On the Bootstrap for Moran's I Test for Spatial Dependence[☆]

Fei Jin[a,1], Lung-fei Lee[b]

[a] School of Economics, Shanghai University of Finance and Economics, Shanghai 200433 China
[b] Department of Economics, The Ohio State University, Columbus, OH 43210 USA

Abstract

This paper is concerned with the use of the bootstrap for statistics in spatial econometric models, with a focus on the test statistic for Moran's I test for spatial dependence. We show that, for many statistics in spatial econometric models, the bootstrap 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 a normal distribution, we show that the bootstrap is generally consistent for test statistics that can be approximated by LQ forms, including Moran's I. Possible asymptotic refinements of the bootstrap are most commonly studied using Edgeworth expansions. For spatial econometric models, we may establish asymptotic refinements of the bootstrap based on asymptotic expansions of LQ forms. When the disturbances are normal, we prove the existence of the usual Edgeworth expansions for LQ forms; when the disturbances are not normal, we establish an asymptotic expansion of LQ forms based on martingales. These results are applied to show the second order correctness of the bootstrap for Moran's I test.

Keywords: Bootstrap, spatial, Moran's I, consistency, asymptotic refinement, linear-quadratic form
JEL classification: C12, C15, C21, R15

1. Introduction

The bootstrap is a statistical procedure that estimates 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 for many econometric estimators and test statistics. In some cases, it can be used as an alternative when evaluating an asymptotic distribution 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 biases of estimators, estimate critical values for hypothesis tests, construct confidence intervals, etc. Useful survey papers on the bootstrap

[☆] We are grateful to an anonymous referee and co-editor, Peter Robinson, of the Journal, whose comments have led to great improvement of this paper.

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

[1] Member of the Key Laboratory of Mathematical Economics (SUFE), Ministry of Education, and supported by the Open Project Program in the Key Laboratory.

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 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 with a Monte Carlo study. Fingleton and Burridge (2010) 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. 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 biases and variances of quasi-maximum likelihood estimators 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. As more researchers see the need to explicitly deal with problems caused by spatial dependence, the bootstrap may be used often for spatial econometric models. Formal justification of the bootstrap for spatial econometric models can help us to reach reliable conclusions. The objective of this paper is to establish a general consistency result of the bootstrap which may be useful for a bunch of statistics—including Moran's I—in spatial econometric models and show that the bootstrap can provide asymptotic refinements for Moran's I. We demonstrate 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, due to the presence of spatial dependence. The bootstrap for many spatial econometric models thus can be studied based on LQ forms in general. Kelejian and Prucha (2001) prove asymptotic normality of LQ forms using a central limit theorem for martingale difference arrays. We shall show that the convergence of the cumulative distribution function (CDF) for an LQ form is uniform under the same conditions. Using this uniform convergence, the bootstrap may generally be consistent for statistics that can be approximated by an LQ form. We apply the result to show formally consistency of the bootstrap for Moran's I.

Possible asymptotic refinements of the bootstrap are most commonly studied using Edgeworth expansions. For statistics in spatial econometric models, we may investigate whether the bootstrap can provide asymptotic refinements by considering Edgeworth expansions of LQ forms. Such expansions, however, have not been proved to exist in the literature. For non-spatial econometric models, the bootstrap is often considered for 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), for which the

Edgeworth expansions are well established. This framework does not apply to LQ forms, which cannot be written as simple averages of disturbances. A related statistic is the U-statistic, for which the Edgeworth expansions have been established (see, e.g., Bickel et al. (1986) and Jing and Wang (2003)). However, in these results, the kernel is a fixed function.[2] For spatial econometric statistics, the involved square matrix in an LQ form changes as the sample size changes. The row and column sum norms of this square matrix are usually bounded, but the value of a particular element is not constrained further. Thus, as shown later, an LQ form can be partly seen as a generalized U-statistic with changing kernels, for which no existing results can be applied. As a result, we might need to study the LQ form as a whole to a large extent. Götze et al. (2007) establishes a one-term asymptotic expansion for a quadratic form based on a symmetrization inequality and the differential inequality method. However, the expansion does not have a remainder term of a desirable order as in a usual Edgeworth expansion.[3]

We shall show the existence of formal Edgeworth expansions for LQ forms of normal disturbances. With normal disturbances, the characteristic functions of LQ forms are available. The expansions are established by using a smoothing inequality that bounds the gap between two functions with related Fourier transforms. The special feature of the square matrix involved in an LQ form for spatial econometric models, i.e., the boundedness in both row and column sum norms, can be used to obtain the order of the bound. When the disturbances are not normal, using the method to establish the usual Edgeworth expansion might not work. We highlight that an LQ form, unlike a linear form, can be neither written as a simple average so that some common structures of elements in the average can be explored, nor written as a U-statistic with a fixed kernel. Alternatively, by decomposing an LQ form into a sum of martingale differences, we show that an LQ form of non-normal disturbances has an asymptotic expansion based on martingales (Mykland, 1993). The expansion is in a test function topology. While it is not in a point-wise topology, it does not require strong conditions such as Cramér-type conditions. We apply these results to show the second order correctness of the bootstrap for Moran's I.[4]

The rest of the paper is organized as follows: Section 2 first introduces Moran's I statistics considered in this paper, then shows the uniform convergence of the CDF for LQ forms, and finally applies the result to show that the bootstrap is consistent for Moran's I. Section 3 establishes the Edgeworth expansion

[2] In the papers mentioned above, expansions for the standardized version of a U-statistic of degree two $U_n = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} h(X_i, X_j)$ is considered. The kernel $h(X_i, X_j)$ is the same for any i, j with $i \neq j$.

[3] In that paper, the sum of all squared elements in the square matrix of a quadratic form is assumed to be finite. The square matrix in an LQ form for spatial econometric models is usually bounded in both row and column sum norms. By dividing each element of a quadratic form for spatial econometric models by the square root of the sample size, the result in that paper can be applied and it is easy to see that the remainder term in the expansion does not have a desirable order as in a usual Edgeworth expansion.

[4] In Mykland (1993), the asymptotic expansion for martingales is called Edgeworth expansion for martingales and is used to show the second order correctness of the bootstrap for an AR(1) process.

of an LQ form with normal disturbances and an asymptotic expansion of an LQ form with non-normal disturbances, which are then applied to show the second-order correctness of the bootstrap for Moran's I. Some Monte Carlo results are reported in Section 4. Section 5 concludes. Lemmas and proofs are collected in the appendices.

2. Consistency of the Bootstrap

As pointed out by Anselin and Bera (1998), a fundamental problem in dealing with spatial dependence is that it is impossible to estimate general $n \times n$ covariance terms or correlations directly from a set of n observations on cross-sectional data. This necessitates the imposition of structure. A main approach is the lattice perspective, where, for each spatial unit, a relevant "neighborhood set" interacting with it is specified with an $n \times n$ spatial weights matrix M_n. With this matrix, spatial dependence is modeled as a functional relationship between a vector of dependent variable y_n (or disturbances ϵ_n) with its associated spatial lag $M_n y_n$ (or $M_n \epsilon_n$). Popular spatial econometric models include SAR models, spatial error (SE) models, spatial moving average (SMA) models, spatial Durbin models, and SAR models with SAR errors (SARAR models). For estimators of spatial econometric models, derivatives of corresponding criterion functions for the quasi-maximum likelihood (QML), instrumental variables and generalized method of moments evaluated at the true parameter vector are often LQ forms of the disturbances, rather than just linear forms, due to the presence of spatial dependence.[5] As a result, the corresponding estimators can be studied using an LQ form. Based on asymptotic normality of these estimators, one can implement classical hypothesis tests such as the Wald, likelihood ratio or Lagrangian multiplier (LM) tests, and tests of non-nested hypotheses such as the spatial J-type tests (Kelejian, 2008; Kelejian and Piras, 2011) or Cox-type tests (Jin and Lee, 2013). Of particular interest, Moran's I (Moran, 1950; Cliff and Ord, 1973, 1981) is a popular test for spatial dependence. Because of its importance, we focus in this paper on using LQ forms to study the bootstrap for Moran's I test.

2.1. Moran's I Test and the Bootstrap

The Moran I statistic is

$$\frac{n}{l_n' M_n l_n} \cdot \frac{\hat{\epsilon}_n' M_n \hat{\epsilon}_n}{\hat{\epsilon}_n' \hat{\epsilon}_n}, \tag{1}$$

where l_n is an n-dimensional vector of ones and $\hat{\epsilon}_n = (\hat{\epsilon}_{n1}, \dots, \hat{\epsilon}_{nn})'$ is the residual vector from the least squares estimation. The test is based on asymptotic normality of a standardized test statistic by deducting the estimated mean and dividing by the standard error. Burridge (1980) shows that for the regressive equation $y_n = X_n \beta + u_n$ with SAR errors $u_n = \rho M_n u_n + \epsilon_n$, or with SMA errors $u_n = \rho M_n \epsilon_n + \epsilon_n$, where

[5] See, e.g., Kelejian and Prucha (1998); Lee (2004a, 2007).

$\epsilon_n \sim N(0, \sigma^2 I_n)$, X_n is an $n \times k_x$ matrix of exogenous variables and I_n is the $n \times n$ identity matrix, the LM test statistic for $\rho = 0$ is proportional to the Moran I statistic, which is

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

Let $H_n = I_n - X_n (X_n' X_n)^{-1} X_n'$. Under the null hypothesis of no spatial dependence, (2) becomes

$$\begin{aligned} \mathbb{I}_n &= \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\epsilon_n' H_n M_n H_n \epsilon_n}{\epsilon_n' H_n \epsilon_n} \\ &= \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)}{(n - k_x)\sigma_0^2} + \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\operatorname{tr}(M_n H_n)}{n - k_x} \\ &\quad - \frac{n}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}} \frac{\epsilon_n' H_n M_n H_n \epsilon_n \big(\epsilon_n' H_n \epsilon_n - (n - k_x)\sigma_0^2\big)}{(n - k_x)\sigma_0^2 \epsilon_n' H_n \epsilon_n}. \end{aligned} \tag{3}$$

Under some regularity assumptions and due to the spatial weights matrix M_n having a zero diagonal, the last two terms on the r.h.s. of (3) have the order $O_P(n^{-1/2})$, thus the LM or Moran I statistic can be approximated by a quadratic form of the disturbances. The $\mathbb{I}_n$ is asymptotically normal under the null hypothesis, then (2) can be used directly for test purposes without adjusting for the mean and variance. 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 an LQ form. Note that when the disturbances are not normal, the proper asymptotic distribution of $\mathbb{I}_n$ in (2) is not standard normal. Thus, following Kelejian and Prucha (2001), we propose to use the following statistic for non-normal ϵ_n:

$$\mathbb{I}_n' = \frac{\hat{\epsilon}_n' M_n \hat{\epsilon}_n - \hat{\sigma}_n^2 \operatorname{tr}(M_n H_n)}{\sqrt{n}\,\underline{\hat{\sigma}_{c_n}}}, \tag{4}$$

where $\hat{\sigma}_n^2 = n^{-1} \hat{\epsilon}_n' \hat{\epsilon}_n$ is an estimate of the variance $\sigma_0^2 = \mathrm{E}(\epsilon_{ni}^2)$, and $\underline{\hat{\sigma}_{c_n}^2}$ is an estimate of the variance $\sigma_{c_n}^2$ of $n^{-1/2} \hat{\epsilon}_n' M_n \hat{\epsilon}_n$ under the null, with $\sigma_{c_n}^2 = n^{-1}(\mu_4 - 3\sigma_0^4) \sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + n^{-1} \sigma_0^4 \operatorname{tr}[H_n M_n H_n (M_n + M_n')]$ and $\mu_4 = \mathrm{E}(\epsilon_{ni}^4)$. For analytical tractability, we let $\underline{\hat{\sigma}_{c_n}^2} = \max\{\hat{\sigma}_{c_n}^2, c_\sigma\}$ for some positive constant c_σ smaller than $\sigma_{c_n}^2$ for any n, so that $\underline{\hat{\sigma}_{c_n}^2}$ is bounded away from zero, 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 $\hat{\mu}_{4n} = n^{-1} \sum_{i=1}^n \hat{\epsilon}_{ni}^4$.[6] Similar to $\mathbb{I}_n$, $\mathbb{I}_n'$ can be approximated by a quadratic form with a remainder of order $O_P(n^{-1/2})$.[7]

To consider the bootstrap for Moran's I, we first briefly discuss the bootstrap for a general 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 $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})'$, exogenous variables

[6] The censored variance estimate $\underline{\hat{\sigma}_{c_n}^2}$ would not be a practical issue once underflowing in machine precision is controlled for.

[7] Compared to $\mathbb{I}_n$, $\mathbb{I}_n'$ has a mean adjustment term $-\frac{\hat{\sigma}_n^2 \operatorname{tr}(M_n H_n)}{\sqrt{n}\underline{\hat{\sigma}_{c_n}}} = O_P(n^{-1/2})$. This is for analytical convenience.

and the true parameter vector θ_0. The t_n may also involve an estimator $\hat{\theta}_n$ of θ_0 and an estimator $\hat{\varsigma}_n$ of another 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 the disturbances ϵ_{ni}'s are assumed to be i.i.d., or one can use the parametric bootstrap if one has a parametric model for the disturbances. We may first derive a consistent estimator of parameters in a spatial econometric model and compute the residual vector $\hat{\epsilon}_n$. For the residual bootstrap, we first deduct the empirical mean of the residual vector from $\hat{\epsilon}_n$ to obtain $\tilde{\epsilon}_n = (I_n - \frac{1}{n} l_n l_n')\hat{\epsilon}_n$ as the $\hat{\epsilon}_n$ may not have mean zero,[8] and then sample randomly with replacement n times from the elements of $\tilde{\epsilon}_n$ to obtain a vector ϵ_n^*.[9] For the parametric bootstrap for residuals, ϵ_n^* can be sampled independently from the estimated model for disturbances. In particular, if we consider the case that ϵ_{ni}'s are normal, then we may sample n times from the normal distribution with mean zero and variance $n^{-1}\hat{\epsilon}_n'\hat{\epsilon}_n$ to obtain a vector ϵ_n^*. With ϵ_n^*, a pseudo data vector y_n^* on the dependent variable can be computed by using the reduced form with the parameter estimate $\hat{\theta}_n$ and disturbances ϵ_n^*. For example, for the SMA model, we have $y_n^* = X_n\hat{\beta}_n + (I_n + \hat{\rho}_n M_n)\epsilon_n^*$. 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. For Moran's I, due to its simplicity, the bootstrapped $\mathbb{I}_n$ and $\mathbb{I}_n'$ can be obtained by replacing $\hat{\epsilon}_n$ with $\hat{\epsilon}_n^*$ everywhere.[10]

2.2. A standard LQ Form

Consider a standardized LQ form c_n/σ_{c_n}, where

$$c_n = n^{-1/2}\big(\epsilon_n' A_n \epsilon_n - \sigma_0^2 \operatorname{tr}(A_n) + b_n'\epsilon_n\big) \tag{5}$$

is an LQ form with mean zero and variance $\sigma_{c_n}^2 = n^{-1}\big[2\sigma_0^4 \operatorname{tr}(A_n^2) + \sigma_0^2 b_n' b_n + \sum_{i=1}^n \big((\mu_4 - 3\sigma_0^4)a_{n,ii}^2 + 2\mu_3 a_{n,ii} b_{ni}\big)\big]$, with $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})'$ being a vector of i.i.d. disturbances which have mean zero, variance σ_0^2, third moment μ_3 and fourth moment μ_4, $A_n = [a_{n,ij}]$ being an $n \times n$ non-stochastic symmetric matrix[11], and $b_n = (b_{n1}, \ldots, b_{nn})'$ being an n-dimensional non-stochastic vector. 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

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

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

[10] This is the case when the bootstrap DGP is the null model with no spatial dependence. We consider only this case in this paper. The bootstrap DGP can also be the unrestricted alternative model when generating the bootstrap data vector. See MacKinnon (2002) for discussions. If the bootstrap DGP is the alternative model, which might be the SAR, SE or SARAR model, similar results can be proved with the lemmas in the appendices.

[11] 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.

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 $\mu_{4n}^* = n^{-1}\sum_{i=1}^n \tilde{\epsilon}_{ni}^4$ for the residual bootstrap, and $\sigma_n^{*2} = n^{-1}\hat{\epsilon}_n'\hat{\epsilon}_n$, $\mu_{3n}^* = 0$ and $\mu_{4n}^* = 3(n^{-1}\hat{\epsilon}_n'\hat{\epsilon}_n)^2$ for the parametric bootstrap with normal disturbances. We assume the following conditions on the components of c_n/σ_{c_n}:

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

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

Assumption 3. *The sequence $\{\sigma_{c_n}^2\}$ 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, 2004a), 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 of c_n/σ_{c_n} to that of a standard normal variable as subsequently shown. As in Kelejian and Prucha (2001), we write c_n as 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.

2.3. Uniform Convergence of CDFs

Let $\Phi(x)$ be the CDF for a standard normal random variable, and P^* and E^* be, respectively, the probability distribution and expectation induced by the bootstrap sampling process. Under Assumption 2, let K_a and K_b be finite constants such that, for any n,

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

Suppose that t_n is a statistic that can be approximated by a standardized LQ form c_n/σ_{c_n} with c_n given in (5) such that $d_n = t_n - c_n/\sigma_{c_n}$ converges to zero in probability. Correspondingly, $d_n^* = t_n^* - c_n^*/\sigma_{c_n}^*$, where, as described earlier, t_n^*, c_n^* and $\sigma_{c_n}^*$ are bootstrapped quantities of, respectively, t_n, c_n and σ_{c_n}.

