"LM Tests of Spatial Dependence Based on Bootstrap Critical Values"

A Supplement to Appendix A: Detailed Proof of Lemma A8

Zhenlin Yang

School of Economics, Singapore Management University,

zlyang@smu.edu.sg

Lemma A8: *For models specified by (9), (16) and (23) with $u_n = \Omega_n^{\frac{1}{2}}(\lambda)\varepsilon_n$, assume (a) Assumptions S1-S3 hold, (b) the unrestricted QMLEs of the parameters that the tests concern are $\sqrt{n_r}$-consistent,*[1] *and (c) the matrices $B_n^{-1}(\rho)$, $A_n^{-1}(\lambda)$ and $\Omega_n^{-\frac{1}{2}}(\lambda)$ defined therein are uniformly bounded in both row and column sums. Then, (i) $\hat\kappa_{3n} = \kappa_3 + O_p(n_r^{-\frac{1}{2}})$ and $\hat\kappa_{jn} = \kappa_j + o_p(1), j = 4, 5, 6,$ (ii) $\tilde\kappa_{jn} = O_p(1), j = 3, 4, 5, 6,$ and (iii) if $\kappa_3 = 0$ and conditions in* (A-7) *hold for model (16), then both $\hat\kappa_{3n}$ and $\tilde\kappa_{3n}$ are $o_p(1)$. Finally, the results remain for Model (23) if instead $u_n = W_n v_n + \varepsilon_n$ such that the jth sample cumulant of $\sigma^{-1}\Omega_n^{-\frac{1}{2}}(\lambda)u_n \xrightarrow{p} \kappa_j,\ j = 1, \ldots, 6.$*

Proof: Note that $\ddot\kappa_{jn}$ is the jth cumulant of $\ddot\sigma_n^{-1}\ddot\varepsilon_n$ where $\ddot{}$ denotes either $\hat{}$ or $\tilde{}$, and that $\ddot\kappa_{1n} = \kappa_1 = 0$ and $\ddot\kappa_{2n} = \kappa_2 = 1$ by construction. Note also that ρ and λ denote the true parameter values and the null hypothesis, $\rho = 0$ or $\lambda = 0$, is not imposed.

Proof for SED Model (9): $Y_n = X_n\beta + B_n^{-1}(\rho)\varepsilon_n$. The unrestricted residuals are

$$\hat\varepsilon_n = B_n(\hat\rho_n)(Y_n - X_n\hat\beta_n) = M_n(\hat\rho_n)B_n(\hat\rho_n)Y_n,$$

where $M_n(\rho) = I_n - B_n(\rho)X_n[X_n'B_n(\rho)'B_n(\rho)X_n]^{-1}X_n'B_n(\rho)'$ is the projection matrix, also defined below (9) of the main text. Since $M_n(\hat\rho_n)B_n(\hat\rho_n)X_n = 0$ and $B_n(\hat\rho_n) = B_n(\rho_n) - (\hat\rho_n - \rho)W_n$, we have

$$\begin{aligned}\hat\varepsilon_n &= M_n(\hat\rho_n)B_n(\hat\rho_n)B_n^{-1}(\rho)\varepsilon_n \\ &= \varepsilon_n - (I_n - M_n(\hat\rho_n))\varepsilon_n - (\hat\rho_n - \rho)M_n(\hat\rho_n)W_nB_n^{-1}(\rho)\varepsilon_n \equiv \varepsilon_n - \varepsilon_n^\dagger. \qquad \text{(A-1)}\end{aligned}$$

