 a_{n,ii}+\sigma_0^2 b_{ni})\Big|^q &\leq \frac{1}{n-1}\sum_{j=1}^{n-1}\Big|\sum_{i=j+1}^{n}|a_{n,ij}|^{1/p}|a_{n,ij}|^{1/q}(|\mu_3 a_{n,ii}|+\sigma_0^2|b_{n,ii}|)\Big|^q\\
&\leq \frac{1}{n-1}\sum_{j=1}^{n-1}\big(2\sum_{i=j+1}^{n}|a_{n,ij}|\big)^{q/p}\sum_{i=j+1}^{n}|a_{n,ij}|(|\mu_3 a_{n,ii}|^q+\sigma_0^{2q}|b_{n,ii}|^q)\\
&\leq \frac{c}{n-1}\sum_{j=1}^{n-1}\sum_{i=j+1}^{n}|a_{n,ij}|(|\mu_3 a_{n,ii}|^q+\sigma_0^{2q}|b_{n,ii}|^q)<\infty,
\end{aligned}
$$

where c is a constant. Thus $n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^{n}\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})-1\big)$ is asymptotically normal. Noting that $n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^{n}\big(c_{ni}^2-\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)$ is a sum of martingale differences, it can be shown to be asymptotically normal by a central limit theorem for martingales. Let

$$
\begin{aligned}
z_{ni}&=n^{1/2}\big(c_{ni}^2-\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)\\
&=n^{-1/2}\Big(a_{n,ii}^2(\epsilon_{ni}^4-\mu_4)+2a_{n,ii}b_{ni}(\epsilon_{ni}^3-\mu_3)+(b_{ni}^2-2\sigma_0^2a_{n,ii}^2)(\epsilon_{ni}^2-\sigma_0^2)-2\sigma_0^2a_{n,ii}b_{ni}\epsilon_{ni}\\
&\quad+4[a_{n,ii}(\epsilon_{ni}^3-\mu_3)+b_{ni}(\epsilon_{ni}^2-\sigma_0^2)-\sigma_0^2a_{n,ii}\epsilon_{ni}]\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{nj}+4(\epsilon_{ni}^2-\sigma_0^2)\big(\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{nj}\big)^2\Big).
\end{aligned}
$$

As $(\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{nj})^2=\sum_{j=1}^{i-1}a_{n,ij}^2\epsilon_{nj}^2+2\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}a_{n,ij}a_{n,ik}\epsilon_{nj}\epsilon_{nk}$, for some $q>2$ and $1/p+1/q=1$,

$$
\begin{aligned}
|z_{ni}|^q&\leq n^{-q/2}\Big[|a_{n,ii}|^{2p}+|2a_{n,ii}|^p+1+|2\sigma_0^2a_{n,ii}^2|^p+|2\sigma_0^2a_{n,ii}|^p+4^p(|a_{n,ii}|^p+1+\sigma_0^{2p}|a_{n,ii}|^p)\sum_{j=1}^{i-1}|a_{n,ij}|\\
&\quad+4^p\Big(\sum_{j=1}^{i-1}|a_{n,ij}|^{p+1}+2^p\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}|a_{n,ij}||a_{n,ik}|\Big)\Big]^{q/p}\Big[|\epsilon_{ni}^4-\mu_4|^q+|b_{ni}|^q|\epsilon_{ni}^3-\mu_3|^q\\
&\quad+(|b_{ni}|^{2q}+1)|\epsilon_{ni}^2-\sigma_0^2|^q+|b_{ni}\epsilon_{ni}|^q+[|\epsilon_{ni}^3-\mu_3|^q+|b_{ni}|^q|\epsilon_{ni}^2-\sigma_0^2|^q+|\epsilon_{ni}|^q]\sum_{j=1}^{i-1}|a_{n,ij}||\epsilon_{nj}|^q\\
&\quad+|\epsilon_{ni}^2-\sigma_0^2|^q\Big(\sum_{j=1}^{i-1}|a_{n,ij}||\epsilon_{nj}|^{2q}+\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}|a_{n,ij}||a_{n,ik}||\epsilon_{nj}|^q|\epsilon_{nk}|^q\Big)\Big].
\end{aligned}
$$