Theorem 1. *Under Assumptions 1–3,*

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

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

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

$$\sup_{x\in\mathbb{R}} \left|\mathrm{P}^*\left(c_n^*/\sigma_{c_n}^* + d_n^* \le x\right) - \Phi(x)\right| \le r_n^* + (2\pi)^{-1/2}\tau + \mathrm{P}^*(|d_n^*| > \tau). \tag{9}$$

where τ is an arbitrary positive number, $r_n = K\sigma_{c_n}^{-2(1+\delta_1)/(3+2\delta_1)} n^{-\delta_1/(3+2\delta_1)}\Big((K_a+1)^{1+2\delta_1}\big(K_a \operatorname{E}|\epsilon_{ni}^2 - \sigma_0^2|^{2+2\delta_1} + 2^{2+2\delta_1} K_a(\operatorname{E}|\epsilon_{ni}|^{2+2\delta_1})^2 + K_b \operatorname{E}|\epsilon_{ni}|^{2+2\delta_1}\big) + 4^{1+\delta_1}\big(\sigma_0^4 K_a^4(\mu_4 - \sigma_0^4) + 4\sigma_0^8 K_a^4 + \sigma_0^2 K_a^2(\mu_3^2 K_a + \sigma_0^4 K_b)(K_a+1) + 2|\mu_3|\sigma_0^2 K_a^3(|\mu_3| K_a + \sigma_0^2 K_b)\big)^{(1+\delta_1)/2}\Big)^{1/(3+2\delta_1)}$ *with* $\delta_1 = \min\{\delta, 1\}$ *and K being a constant depending only on δ_1, and r_n^* is a term obtained from replacing the population moment parameters of ϵ_{ni} in r_n with the corresponding sample moments of $\tilde{\epsilon}_n$.*

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, $O(n^{-\delta_1/(3+2\delta_1)})$, of the CDF of c_n/σ_{c_n} to that of a standard normal random variable. The larger δ_1 is, i.e., the higher moments of ϵ_{ni} assumed to exist, the faster is the convergence. The expression for r_n can be simplified if we let K_a and K_b be $\max\{K_a, K_b\}$. The different K_a and K_b allow for better approximation of the constant factor of the rate of convergence. A similar result for the bootstrapped version of c_n/σ_{c_n} is given in (7). The results in (8) and (9) are shown by using (6) and (7). They imply immediately that

$$\sup_{x\in\mathbb{R}} |\operatorname{P}^*(t_n^* \le x) - \operatorname{P}(t_n \le x)| \le r_n + \operatorname{P}(|d_n| > \tau) + r_n^* + \operatorname{P}^*(|d_n^*| > \tau) + \sqrt{\frac{2}{\pi}}\tau. \tag{10}$$

As τ is arbitrary, to prove consistency of the bootstrapped t_n, we may show that, except for the last one, the r.h.s. terms converge to zero (in probability) for any $\tau > 0$. This type of convergence with respect to the Kolmogorov-Smirnov distance implies asymptotic consistency of confidence intervals. If we can show that the sample moments of $\tilde{\epsilon}_n$ converge in probability to finite constants, then the continuous mapping theorem implies that r_n^* is of order $O_P(n^{-\delta_1/(3+2\delta_1)})$. The remainder term d_n is assumed to converge to zero in probability, so it only remains to show that $\operatorname{P}^*(|d_n^*| > \tau) = o_P(1)$.

For statistics with non-unit asymptotic variances, e.g., various estimators, we may rescale those statistics and apply (6) and (7) for the proof of consistency. Let $t_n e_n$ be a statistic of interest, where e_n is a positive non-stochastic term that may depend on n, θ_0 and moment parameters of ϵ_{ni}. The bootstrapped statistic corresponding to $t_n e_n$ is $t_n^* e_n^*$, where 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 $\tilde{\epsilon}_n$. For any $\eta > 0$,

$$\begin{aligned}
&\operatorname{P}\Big(\sup_{x\in\mathbb{R}}\big|\operatorname{P}^*(t_n^* e_n^* \le x) - \operatorname{P}(t_n e_n \le x)\big| > \eta\Big) \\
&= \operatorname{P}\Big(\sup_{x\in\mathbb{R}}\big|\mathbf{1}(e_n^* > 0)\big(\operatorname{P}^*(t_n^* \le x/e_n^*) - \operatorname{P}(t_n \le x/e_n)\big) + \mathbf{1}(e_n^* \le 0)\big(\operatorname{P}^*(t_n^* e_n^* \le x) - \operatorname{P}(t_n e_n \le x)\big)\big| > \eta\Big) \\
&\le \operatorname{P}\Big(\sup_{x\in\mathbb{R}}\big|\big(\operatorname{P}^*(t_n^* \le x/e_n^*) - \Phi(x/e_n^*)\big) - \big(\operatorname{P}(t_n \le x/e_n) - \Phi(x/e_n)\big) + \big(\Phi(x/e_n^*) - \Phi(x/e_n)\big)\big| > \eta/2\Big) \\
&\quad + \operatorname{P}\big(\mathbf{1}(e_n^* \le 0) > \eta/4\big) \\
&\le \operatorname{P}\Big(r_n + \operatorname{P}(|d_n| > \tau) + r_n^* + \operatorname{P}^*(|d_n^*| > \tau) + \sqrt{\frac{2}{\pi}}\tau + \sup_{x\in\mathbb{R}}\big|\Phi(x/e_n) - \Phi(x/e_n^*)\big| > \eta/2\Big) + \operatorname{P}(e_n^* \le 0),
\end{aligned} \tag{11}$$

where the last inequality follows by (8) and (9), and $\mathbf{1}(\cdot)$ is the set indicator. With consistent estimators for the parameters in e_n, we would have $\mathrm{P}(e_n^* \leq 0) = o(1)$. Then, for the proof of consistency, it remains to show that the first term in the last line of (11) tends to zero as n goes to infinity.

2.4. Consistency of the Bootstrapped Moran's I

Now we apply Theorem 1 to show consistency of the bootstrapped Moran's I. We make the following assumptions:

Assumption I1. *The ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})'$ are i.i.d. $(0, \sigma_0^2)$ and $\mathrm{E}|\epsilon_{ni}|^{4(1+\delta)} < \infty$ for some $\delta > 0$.*

Assumption I2. *The matrices $\{M_n\}$ have zero diagonals and are bounded in both row and column sum norms.*

Assumption I3. *The elements of the regressor matrices $\{X_n\}$ are uniformly bounded constants, and $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ exists and is nonsingular.*

Assumption I4. *The sequence $\{(2n)^{-1}\sigma_0^4 \operatorname{tr}[(M_n + M_n')^2] + n^{-1}(\mu_4 - 3\sigma_0^4)\sum_{i=1}^n (H_n M_n H_n)_{ii}^4\}$ is bounded away from zero.*

Assumption I4 requires the variance of $n^{-1/2}\epsilon_n' H_n M_n H_n \epsilon_n$ to be bounded away from zero. When ϵ_{ni}'s are normal, the relevant sequence in the assumption is simplified to be $\{n^{-1} \operatorname{tr}[(M_n + M_n')^2]\}$. Let $\mathbb{I}_n^*$ and $\mathbb{I}_n'^*$ be, respectively, the bootstrapped $\mathbb{I}_n$ and $\mathbb{I}_n'$ described earlier.

Proposition 1. *Under H_0 and Assumptions I1–I4, the Moran I statistic $\mathbb{I}_n$ in (2) satisfies $\sup_{x\in\mathbb{R}} |\mathrm{P}^*(\mathbb{I}_n^* \leq x) - \mathrm{P}(\mathbb{I}_n \leq x)| = o_P(1)$ for the parametric bootstrap when $\epsilon_n \sim N(0, \sigma_0^2 I_n)$, and $\mathbb{I}_n'$ in (4) satisfies $\sup_{x\in\mathbb{R}} |\mathrm{P}^*(\mathbb{I}_n'^* \leq x) - \mathrm{P}(\mathbb{I}_n' \leq x)| = o_P(1)$ for the residual bootstrap.*

3. Asymptotic Refinements

The bootstrap can provide asymptotic refinements for many statistics whose asymptotic distributions do not depend on any unknown parameters, i.e., asymptotically pivotal statistics. This is usually shown by Edgeworth expansions. For a smooth function of sample averages of independent random vectors and/or stationary dependent random vectors, the Edgeworth expansion has been well established. For statistics in spatial econometric models, possible asymptotic refinements may be studied by means of asymptotic expansions for LQ forms. However, an LQ form involves spatial weights matrices and cannot be written as simple sample averages of disturbances. In addition, note that

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

The first term on the r.h.s. of the above equation can be regarded as a generalized U-statistic with kernels $h_{n,ij}(x, y) = a_{n,ij}xy$. Even for this term, with normalization by its standard deviation, the Edgeworth expansion with a remainder term of desirable order has not been established. The second and third terms add further complication. In this section, we establish the Edgeworth expansion for an LQ form with normal disturbances, and Mykland's asymptotic expansion on martingales with non-normal disturbances. The results are then applied to show the second order correctness of the bootstrap.

3.1. Normal Disturbances

With normal disturbances in an LQ form $c_n/\sigma_{c_n} = n^{-1/2}\big(\epsilon_n' A_n \epsilon_n - \sigma_0^2 \operatorname{tr}(A_n) + b_n'\epsilon_n\big)/\sigma_{c_n}$, we can easily derive the characteristic function of c_n/σ_{c_n}. By a smoothing inequality in Feller (1970), the difference between a CDF and another function with certain properties has an upper bound generated from the Fourier transforms of the derivatives of the two functions. Such an inequality has been used to establish the Berry-Esseen bound for the error in the approximation by a normal distribution or an Edgeworth expansion to the true distribution for a sample mean of i.i.d. disturbances. As we shall show, it can also be used to establish the Edgeworth expansion of an LQ form. Let $f^{(k)}(x)$ be the kth order derivative of a function $f(x)$. We can use the boundedness in both row and column sum norms of the matrix A_n to bound the derivatives of the characteristic function or cumulants for an LQ form. As usual, the Edgeworth expansion is expressed in terms of cumulants, the normal CDF $\Phi(x)$ and the normal density function $\Phi^{(1)}(x)$. In particular, note that the third cumulant of the distribution of c_n/σ_{c_n} is $6\kappa_n = n^{-3/2}\sigma_{c_n}^{-3}[8\sigma_0^6 \operatorname{tr}(A_n^3) + 6\sigma_0^4 b_n' A_n b_n]$ with $\sigma_{c_n}^2 = n^{-1}[2\sigma_0^4 \operatorname{tr}(A_n^2) + \sigma_0^2 b_n' b_n]$. The c_n/σ_{c_n} has mean zero, then its third moment is equal to the third cumulant $6\kappa_n$. As the disturbances are normal, we consider the parametric bootstrap where the elements of ϵ_n^* are drawn from the normal distribution with mean zero and variance $n^{-1}\hat{\epsilon}_n'\hat{\epsilon}_n$.

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

$$\sup_{x\in\mathbb{R}} |\mathrm{P}(c_n/\sigma_{c_n} \le x) - [\Phi(x) + \kappa_n(1-x^2)\Phi^{(1)}(x)]| = O(n^{-1}), \tag{12}$$

and

$$\sup_{x\in\mathbb{R}} |\mathrm{P}^*(c_n^*/\sigma_{c_n}^* \le x) - [\Phi(x) + \kappa_n^*(1-x^2)\Phi^{(1)}(x)]| = O_P(n^{-1}), \tag{13}$$

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* $r \ge 3$,

$$\sup_{x\in\mathbb{R}} \Big|\mathrm{P}(c_n/\sigma_{c_n} \le x) - \Phi(x) - \Phi^{(1)}(x)\sum_{i=3}^{r} n^{-(i-2)/2} P_{ni}(x)\Big| = O(n^{-(r-1)/2}), \tag{14}$$

and

$$\sup_{x\in\mathbb{R}} \Big|\mathrm{P}^*(c_n^*/\sigma_{c_n}^* \le x) - \Phi(x) - \Phi^{(1)}(x)\sum_{i=3}^{r} n^{-(i-2)/2} P_{ni}^*(x)\Big| = O_P(n^{-(r-1)/2}), \tag{15}$$

where $P_{n3}(x), \ldots, P_{nr}(x)$ *are real polynomials with bounded coefficients and* $P_{n3}^*(x), \ldots, P_{nr}^*(x)$ *are the corresponding bootstrapped versions.*

Eqs. (12) and (13) can be used to show that the bootstrap can provide asymptotic refinements for statistics that can be approximated by an LQ form. Eqs. (14) and (15) present general high order Edgeworth expansions for the CDF of an LQ form. Note that κ_n has a relatively simple form. Instead of bootstrapping test statistics, we may correct the bias distortion for test statistics that can be approximated by an LQ form.[12] The above theorem can be applied to show that the bootstrapped Moran's I is often more accurate than the first-order asymptotic approximation. While the approximation of asymptotic theory to the true distribution has the order $O_P(n^{-1/2})$, the following proposition shows that approximation of the bootstrap has the smaller order $O_P(n^{-1})$:

Proposition 2. *Under H_0 and Assumptions I1–I4, the Moran I statistic $\mathbb{I}_n$ with normal disturbances in (2) satisfies $\mathrm{P}^*(\mathbb{I}_n^* \le x) - \mathrm{P}(\mathbb{I}_n \le x) = O_P(n^{-1})$ for the parametric bootstrap.*

3.2. Non-normal Disturbances

For LQ forms with non-normal disturbances, a theorem on an asymptotic expansion for martingales in Mykland (1993) can be applied to establish an expansion, which is called the Edgeworth expansion for martingales by the author. Mykland (1993) considers an asymptotic expansion for the distribution function $F_n(x)$ of a triangular array of normalized zero-mean martingales. The conditions needed are integrability and central limit conditions imposed on the normalizing factor and variation measures associated with martingales, the optional and predictable kth-order variations.[13] Under those conditions, the distribution function $F_n(x)$ has an expansion of the form

$$\int_{-\infty}^{\infty} h(x)\,dF_n(x) = \int_{-\infty}^{\infty} h(x)\,d\Phi(\beta^{-1}x) + \frac{1}{2}r_n \int_{-\infty}^{\infty} [\beta^2\psi(x)h''(\beta x) - \beta x\psi_*(x)h'(\beta x)]\,d\Phi(x) + o(r_n), \quad (16)$$

where β is a parameter, r_n is a nonstochastic sequence going to zero, and $\psi(x)$ and $\psi_*(x)$ are defined in the central limit condition, uniformly over large classes of twice differentiable functions h. Subject to some minimum niceness on the part of ψ and ψ_*, the second term on the r.h.s. of (16) can be shown via integration by parts to equal

$$\frac{1}{2}r_n \int_{-\infty}^{\infty} h(x)d\Big[\big(\psi^{(1)}(\beta^{-1}x) - \psi(\beta^{-1}x)\beta^{-1}x + \psi_*(\beta^{-1}x)\beta^{-1}x\big)\Phi^{(1)}(\beta^{-1}x)\Big].$$

If $o_2(r_n)$ is used to denote the kind of convergence in (16), then a more standard way of stating an expansion is[14]

$$F_n(x) = \Phi(\beta^{-1}x) + \frac{1}{2}r_n\big(\psi^{(1)}(\beta^{-1}x) - \psi(\beta^{-1}x)\beta^{-1}x + \psi_*(\beta^{-1}x)\beta^{-1}x\big)\Phi^{(1)}(\beta^{-1}x) + o_2(r_n),$$

[12] 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.

[13] They are defined as, respectively, the sum of the kth powers of the martingale differences and the sum of expected values of the kth powers conditional on the filtration.

[14] Please refer to Theorem A in Appendix D for more details.

due to Mykland (1993).

In the current case, c_n/σ_{c_n} can be decomposed as a sum of martingale differences that are quadratic in the disturbances. We need the existence of $\mathrm{E}\,|\epsilon_{ni}|^{4(1+\eta)}$ for some $\eta > 0$ to show asymptotic normality of c_n/σ_{c_n}, by means of a central limit theorem for martingales. For the central limit theorem relating to the optional square variations, higher moments for ϵ_{ni} are required to exist. Furthermore, a slightly stronger condition on b_n is also assumed.

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

Assumption 2′. *The symmetric matrices of the sequence $\{A_n\}$ 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*

$$\int_{-\infty}^{+\infty} h(x)\, dF_n(x) = \int_{-\infty}^{+\infty} h(x)\, d\Phi(x) + \frac{1}{2} n^{-1/2} \int_{-\infty}^{+\infty} h(x)\, d\big[\big(\psi^{(1)}(x) - \psi(x)x\big)\Phi^{(1)}(x)\big] + o(n^{-1/2}), \quad (17)$$

and

$$\int_{-\infty}^{+\infty} h(x)\, dF_n^*(x) = \int_{-\infty}^{+\infty} h(x)\, d\Phi(x) + \frac{1}{2} n^{-1/2} \int_{-\infty}^{+\infty} h(x)\, d\big[\big(\psi^{(1)}(x) - \psi(x)x\big)\Phi^{(1)}(x)\big] + o_P(n^{-1/2}), \quad (18)$$

where $F_n(x) = \mathrm{P}(c_n/\sigma_{c_n} \le x)$, $F_n^*(x) = \mathrm{P}^*(c_n^*/\sigma_{c_n}^* \le x)$, and $\psi(x) = \frac{1}{3}\psi_o(x) + \frac{2}{3}\psi_p(x)$ with $\psi_o(x)$ and $\psi_p(x)$ being linear functions given in (D.11)–(D.14), uniformly on the 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 (17) by $o_2(n^{-1/2})$ (Mykland, 1993), then

$$F_n(x) = \Phi(x) + \frac{1}{2} n^{-1/2} \big(\psi^{(1)}(x) - \psi(x)x\big)\Phi^{(1)}(x) + o_2(n^{-1/2}), \quad (19)$$

and

$$F_n^*(x) = \Phi(x) + \frac{1}{2} n^{-1/2} \big(\psi^{(1)}(x) - \psi(x)x\big)\Phi^{(1)}(x) + o_2(n^{-1/2}) \text{ in probability.} \quad (20)$$

As pointed out by Mykland (1993), the expansion as in (17) generally does not hold for an indicator function h of an interval. So (19) is a "smoothed" expansion. Note that $\psi_o(x)$ and $\psi_p(x)$ are linear in x, hence $\psi^{(1)}(x) - \psi(x)x = (1 - x^2)\psi^{(1)}(x)$. In the special case that ϵ_{ni}'s are i.i.d. normal, we can verify that $\frac{1}{2}\psi^{(1)}(x) = \lim_{n\to\infty} n^{1/2}\kappa_n$, thus (19) has similar terms as the usual one-term Edgeworth expansion (12). With this theorem, we can show the second-order correctness with the order $o_2(n^{-1/2})$ of the bootstrap for Moran's I with non-normal disturbances.

Assumption I1′. *The ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \dots, \epsilon_{nn})'$ are i.i.d. $(0, \sigma_0^2)$ and $\mathrm{E}\,|\epsilon_{ni}|^{8(1+\delta)} < \infty$ for some $\delta > 0$.*

Proposition 3. *Under H_0 and Assumptions I1′, I2–I4, the Moran I statistic $\mathbb{I}'_n$ in (4) satisfies*

$$\int_{-\infty}^{+\infty} h(x)\, d[F_n^*(x) - F_n(x)] = o_P(n^{-1/2}), \tag{21}$$

where $F_n^*(x) = \mathrm{P}^*(\mathbb{I}_n^{'*} \le x)$ and $F_n(x) = \mathrm{P}(\mathbb{I}'_n \le x)$, uniformly for the set of functions h in Theorem 3. That is, $F_n^*(x) = F_n(x) + o_2(n^{-1/2})$ in probability, in the sense that (21) holds.