As the elements of X_n are uniformly bounded (Assumption S2) and $B_n(\rho) = I_n - \rho W_n$ is uniformly bounded in both row and column sums (Assumption S3), the elements of $B_n(\rho)X_n$ are uniformly bounded. $B_n'(\rho)B_n(\rho)$ is positive definite by Assumption (c). Thus, $\lim_{n\to\infty}\frac{1}{n}X_n'B_n'(\rho)B_n(\rho)X_n$ exists and is nonsingular, and $M_n(\rho)$ is uniformly bounded in both row and column sums (Lee, 2004b, Lemma A.5). By a result of Kelejian and Prucha (1999) or Lee (2002): product of two conformable square matrices uniformly bounded in both row and column sums is also uniformly bounded in both row and column sums, $M_n(\rho)W_nB_n^{-1}(\rho)$ is uniformly bounded in both row and column sums. Thus, By Lemma A.3 of Lee (2004b), $M_n(\hat\rho_n)W_nB_n^{-1}(\rho)$ uniformly bounded (in probability) in both row and column sums. This shows that the elements of $M_n(\hat\rho_n)W_nB_n^{-1}(\rho)\varepsilon_n$ are $O_p(1)$, and hence the elements of $\varepsilon_n^\dagger$ are $O_p(n_r^{-\frac{1}{2}})$ because the elements of $I_n - M_n(\hat\rho_n)$ are $O_p(n^{-1})$ and $\hat\rho_n - \rho = O_p(n_r^{-\frac{1}{2}})$.

Using $\hat\sigma_n^2 = \sigma^2 + O_p(n_r^{-\frac{1}{2}})$, we have by (A-1) for the 3rd cumulant,

$$\begin{aligned}\hat\kappa_{3n} &= \tfrac{1}{n\hat\sigma_n^3}(\hat\varepsilon_n \odot \hat\varepsilon_n)'\hat\varepsilon_n \\ &= \tfrac{1}{n\sigma^3}(\varepsilon_n \odot \varepsilon_n)'\varepsilon_n - \tfrac{3}{n\sigma^3}(\varepsilon_n \odot \varepsilon_n)'\varepsilon_n^\dagger + \tfrac{3}{n\sigma^3}(\varepsilon_n \odot \varepsilon_n^\dagger)'\varepsilon_n^\dagger - \tfrac{1}{n\sigma^3}(\varepsilon_n^\dagger \odot \varepsilon_n^\dagger)'\varepsilon_n^\dagger + O_p(n_r^{-\frac{1}{2}}),\end{aligned}$$

where $\odot$ denotes the Hadamard product. As the elements of $\varepsilon_n^\dagger$ are $O_p(n_r^{-\frac{1}{2}})$, all terms involving $\varepsilon_n^\dagger$ are $O_p(n_r^{-\frac{1}{2}})$ or smaller. Further, by the generalized Chebyshev inequality and Assumption S3:

[1] The $\sqrt{n_r}$-consistency of $\hat\lambda_n$ for the SLD model is proved by Lee (2004a). Similarly, one can prove the $\sqrt{n_r}$-consistency of $\hat\rho_n$ for the SED model and that of $\hat\lambda_n$ for the SEC model. Following Lee (2004a), it can be proved that $\hat\sigma_n^2$ is always $\sqrt{n}$-consistent, but $\hat\beta_n$ is $\sqrt{n_r}$-consistent in general for the SLD model and $\sqrt{n}$-consistent for the other two models.

$P(\sqrt{n}|\frac{1}{n}\sum_{i=1}^n e_{n,i}^3 - \kappa_3| \geq M) \leq \frac{1}{M^2}\frac{1}{n}\mathrm{Var}(\sum_{i=1}^n e_{n,i}^3) = \frac{1}{M^2}O(1)$. It follows that $\frac{1}{n\sigma^3}(\varepsilon_n \odot \varepsilon_n)'\varepsilon_n = \frac{1}{n}\sum_{i=1}^n e_{n,i}^3 = \kappa_3 + O_p(n^{-\frac{1}{2}})$ and that $\hat{\kappa}_{3n} = \kappa_3 + O_p(n_r^{-\frac{1}{2}})$. Similarly, one shows that $\hat{\kappa}_{4n} = \frac{1}{n\hat{\sigma}_n^4}(\hat{\varepsilon}_n \odot \hat{\varepsilon}_n)'(\hat{\varepsilon}_n \odot \hat{\varepsilon}_n) - 3 = \frac{1}{n}\sum_{i=1}^n e_{n,i}^4 - 3 + O_p(n_r^{-\frac{1}{2}}) = \kappa_4 + o_p(1)$, where the last step follows Kolmogorov law of large numbers: $\frac{1}{n}\sum_{i=1}^n e_{n,i}^4 = \mathrm{E}(e_{n,i}^4) + o_p(1)$, and that $\hat{\kappa}_{jn} = \kappa_j + o_p(1), j = 5, 6$.[2]