Then $\sigma_{c_n}^{-2q}\sum_{i=1}^{n}\mathrm{E}[\mathrm{E}(|z_{ni}|^q|\mathscr{F}_{n,i-1})]=o(1)$. Next we show that $\sum_{i=1}^{n}\big(\mathrm{E}(z_{ni}^2|\mathscr{F}_{n,i-1})-\mathrm{E}(z_{ni}^2)\big)=o_P(1)$. Let $e_{1n,i}=a_{n,ii}^2(\epsilon_{ni}^4-\mu_4)+2a_{n,ii}b_{n,ii}(\epsilon_{ni}^3-\mu_3)+(b_{ni}^2-2\sigma_0^2a_{n,ii}^2)(\epsilon_{ni}^2-\sigma_0^2)-2\sigma_0^2a_{n,ii}b_{ni}\epsilon_{ni}$, $e_{2n,i}=4[a_{n,ii}(\epsilon_{ni}^3-\mu_3)+b_{n,ii}(\epsilon_{ni}^2-\sigma_0^2)-\sigma_0^2a_{n,ii}\epsilon_{ni}]$, $e_{3n,i}=4(\epsilon_{ni}^2-\sigma_0^2)$ and $e_{4n,i}=\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{nj}$. Then $z_{ni}=n^{-1/2}(e_{1n,i}+$

$e_{2n,i}e_{4n,i} + e_{3n,i}e^2_{4n,i})$ and

$$\sum_{i=1}^{n}\big(\mathrm{E}(z^2_{ni}|\mathscr{F}_{n,i-1}) - \mathrm{E}(z^2_{ni})\big) = \frac{1}{n}\sum_{i=1}^{n}\big((e^2_{4n,i} - \mathrm{E}\, e^2_{4n,i})\,\mathrm{E}(e^2_{2n,i} + 2e_{1n,i}e_{3n,i}) + (e^4_{4n,i} - \mathrm{E}\, e^4_{4n,i})\,\mathrm{E}\, e^2_{3n,i} + 2e_{4n,i}\,\mathrm{E}(e_{1n,i}e_{2n,i}) + 2(e^3_{4n,i} - \mathrm{E}\, e^3_{4n,i})\,\mathrm{E}(e_{2n,i}e_{3n,i})\big). \tag{C.15}$$

For $r = 1, 2, 3$, $n^{-1}\sum_{i=1}^{n} b^r_{ni}e_{4n,i} = o_P(1)$, since

$$\begin{aligned}\mathrm{E}\big(\frac{1}{n}\sum_{i=1}^{n} b^r_{ni}e_{4n,i}\big)^2 &= \frac{\sigma_0^2}{n^2}\sum_{j=1}^{n-1}\big(\sum_{i=j+1}^{n} a_{n,ij}b^r_{ni}\big)^2 \\ &\le \frac{\sigma_0^2}{n^2}\sum_{j=1}^{n-1}\big(\sum_{i=j+1}^{n} |a_{n,ij}|b^{2r}_{ni}\big)\big(\sum_{i=j+1}^{n} |a_{n,ij}|\big) \\ &\le \frac{h}{n^2}\sum_{i=1}^{n} b^{2r}_{ni}\big(\sum_{j=1}^{i-1} |a_{n,ij}|\big) = O(n^{-1}),\end{aligned} \tag{C.16}$$

where h is a constant. For $r = 1, 2$, $\frac{1}{n}\sum_{i=1}^{n} b^r_{ni}(e^2_{4n,i} - \mathrm{E}\, e^2_{4n,i}) = o_P(1)$, since

$$\frac{1}{n}\sum_{i=1}^{n} b^r_{ni}(e^2_{4n,i} - \mathrm{E}\, e^2_{4n,i}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1} b^r_{ni}a^2_{n,ij}(\epsilon^2_{nj} - \sigma_0^2) + \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} b^r_{ni}a_{n,ij}a_{n,ik}\epsilon_{nj}\epsilon_{nk},$$

where

$$\mathrm{E}\big(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1} b^r_{ni}a^2_{n,ij}(\epsilon^2_{nj} - \sigma_0^2)\big)^2 = \frac{\mathrm{E}\,|\epsilon^2_{nj} - \sigma_0^2|^2}{n^2}\sum_{j=1}^{n-1}\big(\sum_{i=j+1}^{n} b^r_{ni}a^2_{n,ij}\big)^2 = O(n^{-1})$$