4. Monte Carlo Study

We conduct some Monte Carlo experiments to compare the finite sample performance of the bootstrap and first order asymptotic theory for Moran's I test. The data generating process (DGP) is either the SAR model

$$y_n = \rho M_n y_n + X_n\beta + \epsilon_n, \tag{22}$$

or the SE model

$$y_n = X_n\beta + u_n, \quad u_n = \rho M_n u_n + \epsilon_n, \tag{23}$$

where the exogenous variable matrix X_n consists of an intercept term, a variable generated from the standard normal distribution and a variable generated from the uniform distribution $U[0,1]$. Several spatial weights matrices are considered: the first one, denoted by "bd49", is a block diagonal matrix with the diagonal blocks being the continuity matrix for 49 neighborhoods in Columbus, OH from Anselin (1988); the second and third ones are generated according to, respectively, the queen and rook criteria; several "circular" world matrices as in Kelejian and Prucha (1999) are also considered. In a "circular" world matrix, each spatial unit is related to k spatial units immediately before it and k immediately after it for some positive integer k, and all nonzero elements in a row are equal. As k increases, the density of a "circular" world matrix increases. We consider $k = 1, 3, 5$ and denote the corresponding matrices by "cir2", "cir6" and "cir10", respectively. Every spatial weights matrix is row normalized to have row sum equal to 1. The sample sizes considered are $n = 49$, 98, 147 and 196.[15] Two designs of disturbances are considered: in the first one, the disturbances are randomly drawn from the standard normal distribution and the Moran I statistic is computed using (2); in the second one, the disturbances are randomly drawn from the normalized chi-square distribution $(\chi^2(3) - 3)/\sqrt{6}$ and the Moran I statistic is computed using (4) with $c_\sigma = 0.0001$. The parameter ρ in the SAR and SE models takes values 0, 0.1, 0.3, 0.5, 0.7, −0.1, −0.3, −0.5, and −0.7. The nominal level of significance for Moran's I test is set to 5%. The number of Monte Carlo repetitions is 5000 and the number of bootstrap samples is 399.

[15] For $n = 98$, the second and third matrices are first generated on a 10×10 lattice, then the last 2 rows are deleted from these matrices. Similarly, for $n = 147$, the second and third matrices are first generated on a 13×13 lattice, then the last 22 rows are deleted.

Table 1: Empirical size and power of the asymptotic and bootstrap tests with normal disturbances

		size	power (SAR model)								power (SE model)							
	ρ =	0.0	0.1	0.3	0.5	0.7	−0.1	−0.3	−0.5	−0.7	0.1	0.3	0.5	0.7	−0.1	−0.3	−0.5	−0.7
	n=49																	
bd49	asym	4.1	5.8	28.7	71.9	96.5	6.1	19.1	41.6	66.2	5.7	26.0	66.6	94.1	6.4	21.2	51.0	78.5
	bs	5.0	8.0	34.2	75.2	97.2	6.6	19.8	42.4	66.0	7.7	30.5	70.9	95.2	6.8	21.4	51.1	78.4
rook	asym	4.5	4.8	24.6	66.2	94.3	9.4	39.7	79.6	98.1	5.4	24.4	64.0	93.1	9.2	39.4	78.7	98.0
	bs	4.9	7.7	32.6	73.8	96.4	7.9	34.3	75.4	97.3	7.9	32.6	72.1	95.5	7.6	34.1	74.8	97.2
queen	asym	3.7	5.3	20.1	58.7	89.9	4.9	11.9	27.2	44.5	5.0	19.8	50.8	85.0	4.7	12.7	31.0	56.3
	bs	5.2	7.2	24.8	64.0	91.9	6.1	13.5	29.2	47.1	7.1	24.4	55.9	87.5	5.7	14.4	32.7	59.0
cir2	asym	4.7	6.7	40.9	86.8	99.4	11.7	53.2	92.1	99.7	8.0	42.9	89.3	99.6	12.3	54.9	93.9	99.9
	bs	4.9	9.2	48.8	90.5	99.6	10.3	48.2	90.4	99.6	10.6	49.1	92.0	99.8	10.3	50.1	92.0	99.8
cir6	asym	3.8	5.3	23.8	64.1	94.4	4.2	9.9	20.2	34.4	5.2	22.5	57.1	89.4	3.8	10.4	25.0	47.1
	bs	5.2	7.1	27.4	67.3	95.0	6.2	13.3	24.7	40.1	6.9	26.0	60.4	90.7	5.4	14.1	30.4	52.8
cir10	asym	2.5	4.5	15.3	44.7	79.5	2.5	3.9	7.0	11.0	4.3	14.6	37.8	71.4	2.1	4.0	9.5	14.9
	bs	4.8	6.8	19.0	49.0	82.1	5.1	8.4	14.0	20.2	6.6	17.9	42.3	74.8	4.8	9.1	17.3	27.4
	n=98																	
bd49	asym	4.3	9.9	58.3	96.4	100.0	8.7	40.2	76.2	94.3	9.4	54.3	94.1	99.9	9.0	43.5	85.2	98.8
	bs	5.0	10.9	61.2	96.9	100.0	8.9	39.2	74.8	93.7	10.8	57.3	94.7	99.9	9.2	42.7	84.2	98.8
rook	asym	4.4	8.4	53.1	94.6	99.9	13.6	63.4	97.2	100.0	8.3	51.2	94.5	99.9	12.7	64.5	96.6	100.0
	bs	4.5	10.7	58.6	95.9	100.0	11.3	58.5	96.4	100.0	10.5	56.9	95.9	99.9	10.7	59.3	95.9	100.0
queen	asym	4.5	7.6	40.2	88.0	99.7	7.0	25.3	53.4	79.5	7.1	37.4	81.9	98.8	7.3	27.5	63.6	89.4
	bs	5.1	9.2	43.4	89.4	99.7	7.7	25.5	53.8	79.3	8.5	40.3	83.3	99.0	7.7	27.8	62.8	89.2
cir2	asym	4.7	12.5	78.7	99.7	100.0	17.4	83.4	99.9	100.0	12.7	79.9	99.7	100.0	17.6	85.3	99.9	100.0
	bs	4.8	15.5	81.6	99.8	100.0	15.1	81.0	99.8	100.0	15.2	82.5	99.8	100.0	15.6	82.6	99.9	100.0
cir6	asym	4.2	8.2	48.9	93.7	100.0	6.6	22.1	48.0	73.1	9.0	44.0	89.0	99.7	6.4	26.3	61.5	88.6
	bs	4.7	9.2	50.7	93.8	100.0	7.5	24.0	49.7	74.9	10.1	45.8	89.5	99.7	7.4	28.3	63.8	89.7
cir10	asym	3.7	7.1	34.0	81.2	98.5	3.7	10.5	22.3	35.8	6.8	30.3	71.0	96.5	3.8	11.8	29.1	54.6
	bs	5.1	8.4	36.1	82.3	98.6	5.8	14.3	28.3	43.6	8.1	32.3	72.3	96.8	5.7	16.3	36.7	61.9
	n=147																	
bd49	asym	4.4	12.2	75.4	99.8	100.0	12.3	57.3	91.7	99.2	12.0	72.1	99.3	100.0	12.5	62.6	96.5	100.0
	bs	4.7	13.4	76.7	99.8	100.0	11.7	56.5	91.2	99.2	13.5	73.7	99.2	100.0	11.9	61.2	96.0	99.9
rook	asym	4.6	11.9	72.7	99.4	100.0	17.0	79.4	99.7	100.0	11.7	72.0	99.2	100.0	16.4	78.6	99.7	100.0
	bs	4.5	14.5	76.6	99.5	100.0	14.7	76.2	99.5	100.0	14.4	75.5	99.5	100.0	14.0	75.7	99.5	100.0
queen	asym	4.2	9.6	58.0	96.6	100.0	7.1	36.7	74.1	92.8	9.8	51.6	94.1	100.0	8.3	41.8	81.8	98.1
	bs	4.6	10.7	60.2	96.7	100.0	7.7	36.3	73.6	92.3	10.9	53.8	94.4	100.0	8.3	41.4	81.0	97.9
cir2	asym	5.1	18.8	93.3	100.0	100.0	23.6	95.2	100.0	100.0	19.7	94.0	100.0	100.0	23.9	95.9	100.0	100.0
	bs	5.2	21.6	94.4	100.0	100.0	21.7	94.0	100.0	100.0	22.7	94.8	100.0	100.0	21.9	95.2	100.0	100.0
cir6	asym	4.6	10.8	67.1	99.0	100.0	7.7	37.2	70.4	90.3	10.7	60.8	97.4	100.0	7.9	43.2	83.7	98.5
	bs	5.1	12.1	68.0	99.0	100.0	8.4	38.2	70.9	90.5	11.7	61.7	97.4	100.0	8.6	44.3	83.8	98.5
cir10	asym	3.8	8.4	48.9	93.3	100.0	5.2	18.2	37.4	58.6	8.8	43.3	88.5	99.5	5.6	21.1	51.2	79.8
	bs	4.9	9.2	49.6	93.4	100.0	6.7	21.5	42.1	63.0	9.8	44.0	88.7	99.6	7.2	24.4	55.7	82.8
	n=196																	
bd49	asym	4.4	15.6	87.5	100.0	100.0	14.7	70.7	97.4	100.0	14.7	84.6	99.9	100.0	15.6	76.8	99.6	100.0
	bs	4.8	17.1	87.8	100.0	100.0	14.2	69.7	97.3	100.0	16.3	85.4	99.9	100.0	15.0	75.4	99.4	100.0
rook	asym	4.6	13.9	84.6	100.0	100.0	19.9	89.3	100.0	100.0	16.2	84.1	99.9	100.0	19.0	89.2	100.0	100.0
	bs	4.7	16.2	86.4	100.0	100.0	17.0	87.2	99.9	100.0	18.5	85.9	99.9	100.0	17.0	87.4	100.0	100.0
queen	asym	4.6	11.1	70.5	99.2	100.0	9.3	47.7	85.4	98.3	11.7	66.6	98.0	100.0	9.7	53.5	91.7	99.7
	bs	4.9	12.2	71.6	99.4	100.0	9.8	47.0	85.0	97.9	12.6	68.1	98.2	100.0	10.1	52.6	91.5	99.7
cir2	asym	5.2	25.4	98.1	100.0	100.0	30.2	98.8	100.0	100.0	25.3	98.4	100.0	100.0	31.0	98.7	100.0	100.0
	bs	5.4	27.8	98.5	100.0	100.0	27.6	98.3	100.0	100.0	27.6	98.7	100.0	100.0	28.3	98.5	100.0	100.0
cir6	asym	5.0	14.1	78.9	99.9	100.0	9.7	46.9	86.1	96.9	12.9	74.6	99.5	100.0	9.9	55.9	93.9	99.8
	bs	5.4	14.8	79.3	99.9	100.0	10.0	48.1	86.1	97.1	13.4	75.3	99.5	100.0	10.4	55.9	93.5	99.8
cir10	asym	4.4	10.6	61.8	98.2	100.0	7.2	26.1	53.1	75.2	10.2	54.3	95.2	100.0	5.9	28.8	68.2	92.0
	bs	4.9	10.9	62.5	98.1	100.0	8.5	29.1	56.6	77.2	10.8	55.2	95.2	100.0	7.4	31.8	71.2	93.3

"asym" and "bs" mean that the rejection probabilities reported are computed using the asymptotic and bootstrap critical values respectively. The sign "%" is omitted. The first column shows the spatial weights matrices in the DGP: "bd49" means a block diagonal matrix with each diagonal block being the continuity matrix for 49 neighborhoods in Columbus, OH; "rook" and "queen" mean the matrices generated according to the rook and queen criteria respectively; "cir2", "cir6" and "cir10" mean the "circular" world matrices with, respectively, 2, 6 and 10 nonzero elements in each row.

Table 2: Empirical size and power of the asymptotic and bootstrap tests with chi-square disturbances

		size	power (SAR model)								power (SE model)							
	ρ =	0.0	0.1	0.3	0.5	0.7	−0.1	−0.3	−0.5	−0.7	0.1	0.3	0.5	0.7	−0.1	−0.3	−0.5	−0.7
	n=49																	
bd49	asym	5.3	9.3	38.3	80.2	98.2	5.6	16.0	39.5	63.1	8.8	36.1	76.0	96.9	5.4	18.9	48.0	76.8
	bs	5.0	6.9	31.6	75.2	97.5	6.7	19.9	44.2	66.8	7.0	29.9	69.6	95.2	6.6	23.2	53.2	79.6
rook	asym	5.3	8.1	34.2	76.9	97.1	8.8	37.9	79.9	97.8	8.2	34.8	76.1	97.2	8.3	37.9	79.1	97.9
	bs	5.0	6.9	30.6	73.7	96.3	8.5	37.1	78.2	97.4	6.9	31.7	73.3	96.7	8.4	37.0	77.9	97.7
queen	asym	5.9	9.2	30.9	70.5	94.1	4.6	10.1	24.2	41.7	8.9	28.5	63.4	90.8	5.2	11.5	28.7	54.3
	bs	5.9	7.1	24.7	63.4	92.1	5.8	14.6	30.2	48.2	6.4	22.5	56.7	87.8	6.3	16.1	35.8	61.0
cir2	asym	5.4	11.1	51.6	92.3	99.7	10.0	53.1	91.2	99.5	10.4	51.6	94.7	100.0	11.2	54.0	93.9	99.8
	bs	4.9	9.0	46.5	90.5	99.5	10.4	53.5	90.4	99.4	8.6	46.7	93.0	100.0	11.9	53.9	93.5	99.8
cir6	asym	5.7	8.6	35.5	75.6	96.9	4.6	7.9	17.1	30.7	9.1	32.4	67.9	94.6	4.6	8.5	23.3	44.9
	bs	5.1	6.8	28.7	70.0	95.5	6.3	13.5	26.4	40.4	6.7	25.4	61.0	92.2	6.1	14.2	33.6	56.3
cir10	asym	5.0	8.6	26.5	59.0	88.0	3.0	3.7	5.4	10.6	8.6	23.6	50.8	79.4	2.9	3.6	7.5	14.9
	bs	4.7	6.9	20.4	51.4	83.8	4.4	8.4	14.3	21.8	6.6	17.0	42.2	73.9	5.1	8.7	17.7	30.1
	n=98																	
bd49	asym	4.6	11.9	63.6	98.3	100.0	7.3	37.4	73.3	93.6	12.5	61.2	96.8	100.0	8.0	41.8	83.1	98.3
	bs	4.7	9.7	57.4	97.5	100.0	9.0	42.3	76.8	94.4	10.1	55.2	95.4	99.9	10.1	47.1	86.0	98.6
rook	asym	5.3	12.0	60.9	96.9	100.0	12.4	61.0	96.5	100.0	11.9	59.1	96.9	100.0	11.1	63.7	96.3	100.0
	bs	5.0	11.0	57.5	96.4	100.0	13.1	61.2	96.4	99.9	10.6	56.0	96.1	100.0	11.8	64.1	96.4	100.0
queen	asym	4.4	10.1	49.5	92.2	99.8	5.3	21.5	49.3	76.8	10.4	46.3	88.4	99.5	6.4	23.9	60.1	87.9
	bs	4.6	8.2	44.0	90.2	99.7	7.3	27.6	55.6	80.6	8.7	40.8	84.8	99.3	8.4	30.0	65.5	90.4
cir2	asym	5.1	15.8	82.6	100.0	100.0	16.2	82.8	99.9	100.0	15.9	85.0	100.0	100.0	16.9	84.8	99.7	100.0
	bs	5.1	13.7	79.5	99.8	100.0	17.7	83.4	99.8	100.0	13.8	81.5	99.9	100.0	18.4	85.4	99.8	100.0
cir6	asym	5.4	12.3	56.5	95.7	100.0	5.2	18.7	45.0	68.4	12.2	52.4	92.7	99.9	5.3	22.6	58.6	86.5
	bs	5.1	9.8	50.1	94.6	99.9	7.4	25.7	54.1	74.9	9.5	45.2	90.2	99.8	7.7	30.6	67.5	90.4
cir10	asym	4.7	10.7	43.6	85.4	99.2	3.8	7.4	17.3	31.7	10.4	37.0	78.5	97.7	3.5	8.8	24.5	48.1
	bs	5.3	8.2	36.1	81.6	99.0	6.5	14.7	28.9	45.7	8.3	30.6	73.7	96.8	6.2	17.3	39.0	64.2
	n=147																	
bd49	asym	5.5	15.5	81.7	99.8	100.0	9.8	54.3	90.4	99.1	15.4	77.0	99.5	100.0	10.9	61.2	95.3	99.9
	bs	5.3	13.2	78.1	99.7	100.0	12.1	58.6	91.5	99.3	12.8	72.5	99.4	100.0	13.0	65.6	96.2	100.0
rook	asym	5.3	14.8	77.2	99.8	100.0	14.7	79.1	99.6	100.0	14.2	76.5	99.7	100.0	14.6	78.7	99.6	100.0
	bs	5.1	13.0	74.7	99.7	100.0	15.6	79.3	99.6	100.0	12.7	74.2	99.7	100.0	15.4	79.0	99.5	100.0
queen	asym	4.9	13.2	63.8	98.2	100.0	7.5	33.4	71.3	91.8	12.1	59.9	96.7	100.0	6.9	37.1	79.4	97.2
	bs	4.9	10.7	58.0	97.7	100.0	9.4	39.5	75.7	93.4	9.9	54.5	95.3	99.9	9.3	43.0	83.2	97.9
cir2	asym	5.6	21.7	95.7	100.0	100.0	22.8	94.8	100.0	100.0	21.5	96.1	100.0	100.0	23.6	94.3	100.0	100.0
	bs	5.6	19.8	94.6	100.0	100.0	24.7	95.1	100.0	100.0	18.8	94.8	100.0	100.0	25.5	94.7	100.0	100.0
cir6	asym	5.0	13.8	73.4	99.4	100.0	6.3	33.2	67.3	87.9	13.4	67.6	98.4	100.0	7.3	37.7	80.6	98.1
	bs	5.0	11.3	68.2	99.1	100.0	9.1	40.4	73.5	91.2	11.0	62.1	97.7	100.0	9.8	45.7	85.1	98.9
cir10	asym	5.0	11.6	56.5	96.1	99.9	4.4	14.2	32.8	51.9	11.6	50.9	91.5	99.8	3.6	16.3	45.6	75.1
	bs	5.3	9.0	50.0	94.2	99.9	6.7	23.2	44.3	63.5	9.8	44.4	89.2	99.7	5.9	24.8	57.9	84.4
	n=196																	
bd49	asym	5.0	19.3	89.8	100.0	100.0	13.2	68.5	96.7	99.9	17.9	87.9	99.9	100.0	12.8	75.1	99.2	100.0
	bs	5.2	16.2	87.3	100.0	100.0	15.6	71.5	97.3	99.9	15.1	84.8	99.9	100.0	15.5	77.7	99.3	100.0
rook	asym	4.9	17.6	87.2	100.0	100.0	17.7	87.4	100.0	100.0	17.6	86.7	99.9	100.0	17.2	88.0	100.0	100.0
	bs	4.9	16.1	85.4	99.9	100.0	19.1	87.6	100.0	100.0	16.4	85.0	99.9	100.0	17.8	88.1	100.0	100.0
queen	asym	5.5	15.0	75.1	99.6	100.0	8.0	45.1	81.8	97.1	14.2	70.5	98.8	100.0	8.2	49.1	90.9	99.5
	bs	5.3	12.7	70.7	99.3	100.0	10.6	50.8	84.7	97.5	11.9	66.1	98.4	100.0	10.5	54.3	92.7	99.7
cir2	asym	5.3	29.4	98.7	100.0	100.0	29.3	98.1	100.0	100.0	27.5	99.2	100.0	100.0	29.1	98.8	100.0	100.0
	bs	5.3	26.1	98.4	100.0	100.0	31.2	98.2	100.0	100.0	24.5	98.9	100.0	100.0	30.8	98.8	100.0	100.0
cir6	asym	4.8	16.9	84.8	99.9	100.0	7.7	44.7	81.5	96.1	16.1	78.0	99.6	100.0	8.8	51.3	92.1	99.7
	bs	5.0	13.7	81.0	99.9	100.0	10.7	51.3	85.1	97.0	13.2	73.1	99.5	100.0	11.8	57.9	93.9	99.7
cir10	asym	4.3	13.7	67.4	98.7	100.0	5.3	21.0	47.3	70.0	13.1	62.2	96.9	100.0	4.8	25.1	63.2	90.3
	bs	4.6	10.1	61.6	98.2	100.0	8.7	29.5	57.5	77.9	10.5	56.2	95.3	100.0	7.6	34.6	72.8	93.9

"asym" and "bs" mean that the rejection probabilities reported are computed using the asymptotic and bootstrap critical values respectively. The sign "%" is omitted. The first column shows the spatial weights matrices in the DGP: "bd49" means a block diagonal matrix with each diagonal block being the continuity matrix for 49 neighborhoods in Columbus, OH; "rook" and "queen" mean the matrices generated according to the rook and queen criteria respectively; "cir2", "cir6" and "cir10" mean the "circular" world matrices with, respectively, 2, 6 and 10 nonzero elements in each row.