To prove (ii) and (iii), note that the restricted residuals $\tilde{\varepsilon}_n = M_n Y_n = M_n B_n^{-1}(\rho)\varepsilon_n \equiv G_n\varepsilon_n$, where $M_n = M_n(0)$. We have for 3rd cumulant, $\tilde{\kappa}_{3n} = \frac{1}{n\tilde{\sigma}_n^3}(\tilde{\varepsilon}_n \odot \tilde{\varepsilon}_n)'\tilde{\varepsilon}_n = \frac{1}{n\tilde{\sigma}_n^3}\sum_{i=1}^n (g_{n,i}'\varepsilon_n)^3$, where $g_{n,i}'$ is the ith row of G_n. Let $g_{n,ij}$ be the jth element of $g_{n,i}$. Obviously, G_n is uniformly bounded in both absolute row and column sums, i.e., $\max_j \sum_i |g_{n,ij}| \leq c_1$ and $\max_i \sum_j |g_{n,ij}| \leq c_\infty$, and hence $|g_{n,ij}| \leq c_0, \forall i, j$, for finite positive constants c_0, c_1 and c_∞. We have,

$$\mathrm{E}[\tfrac{1}{n}\textstyle\sum_{i=1}^n (g_{n,i}'\varepsilon_n)^3] = \frac{1}{n}\sigma^3\kappa_3\sum_{i=1}^n\sum_{j=1}^n g_{n,ij}^3 = O(1), \quad \text{(A-2)}$$

since $\frac{1}{n}|\sum_{i=1}^n\sum_{j=1}^n g_{n,ij}^3| \leq \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n |g_{n,ij}^3| \leq \frac{1}{n}c_0^2\sum_{i=1}^n\sum_{j=1}^n |g_{n,ij}| \leq c_0^2 c_\infty = O(1)$; and

$$\tilde{\sigma}_n^2 = \frac{1}{n}\tilde{\varepsilon}_n'\tilde{\varepsilon}_n = \frac{1}{n}\sigma^2\mathrm{tr}(G_n'G_n) + O_p(n^{-\frac{1}{2}}) \equiv \bar{\sigma}_n^2 + O_p(n^{-\frac{1}{2}}), \quad \text{(A-3)}$$

by generalized Chebyshev inequality, where $\bar{\sigma}_n^2 = O(1)$ as $\frac{1}{n}\mathrm{tr}(G_n'G_n) = O(1)$. Further,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n (g_{n,i}'\varepsilon_n)^3 &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n g_{n,ij}g_{n,ik}g_{n,il}\varepsilon_{n,j}\varepsilon_{n,k}\varepsilon_{n,l} \\
&= \frac{1}{n}\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n h_{n,jkl}\,\varepsilon_{n,j}\varepsilon_{n,k}\varepsilon_{n,l}, \quad \text{where } h_{n,jkl} = \sum_{i=1}^n g_{n,ij}g_{n,ik}g_{n,il} \\
&= \frac{1}{n}\sum_{j=1}^n h_{n,jjj}\,\varepsilon_{n,j}^3 + \frac{6}{n}\sum_{j>k}\sum h_{n,jjk}\varepsilon_{n,j}\varepsilon_{n,k}^2 + \frac{4}{n}\sum_{j>k>l}\sum\sum h_{n,jkl}\varepsilon_{n,j}\varepsilon_{n,k}\varepsilon_{n,l}, \\
&= \frac{1}{n}\sum_{j=1}^n h_{n,jjj}\,\varepsilon_{n,j}^3 + \frac{6}{n}\sum_{j=1}^n \varepsilon_{n,j}u_{n,j} + \frac{4}{n}\sum_{j=l}^n \varepsilon_{n,j}v_{n,j},
\end{aligned}$$