and

$$\mathrm{E}\big(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} b^r_{ni}a_{n,ij}a_{n,ik}\epsilon_{nj}\epsilon_{nk}\big)^2 = \frac{\sigma_0^4}{n^2}\sum_{j=1}^{n-1}\sum_{k=1}^{j-1}\big(\sum_{i=j+1}^{n} b^r_{ni}a_{n,ij}a_{n,ik}\big)^2 = O(n^{-1}),$$

as in (C.16). Similarly, $n^{-1}\sum_{i=1}^{n} b_{ni}(e^3_{4n,i} - \mathrm{E}\, e^3_{4n,i}) = o_P(1)$ and $n^{-1}\sum_{i=1}^{n}(e^4_{4n,i} - \mathrm{E}\, e^4_{4n,i}) = o_P(1)$, since they can be decomposed as

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} b_{ni}(e^3_{4n,i} - \mathrm{E}\, e^3_{4n,i}) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1} b_{ni}a^3_{n,ij}(\epsilon^3_{nj} - \mu_3) + \frac{3}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} b_{ni}a^2_{n,ij}a_{n,ik}\epsilon^2_{nj}\epsilon_{nk} \\ &\quad + \frac{3}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} b_{ni}a_{n,ij}a^2_{n,ik}\epsilon_{nj}\epsilon^2_{nk} + \frac{6}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1} b_{ni}a_{n,ij}a_{n,ik}a_{n,il}\epsilon_{nj}\epsilon_{nk}\epsilon_{nl},\end{aligned}$$

and

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}(e^4_{4n,i} - \mathrm{E}\, e^4_{4n,i}) \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1} a^4_{n,ij}(\epsilon^4_{nj} - \mu_4) + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\big(4a^3_{n,ij}a_{n,ik}\epsilon^3_{nj}\epsilon_{nk} + 4a_{n,ij}a^3_{n,ik}\epsilon_{nj}\epsilon^3_{nk} + 6a^2_{n,ij}a^2_{n,ik}(\epsilon^2_{nj}\epsilon^2_{nk} - \sigma_0^4)\big)\end{aligned}$$

$$
\begin{aligned}
&+\frac{12}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1}(a_{n,ij}^2 a_{n,ik}a_{n,il}\epsilon_{nj}^2\epsilon_{nk}\epsilon_{nl}+a_{n,ij}a_{n,ik}^2 a_{n,il}\epsilon_{nj}\epsilon_{nk}^2\epsilon_{nl}+a_{n,ij}a_{n,ik}a_{n,il}^2\epsilon_{nj}\epsilon_{nk}\epsilon_{nl}^2)\\
&+\frac{24}{n}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1}\sum_{m=1}^{l-1}a_{n,ij}a_{n,ik}a_{n,il}a_{im,n}\epsilon_{nj}\epsilon_{nk}\epsilon_{nl}\epsilon_{m,n},
\end{aligned}
$$

where each term on the r.h.s. of above equations converges to zero in probability since its variance has the order $O(n^{-1})$ as in (C.16). Note that in (C.15), $\mathrm{E}(e_{2n,i}^2+2e_{1n,i}e_{3n,i})$, $\mathrm{E}(e_{1n,i}e_{2n,i})$ and $\mathrm{E}(e_{2n,i}e_{3n,i})$ are polynomials of b_{ni}'s with bounded constants. Then $\sum_{i=1}^{n}\big(\mathrm{E}(z_{ni}^2|\mathscr{F}_{n,i-1})-\mathrm{E}(z_{ni}^2)\big)=o_P(1)$. Under the assumption that the limit of $\sum_{i=1}^{n}\mathrm{E}(z_{ni}^2)$ exists, $n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^{n}\big(c_{ni}^2-\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)$ is asymptotically normal by Corollary 3.1 in Hall and Heyde (1980). Then (C.14) is asymptotically trivariate normal by the Cramér-Wold device. Since $\sigma_{c_n}^{-2}\,\mathrm{E}(c_n^2)=1$, we can take

$$
\psi_o(x)-\psi_p(x)=x\lim_{n\to\infty}\sigma_{c_n}^{-3}n^{1/2}\,\mathrm{E}\Big(c_n\sum_{i=1}^{n}\big(c_{ni}^2-\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)\Big) \tag{C.17}
$$