Table 1 reports Monte Carlo results for the case with normal disturbances. For $n = 49$, the asymptotic tests, except the ones with "cir2" and "rook", significantly under-reject the true null hypotheses. For the real world matrix "bd49", the size distortion is 0.009. Between "rook" and "queen", the denser one, "queen", has a larger size distortion, being equal to 0.013. Similarly, among "cir2", "cir6" and "cir10", denser matrices have larger size distortions. The size distortion of "cir10" is as high as 0.025. For all the matrices considered, the empirical sizes of the bootstrap tests are all very close to the nominal 5%. For the empirical power, the asymptotic and bootstrap tests both have higher powers as ρ moves away from 0. There are gaps between the powers of the asymptotic and bootstrap tests, which are often large and can be larger than 10%. The gaps for denser matrices are not necessarily larger. For example, for $\rho = 0.3$, the gap is 7.9% for "cir2", but only 3.7% for "cir10". With larger sample sizes, the patterns of empirical sizes and powers remain the same. The empirical sizes of the bootstrap tests are still close to the nominal size for different sample sizes. When the sample sizes become larger, the size distortions of asymptotic tests become less severe, the powers of asymptotic and bootstrap tests become higher, and the gaps between the powers of the asymptotic and bootstrap tests become smaller. With a sample size of 196, the asymptotic tests almost have no size distortion.

The Monte Carlo results for the case with chi-square disturbances are reported in Table 2. For $n = 49$, the asymptotic tests generally over-reject the true null hypotheses, but the size distortions are not large, with the largest one being 0.009. With sample sizes larger than 49, the asymptotic tests have very small size distortions. The bootstrap tests have smaller size distortions than the asymptotic tests in almost all cases. The patterns of the empirical powers are similar to the case with normal disturbances.

5. Conclusion

In this paper, we consider the use of the bootstrap in spatial econometric models, with a focus on Moran's I test. We demonstrate 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 an LQ form to that of the standard normal random variable. Based on this result, we show that the bootstrap is generally consistent for statistics that can be approximated by an LQ form. In particular, we show that the bootstrap for Moran's I is consistent. Furthermore, we establish the Edgeworth expansion for LQ forms with normal disturbances and an asymptotic expansion for LQ forms with non-normal disturbances based on martingales. These results are applied to show the second order correctness of the bootstrap for Moran's I. Our Monte Carlo results show that, in finite samples, the empirical size of the bootstrapped Moran's I test is usually very close to the nominal level of significance and the bootstrap test generally has smaller size distortion than the asymptotic test.

The results in this paper can also be used to study the bootstrap for other statistics, e.g., the popular spatial J-type tests. 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.[16]

Appendix A. Lemmas

For generality, the following lemmas are proved for the popular SARAR model $y_n = \lambda W_n y_n + X_n\beta + u_n$, $u_n = \rho M_n u_n + \epsilon_n$, where W_n and M_n are $n \times n$ spatial weights matrices which may or may not be different, and ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})'$ are i.i.d. with mean zero, variance σ_0^2, third moment μ_3 and finite fourth moment μ_4. Let $\theta = (\lambda, \rho, \beta')'$, $S_n(\lambda) = I_n - \lambda W_n$ and $R_n(\rho) = I_n - \rho M_n$. The true parameter vector is θ_0. Denote $S_n = S_n(\lambda_0)$ and $R_n = R_n(\rho_0)$ for short. Assume that $S_n^{-1}(\lambda)$ and $R_n^{-1}(\rho)$ are bounded in both row and column sum norms uniformly in their compact parameter spaces, W_n and M_n are bounded in both row and column sum norms, and the matrix X_n satisfies regularity conditions as in the main text. Let $\hat{\epsilon}_n = R_n(\hat{\rho}_n)[S_n(\hat{\lambda}_n)y_n - X_n\hat{\beta}_n]$ with $\hat{\theta}_n$ being $n^{1/2}$-consistent, i.e., $n^{1/2}(\hat{\theta}_n - \theta_0) = O_P(1)$.[17] The ϵ_n^*, y_n^* and $\hat{\theta}_n^*$ are derived by the residual bootstrap or parametric bootstrap when $\epsilon_n \sim N(0, \sigma_0^2 I_n)$, as described in Section 2. Let $\|\cdot\|$ be the Euclidean vector norm.

Lemma 1. *Let $P_{ln} = [p_{ln,ij}]$ be $n \times n$ matrices which are bounded in row sum norms, for $l = 1, \ldots, s$. If $\mathrm{E}|\epsilon_{nj}|^s < \infty$, 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. For $s > 1$, there exists a finite r such that $\frac{1}{r} + \frac{1}{s} = 1$. Hölder's inequality implies that

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

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

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

[16] For a vector of LQ forms with non-normal disturbances, an asymptotic expansion based on martingales can be established by using the results in Mykland (1995).

[17] For some other spatial econometric models, e.g., the SMA or spatial Durbin model, the results in these lemmas can be similarly proved.

Hence,

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

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

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

Proof. We first consider the residual bootstrap. Let $J_n=I_n-\frac{1}{n}l_nl_n'$. As $y_n=S_n^{-1}(X_n\beta_0+R_n^{-1}\epsilon_n)$,

$$\begin{aligned}\tilde{\epsilon}_n&=J_n\hat{\epsilon}_n\\&=J_n\big(R_n+(\rho_0-\hat{\rho}_n)M_n\big)\big(S_ny_n-X_n\beta_0+(\lambda_0-\hat{\lambda}_n)W_ny_n+X_n(\beta_0-\hat{\beta}_n)\big)\\&=\epsilon_n-\frac{l_n'\epsilon_n}{n}l_n+(\lambda_0-\hat{\lambda}_n)\big(J_nR_n+(\rho_0-\hat{\rho}_n)J_nM_n\big)W_nS_n^{-1}X_n\beta_0\\&\quad+\big(J_nR_nX_n+(\rho_0-\hat{\rho}_n)J_nM_nX_n\big)(\beta_0-\hat{\beta}_n)\\&\quad+(\lambda_0-\hat{\lambda}_n)\big(J_nR_n+(\rho_0-\hat{\rho}_n)J_nM_n\big)W_nS_n^{-1}R_n^{-1}\epsilon_n+(\rho_0-\hat{\rho}_n)J_nM_nR_n^{-1}\epsilon_n.\end{aligned}\tag{A.1}$$

Write $\tilde{\epsilon}_n=\epsilon_n+\sum_{j=1}^r\zeta_{1n,j}p_{nj}+\sum_{j=1}^s\zeta_{2n,j}Q_{nj}\epsilon_n$, 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 $\tilde{\epsilon}_{ni}^r$ 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 $n^{-1}\sum_{i=1}^n\tilde{\epsilon}_{ni}^r-n^{-1}\sum_{i=1}^n\epsilon_{ni}^r$, 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. When $\mathrm{E}\,|\epsilon_{ni}|^r<\infty$, T_{2n} is either bounded or randomly bounded by Lemma 1. It follows that $n^{1/2}(n^{-1}\sum_{i=1}^n\tilde{\epsilon}_{ni}^r-n^{-1}\sum_{i=1}^n\epsilon_{ni}^r)=O_P(1)$. By the law of large numbers, $n^{-1}\sum_{i=1}^n\epsilon_{ni}^r-\mathrm{E}\,\epsilon_{ni}^r=o_P(1)$. Then $\mathrm{E}^*\,\epsilon_{ni}^{*r}=\mathrm{E}\,\epsilon_{ni}^r+o_P(1)$ for the residual bootstrap, when $\mathrm{E}\,|\epsilon_{ni}|^r<\infty$. By Chebyshev's inequality, $\mathrm{P}(|n^{1/2}(n^{-1}\sum_{i=1}^n\epsilon_{ni}^r-\mathrm{E}\,\epsilon_{ni}^r)|>\eta)=\mathrm{P}(|n^{-1/2}\sum_{i=1}^n(\epsilon_{ni}^r-\mathrm{E}\,\epsilon_{ni}^r)|>\eta)\leq\eta^{-2}\,\mathrm{E}[(\epsilon_{ni}^r-\mathrm{E}\,\epsilon_{ni}^r)^2]=O(\eta^{-2})$ for any $\eta>0$ when $\mathrm{E}\,\epsilon_{ni}^{2r}<\infty$. Thus, $n^{1/2}(n^{-1}\sum_{i=1}^n\epsilon_{ni}^r-\mathrm{E}\,\epsilon_{ni}^r)=O_P(1)$. Then $n^{1/2}[\mathrm{E}^*\,\epsilon_{ni}^{*r}-\mathrm{E}\,\epsilon_{ni}^r]=O_P(1)$ for the residual bootstrap, when $\mathrm{E}\,\epsilon_{ni}^{2r}<\infty$.

For the parametric bootstrap, $\mathrm{E}^*\,\epsilon_{ni}^{*r}=0$ when r is odd and $\mathrm{E}^*\,\epsilon_{ni}^{*r}=(n^{-1}\hat{\epsilon}_n'\hat{\epsilon}_n)^{r/2}(r-1)!!$ when r is even, where $(r-1)!!$ denotes the double factorial of $r-1$. By a similar argument as above, $n^{-1}\hat{\epsilon}_n'\hat{\epsilon}_n=\sigma_0^2+O_P(n^{-1/2})$ when $\mathrm{E}\,\epsilon_{ni}^4=3(\mathrm{E}\,\epsilon_{ni}^2)^2<\infty$. Then $\mathrm{E}^*\,\epsilon_{ni}^{*r}=\mathrm{E}\,\epsilon_{ni}^r+O_P(n^{-1/2})$ for the parametric bootstrap as long as $\mathrm{E}\,\epsilon_{ni}^2<\infty$. Other results are similarly derived. □

Lemma 3. *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 1 except for the application of Lemma 2. □

Lemma 4. *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 elements (or randomly bounded if $\hat\beta_n$ is involved), $Q_{nj} = [q_{nj,il}]$ 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 2 except for the application of Lemma 3. □

Appendix B. Proofs for Results in Section 2

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

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

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_1 with $0 < \delta_1 \le 1$ such that

$$\mathrm{E}|c_{ni}|^{2+2\delta_1} < \infty, \tag{B.2}$$

then there exists a finite constant K depending only on δ_1, such that[18]

$$\sup_x |\mathrm{P}(c_n \le \sigma_{c_n}x) - \Phi(x)| \le K\Big\{\sigma_{c_n}^{-2-2\delta_1}\Big(\sum_{i=1}^n \mathrm{E}|c_{ni}|^{2+2\delta_1} + \mathrm{E}\Big|\sum_{i=1}^n \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \sigma_{c_n}^2\Big|^{1+\delta_1}\Big)\Big\}^{1/(3+2\delta_1)}. \tag{B.3}$$

[18]Note that the result in Heyde and Brown (1970) is on a fixed square integrable martingale difference sequence with $0 < \delta_1 \le 1$, but the result also applies to a triangular array of martingale differences with $\delta_1 > 1$ (Haeusler, 1988).

Thus if

$$\lim_{n\to\infty} \sigma_{c_n}^{-2-2\delta_1} \sum_{i=1}^{n} \mathrm{E}\,|c_{ni}|^{2+2\delta_1} = 0, \tag{B.4}$$

and

$$\lim_{n\to\infty} \mathrm{E}\Big|\Big(\sigma_{c_n}^{-2} \sum_{i=1}^{n} \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\Big) - 1\Big|^{1+\delta_1} = 0, \tag{B.5}$$

then $\mathrm{P}(c_n \le \sigma_{c_n} x)$ converges uniformly to $\Phi(x)$ and a bound on the rate of convergence is given by (B.3). Now we check that (B.2), (B.4) and (B.5) hold. Let $2 \le q \le 2(1+\delta)$ and $1/p + 1/q = 1$. By the triangle and Hölder's inequalities,

$$\begin{aligned}
\sum_{i=1}^{n} \mathrm{E}\,|c_{ni}|^q &\le n^{-q/2}\,\mathrm{E}\sum_{i=1}^{n}\Big(|a_{n,ii}|^{1/p}|a_{n,ii}|^{1/q}|\epsilon_{ni}^2 - \sigma_0^2| + \sum_{j=1}^{i-1} |a_{n,ij}|^{1/p} 2|a_{n,ij}|^{1/q}|\epsilon_{ni}||\epsilon_{nj}| + |b_{ni}||\epsilon_{ni}|\Big)^q \\
&\le n^{-q/2}\,\mathrm{E}\sum_{i=1}^{n}\Big(\sum_{j=1}^{i} |a_{n,ij}| + 1\Big)^{q/p}\Big(|a_{n,ii}||\epsilon_{ni}^2 - \sigma_0^2|^q + \sum_{j=1}^{i-1} 2^q|a_{n,ij}||\epsilon_{ni}|^q|\epsilon_{nj}|^q + |b_{ni}|^q|\epsilon_{ni}|^q\Big) \\
&\le n^{(2-q)/2}(K_a+1)^{q/p}\Big(K_a\,\mathrm{E}\,|\epsilon_{ni}^2 - \sigma_0^2|^q + 2^q K_a(\mathrm{E}\,|\epsilon_{ni}|^q)^2 + K_b\,\mathrm{E}\,|\epsilon_{ni}|^q\Big).
\end{aligned} \tag{B.6}$$