where $u_{nj} = \sum_{k=1}^{j-1} h_{n,jjk}\varepsilon_{n,k}^2$ and $v_{n,j} = \sum_{k=1}^{j-1}\sum_{l=1}^{k-1} h_{n,jkl}\varepsilon_{n,k}\varepsilon_{n,l}$. Let $\mathcal{S}_{n,j}$ be the increasing σ-field generated by $\{\varepsilon_{n,1}, \ldots, \varepsilon_{n,j}\}$. As both $u_{n,j}$ and $v_{n,j}$ are $\mathcal{S}_{n,j-1}$-measurable, they are independent of $\varepsilon_{n,j}$. Thus, $\{\varepsilon_{n,j}^3 - \mathrm{E}\varepsilon_{n,j}^3, 1 \leq j \leq n\}$, $\{\varepsilon_{n,j}u_{n,j}, 1 \leq j \leq n\}$, and $\{\varepsilon_{n,j}v_{n,j}, 1 \leq j \leq n\}$ each forms a martingale difference (m.d.) array, and the weak law of large numbers (WLLN) for m.d. arrays (see, e.g., Davidson, 1994, p. 299) can be applied to show that:

$$(a)\ \tfrac{1}{n}\textstyle\sum_{j=1}^n h_{n,jjj}\,(\varepsilon_{n,j}^3 - \mathrm{E}\varepsilon_{n,j}^3) \xrightarrow{p} 0,\ (b)\ \tfrac{1}{n}\sum_{j=1}^n \varepsilon_{n,j}u_{n,j} \xrightarrow{p} 0, \ \text{ and } \ (c)\ \tfrac{1}{n}\sum_{j=l}^n \varepsilon_{n,j}v_{n,j} \xrightarrow{p} 0.$$

For (a), first $|\varepsilon_{n,j}^3 - \mathrm{E}\varepsilon_{n,j}^3|$ is uniformly integrable. It is easy to see that $h_{n,jjj} = \sum_{i=1}^n g_{n,ij}^3 > 0$, $\frac{1}{n}\sum_{j=1}^n h_{n,jjj} \leq \frac{1}{n}\sum_{j=1}^n\sum_{i=1}^n |g_{n,ij}|^3 \leq c_0^2\frac{1}{n}\sum_{j=1}^n\sum_{i=1}^n |g_{n,ij}| \leq c_0^2 c_1 < \infty$, and $\frac{1}{n^2}\sum_j h_{n,jjj}^2 = \frac{1}{n^2}\sum_j(\sum_i g_{n,ij}^3)^2 \leq \frac{1}{n^2}\sum_j(\sum_i |g_{n,ij}^3|)^2 \leq \frac{1}{n^2}c_0^4\sum_j(\sum_i |g_{n,ij}|)^2 \leq \frac{1}{n}c_0^4 c_1^2$. Thus, (a) follows by the WLLN for m.d. arrays. For (b), we have $\mathrm{E}|\varepsilon_{n,j}u_{n,j}| = \mathrm{E}|\varepsilon_{n,j}|\ \mathrm{E}|u_{n,j}|$, and

$$\sigma^{-2}\mathrm{E}|u_{n,j}| \leq \textstyle\sum_{k=1}^{j-1}|h_{n,jjk}| \leq \sum_{k=1}^{j-1}\sum_{i=1}^n g_{n,ij}^2|g_{n,ik}| = \sum_{i=1}^n g_{n,ij}^2\sum_{k=1}^{j-1}|g_{n,ik}| \leq c_\infty\sum_{i=1}^n g_{n,ij}^2$$

where $\{\sum_{i=1}^n g_{n,ij}^2\}$ are the diagonal elements of $G_n'G_n$ and thus are uniformly bounded. It is follows that the sequence $\{|\varepsilon_{n,j}u_{n,j}|\}$ are uniformly integrable, and thus the result (b) follows by the WLLN for m.d. arrays. Similarly, (c) follows due to the uniform integrability of $\{|\varepsilon_{n,j}v_{n,j}|\}$, because,