and

$$
\psi_p(x)=x\lim_{n\to\infty}\sigma_{c_n}^{-1}n^{1/2}\,\mathrm{E}\Big(c_n\big(\sigma_{c_n}^{-2}\sum_{i=1}^{n}\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})-1\big)\Big), \tag{C.18}
$$

where

$$
\begin{aligned}
&\sigma_{c_n}^{-3}n^{1/2}\,\mathrm{E}\Big(c_n\sum_{i=1}^{n}\big(c_{ni}^2-\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)\Big)\\
&=\sigma_{c_n}^{-3}n^{1/2}\sum_{i=1}^{n}\mathrm{E}\big(\mathrm{E}(c_{ni}^3|\mathscr{F}_{n,i-1})\big)\\
&=\sigma_{c_n}^{-3}n^{-1}\sum_{i=1}^{n}\Big[a_{n,ii}^3\,\mathrm{E}(\epsilon_{ni}^2-\sigma_0^2)^3+8\mu_3^2\sum_{j=1}^{i-1}a_{n,ij}^3+\mu_3 b_{ni}^3+12\sigma_0^2[(\mu_4-\sigma_0^4)a_{n,ii}+\mu_3 b_{ni}]\sum_{j=1}^{i-1}a_{n,ij}^2\\
&\quad+3(\mu_4-\sigma_0^4)a_{n,ii}b_{ni}^2+3a_{n,ii}^2 b_{ni}\,\mathrm{E}[\epsilon_{ni}(\epsilon_{ni}^2-\sigma_0^2)^2]\Big]
\end{aligned} \tag{C.19}
$$

and

$$
\begin{aligned}
&\sigma_{c_n}^{-1}n^{1/2}\,\mathrm{E}\Big(c_n\big(\sigma_{c_n}^{-2}\sum_{i=1}^{n}\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})-1\big)\Big)\\
&=\sigma_{c_n}^{-3}n^{1/2}\,\mathrm{E}\Big(c_n\sum_{i=1}^{n}\big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})-\mathrm{E}(c_{ni}^2)\big)\Big)\\
&=\sigma_{c_n}^{-3}n^{-1}\,\mathrm{E}\Big[\Big(\sum_{i=1}^{n}[a_{n,ii}(\epsilon_{ni}^2-\sigma_0^2)+b_{ni}\epsilon_{ni}]+2\sum_{i=1}^{n}\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{ni}\epsilon_{nj}\Big)\cdot\\
&\quad\Big(8\sigma_0^2\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}a_{n,ij}a_{n,ik}\epsilon_{nj}\epsilon_{nk}+4\sigma_0^2\sum_{i=1}^{n}\sum_{j=1}^{i-1}a_{n,ij}^2(\epsilon_{nj}^2-\sigma_0^2)+4\sum_{i=1}^{n}(\mu_3 a_{n,ii}+\sigma_0^2 b_{ni})\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{nj}\Big)\Big]\\
&=\sigma_{c_n}^{-3}n^{-1}\Big[4\sum_{i=1}^{n}\sum_{j=1}^{i-1}\Big[\sigma_0^2 a_{n,ij}^2[a_{n,jj}(\mu_4-\sigma_0^4)+\mu_3 b_{nj}]+a_{n,ij}(\mu_3 a_{n,ii}+\sigma_0^2 b_{ni})(\mu_3 a_{n,jj}+\sigma_0^2 b_{nj})\Big]
\end{aligned}
$$

$$+ 16\sigma_0^6 \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} a_{n,ij}a_{n,ik}a_{n,jk}\Bigg]. \tag{C.20}$$

Therefore, (29) holds by Theorem 1 in Mykland (1993). Eq. (30) follows by (2.20) in Mykland (1993).

□

References

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Boston: Kluwer Academic Publishers.

Anselin, L., 1990. Some robust approaches to testing and estimation in spatial econometrics. Regional Science and Urban Economics 20, 141–163.

Burridge, P., 1980. On the Cliff-Ord test for spatial autocorrelation. Journal of the Royal Statistical Society, Series B 42, 107–108.

Burridge, P., 2012. Improving the J test in the SARAR model by likelihood-based estimation. Spatial Economic Analysis 7, 75–107.

Can, A., 1992. Specification and estimation of hedonic housing price models. Regional science and urban economics 22, 453–474.