Thus (B.2) holds. As $\sigma_{c_n}^2 = \sum_{i=1}^{n} \mathrm{E}\,c_{ni}^2$, (B.6) implies that $\sigma_{c_n}^2$ is bounded. Then (B.4) holds by Assumption 3. For $0 < \delta_1 \le 1$,

$$\begin{aligned}
&\mathrm{E}\Big|\Big(\sigma_{c_n}^{-2}\sum_{i=1}^{n}\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\Big) - 1\Big|^{1+\delta_1} \\
&= \sigma_{c_n}^{-2(1+\delta_1)}\,\mathrm{E}\Big|\sum_{i=1}^{n}\Big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}\,c_{ni}^2\Big)\Big|^{1+\delta_1} \\
&\le \sigma_{c_n}^{-2(1+\delta_1)}\Big(\mathrm{E}\Big|\sum_{i=1}^{n}\Big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}\,c_{ni}^2\Big)\Big|^2\Big)^{(1+\delta_1)/2} \\
&= 4^{1+\delta_1} n^{-1-\delta_1}\sigma_{c_n}^{-2(1+\delta_1)}\Big(\mathrm{E}\Big(\sigma_0^2\sum_{i=1}^{n}\sum_{j=1}^{i-1} a_{n,ij}^2(\epsilon_{nj}^2 - \sigma_0^2) + 2\sigma_0^2\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} a_{n,ij}a_{n,ik}\epsilon_{nj}\epsilon_{nk} \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{i-1}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}\epsilon_{nj}\Big)^2\Big)^{(1+\delta_1)/2} \\
&= 4^{1+\delta_1} n^{-1-\delta_1}\sigma_{c_n}^{-2(1+\delta_1)}\Big(\sigma_0^4(\mu_4 - \sigma_0^4)\sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^{n} a_{n,ij}^2\Big)^2 + 4\sigma_0^8\sum_{j=1}^{n-1}\sum_{k=1}^{j-1}\Big(\sum_{i=j+1}^{n} a_{n,ij}a_{n,ik}\Big)^2 \\
&\quad + \sigma_0^2\sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^{n}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}\Big)^2 + 2\mu_3\sigma_0^2\sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^{n} a_{n,ij}^2\Big)\Big(\sum_{i=j+1}^{n}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}\Big)\Big)^{(1+\delta_1)/2} \\
&\le 4^{1+\delta_1} n^{-1-\delta_1}\sigma_{c_n}^{-2(1+\delta_1)}\Big(n\sigma_0^4 K_a^4(\mu_4 - \sigma_0^4) + 4\sigma_0^8 K_a^2\sum_{j=1}^{n-1}\sum_{k=1}^{j-1}\sum_{i=j+1}^{n}|a_{n,ij}||a_{n,ik}| \\
&\quad + \sigma_0^2\sum_{j=1}^{n-1}\sum_{i=j+1}^{n}(|\mu_3^2 a_{n,ii}a_{n,ij}| + \sigma_0^4 b_{ni}^2|a_{n,ij}|)\sum_{i=j+1}^{n}(|a_{n,ii}a_{n,ij}| + |a_{n,ij}|) \\
&\quad + 2|\mu_3|\sigma_0^2 K_a^2\sum_{j=1}^{n-1}\sum_{i=j+1}^{n}|(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})a_{n,ij}|\Big)^{(1+\delta_1)/2}
\end{aligned}$$

$$
\begin{aligned}
&\le 4^{1+\delta_1} n^{-\delta_1} \sigma_{c_n}^{-2(1+\delta_1)} \Big(\sigma_0^4 K_a^4 (\mu_4 - \sigma_0^4) + 4\sigma_0^8 K_a^4 + \sigma_0^2 K_a^2 (\mu_3^2 K_a + \sigma_0^4 K_b)(K_a + 1) \\
&\quad + 2|\mu_3| \sigma_0^2 K_a^3 (|\mu_3| K_a + \sigma_0^2 K_b)\Big)^{(1+\delta_1)/2}, \qquad \text{(B.7)}
\end{aligned}
$$

where we have used the Cauchy-Schwarz inequality to derive $\big(\sum_{i=j+1}^n (\mu_3 a_{n,ii} a_{n,ij} + \sigma_0^2 b_{ni} a_{n,ij})\big)^2 \le \sum_{i=j+1}^n (|\mu_3^2 a_{n,ii} a_{n,ij}| + \sigma_0^4 b_{ni}^2 |a_{n,ij}|) \sum_{i=j+1}^n (|a_{n,ii} a_{n,ij}| + |a_{n,ij}|) \le \sum_{i=j+1}^n (|\mu_3^2 a_{n,ii} a_{n,ij}| + \sigma_0^4 b_{ni}^2 |a_{n,ij}|) K_a (K_a + 1)$. Thus (B.5) holds. Using (B.3), (B.6) with $q = 2 + 2\delta_1$ and (B.7), we have

$$
\begin{aligned}
&\sup_x |\mathrm{P}(c_n \le \sigma_{c_n} x) - \Phi(x)| \\
&\le K \sigma_{c_n}^{-2(1+\delta_1)/(3+2\delta_1)} n^{-\delta_1/(3+2\delta_1)} \Big((K_a + 1)^{1+2\delta_1} \big(K_a \mathrm{E}|\epsilon_{ni}^2 - \sigma_0^2|^{2+2\delta_1} + 2^{2+2\delta_1} K_a (\mathrm{E}|\epsilon_{ni}|^{2+2\delta_1})^2 + K_b \mathrm{E}|\epsilon_{ni}|^{2+2\delta_1}\big) \\
&\quad + 4^{1+\delta_1} \big(\sigma_0^4 K_a^4 (\mu_4 - \sigma_0^4) + 4\sigma_0^8 K_a^4 + \sigma_0^2 K_a^2 (\mu_3^2 K_a + \sigma_0^4 K_b)(K_a + 1) \\
&\quad + 2|\mu_3| \sigma_0^2 K_a^3 (|\mu_3| K_a + \sigma_0^2 K_b)\big)^{(1+\delta_1)/2} \Big)^{1/(3+2\delta_1)} = r_n,
\end{aligned}
$$

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

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

and similarly because $c_n/\sigma_{c_n} \le x - \tau$ and $|d_n| \le \tau$ imply $c_n/\sigma_{c_n} + d_n \le x$,

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

we have

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

i.e., (8) holds. Similarly, (9) holds. □

Proof of Proposition 1. We first show the result for $\mathbb{I}_n$ in (2) with normal disturbances. As $\mathbb{I}_n$ only has unit asymptotic variance but not unit variance for finite samples, it is convenient to use (11) to prove the result. Note that as 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

$$
\begin{aligned}
c_n &= n^{-1/2} \big(\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)\big), \\
\sigma_{c_n}^2 &= n^{-1} \sigma_0^4 \operatorname{tr}[H_n M_n H_n (M_n + M_n')], \\
e_n &= \frac{n}{n - k_x} \frac{\sqrt{\operatorname{tr}[H_n M_n H_n (M_n + M_n')]}}{\sqrt{\operatorname{tr}(M_n^2 + M_n' M_n)}},
\end{aligned}
$$

and

$$
\begin{aligned}
d_n &= \mathbb{I}_n/e_n - c_n/\sigma_{c_n} \\
&= \frac{\sigma_0^2 \epsilon_n' H_n \epsilon_n \operatorname{tr}(M_n H_n) - \epsilon_n' H_n M_n H_n \epsilon_n [\epsilon_n' H_n \epsilon_n - (n-k_x)\sigma_0^2]}{\sigma_0^2 \epsilon_n' H_n \epsilon_n \sqrt{\operatorname{tr}[H_n M_n H_n (M_n + M_n')]}} \\
&= \frac{\left[\frac{1}{n}\left(\epsilon_n' H_n \epsilon_n - (n-k_x)\sigma_0^2\right) + \frac{n-k_x}{n}\sigma_0^2\right] \operatorname{tr}(M_n H_n)}{\left[\frac{1}{n}\left(\epsilon_n' H_n \epsilon_n - (n-k_x)\sigma_0^2\right) + \frac{n-k_x}{n}\sigma_0^2\right] \sqrt{\operatorname{tr}[H_n M_n H_n (M_n + M_n')]}} \\
&\quad - \frac{\left[\frac{1}{\sqrt{n}}\left(\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)\right) + \frac{\sigma_0^2}{\sqrt{n}} \operatorname{tr}(M_n H_n)\right] \frac{1}{\sqrt{n}} [\epsilon_n' H_n \epsilon_n - (n-k_x)\sigma_0^2]}{\sigma_0^2 \left[\frac{1}{n}\left(\epsilon_n' H_n \epsilon_n - (n-k_x)\sigma_0^2\right) + \frac{n-k_x}{n}\sigma_0^2\right] \sqrt{\operatorname{tr}[H_n M_n H_n (M_n + M_n')]}}.
\end{aligned}
$$

By Lemma A.9 in Lee (2004b) and Lemma A.3 in Lin and Lee (2010), $\operatorname{tr}(M_n H_n) = \operatorname{tr}(M_n) + O(1) = O(1)$, $n^{-1/2}[\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)] = O_P(1)$, $n^{-1/2}[\epsilon_n' H_n \epsilon_n - (n-k_x)\sigma_0^2] = O_P(1)$, and $n^{-1} \operatorname{tr}[H_n M_n H_n (M_n + M_n')] = (2n)^{-1} \operatorname{tr}[(M_n + M_n')^2] + o(1) = O(1)$ is bounded away from zero by Assumption I4, then $d_n = O_P(n^{-1/2})$. Note that in (11), $r_n = O(n^{-\delta_1/(3+2\delta_1)})$, and $r_n^* = O_P(n^{-\delta_1/(3+2\delta_1)})$ by Lemma 2. As $e_n = e_n^*$, it remains to show that $\mathrm{P}^*(|d_n^*| > \tau) = o_P(1)$ for $\tau > 0$. For large enough n,

$$
\begin{aligned}
&\mathrm{P}^*(|d_n^*| > \tau) \\
&\leq \mathrm{P}^*\Big(|d_n^*| > \tau, \frac{1}{n}|\epsilon_n^{*\prime} H_n \epsilon_n^* - (n-k_x)\sigma_n^{*2}| \leq \eta_n\Big) + \mathrm{P}^*\Big(\frac{1}{n}|\epsilon_n^{*\prime} H_n \epsilon_n^* - (n-k_x)\sigma_n^{*2}| > \eta_n\Big) \\
&\leq \mathrm{P}^*\Bigg(\frac{|\epsilon_n^{*\prime} H_n M_n H_n \epsilon_n^*| \eta_n}{\sigma_n^{*2}\left(\frac{n-k_x}{n}\sigma_n^{*2} - \eta_n\right)\sqrt{\operatorname{tr}\big(H_n (M_n + M_n') H_n M_n\big)}} \geq \tau - \frac{|\operatorname{tr}(M_n H_n)|}{\sqrt{\operatorname{tr}\big(H_n (M_n + M_n') H_n M_n\big)}}\Bigg) \\
&\quad + \mathrm{P}^*\Big(\frac{1}{n}|\epsilon_n^{*\prime} H_n \epsilon_n^* - (n-k_x)\sigma_n^{*2}| > \eta_n\Big)
\end{aligned}
$$

Note that $\mathrm{E}^* |\epsilon_n^{*\prime} H_n M_n H_n \epsilon_n^*|^2 = (\mathrm{E}^* \epsilon_{ni}^{*4} - 3\sigma_n^{*4}) \sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + \sigma_n^{*4} \operatorname{tr}^2(H_n M_n) + \sigma_n^{*4} \operatorname{tr}(H_n M_n H_n (M_n + M_n')) = O_P(n)$, then by Chebyshev's inequality,

$$
\begin{aligned}
&\mathrm{P}^*(|d_n^*| > \tau) \\
&\leq \frac{\eta_n^2 \, \mathrm{E}^* |\epsilon_n^{*\prime} H_n M_n H_n \epsilon_n^*|^2}{\sigma_n^{*4}\left(\frac{n-k_x}{n}\sigma_n^{*2} - \eta_n\right)^2 \operatorname{tr}\big(M_n H_n (M_n + M_n') H_n\big)\Big(\tau - |\operatorname{tr}(M_n H_n)| \operatorname{tr}^{-1/2}\big(H_n (M_n + M_n') H_n M_n\big)\Big)^2} \\
&\quad + \frac{(\mathrm{E}^* \epsilon_{ni}^{*4} - 3\sigma_n^{*4}) \sum_{i=1}^n (H_n)_{ii}^2 + 2(n-k_x)\sigma_n^{*4}}{n^2 \eta_n^2} \\
&= o_P(1),
\end{aligned}
$$

if $\eta_n = n^{-1/4}$. Thus the result follows.

For $\mathbb{I}_n'$ in (4), let

$$
\begin{aligned}
c_n &= n^{-1/2}\big(\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(M_n H_n)\big), \\
\sigma_{c_n}^2 &= n^{-1}(\mu_4 - 3\sigma_0^4) \sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + n^{-1}\sigma_0^4 \operatorname{tr}[H_n M_n H_n (M_n + M_n')],
\end{aligned}
$$

and

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

Note that $\underline{\hat{\sigma}_{c_n}}^2 = \hat{\sigma}_{c_n}^2 + (c_\sigma - \hat{\sigma}_{c_n}^2)\mathbf{1}(\hat{\sigma}_{c_n}^2 \le c_\sigma)$. By Lemma 2, $\hat{\sigma}_n^2 - \sigma_0^2 = o_P(1)$, $\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2 = o_P(1)$, and $r_n^* = O_P(n^{-\delta_1/(3+2\delta_1)})$ in (10), as $\mathrm{E}(|\epsilon_{ni}|^{4(1+\delta)}) < \infty$ for some $\delta > 0$. Furthermore, for $\eta > 0$, $\mathrm{P}(|(c_\sigma - \hat{\sigma}_{c_n}^2)\mathbf{1}(\hat{\sigma}_{c_n}^2 \le c_\sigma)| \ge \eta) \le \mathrm{P}(\hat{\sigma}_{c_n}^2 \le c_\sigma) \le \mathrm{P}(|\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2| \ge \sigma_{c_n}^2 - c_\sigma) = o(1)$. Thus $(c_\sigma - \hat{\sigma}_{c_n}^2)\mathbf{1}(\hat{\sigma}_{c_n}^2 \le c_\sigma) = o_P(1)$. Then, by the continuous mapping theorem, $\underline{\hat{\sigma}_{c_n}} - \sigma_{c_n} = o_P(1)$. Hence $d_n = o_P(1)$. By (10), it remains to show that $\mathrm{P}^*(|d_n^*| > \tau) = o_P(1)$ for $\tau > 0$. The d_n^* is

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

By Lemma 4, $\mathrm{P}^*(|\hat{\sigma}_n^{*2} - \sigma_n^{*2}| > \tau) = o_P(1)$ and $\mathrm{P}^*(|\underline{\hat{\sigma}_{c_n}^*}^2 - \sigma_{c_n}^{*2}| > \tau) = o_P(1)$ for $\tau > 0$, as $\mathrm{E}(\epsilon_{ni}^4) < \infty$. By Chebyshev's inequality,

$$\mathrm{P}^*(n^{-1/2}|\epsilon^{*\prime} H_n M_n H_n \epsilon_n^* - \sigma_n^{*2}\operatorname{tr}(M_n H_n)| > \tau) \le \tau^{-2} n^{-1}\,\mathrm{E}^*|\epsilon^{*\prime} H_n M_n H_n \epsilon_n^* - \sigma_n^{*2}\operatorname{tr}(M_n H_n)|^2 = O_P(\tau^{-2}).$$

Then $\mathrm{P}^*(|d_n^*| > \tau) = o_P(1)$. □

Appendix C. Proofs for Results in Subsection 3.1

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

$$\begin{aligned}
\varphi_n(t) &= \mathrm{E}\exp(itc_n/\sigma_{c_n}) \\
&= \exp\Big(-\frac{it\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}(A_n)\Big)\,\mathrm{E}\exp\Big(it\frac{(\sigma_0^{-1}\epsilon_n')\sigma_0^2 A_n(\sigma_0^{-1}\epsilon_n) + \sigma_0 b_n'(\sigma_0^{-1}\epsilon_n)}{\sqrt{n}\sigma_{c_n}}\Big) \\
&= \exp\Big(-\frac{it\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}(A_n)\Big)\Big|I_n - 2it\frac{\sigma_0^2 A_n}{\sqrt{n}\sigma_{c_n}}\Big|^{-1/2}\exp\Big(-\frac{\sigma_0^2}{2n\sigma_{c_n}^2}t^2 b_n'\Big(I_n - 2it\frac{\sigma_0^2 A_n}{\sqrt{n}\sigma_{c_n}}\Big)^{-1} b_n\Big) \\
&= \exp\Big(g_n(t) - \frac{1}{2}t^2\Big),
\end{aligned}$$

where the third line follows from (3.2.11) on p. 53 of Lukacs and Laha (1964), and $g_n(t) = -\frac{n}{2}\ln\sigma_0^2 - \frac{1}{2}\ln|B_n(t)| - \frac{t^2}{2n\sigma_{c_n}^2}b_n' B_n(t)^{-1}b_n - \frac{it\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\operatorname{tr}(A_n) + \frac{1}{2}t^2$ with $B_n(t) = \frac{I_n}{\sigma_0^2} - \frac{2itA_n}{\sqrt{n}\sigma_{c_n}}$. The derivatives of $g_n(t)$ are

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

$$+ \frac{24it^2}{n^{5/2}\sigma_{c_n}^5} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^3 b_n,$$

$$g_n^{(4)}(t) = \frac{48}{n^2\sigma_{c_n}^4}\operatorname{tr}\big[\big(A_n B_n(t)^{-1}\big)^4\big] + \frac{48}{n^2\sigma_{c_n}^4} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^2 b_n + \frac{192it}{n^{5/2}\sigma_{c_n}^5} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^3 b_n$$

$$- \frac{192t^2}{n^3\sigma_{c_n}^6} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^4 b_n,$$

$$g_n^{(k)}(t) = \frac{c_{k1} i^k}{n^{k/2}\sigma_{c_n}^k}\operatorname{tr}\big[\big(A_n B_n(t)^{-1}\big)^k\big] + \frac{c_{k2} i^k}{n^{k/2}\sigma_{c_n}^k} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^{k-2} b_n$$

$$+ \frac{c_{k3} i^{k+1} t}{n^{(k+1)/2}\sigma_{c_n}^{k+1}} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^{k-1} b_n + \frac{c_{k4} i^{k+2} t^2}{n^{(k+2)/2}\sigma_{c_n}^{k+2}} b_n' B_n(t)^{-1}\big(A_n B_n(t)^{-1}\big)^{k} b_n, \text{ for } k \geq 3,$$

$$g_n(0) = g_n^{(1)}(0) = g_n^{(2)}(0) = 0,$$

$$g_n^{(3)}(0) = -\frac{8i\sigma_0^6}{n^{3/2}\sigma_{c_n}^3}\operatorname{tr}(A_n^3) - \frac{6i\sigma_0^4}{n^{3/2}\sigma_{c_n}^3} b_n' A_n b_n,$$

$$g_n^{(k)}(0) = \frac{c_{k1}\sigma_0^{2k} i^k}{n^{k/2}\sigma_{c_n}^k}\operatorname{tr}(A_n^k) + \frac{c_{k2}\sigma_0^{2(k-1)} i^k}{n^{k/2}\sigma_{c_n}^k} b_n' A_n^{k-2} b_n, \text{ for } k \geq 3, \tag{C.1}$$

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

$$|g_n^{(4)}(t)| \leq \frac{48\iota_n^2\sigma_0^8}{n^2\sigma_{c_n}^4}\sum_{j=1}^n \iota_{nj}^2 + \frac{48\iota_n^2\sigma_0^6}{n^2\sigma_{c_n}^4} b_n' b_n + \frac{96\iota_n^2\sigma_0^6}{n^2\sigma_{c_n}^4} b_n' b_n + \frac{48\iota_n^2\sigma_0^6}{n^2\sigma_{c_n}^4} b_n' b_n \leq \frac{192\iota_n^2\sigma_0^4}{n\sigma_{c_n}^2},$$

$$|g_n^{(k)}(t)| \leq \frac{|c_{k1}|\iota_n^{k-2}\sigma_0^{2k}}{n^{k/2}\sigma_{c_n}^k}\sum_{j=1}^n \iota_{nj}^2 + \frac{|c_{k2}|\iota_n^{k-2}\sigma_0^{2k-2}}{n^{k/2}\sigma_{c_n}^k} b_n' b_n + \frac{|c_{k3}|\iota_n^{k-2}\sigma_0^{2k-2}}{2n^{k/2}\sigma_{c_n}^k} b_n' b_n + \frac{|c_{k4}|\iota_n^{k-2}\sigma_0^{2k-2}}{4n^{k/2}\sigma_{c_n}^k} b_n' b_n$$

$$\leq \frac{c_{k5}(\iota_n\sigma_0^2)^{k-2}}{n^{(k-2)/2}\sigma_{c_n}^{k-2}}. \text{ for } k \geq 3. \tag{C.2}$$