[2] Note that the results become $\hat{\kappa}_{jn} = \kappa_j + O_p(n_r^{-\frac{1}{2}})$ if $(2j)$th moment of $e_{n,i}$ exists, $j = 4, 5, 6$.

$$
\begin{aligned}
\mathrm{E}|v_{n,j}| &= \mathrm{E}\big|\textstyle\sum_{k=1}^{j-1}\sum_{l=1}^{k-1} h_{n,jkl}\varepsilon_{n,k}\varepsilon_{n,l}\big| \le \sum_{k=1}^{j-1}\sum_{l=1}^{k-1} |h_{n,jkl}|\mathrm{E}|\varepsilon_{n,k}|\mathrm{E}|\varepsilon_{n,l}| \\
&\le (\mathrm{E}|\varepsilon_{n,k}|)^2 \textstyle\sum_{k=1}^{j-1}\sum_{l=1}^{k-1}\sum_{i=1}^{n} |g_{n,ij}||g_{n,ik}||g_{n,il}| \\
&\le (\mathrm{E}|\varepsilon_{n,k}|)^2 \big(\textstyle\sum_{i=1}^{n}|g_{n,ij}|\big)\big(\sum_{k=1}^{n}|g_{n,ik}|\big)\big(\sum_{l=1}^{n}|g_{n,il}|\big) \le (\mathrm{E}|\varepsilon_{n,k}|)^2 c_1 c_\infty^2.
\end{aligned}
$$

Hence, the results $(a)-(c)$ together with (A-2) and (A-3) lead to

$$
\tilde{\kappa}_{3n} = \kappa_3\left(\tfrac{\sigma^3}{n\bar{\sigma}^3}\textstyle\sum_{i=1}^{n}\sum_{j=1}^{n} g_{n,ij}^3\right) + o_p(1) = O_p(1), \tag{A-4}
$$

where the leading term equals to κ_3 if $\rho = 0$ but not in general. The results for the other terms in (ii) can be proved in the same manner, but obviously much more tedious. The results in (iii) follow immediate from a result in (i) for $\hat{\kappa}_{3n}$, and from (A-4) for $\tilde{\kappa}_{3n}$.

Proof for SLD Model (16): $A_n(\lambda)Y_n = X_n\beta + \varepsilon_n$. The unrestricted residuals are $\hat{\varepsilon}_n = M_nA_n(\hat{\lambda}_n)Y_n = M_nA_n(\lambda)Y_n - (\hat{\lambda}_n - \lambda)M_nW_nY_n = M_n\varepsilon_n - (\hat{\lambda}_n - \lambda)M_nW_nY_n = \varepsilon_n - \varepsilon_n^\dagger$, where $\varepsilon_n^\dagger = (I_n - M_n)\varepsilon_n + (\hat{\lambda}_n - \lambda)M_nW_nY_n$. It is easy to see that the elements of $\varepsilon_n^\dagger$ are $O_P(n_r^{-\frac{1}{2}})$, and thus similar arguments as for the SED model lead to the results in (i).