Cliff, A., Ord, J.K., 1973. Spatial Autocorrelation. Pion, London.

Cliff, A., Ord, J.K., 1981. Spatial Process: Models and Applications. Pion, London.

Davison, A.C., Hinkley, D.V., Young, G.A., 2003. Recent developments in bootstrap methodology. Statistical Science 18, 141–157.

DiCiccio, T.J., Efron, B., 1996. Bootstrap confidence intervals. Statistical Science , 189–212.

Feller, W., 1970. An introduction to probability theory and its applications. volume II. John Wiley & Sons, Inc., New York, US.

Fingleton, B., 2008. A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors. Spatial Economic Analysis 3, 27–44.

Fingleton, B., Burridge, P., 2010. Bootstrap inference in spatial econometrics: the J test. Spatial Economic Analysis 5, 93–119.

Fingleton, B., Le Gallo, J., 2008. Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: Finite sample properties. Papers in Regional Science 87, 319–339.

Freedman, D.A., 1981. Bootstrapping regression models. Annals of Statistics 9, 1218–1228.

Götze, F., Hipp, C., 1983. Asymptotic expansions for sums of weakly dependent random vectors. Probability Theory and Related Fields 64, 211–239.

Götze, F., Hipp, C., 1994. Asymptotic distribution of statistics in time series. Annals of Statistics 22, 2062–2088.

Götze, F., Tikhomirov, A.N., Yurchenko, V.A., 2007. Asymptotic expansion in the central limit theorem for quadratic forms. Journal of Mathematical Sciences 147, 6891–6911.

Haeusler, E., 1988. On the rate of convergence in the central limit theorem for martingales with discrete and continuous time. Annals of Probability 16, 275–299.

Hall, P., 1997. The bootstrap and Edgeworth expansion. Springer Verlag.

Hall, P., Heyde, C., 1980. Martingale limit theory and its applications. Academic Press, New York.

Heyde, C.C., Brown, B.M., 1970. On the departure from normality of a certain class of martingales. Annals of mathematical statistics 41, 2161–2165.

Horowitz, J.L., 2001. The bootstrap, in: Heckman, J.J., Leamer, E. (Eds.), Handbook of econometrics. Elsevier. volume 5, pp. 3159–3228.

Horowitz, J.L., 2003. The bootstrap in econometrics. Statistical Science 18, 211–218.

Jin, F., Lee, L.F., 2012. Cox-type tests for competing spatial autoregressive models with spatial autoregressive disturbances. Working Paper, Department of Economics, Ohio State University.

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

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

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

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

Lee, L.F., 2001. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. Manuscript. Department of Economics, Ohio State University.

Lee, L.F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Econometrica 72, 1899–1925.

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

Lee, L.F., Liu, X., 2010. Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. Econometric Theory 26, 187–230.

Lin, K.P., Long, Z., Ou, B., 2011. The size and power of bootstrap tests for spatial dependence in a linear regression model. Computational Economics 38, 153–171.

Lin, X., Lee, L.F., 2010. GMM estimation of spatial autoregressive models with unknown heteroskedasticity. Journal of Econometrics 157, 34–52.

MacKinnon, J.G., 2002. Bootstrap inference in econometrics. Canadian Journal of Economics/Revue canadienne d'économique 35, 615–645.

Monchuk, D.C., Hayes, D.J., Miranowski, J.A., Lambert, D.M., 2011. Inference based on alternative bootstrapping methods in spatial models with an application to county income growth in the United States. Journal of Regional Science 51, 880–896.

Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika 35, 255–260.

Mykland, P.A., 1993. Asymptotic expansions for martingales. Annals of Probability 21, 800–818.

Robinson, P.M., Rossi, F., 2010. Finite sample corrections for the Lagrange multiplier test in spatial autoregressive models. Working paper.

Su, L., Yang, Z., 2008. Asymptotics and bootstrap for transformed panel data regressions. Working paper, School of Economics, Singapore Management University.

White, H., 1994. Estimation, inference and specification analysis. New York: Cambridge University Press.

Yang, Z., 2011. LM tests of spatial dependence based on bootstrap critical values. Working paper, School of Economics, Singapore Management University.

Yang, Z., 2012. A general method for third-order bias and variance correction on a nonlinear estimator. Working paper, School of Economics, Singapore Management University.