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 $\kappa_n = \frac{1}{6}(-i)^3 g_n^{(3)}(0) = \frac{4\sigma_0^6}{3n^{3/2}\sigma_{c_n}^3}\operatorname{tr}(A_n^3) + \frac{\sigma_0^4}{n^{3/2}\sigma_{c_n}^3} b_n' A_n b_n$ is one sixth of the third moment of c_n/σ_{c_n} or one sixth of the third cumulant as c_n/σ_{c_n} has mean zero. By a smoothing inequality in Feller (1970, p. 538), for all $T > 0$,

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

where $\Phi^{(3)}(x) = (x^2 - 1)\Phi^{(1)}(x)$ and

$$|\varphi_n(t) - \gamma_n(t)| = \exp(-t^2/2)|\exp(g_n(t)) - (1 - i\kappa_n t^3)|$$

$$= \exp(-t^2/2)|\exp(g_n(t)) - \exp(-i\kappa_n t^3) + \exp(-i\kappa_n t^3) - (1 - i\kappa_n t^3)|$$

$$\leq \exp(-t^2/2)[|\exp(g_n(t)) - \exp(-i\kappa_n t^3)| + |\exp(-i\kappa_n t^3) - (1 - i\kappa_n t^3)|].$$

As

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

$|\exp(-i\kappa_n t^3) - (1 - i\kappa_n t^3)| \le |i\kappa_n t^3|^2/2 = \kappa_n^2 t^6/2 \le \iota_n^2 \sigma_0^4 t^6/(2n\sigma_{c_n}^2)$ (Feller, 1970, (4.13) on p. 514). By a four-term Taylor expansion of $g_n(t)$ and with the bound in (C.2), $|g_n(t) + i\kappa_n t^3| \le \frac{8\iota_n^2\sigma_0^4 t^4}{n\sigma_{c_n}^2} \le \frac{1}{4}t^2$, when $|t| \le \frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n \sigma_0^2}$. Then $|\exp(g_n(t)) - \exp(-i\kappa_n t^3)| = |\exp(g_n(t) + i\kappa_n t^3) - 1| \le |g_n(t) + i\kappa_n t^3| \exp(|g_n(t) + i\kappa_n t^3|) \le \frac{8\iota_n^2\sigma_0^4 t^4}{n\sigma_{c_n}^2}\exp(\frac{1}{4}t^2)$ by (2.8) on p. 534 of Feller (1970), and

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

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

Note that $\varphi_n(t) = (\sigma_0^2)^{-n/2}|B_n(t)|^{-1/2}\exp\{-\frac{t^2}{2n\sigma_{c_n}^2} b_n' B_n^{-1}(t) b_n - \frac{it\sigma_0^2}{\sqrt{n}\sigma_{c_n}}\mathrm{tr}(A_n)\}$ and $B_n(t)$ has eigenvalues $\frac{1}{\sigma_0^2} - \frac{2it\iota_{nj}}{\sqrt{n}\sigma_{c_n}}$, $j = 1, \ldots, n$. We have $|(\sigma_0^2)^{-n/2}|B_n(t)|^{-1/2}| = \prod_{j=1}^n \big(1 + \frac{4t^2\sigma_0^4\iota_{nj}^2}{n\sigma_{c_n}^2}\big)^{-1/4}$, and the real part of $[-\frac{t^2}{2n\sigma_{c_n}^2} b_n' B_n^{-1}(t) b_n - \frac{it\sigma_0^2}{\sqrt{n}\sigma_{cn}} tr(A_n)]$ is $-\frac{t^2\sigma_0^2}{2} b_n' P_n \mathrm{Diag}\{(n\sigma_{c_n}^2 + 4\sigma_0^4 t^2 \iota_{n1}^2)^{-1}, \ldots, (n\sigma_{c_n}^2 + 4\sigma_0^4 t^2 \iota_{nn}^2)^{-1}\} P_n' b_n$, where P_n is the orthonormal matrix of eigenvectors of A_n and $\mathrm{Diag}(a_n)$ denotes a diagonal matrix with the diagonal elements being those of the vector a_n. Hence

$$|\varphi_n(t)| \le \exp\Big(-\frac{\sigma_0^2 t^2 b_n' b_n}{2n\sigma_{c_n}^2 + 8\iota_n^2\sigma_0^4 t^2}\Big)\prod_{j=1}^n \Big(1 + \frac{4t^2\sigma_0^4\iota_{nj}^2}{n\sigma_{c_n}^2}\Big)^{-1/4}.$$

When $|t| > \frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n\sigma_0^2}$,

$$\begin{aligned}
|\varphi_n(t)| &< \exp\Big(-\frac{b_n' b_n}{72\iota_n^2\sigma_0^2} - \frac{1}{4}\sum_{j=1}^n \ln\Big((1 + \frac{\iota_{nj}^2}{8\iota_n^2})\Big)\Big) \\
&\le \exp\Big(-\frac{b_n' b_n}{72\iota_n^2\sigma_0^2} - \frac{1}{32\iota_n^2}\sum_{j=1}^n \iota_{nj}^2 + \frac{1}{512\iota_n^4}\sum_{j=1}^n \iota_{nj}^4\Big) \\
&\le \exp\Big(-\frac{b_n' b_n}{72\iota_n^2\sigma_0^2} - \frac{15}{512\iota_n^2}\sum_{j=1}^n \iota_{nj}^2\Big) \\
&\le \exp\Big(-\frac{n\sigma_{c_n}^2}{72\iota_n^2\sigma_0^4}\Big),
\end{aligned} \tag{C.5}$$

where the second inequality follows by $\ln(1+x) \ge x - \frac{x^2}{2}$ for $x \ge 0$[19]. In (C.3), let $T = n\sigma_{c_n}^2$. By (C.4), the contribution of the integral in (C.3) 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^{-1} by $|\varphi_n(t) - \gamma_n(t)|/|t| \le (|\varphi_n(t)| + |\gamma_n(t)|)/|t|$, where $|\varphi_n(t)|$ satisfies (C.5) and $|\gamma_n(t)|/|t| = \exp(-\frac{1}{2}t^2)(1 + \kappa_n^2 t^6)^{1/2}/|t| < \exp(-\frac{n\sigma_{c_n}^2}{64\iota_n^2\sigma_0^4})(1 + \kappa_n^2 n^6 \sigma_{c_n}^{12})^{1/2}/(\frac{\sqrt{2n}\sigma_{c_n}}{8\iota_n\sigma_0^2})$. The second term on the r.h.s. of (C.3) has the order $O(n^{-1})$. Therefore, $\sup_{x\in\mathbb{R}}|\mathrm{P}(c_n/\sigma_{c_n} \le x) - (\Phi(x) - \kappa_n\Phi^{(3)}(x))| = O(n^{-1})$. Since we are considering a parametric bootstrap where elements of ϵ_n^* are drawn

[19]This inequality holds because $\ln(1+x) - x + x^2/2$ is increasing for $x > -1$ and it takes the value 0 at $x = 0$.

from the normal distribution with mean zero and variance $n^{-1}\hat{\epsilon}_n'\hat{\epsilon}_n$, (13) holds by a similar argument and Lemma 2.

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 $p_{n1},\dots,p_{nm}$, where $m = r(r-2)$. Note that $p_{n1},\dots,p_{nm}$ are real by (C.1). 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 $T>0$,

$$\sup_{x\in\mathbb{R}}\Big|\mathrm{P}(c_n/\sigma_{c_n}\le x)-\int_{-\infty}^{x}\omega_n(t)\,dt\Big| \le \frac{1}{\pi}\int_{-T}^{T}\Big|\frac{\varphi_n(t)-\exp(-t^2/2)(1+p_n(it))}{t}\Big|\,dt + \frac{24\sup_x|\omega_n(x)|}{\pi T}. \tag{C.6}$$

Using the inequality that $|e^{\alpha}-1-\sum_{k=1}^{r-2}\beta^k/k!| = |(e^{\alpha}-e^{\beta})+(e^{\beta}-1-\sum_{k=1}^{r-2}\beta^k/k!)| \le \exp(\max\{|\alpha|,|\beta|\})(|\alpha-\beta|+\frac{1}{(r-1)!}|\beta|^{r-1})$, which is (2.17) on p. 535 of Feller (1970), we have

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

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

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

where $|n^{\frac{1}{2}}t^3\tau_{nr}(t)| \le \sum_{k=3}^{r}\frac{c_{k5}(\iota_n\sigma_0^2)^{k-2}}{n^{(k-3)/2}\sigma_{c_n}^{k-2}}\frac{|t|^k}{k!}$. Now let $T = n^{(r-1)/2}$, then the contribution of the integral on the r.h.s. of (C.6) when $t\le cn^{1/2}$ is $O(n^{-(r-1)/2})$, and the contribution of the integral when $t>cn^{1/2}$ tends to zero faster than any power of n^{-1} similar to the above. In addition, the second term on the r.h.s. of (C.6) has the order $O(n^{-(r-1)/2})$. Hence, $\sup_{x\in\mathbb{R}}|\mathrm{P}(c_n/\sigma_{c_n}\le x)-\int_{-\infty}^{x}\omega_n(t)\,dt| = O(n^{-(r-1)/2})$. 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.2). Rearranging it according to ascending powers of $n^{-1/2}$ and dropping the terms involving powers $n^{-k/2}$ with $k>r-1$ yields (14). Eq. (15) follows similarly. □

Proof of Proposition 2. Let $A_n = [H_nM_nH_n - n^{-1}xH_n\,\mathrm{tr}^{1/2}(M_n^2+M_n'M_n)]$, $\sigma_{a_n}^2 = n^{-1}\,\mathrm{E}[\epsilon_n'A_n\epsilon_n-\sigma_0^2\,\mathrm{tr}(A_n)]^2 = 2n^{-1}\sigma_0^4\,\mathrm{tr}(A_n^2)$, $z_n = -\sigma_0^2n^{-1/2}\sigma_{a_n}^{-1}\,\mathrm{tr}(A_n) = -2^{-1/2}\,\mathrm{tr}(A_n)\,\mathrm{tr}^{-1/2}(A_n^2)$ and

$$\kappa_n = 4\sigma_0^6n^{-3/2}\sigma_{a_n}^{-3}\,\mathrm{tr}(A_n^3)/3 = 4[2\,\mathrm{tr}(A_n^2)]^{-3/2}\,\mathrm{tr}(A_n^3)/3.$$

Then

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

by Theorem 2. Similarly, as z_n and κ_n do not involve any population parameter, neither do the corresponding bootstrap versions z_n^* and κ_n^*. Thus $z_n^* = z_n$ and $\kappa_n^* = \kappa_n$, and

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

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

Appendix D. Proofs for Results in Subsection 3.2

The following theorem from Mykland (1993) gives an asymptotic expansion for martingales. We shall use it to prove Theorem 3 and Proposition 3.

Theorem A. *Consider a triangular array of normalized martingales $c_{T_n}/\hat{\sigma}_{c_{T_n}} = \sum_{i=1}^{T_n} c_{ni}/\hat{\sigma}_{c_{T_n}}$ for $n = 1, \dots,$ where $\hat{\sigma}_{c_{T_n}}$ is a normalizing factor, which can be stochastic, and c_{ni}'s are martingale differences with the filtration $\mathscr{F}_{ni}$. Suppose that the following conditions are satisfied:*

(i) (Integrability condition for the fourth-order variation.) For nonrandom sequences $\{\sigma_{c_{T_n}}^2\}$ and $\{r_n^2\}$ with $r_n = o(1)$, we have

$$\sum_{i=1}^{T_n} \mathrm{E}(c_{ni}^4) = O(\sigma_{c_{T_n}}^4 r_n^2). \tag{D.1}$$

(ii) (Integrability conditions for the square variation.) There are constants b^2, $\underline{k}$ and $\bar{k}$ so that $0 \le \underline{k} < b^2 < \bar{k} \le \infty$, and, for z_{T_n} being either $\sum_{i=1}^{T_n} c_{ni}^2$ or $\sum_{i=1}^{T_n} \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})$,

$$r_n^{-1}(\sigma_{c_{T_n}}^{-2} z_{T_n} - b^2)\mathbf{1}(\underline{k} \le \sigma_{c_{T_n}}^{-2} z_{T_n} \le \bar{k}) \text{ is uniformly integrable,} \tag{D.2}$$

with $\mathrm{P}(\underline{k} \le \sigma_{c_{T_n}}^{-2} z_{T_n} \le \bar{k}) = 1 - o(r_n)$.[20]

(iii) (Integrability conditions for $\hat{\sigma}_{c_{T_n}}^2$.) There are measurable sets Q_n and constants b_*^2 and $\delta > 0$ so that $\mathrm{P}(Q_n) = 1 - o(r_n)$ and

$$\sup_n \mathrm{E}(r_n^{-1}|\sigma_{c_{T_n}}^{-2}\hat{\sigma}_{c_{T_n}}^2 - b_*^2|)^{1+\delta}\mathbf{1}(Q_n) < \infty. \tag{D.3}$$

(iv) (The central limit condition.) There are Borel-measurable functions ψ_o, ψ_p and ψ_*, so that, whenever

$$\Big(b^{-1}\frac{c_{T_n}}{\sigma_{c_{T_n}}}, r_n^{-1}\big(\frac{1}{\sigma_{c_{T_n}}^2}\sum_{i=1}^{T_n} c_{ni}^2 - b^2\big), r_n^{-1}\big(\frac{1}{\sigma_{c_{T_n}}^2}\sum_{i=1}^{T_n}\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - b^2\big), r_n^{-1}\big(\frac{\hat{\sigma}_{c_{T_n}}^2}{\sigma_{c_{T_n}}^2} - b_*^2\big)\Big) \xrightarrow{d} (Z, \xi_o, \xi_p, \xi_*), \tag{D.4}$$

as $n \to \infty$ through a subset of the integers, then $\mathrm{E}(\xi_o|Z) = b^2\psi_o(Z)$ a.s., $\mathrm{E}(\xi_p|Z) = b^2\psi_p(Z)$ a.s. and $\mathrm{E}(\xi_*|Z) = b_*^2\psi_*(Z)$ a.s..

[20] With (D.1), the condition on $\sum_{i=1}^{T_n} c_{ni}^2$ is equivalent to that on $\sum_{i=1}^{T_n} \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})$ (Mykland, 1993).

Then,

$$\int_{-\infty}^{\infty} h(x)\,dF_n(x) = \int_{-\infty}^{\infty} h(x)\,d\Phi(\beta^{-1}x) + \frac{1}{2}r_n\,\mathrm{E}[\beta^2\psi(x)h^{(2)}(\beta Z) - \psi_*(Z)\beta Z h^{(1)}(\beta Z)] + o(r_n), \tag{D.5}$$

where $F_n(x) = \mathrm{P}(\hat{\sigma}_{c_{T_n}}^{-1} c_{T_n} \le x | \hat{\sigma}_{c_{T_n}} > 0)$, $\beta = bb_*^{-1}$ *and* $\psi = \frac{1}{3}\psi_o + \frac{2}{3}\psi_p$, *uniformly on the 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.*

Note that Z is standard normal, because the conditions in the theorem imply those for the martingale central limit theorem.[21] Subject to some minimum niceness on the part of ψ and ψ_*, the second term on the r.h.s. of (D.5) can be shown via integration by parts to equal

$$\frac{1}{2}r_n \int_{-\infty}^{\infty} h(x)d\big[\big(\psi^{(1)}(\beta^{-1}x) - \psi(\beta^{-1}x)\beta^{-1}x + \psi_*(\beta^{-1}x)\beta^{-1}x\big)\Phi^{(1)}(\beta^{-1}x)\big]. \tag{D.6}$$

Mykland (1993) introduces the notion of order $o_2(r_n)$ to denote the kind of convergence in (D.5), stating an expansion in a more standard way as

$$F_n(x) = \Phi(\beta^{-1}x) + \frac{1}{2}r_n\big(\psi^{(1)}(\beta^{-1}x) - \psi(\beta^{-1}x)\beta^{-1}x + \psi_*(\beta^{-1}x)\beta^{-1}x\big)\Phi^{(1)}(\beta^{-1}x) + o_2(r_n). \tag{D.7}$$

Proof of Theorem 3. We shall first establish the asymptotic expansion for $F_n(x)$ and then show that a similar expansion for $F_n^*(x)$ exists.

For (17) and (19), using the same σ-fields and decomposition of c_n as a sum of martingale differences in the proof of Theorem 1, we shall show that the conditions in Theorem A are met with $T_n = n$, $\hat{\sigma}_{c_{T_n}} = \sigma_{c_n}$ and $r_n = n^{-1/2}$. By (B.6), $\sum_{i=1}^n \mathrm{E}(c_{ni}^4/\sigma_{c_n}^4) = O(n^{-1})$. By (B.7), $\mathrm{E}\big|n^{1/2}\big[\big(\sigma_{c_n}^{-2}\sum_{i=1}^n \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big) - 1\big]\big|^2 = O(1)$. 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. We shall show that

$$\Big(c_n/\sigma_{c_n}, n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^n\big(c_{ni}^2 - \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big), n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^n\big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}(c_{ni}^2)\big)\Big) \tag{D.8}$$

is asymptotically trivariate normal. Note that each term in (D.8) is a martingale. Under Assumptions 1′, 2′ and 3, the LQ form c_n/σ_{c_n} is asymptotically normal by Theorem 1 in Kelejian and Prucha (2001). Explicitly, this result follows by Corollary 3.1 in Hall and Heyde (1980), after proving that $\sigma_{c_n}^{-q}\sum_{i=1}^n \mathrm{E}|c_{ni}|^q = o(1)$ for some $q > 2$ and $\sigma_{c_n}^{-2}\sum_{i=1}^n[\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}(c_{ni}^2)] = o(1)$. The asymptotic normality of the second term in (D.8) can be analyzed by the same argument and that of the last term by writing it as an LQ form. The joint asymptotic normality will follow by the Cramér-Wold device with similar arguments but applied to the combined terms.

Note that

$$n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^n\big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}(c_{ni}^2)\big) = 4\sigma_{c_n}^{-2}n^{-1/2}\big[\sigma_0^2\sum_{j=1}^{n-1}\sum_{k=1}^{n-1}[\epsilon_{nj}\epsilon_{nk} - \mathrm{E}(\epsilon_{nj}\epsilon_{nk})]\sum_{i=\max\{j,k\}+1}^{n} a_{n,ij}a_{n,ik}$$

[21] See, e.g., Theorem 3.2 and Corrollary 3.1 on p. 58 of Hall and Heyde (1980).

$$+\sum_{j=1}^{n-1}\epsilon_{nj}\sum_{i=j+1}^{n}a_{n,ij}(\mu_3 a_{n,ii}+\sigma_0^2 b_{ni})\Big],$$

which is an LQ form. The involved matrix in this LQ form is bounded in both row and column sum norms, since the matrix is symmetric and $\sum_{j=1}^{n-1}|\sum_{i=\max\{j,k\}+1}^{n}a_{n,ij}a_{n,ik}|\le\sum_{i=k+1}^{n}|a_{n,ik}|\sum_{j=1}^{n-1}|a_{n,ij}|<\infty$. In addition, for $q=2+\delta>2$ and $1/p+1/q=1$, by the Hölder and c_r inequalities,

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

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.