To show (ii) and (iii), the restricted residuals are $\tilde{\varepsilon}_n = M_nY_n = M_nA_n^{-1}(\lambda)(X_n\beta+\varepsilon_n) = \mu_n + G_n\varepsilon_n$, where $\mu_n = G_nX_n\beta$ and $G_n = M_nA_n^{-1}(\lambda)$ as that for the SED model. We have for the 3rd cumulant,

$$
\begin{aligned}
\tilde{\sigma}_n^3\tilde{\kappa}_{3n} &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n}(\mu_{n,i} + g_{n,i}'\varepsilon_n)^3 \\
&= \tfrac{1}{n}\textstyle\sum_{i=1}^{n}\mu_{n,i}^3 + \tfrac{3}{n}\sum_{i=1}^{n}\mu_{n,i}^2(g_{n,i}'\varepsilon_n) + \tfrac{3}{n}\sum_{i=1}^{n}\mu_{n,i}(g_{n,i}'\varepsilon_n)^2 + \tfrac{1}{n}\sum_{i=1}^{n}(g_{n,i}'\varepsilon_n)^3,
\end{aligned}
$$

where $\frac{1}{n}\sum_{i=1}^{n}(g_{n,i}'\varepsilon_n)^3 = \kappa_3\big(\frac{\sigma^3}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}g_{n,ij}^3\big) + o_p(1)$ as for the SED model; $\frac{1}{n}\sum_{i=1}^{n}\mu_{n,i}^2(g_{n,i}'\varepsilon_n) = \frac{1}{n}\sum_{j=1}^{n}\psi_{n,j}\varepsilon_{n,j} = o_p(1)$ by the WLLN for m.d. arrays where $\psi_{n,j} = \sum_{i=1}^{n}\mu_{n,i}^2 g_{n,ij}$; and

$$
\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\mu_{n,i}(g_{n,i}'\varepsilon_n)^2 = \tfrac{1}{n}\sum_{j=1}^{n}\zeta_{n,j}\varepsilon_{n,j}^2 + \tfrac{2}{n}\sum_{j=1}^{n}u_{n,j}\varepsilon_{n,j} = \tfrac{\sigma^2}{n}\sum_{j=1}^{n}\zeta_{n,j} + o_p(1),
$$

by the WLLN for m.d., arrays, where $\zeta_{n,j} = \sum_{i=1}^{n}\mu_{n,i}g_{n,ij}^2$ and $u_{n,j} = \sum_{k=1}^{j-1}\big(\sum_{i=1}^{n}\mu_{n,i}g_{n,ij}g_{n,ik}\big)\varepsilon_{n,k}$. Finally by generalized Chebyshev inequality, we show that

$$
\tilde{\sigma}_n^2 = \tfrac{1}{n}\tilde{\varepsilon}_n'\tilde{\varepsilon}_n = \tfrac{1}{n}\mu_n'\mu_n + \tfrac{1}{n}\sigma^2\mathrm{tr}(G_n'G_n) + O_p(n^{-\frac{1}{2}}) \equiv \bar{\sigma}_n^2 + O_p(n^{-\frac{1}{2}}), \tag{A-5}
$$

where it is clear that $\bar{\sigma}_n^2$ is bounded away from zero. Putting everything together, we obtain,

$$
\tilde{\kappa}_{3n} = \tfrac{1}{n\bar{\sigma}_n^3}\textstyle\sum_{i=1}^{n}\mu_{n,i}^3 + \tfrac{\sigma^2}{n\bar{\sigma}_n^3}\sum_{j=1}^{n}\zeta_{n,j} + \kappa_3\big(\tfrac{\sigma^3}{n\bar{\sigma}_n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}g_{n,ij}^3\big) + o_p(1) = O_p(1). \tag{A-6}
$$

The results $\tilde{\kappa}_{jn} = O_p(1), j = 4, 5, 6$, can each be shown in a similar fashion as for $\tilde{\kappa}_{3n}$, but obviously at the cost of a much more tedious algebra.

Finally, when $\kappa_3 = 0$, $\hat{\kappa}_{3n} = o_p(1)$ as seen from the result in (i). Now, from (A-6), $\tilde{\kappa}_{3n} = \frac{1}{n\bar{\sigma}_n^3}\sum_{i=1}^{n}\mu_{n,i}^3 + \frac{\sigma^2}{n\bar{\sigma}_n^3}\sum_{j=1}^{n}\zeta_{n,j} + o_p(1)$. From (A-5), it is easy to show that $\bar{\sigma}_n^{-2} = O(1)$. It follows that $\tilde{\kappa}_{3n} = o_p(1)$, provided that

$$
\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mu_{n,i}^3 = 0 \text{ and } \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\zeta_{n,i} = 0. \tag{A-7}
$$

While these conditions are not restrictive, as average of of $\{\mu_{n,i}\}$ is zero, $\frac{1}{n}\sum_{i=1}^{n}\mu_{n,i}^3$ represents the sample skewness, and $\frac{1}{n}\sum_{j=1}^{n}\zeta_{n,j}$ is the sum of the sample covariances between μ_n and $g_{n,j}^2, j = 1, \ldots, n$, they do show that it is less reliable to use the restricted residuals than the unrestricted ones.

Proof for SEC Model (23). The unrestricted QML residuals for the SEC model are

$$\hat{\varepsilon}_n = \Omega_n^{-\frac{1}{2}}(\hat{\lambda}_n)(Y_n - X_n\hat{\beta}_n) = M_n(\hat{\lambda}_n)\Omega_n^{-\frac{1}{2}}(\hat{\lambda}_n)Y_n,$$

where $M_n(\lambda) = I_n - \Omega_n^{-\frac{1}{2}}(\lambda)X_n[X_n'\Omega_n^{-1}(\lambda)X_n]^{-1}X_n'\Omega_n^{-\frac{1}{2}}(\lambda)$, a projection matrix.

First, suppose the SEC model can be written in the form: $Y_n = X_n\beta + \Omega_n^{\frac{1}{2}}(\lambda)\varepsilon_n$, satisfying Assumptions S1-S3. Then, $\hat{\varepsilon}_n = M_n(\hat{\lambda}_n)\Omega_n^{-\frac{1}{2}}(\hat{\lambda}_n)\Omega_n^{\frac{1}{2}}(\lambda)\varepsilon_n = \varepsilon_n - (I_n - M_n(\hat{\lambda}_n)) + (\bar{\lambda}_n - \lambda)[\frac{d}{d\lambda}\Omega_n^{-\frac{1}{2}}(\lambda)]\Omega_n^{\frac{1}{2}}(\lambda)\varepsilon_n \equiv \varepsilon_n - \varepsilon_n^{\dagger}$, where $\bar{\lambda}_n$ lies between $\hat{\lambda}_n$ and λ and thus $\bar{\lambda}_n - \lambda = O_p(n_r^{-\frac{1}{2}})$. It is easy to see that the elements of $\varepsilon_n^{\dagger}$ are $O_p(n_r^{-\frac{1}{2}})$, and the rest of the proof for (i) follows that of the SED model. The proofs of (ii) and (iii) are the same as those of the SED model as $\tilde{\varepsilon}_n = M_nY_n = M_n\Omega_n^{\frac{1}{2}}(\lambda)\varepsilon_n$. It is easy to see that the results remain valid with a more general $\Omega_n(\lambda)$ matrix as discussed in Footnote 4.