For the second term in (D.8), 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 $1/p+1/q=1$,

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

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 $e_{1n,i}=a_{n,ii}^2(\epsilon_{ni}^4-\mu_4)+2a_{n,ii}b_{n,ii}(\epsilon_{ni}^3-\mu_3)+(b_{ni}^2-2\sigma_0^2a_{n,ii}^2)(\epsilon_{ni}^2-\sigma_0^2)-2\sigma_0^2a_{n,ii}b_{ni}\epsilon_{ni}$, $e_{2n,i}=4\big[a_{n,ii}(\epsilon_{ni}^3-\mu_3)+b_{n,ii}(\epsilon_{ni}^2-\sigma_0^2)-\sigma_0^2a_{n,ii}\epsilon_{ni}\big]$, $e_{3n,i}=4(\epsilon_{ni}^2-\sigma_0^2)$ and $e_{4n,i}=\sum_{j=1}^{i-1}a_{n,ij}\epsilon_{nj}$. Then $z_{ni}=n^{-1/2}(e_{1n,i}+e_{2n,i}e_{4n,i}+e_{3n,i}e_{4n,i}^2)$ and

$$\begin{aligned}\sum_{i=1}^{n}\big(\mathrm{E}(z_{ni}^2|\mathscr{F}_{n,i-1})-\mathrm{E}(z_{ni}^2)\big)&=\frac{1}{n}\sum_{i=1}^{n}\Big((e_{4n,i}^2-\mathrm{E}\,e_{4n,i}^2)\,\mathrm{E}(e_{2n,i}^2+2e_{1n,i}e_{3n,i})+(e_{4n,i}^4-\mathrm{E}\,e_{4n,i}^4)\,\mathrm{E}\,e_{3n,i}^2\\ &\quad+2e_{4n,i}\,\mathrm{E}(e_{1n,i}e_{2n,i})+2(e_{4n,i}^3-\mathrm{E}\,e_{4n,i}^3)\,\mathrm{E}(e_{2n,i}e_{3n,i})\Big).\end{aligned}\tag{D.9}$$

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

$$\begin{aligned}\mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^n b_{ni}^r e_{4n,i}\Big)^2 &= \frac{\sigma_0^2}{n^2}\sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^n a_{n,ij}b_{ni}^r\Big)^2 \\ &\le \frac{\sigma_0^2}{n^2}\sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^n |a_{n,ij}|b_{ni}^{2r}\Big)\Big(\sum_{i=j+1}^n |a_{n,ij}|\Big) \\ &\le \frac{h}{n^2}\sum_{i=1}^n b_{ni}^{2r}\Big(\sum_{j=1}^{i-1}|a_{n,ij}|\Big) = O(n^{-1}),\end{aligned} \tag{D.10}$$

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

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

where

$$\mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^{i-1} b_{ni}^r a_{n,ij}^2(\epsilon_{nj}^2 - \sigma_0^2)\Big)^2 = \frac{\mathrm{E}\,|\epsilon_{nj}^2 - \sigma_0^2|^2}{n^2}\sum_{j=1}^{n-1}\Big(\sum_{i=j+1}^n b_{ni}^r a_{n,ij}^2\Big)^2 = O(n^{-1})$$

and

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

as in (D.10). Similarly, $n^{-1}\sum_{i=1}^n b_{ni}(e_{4n,i}^3 - \mathrm{E}\, e_{4n,i}^3) = o_P(1)$ and $n^{-1}\sum_{i=1}^n(e_{4n,i}^4 - \mathrm{E}\, e_{4n,i}^4) = o_P(1)$, since they can be decomposed as

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

and

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^n(e_{4n,i}^4 - \mathrm{E}\, e_{4n,i}^4) \\ &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^{i-1} a_{n,ij}^4(\epsilon_{nj}^4 - \mu_4) + \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\big(4a_{n,ij}^3 a_{n,ik}\epsilon_{nj}^3\epsilon_{nk} + 4a_{n,ij}a_{n,ik}^3\epsilon_{nj}\epsilon_{nk}^3 + 6a_{n,ij}^2 a_{n,ik}^2(\epsilon_{nj}^2\epsilon_{nk}^2 - \sigma_0^4)\big) \\ &\quad + \frac{12}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1}(a_{n,ij}^2 a_{n,ik}a_{n,il}\epsilon_{nj}^2\epsilon_{nk}\epsilon_{nl} + a_{n,ij}a_{n,ik}^2 a_{n,il}\epsilon_{nj}\epsilon_{nk}^2\epsilon_{nl} + a_{n,ij}a_{n,ik}a_{n,il}^2\epsilon_{nj}\epsilon_{nk}\epsilon_{nl}^2) \\ &\quad + \frac{24}{n}\sum_{i=1}^n\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1}\sum_{m=1}^{l-1} a_{n,ij}a_{n,ik}a_{n,il}a_{n,im}\epsilon_{nj}\epsilon_{nk}\epsilon_{nl}\epsilon_{nm},\end{aligned}$$

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 (D.10). Note that in (D.9), $\mathrm{E}(e_{2n,i}^2 + 2e_{1n,i}e_{3n,i})$, $\mathrm{E}(e_{1n,i}e_{2n,i})$ and $\mathrm{E}(e_{2n,i}e_{3n,i})$ are polynomials of b_{ni}'s with bounded coefficients. Then $\sum_{i=1}^n\big(\mathrm{E}(z_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}(z_{ni}^2)\big) = o_P(1)$. Thus $n^{1/2}\sigma_{c_n}^{-2}\sum_{i=1}^n\big(c_{ni}^2 - \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)$ is asymptotically normal by Corollary 3.1 in Hall and Heyde (1980).

As each term in (D.8) is a martingale with the same filtration, an arbitrary linear combination of the terms in (D.8) is also a martingale. Furthermore, the first and third terms in (D.8) are both LQ forms, with $c_{ni} = n^{-1/2}[a_{n,ii}(\epsilon_{ni}^2 - \sigma_0^2) + 2\epsilon_{ni}\sum_{j=1}^{i-1} a_{n,ij}\epsilon_{nj} + b_{ni}\epsilon_{ni}]$ being a typical martingale difference for an LQ form. The terms in c_{ni} are similar to some terms in z_{ni}. Thus, an arbitrary linear combination of the terms in (D.8) can be shown to be asymptotically normal in a way similar to that for the second term in (D.8). It follows that (D.8) is asymptotically trivariate normal by the Cramér-Wold device. Now consider a product of any two elements in (D.8). By the continuous mapping theorem, the product converges in distribution; by Chebyshev's inequality that $[\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$ for any two random variables and the finite variance of each element in (D.8), the product is uniformly integrable. Hence, for the central limit condition in Theorem A,

$$\psi_o(x) - \psi_p(x) = x \lim_{n\to\infty} \sigma_{c_n}^{-3} n^{1/2}\,\mathrm{E}\Big(c_n \sum_{i=1}^{n}\big(c_{ni}^2 - \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)\Big) \tag{D.11}$$

and similarly,

$$\psi_p(x) = x \lim_{n\to\infty} \sigma_{c_n}^{-3} n^{1/2}\,\mathrm{E}\Big(c_n \sum_{i=1}^{n}\big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}(c_{ni}^2)\big)\Big), \tag{D.12}$$

where

$$\begin{aligned}
&\sigma_{c_n}^{-3} n^{1/2}\,\mathrm{E}\Big(c_n \sum_{i=1}^{n}\big(c_{ni}^2 - \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1})\big)\Big) \\
&= \sigma_{c_n}^{-3} n^{1/2} \sum_{i=1}^{n} \mathrm{E}\big(\mathrm{E}(c_{ni}^3|\mathscr{F}_{n,i-1})\big) \\
&= \sigma_{c_n}^{-3} n^{-1} \sum_{i=1}^{n}\Big[a_{n,ii}^3\,\mathrm{E}(\epsilon_{ni}^2 - \sigma_0^2)^3 + 8\mu_3^2 \sum_{j=1}^{i-1} a_{n,ij}^3 + \mu_3 b_{ni}^3 + 12\sigma_0^2[(\mu_4 - \sigma_0^4)a_{n,ii} + \mu_3 b_{ni}] \sum_{j=1}^{i-1} a_{n,ij}^2 \\
&\quad + 3(\mu_4 - \sigma_0^4)a_{n,ii} b_{ni}^2 + 3a_{n,ii}^2 b_{ni}\,\mathrm{E}[\epsilon_{ni}(\epsilon_{ni}^2 - \sigma_0^2)^2]\Big]
\end{aligned} \tag{D.13}$$

and

$$\begin{aligned}
&\sigma_{c_n}^{-3} n^{1/2}\,\mathrm{E}\Big(c_n \sum_{i=1}^{n}\big(\mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - \mathrm{E}(c_{ni}^2)\big)\Big) \\
&= \sigma_{c_n}^{-3} n^{-1}\,\mathrm{E}\Bigg[\Big(\sum_{i=1}^{n}[a_{n,ii}(\epsilon_{ni}^2 - \sigma_0^2) + b_{ni}\epsilon_{ni}] + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} a_{n,ij}\epsilon_{ni}\epsilon_{nj}\Big)\cdot \\
&\quad \Big(8\sigma_0^2 \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} a_{n,ij}a_{n,ik}\epsilon_{nj}\epsilon_{nk} + 4\sigma_0^2 \sum_{i=1}^{n}\sum_{j=1}^{i-1} a_{n,ij}^2(\epsilon_{nj}^2 - \sigma_0^2) + 4\sum_{i=1}^{n}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni}) \sum_{j=1}^{i-1} a_{n,ij}\epsilon_{nj}\Big)\Bigg] \\
&= \sigma_{c_n}^{-3} n^{-1}\Bigg[4\sum_{i=1}^{n}\sum_{j=1}^{i-1}\Big[\sigma_0^2 a_{n,ij}^2[a_{n,jj}(\mu_4 - \sigma_0^4) + \mu_3 b_{nj}] + a_{n,ij}(\mu_3 a_{n,ii} + \sigma_0^2 b_{ni})(\mu_3 a_{n,jj} + \sigma_0^2 b_{nj})\Big] \\
&\quad + 16\sigma_0^6 \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} a_{n,ij}a_{n,ik}a_{n,jk}\Bigg].
\end{aligned} \tag{D.14}$$

Therefore, all conditions in Theorem A are met. Thus (17) holds by (D.6) and (19) holds by (D.7).

To establish the expansion of the bootstrapped version, we shall follow the arguments in Mykland (1992) [p. 8, below (3.5); p. 10, below (3.13)], and Proposition 3 of Mykland (1993). The asymptotic expansion is first extended for non-random bootstrap parameters, and then the use of a representation of weak convergence of distributions by a sequence of almost surely convergent random variables. For (18) and (20), first consider, instead of Assumption 1′, the case that ϵ_{ni}'s in $\epsilon_n = (\epsilon_{n1}, \ldots, \epsilon_{nn})'$ are randomly drawn from the distribution with mean zero, and moments μ_{jn} for $j = 2, \ldots, 8$ and $\mathrm{E}|\epsilon_{ni}|^{8(1+\delta)} < \infty$ for some $\delta > 0$, all non-stochastic but depending on n, such that $\mu_{jn} = \mu_j + o(1)$ for $j = 2, \ldots, 8$. Also the bootstrap distribution depends on a nonstochastic sequence of β_n for our LQ form such that $\beta_n = \beta_0 + O(n^{-1/2})$, as the residual vector is computed as $y_n - X_n\beta_n$. For such ϵ_{ni}'s, with only adjustments of notations and bounds of some inequalities in the proof above, we have the expansions (17) and (19), where $\psi_o(x)$ and $\psi_p(x)$ do not change by inspecting (D.11)–(D.14). For the bootstrap with ϵ_n^* sampled from the the recentered residuals $(I_n - \frac{1}{n}l_n l_n')(y_n - X_n\hat{\beta}_n)$, where $\hat{\beta}_n = (X_n'X_n)^{-1}X_n'y_n$, we have $\mu_{jn}^* = \mu_j + o_P(1)$ for $j = 2, \ldots, 8$ by Lemma 2 and $\hat{\beta}_n = \beta_0 + O_P(n^{-1/2})$ by Chebyshev's inequality. Then (18) and (20) follow by Theorem IV.13 on p. 71 of Pollard (1984), where a weakly convergent sequence of probability measures can be represented by an almost surely convergent sequence of random variables. □

Proof of Proposition 3. The $\mathbb{I}_n'$ may be written as $\mathbb{I}_n' = B_{1n} + B_{2n}$, where $B_{1n} = \frac{\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(H_n M_n)}{\sqrt{n}\underline{\hat{\sigma}_{c_n}}}$ and $B_{2n} = \frac{(\sigma_0^2 - \hat{\sigma}_n^2)\operatorname{tr}(H_n M_n)}{\sqrt{n}\underline{\hat{\sigma}_{c_n}}}$. By the mean value theorem,

$$\int_{-\infty}^{\infty} h(x)\, dF_n(x) = \mathrm{E}\, h(B_{1n} + B_{2n}) = \mathrm{E}\, h(B_{1n}) + \mathrm{E}[h'(\bar{B}_{1n})B_{2n}] = \int_{-\infty}^{\infty} h(x)\, dG_n(x) + \mathrm{E}[h'(\bar{B}_{1n})B_{2n}],$$

where $\bar{B}_{1n}$ is between B_{1n} and $B_{1n} + B_{2n}$, and $G_n(x)$ is the distribution function of B_{1n}. In the following, we first prove an asymptotic expansion for $\int_{-\infty}^{\infty} h(x)\, dF_n(x)$ by showing that (i) $\int_{-\infty}^{\infty} h(x)\, dG_n(x)$ has an asymptotic expansion based on martingales using Theorem 3 and (ii) $\mathrm{E}[h'(\bar{B}_{1n})B_{2n}] = o(n^{-1/2})$. Then $\int_{-\infty}^{\infty} h(x)\, dF_n^*(x)$ is shown to have a similar expansion. The result in the proposition follows by the expansions for $\int_{-\infty}^{\infty} h(x)\, dF_n(x)$ and $\int_{-\infty}^{\infty} h(x)\, dF_n^*(x)$.

(i) The distribution function of $D_n = \frac{\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(H_n M_n)}{\sqrt{n}\sigma_{c_n}}$ has an asymptotic expansion by Theorem 3. The D_n and B_{1n} only differ in the denominators: while D_n has σ_{c_n}, B_{1n} has an estimate of σ_{c_n}. By Theorem A and the proof of Theorem 3, for the asymptotic expansion of $\int_{-\infty}^{\infty} h(x)\, dG_n(x)$, we only need to show that there are sets Q_n with $\mathrm{P}(Q_n) = 1 - o(n^{-1/2})$ such that

$$\sup_n \mathrm{E}[\sqrt{n}|\underline{\hat{\sigma}_{c_n}^2}/\sigma_{c_n}^2 - 1|]^{1+\delta}\mathbf{1}(Q_n) < \infty \tag{D.15}$$

for some $\delta > 0$, and the central limit condition, for which we shall show that

$$\Big(c_n/\sigma_{c_n}, n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^n c_{ni}^2 - 1\big), n^{1/2}\big(\sigma_{c_n}^{-2}\sum_{i=1}^n \mathrm{E}(c_{ni}^2|\mathscr{F}_{n,i-1}) - 1\big), n^{1/2}\sigma_{c_n}^{-2}\big(\underline{\hat{\sigma}_{c_n}^2} - \sigma_{c_n}^2\big)\Big) \tag{D.16}$$

is asymptotically jointly normal.

We shall show that (D.15) holds with $\delta = 1$ and $Q_n = (\|\xi_n\| \le \delta, |\hat{\sigma}_n^2 - \sigma_0^2| \le \delta)$, where $\xi_n = (\xi_{n1}, \ldots, \xi_{n,k_x})' = -(X_n'X_n)^{-1}X_n'\epsilon_n = O_P(n^{-1/2})$. Note that $\mathrm{P}(\|\xi_n\| > \delta) \le \delta^{-2}\,\mathrm{E}(\xi_n'\xi_n) = O(n^{-1})$ and $\mathrm{P}(|\hat{\sigma}_n^2 - \sigma_0^2| > \delta) \le$

$\mathrm{P}(|n^{-1}(\epsilon_n'\epsilon_n - n\sigma_0^2)| > \frac{1}{2}\delta) + \mathrm{P}(|n^{-1}\epsilon_n' X_n(X_n'X_n)^{-1}X_n'\epsilon_n > \frac{1}{2}\delta) = O(n^{-1})$, since $\mathrm{P}(|n^{-1}(\epsilon_n'\epsilon_n - n\sigma_0^2)| > \frac{1}{2}\delta) = O(n^{-1})$ by Chebyshev's inequality and $\mathrm{P}(|n^{-1}\epsilon_n' X_n(X_n'X_n)^{-1}X_n'\epsilon_n > \frac{1}{2}\delta) = O(n^{-1})$ by Markov's inequality. Then $\mathrm{P}(Q_n) \geq 1 - \mathrm{P}(\|\xi_n\| > \delta) - \mathrm{P}(|\hat{\sigma}_n^2 - \sigma_0^2| > \delta) = 1 - O(n^{-1})$. Thus $\mathrm{P}(Q_n) = 1 - O(n^{-1})$. Since $\underline{\hat{\sigma}_{c_n}^2} = \hat{\sigma}_{c_n}^2 + (c_\sigma - \hat{\sigma}_{c_n}^2)\mathbf{1}(\hat{\sigma}_{c_n}^2 \leq c_\sigma)$, we have $\mathrm{E}[n(\underline{\hat{\sigma}_{c_n}^2} - \sigma_{c_n}^2)^2]\mathbf{1}(Q_n) \leq \mathrm{E}[n(\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2)^2]\mathbf{1}(Q_n) + \mathrm{E}[n(c_\sigma - \sigma_{c_n}^2)^2]\mathbf{1}(\hat{\sigma}_{c_n}^2 \leq c_\sigma)$. As $\sigma_{c_n}^2$ is bounded away from zero, we shall show that $\mathrm{E}[n(\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2)^2]\mathbf{1}(Q_n) < \infty$ and $\mathrm{E}[n(c_\sigma - \sigma_{c_n}^2)^2]\mathbf{1}(\hat{\sigma}_{c_n}^2 \leq c_\sigma) < \infty$ for (D.15). Note that $\hat{\epsilon}_n = \epsilon_n + X_n\xi_n$, $\hat{\epsilon}_{ni}^4$ may be expanded by writing it as $\hat{\epsilon}_{ni}^4 = (\epsilon_{ni} + \sum_{j=1}^{k_x} x_{n,ij}\xi_{nj})^4$. Using the expansion form of $\hat{\epsilon}_{ni}^4$, $\sqrt{n}(\hat{\mu}_{4n} - \mu_4)$ is the sum of $n^{-1/2}\sum_{i=1}^n(\epsilon_{ni}^4 - \mu_4)$ and terms with the form $n^{-1/2}(\sum_{i=1}^n \epsilon_{ni}^{r_0} x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}})\xi_{n1}^{r_1}\dots\xi_{n,k_x}^{r_{k_x}}$, where integers $r_j \geq 0$ for $j = 0, 1, \dots, k_x$ and $\sum_{j=0}^{k_x} r_j = 4$. Because $\mathrm{E}[n^{-1/2}\sum_{i=1}^n(\epsilon_{ni}^4 - \mu_4)]^2 < \infty$,

$$\begin{aligned}
&\mathrm{E}\{n^{-1/2}[\sum_{i=1}^n(\epsilon_{ni}^{r_0} - \mathrm{E}\,\epsilon_{ni}^{r_0})x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}}]\xi_{n1}^{r_1}\dots\xi_{n,k_x}^{r_{k_x}}\}^2\mathbf{1}(\|\xi_n\| \leq \delta) \\
&\leq \frac{1}{n}\delta^{2(4-r_0)}\sum_{i=1}^n \mathrm{E}(\epsilon_{ni}^{r_0} - \mathrm{E}\,\epsilon_{ni}^{r_0})^2 x_{n,i1}^{2r_1}\dots x_{n,i,k_x}^{2r_{k_x}} \\
&< \infty,
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{E}\big\{n^{-1/2}[\sum_{i=1}^n(\mathrm{E}\,\epsilon_{ni}^{r_0})x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}}]\xi_{n1}^{r_1}\dots\xi_{n,k_x}^{r_{k_x}}\big\}^2\mathbf{1}(\|\xi_n\| \leq \delta) \\
&\leq \delta^{2(4-r_0)-2}(\frac{1}{n}\sum_{i=1}^n|\mathrm{E}(\epsilon_{ni}^{r_0})x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}}|)^2 n\,\mathrm{E}(\xi_{nj}^2) \\
&\leq \delta^{2(4-r_0)-2}(\frac{1}{n}\sum_{i=1}^n|\mathrm{E}(\epsilon_{ni}^{r_0})x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}}|)^2 n\,\mathrm{E}(\xi_n'\xi_n) \\
&< \infty,
\end{aligned}$$

for some $j > 0$ such that $r_j > 0$, we have $\sup_n \mathrm{E}[\sqrt{n}(\hat{\mu}_4 - \mu_4)]^2\mathbf{1}(Q_n) < \infty$. Similarly, $\sup_n \mathrm{E}[\sqrt{n}(\hat{\sigma}_n^2 - \sigma_0^2)]^2\mathbf{1}(Q_n) < \infty$. Hence, $\mathrm{E}[\sqrt{n}(\hat{\sigma}_n^4 - \sigma_0^4)]^2\mathbf{1}(Q_n) = \mathrm{E}[n(\hat{\sigma}_n^2 - \sigma_0^2)^2(\hat{\sigma}_n^2 - \sigma_0^2 + 2\sigma_0^2)^2]\mathbf{1}(Q_n) \leq (2\delta^2 + 8\sigma_0^4)\,\mathrm{E}[n(\hat{\sigma}_n^2 - \sigma_0^2)^2]\mathbf{1}(Q_n) < \infty$. Therefore, $\sup_n \mathrm{E}[\sqrt{n}(\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2)]^2\mathbf{1}(Q_n) < \infty$. Furthermore, $\mathrm{E}[n(c_\sigma - \sigma_{c_n}^2)^2]\mathbf{1}(\hat{\sigma}_{c_n}^2 \leq c_\sigma) = n(c_\sigma - \sigma_{c_n}^2)^2\,\mathrm{P}(\hat{\sigma}_{c_n}^2 \leq c_\sigma) \leq n(c_\sigma - \sigma_{c_n}^2)^2\,\mathrm{P}(|\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2| \geq \sigma_{c_n}^2 - c_\sigma)$. Then $\mathrm{E}[n(c_\sigma - \sigma_{c_n}^2)^2]\mathbf{1}(\hat{\sigma}_{c_n}^2 \leq c_\sigma)$ is bounded if $\mathrm{P}(|\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2| \geq \sigma_{c_n}^2 - c_\sigma) = O(n^{-1})$. Since $\mathrm{P}(|n^{-1}\sum_{i=1}^n(\epsilon_{ni}^4 - \mu_4)| \geq \eta) = O(n^{-1})$ for $\eta > 0$ and

$$\begin{aligned}
&\mathrm{P}(|n^{-1}(\sum_{i=1}^n \epsilon_{ni}^{r_0}x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}})\xi_{n1}^{r_1}\dots\xi_{n,k_x}^{r_{k_x}}| \geq \eta) \\
&\leq \mathrm{P}(|n^{-1}(\sum_{i=1}^n \epsilon_{ni}^{r_0}x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}})\xi_{n1}^{r_1}\dots\xi_{n,k_x}^{r_{k_x}}| \geq \eta, \|\xi_n\| \leq \eta) + \mathrm{P}(\|\xi_n\| > \eta) \\
&\leq \mathrm{P}(|n^{-1}[\sum_{i=1}^n(\epsilon_{ni}^{r_0} - \mathrm{E}\,\epsilon_{ni}^{r_0})x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}}]|\eta^{4-r_0} \geq \frac{\eta}{2}) + \mathrm{P}(n^{-1}\sum_{i=1}^n|(\mathrm{E}\,\epsilon_{ni}^{r_0})x_{n,i1}^{r_1}\dots x_{n,i,k_x}^{r_{k_x}}|\cdot\eta^{3-r_0}\|\xi_n\| \geq \frac{\eta}{2}) \\
&\quad + \mathrm{P}(\|\xi_n\| > \eta) \\
&= O(n^{-1}),
\end{aligned}$$

by Chebyshev's inequality, we have $\mathrm{P}(|\hat{\mu}_{4n} - \mu_4| \geq \eta) = O(n^{-1})$. As $\mathrm{P}(|\hat{\sigma}_n^2 - \sigma_0^2| \geq \eta) = O(n^{-1})$, it follows that $\mathrm{P}(|\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2| \geq \sigma_{c_n}^2 - c_\sigma) = O(n^{-1})$.

For (D.16), the first three terms have been shown to be asymptotically trivariate normal by a martingale central limit theorem. We shall show that the last term is asymptotically normal also by the martingale central limit theorem, and the joint asymptotic normality of (D.16) follows by the Cramér-Wold device by similar arguments to any linear combination of terms in (D.16). Note that $n^{1/2}(\underline{\hat{\sigma}_{c_n}^2} - \sigma_{c_n}^2) = n^{1/2}(\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2) + n^{1/2}(c_\sigma - \hat{\sigma}_{c_n}^2)\mathbf{1}(\hat{\sigma}_{c_n}^2 \leq c_\sigma)$, where $n^{1/2}(c_\sigma - \hat{\sigma}_{c_n}^2)\mathbf{1}(\hat{\sigma}_{c_n}^2 \leq c_\sigma) = o_P(1)$ as in the proof of Proposition 1, then $n^{1/2}(\underline{\hat{\sigma}_{c_n}^2} - \sigma_{c_n}^2)$ has the same asymptotic distribution as $n^{1/2}(\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2)$. As $\hat{\epsilon}_n = H_n\epsilon_n$, we have

$$\sqrt{n}(\hat{\sigma}_n^2 - \sigma_0^2) = \frac{1}{\sqrt{n}}(\epsilon_n'\epsilon_n - n\sigma_0^2) - \frac{1}{\sqrt{n}}\epsilon_n' X_n(\frac{1}{n}X_n'X_n)^{-1}\frac{1}{n}X_n'\epsilon_n = \frac{1}{\sqrt{n}}(\epsilon_n'\epsilon_n - n\sigma_0^2) + o_P(1).$$

Then

$$\sqrt{n}(\hat{\sigma}_n^4 - \sigma_0^4) = \sqrt{n}(\hat{\sigma}_n^2 + \sigma_0^2)(\hat{\sigma}_n^2 - \sigma_0^2) = 2\sigma_0^2\frac{1}{\sqrt{n}}(\epsilon_n'\epsilon_n - n\sigma_0^2) + o_P(1). \tag{D.17}$$

Let e_{ni} be the ith column of the n-dimensional identity matrix. Since $\hat{\epsilon}_n = \epsilon_n + X_n\xi_n$, where $\xi_n = O_p(n^{-1/2})$, by an argument as for the proof of Lemma 2, we have

$$\begin{aligned}\sqrt{n}(\hat{\mu}_4 - \mu_4) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n(\epsilon_{ni}^4 - \mu_4) - \frac{4}{\sqrt{n}}\sum_{i=1}^n \epsilon_{ni}^3 e_{ni}' X_n(X_n'X_n)^{-1}X_n'\epsilon_n + o_P(1)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n(\epsilon_{ni}^4 - \mu_4) - \frac{4\mu_3}{\sqrt{n}}l_n'X_n(X_n'X_n)^{-1}X_n'\epsilon_n + o_P(1),\end{aligned} \tag{D.18}$$

where the second equality holds because

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n(\epsilon_{ni}^3 - \mu_3)e_{ni}'X_n(X_n'X_n)^{-1}X_n'\epsilon_n = \sum_{j=1}^{k_x}\frac{1}{\sqrt{n}}\sum_{i=1}^n(\epsilon_{ni}^3 - \mu_3)e_{ni}'X_n e_{k_x,j}e_{k_x,j}'(X_n'X_n)^{-1}X_n'\epsilon_n$$

with $\frac{1}{\sqrt{n}}\sum_{i=1}^n(\epsilon_{ni}^3 - \mu_3)e_{ni}'X_n e_{k_x,j} = O_P(1)$ by Chebyshev's inequality. Hence,

$$\begin{aligned}\sqrt{n}(\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2) &= \frac{1}{n}\sum_{i=1}^n(H_nM_nH_n)_{ii}^2[\sqrt{n}(\hat{\mu}_4 - \mu_4) - 3\sqrt{n}(\hat{\sigma}_n^4 - \sigma_0^4)] + \frac{1}{n}\operatorname{tr}[H_nM_nH_n(M_n + M_n')]\sqrt{n}(\hat{\sigma}_n^4 - \sigma_0^4)\\ &= \sum_{j=1}^n v_{nj} + o_P(1),\end{aligned}$$

where

$$\begin{aligned}v_{nj} = &\frac{1}{n}\sum_{i=1}^n(H_nM_nH_n)_{ii}^2\frac{1}{\sqrt{n}}[(\epsilon_{nj}^4 - \mu_4) - 4\mu_3 l_n'X_n(X_n'X_n)^{-1}X_n'e_{nj}\epsilon_{nj} - 6\sigma_0^2(\epsilon_{nj}^2 - \sigma_0^2)]\\ &+ \frac{2\sigma_0^2}{n}\operatorname{tr}[H_nM_nH_n(M_n + M_n')]\frac{1}{\sqrt{n}}(\epsilon_{nj}^2 - \sigma_0^2).\end{aligned}$$

Under Assumption I1′, $\sum_{j=1}^n \mathrm{E}|v_{nj}|^q = o(1)$ for some $q > 2$. Note that $\mathrm{E}(v_{nj}) = 0$ and v_{ni} and v_{nj} are independent for $i \neq j$, then $\sqrt{n}\sigma_{c_n}^{-2}(\hat{\sigma}_{c_n}^2 - \sigma_{c_n}^2)$ is asymptotically normal. As explained in the proof of Theorem 3, (D.16) is asymptotically normal by the Cramér-Wold device, and in the central limit condition of Theorem A, $\psi_o(x)$ and $\psi_p(x)$ are as given in Theorem 3 for $c_n = n^{-1/2}[\epsilon_n' H_n M_n H_n \epsilon_n - \sigma_0^2 \operatorname{tr}(H_n M_n)]$ and $\sigma_{c_n}^2 = n^{-1}(\mu_4 - 3\sigma_0^4)\sum_{i=1}^n (H_n M_n H_n)_{ii}^2 + n^{-1}\sigma_0^4 \operatorname{tr}[H_n M_n H_n (M_n + M_n')]$, and

$$\begin{aligned}\psi_*(x) &= x \lim_{n\to\infty} \sigma_{c_n}^{-3} \sum_{j=1}^n \mathrm{E}(c_{nj} v_{nj}) \\ &= x \lim_{n\to\infty} \frac{1}{n\sigma_{c_n}^3} \sum_{j=1}^n \Big\{ [a_{n,jj}(\mu_6 - \mu_4\sigma_0^2) - 4a_{n,jj}\mu_3^2 l_n' X_n (X_n'X_n)^{-1} X_n' e_{nj} - 6a_{n,jj}\sigma_0^2(\mu_4 - \sigma_0^4) \\ &\quad + \mu_5 b_{nj} - 4\mu_3\sigma_0^2 b_{nj} l_n' X_n (X_n'X_n)^{-1} X_n' e_{nj} - 6\mu_3\sigma_0^2 b_{nj}] \frac{1}{n}\sum_{i=1}^n (H_n M_n H_n)_{ii}^2 \\ &\quad + [2a_{n,jj}\sigma_0^2(\mu_4 - \sigma_0^4) + 2\sigma_0^2\mu_3 b_{nj}] \frac{1}{n} \operatorname{tr}[H_n M_n H_n (M_n + M_n')] \Big\}.\end{aligned}$$

Hence, by Theorem A and the proof of Theorem 3, we have

$$\int_{-\infty}^{+\infty} h(x)\, dG_n(x) = \int_{-\infty}^{+\infty} h(x)\, d\Phi(x) + \frac{1}{2} n^{-1/2} \mathrm{E}\{[\frac{1}{3}\psi_0(Z) + \frac{2}{3}\psi_p(Z)] h^{(2)}(Z) - \psi_*(Z) Z h^{(1)}(Z)\} + o(n^{-1/2}),$$

where Z is a standard normal random variable. This finishes the proof of (i).

(ii) As $\underline{\hat{\sigma}_{c_n}^2} \geq c_\sigma > 0$ and $h'(x)$ is bounded,

$$|\mathrm{E}[h'(\bar{B}_{1n}) B_{2n}]| \leq c |n^{-1/2} \operatorname{tr}(H_n M_n)| \, \mathrm{E}|\sigma_0^2 - \hat{\sigma}_n^2|,$$

for some constant c. Since $\operatorname{tr}(H_n M_n) = O(1)$, we have $\mathrm{E}[h'(\bar{B}_{1n}) B_{2n}] = o(n^{-1/2})$ if $\mathrm{E}|\sigma_0^2 - \hat{\sigma}_n^2| = o(1)$. The $\hat{\sigma}_n^2 - \sigma_0^2 = \frac{1}{n}(\epsilon_n'\epsilon_n - n\sigma_0^2) - \frac{1}{n}\epsilon_n' X_n (X_n'X_n)^{-1} X_n' \epsilon_n$, then

$$\begin{aligned}\mathrm{E}|\sigma_0^2 - \hat{\sigma}_n^2| &\leq \mathrm{E}|\frac{1}{n}(\epsilon_n'\epsilon_n - n\sigma_0^2)| + \mathrm{E}[\frac{1}{n}\epsilon_n' X_n (X_n'X_n)^{-1} X_n' \epsilon_n] \\ &\leq [\mathrm{E}|\frac{1}{n}(\epsilon_n'\epsilon_n - n\sigma_0^2)|^2]^{1/2} + \frac{k_x}{n}\sigma_0^2 \\ &= [\frac{1}{n}(\mu_4 - \sigma_0^4)]^{1/2} + \frac{k_x}{n}\sigma_0^2.\end{aligned}$$

Thus $\mathrm{E}|\sigma_0^2 - \hat{\sigma}_n^2| = o(1)$ and $\mathrm{E}[h'(\bar{B}_{1n}) B_{2n}] = o(n^{-1/2})$.

With (i) and (ii), we have

$$\int_{-\infty}^{+\infty} h(x)\, dF_n(x) = \int_{-\infty}^{+\infty} h(x)\, d\Phi(x) + \frac{1}{2} n^{-1/2} \mathrm{E}\{[\frac{1}{3}\psi_0(Z) + \frac{2}{3}\psi_p(Z)] h^{(2)}(Z) - \psi_*(Z) Z h^{(1)}(Z)\} + o(n^{-1/2}).$$

By an argument similar to that in the proof of Theorem 3, we have

$$\int_{-\infty}^{+\infty} h(x)\, dF_n^*(x) = \int_{-\infty}^{+\infty} h(x)\, d\Phi(x) + \frac{1}{2} n^{-1/2} \mathrm{E}\{[\frac{1}{3}\psi_0(Z) + \frac{2}{3}\psi_p(Z)] h^{(2)}(Z) - \psi_*(Z) Z h^{(1)}(Z)\} + o_P(n^{-1/2}).$$

Therefore,

$$\int_{-\infty}^{+\infty} h(x)\, d[F_n^*(x) - F_n(x)] = o_P(n^{-1/2}). \qquad \square$$

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.

Anselin, L., Bera, A., 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics, in: Ullah, A., Giles, D.E. (Eds.), Handbook of Applied Economic Statistics. Marcel Dekker, New York, pp. 237–289.

Bickel, P.J., Götze, F., van Zwet, W.R., 1986. The Edgeworth expansion for U-statistics of degree two. Annals of Statistics 14, 1463–1484.

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

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., 2013. Cox-type tests for competing spatial autoregressive models with spatial autoregressive disturbances. Regional Science and Urban Economics 43, 590–616.

Jing, B.Y., Wang, Q., 2003. Edgeworth expansion for U-statistics under minimal conditions. Annals of Statistics 31, 1376–1391.

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., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review 40, 509–533.

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

Lee, L.F., 2004b. A supplement to asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Department of Economics, The Ohio State University, Columbus, OH.

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

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.

Lukacs, E., Laha, R.G., 1964. Applications of characteristic functions. Hafner Publishing Company: New York.

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., 1992. Asymptotic expansions and bootstrapping distributions for dependent variables: A martingale approach. Annals of Statistics 20, 623–654.

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

Mykland, P.A., 1995. Martingale expansions and second order inference. Annals of Statistics 23, 707–731.

Pollard, D., 1984. Convergence of stochastic processes. Springer, New York.

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.

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.