Now, for the true SEC model: $Y_n = X_n\beta + u_n$, with $u_n = W_nv_n + \varepsilon_n$ and $\varepsilon_n = \sigma e_n$, we have, $\hat{\varepsilon}_n = M_n(\hat{\lambda}_n)\Omega_n^{-\frac{1}{2}}(\hat{\lambda}_n)Y_n = M_n(\hat{\lambda}_n)\Omega_n^{-\frac{1}{2}}(\hat{\lambda}_n)(W_nv_n + \varepsilon_n) = \Omega_n^{-\frac{1}{2}}(\lambda)(W_nv_n + \varepsilon_n) + O_p(n_r^{-\frac{1}{2}}) \equiv \epsilon_n + O_p(n_r^{-\frac{1}{2}})$. This shows that the EDF $\hat{\mathcal{F}}_n$ of $\hat{\sigma}_n^{-1}\hat{\varepsilon}_n$ and the EDF of $\sigma_n^{-1}\epsilon_n$, say $\mathcal{F}_n$, agree asymptotically, and hence their cumulants, where σ_n^2 is the sample variance of ϵ_n. By the assumption given in Proposition 3.3 and Lemma A8, $\kappa_{jn} = \kappa_j(\mathcal{F}_n) \xrightarrow{p} \kappa_j = \kappa_j(\mathcal{F})$, $j = 1, \ldots, 6$. It follows that $\hat{\kappa}_{jn} \xrightarrow{p} \kappa_j$, $j = 1, \ldots, 6$. Thus the results remain valid for SEC model with $u_n = W_nv_n + \varepsilon_n$.

A discussion on the plausibility of the underlining assumption is as follows. It is clear that $\kappa_{jn} \xrightarrow{p} \kappa_j$ for $j = 1, 2$. For $j \geq 3$, denote $\epsilon_n = H_nv_n + G_n\varepsilon_n$ where $H_n = \Omega_n^{-\frac{1}{2}}(\lambda)W_n$ with its elements denoted by $\{h_{n,ij}\}$ and $G_n = \Omega_n^{-\frac{1}{2}}(\lambda)$ with its elements denote by $\{g_{n,ij}\}$. Let κ_{vr} be the rth cumulant of $v_{n,i}$. By repeatedly using the WLLN for m.d. arrays, it is straightforward, though tedious, to prove the following useful result:

$$\kappa_{jn} = \kappa_{vj}\frac{\lambda^{j/2}}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}h_{n,it}^{j} + \kappa_j\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}g_{n,it}^{j} + o_p(1), \tag{A-8}$$

for $j \geq 3$. Thus, if $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}h_{n,it}^{j} \to 0$ and $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}g_{n,it}^{j} \to 1$, then, $\kappa_{jn} \to \kappa_j, j \geq 3$. It can be shown that these conditions are true for spatial layouts with unbounded h_n. In this case, the elements of W_nW_m are $O(h_n^{-1})$, and thus the diagonal elements of $\Omega_n(\lambda)$ are $1 + O(h_n^{-1})$ and off-diagonal elements are $O(h_n^{-1})$. This leads to that the diagonal elements of $\Omega_n^{-\frac{1}{2}}(\lambda)$ are $1+O(h_n^{-1})$ and off-diagonal elements are $O(h_n^{-1})$; and the elements of $\Omega_n^{-\frac{1}{2}}(\lambda)W_n$ are all $O(h_n^{-1})$ as the diagonal elements of W_n are zero and off-diagonal elements are $O(h_n^{-1})$. It follows that $\sum_{t=1}^{n}h_{n,it}^{j} = o(1)$ and $\sum_{t=1}^{n}g_{n,it}^{j} = 1 + o(1)$ for $j \geq 3$, and hence $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}h_{n,it}^{j} \to 0$ and $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}g_{n,it}^{j} \to 1$.

Obviously, when σ_v^2 is small relative to σ^2, $\lambda^{j/2}\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}h_{n,it}^{j} \approx 0$ and $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}g_{n,it}^{j} \approx 1$ for any spatial layouts as in this case, $H_n \approx W_n$ and $G_n \approx I_n$. An accurate approximation to the finite sample critical values when λ is close to its null value is clearly more important than when it is far away. Further, κ_{vj}, κ_j and κ_{jn}, $j \geq 3$, are all zero when v_n and e_n are both normally distributed.

Additional References

Davidson, J., 1994. *Stochastic Limit Theory.* Oxford University Press.

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.

Lee, L. F., 2002. Consistency and efficiency of least squares estimation for mixed regressive spatial autoregressive models. *Econometric Theory* 18, 252-277.