Supplement to "GEL estimation and tests of spatial autoregressive models"

Fei Jin[1] and Lung-fei Lee[*2]

[1]School of Economics, Fudan University, Shanghai 200433, China

[2]Department of Economics, The Ohio State University, Columbus, OH 43210, USA

November 13, 2018

Note: This supplement includes

(1) sufficient identification conditions for the GMM and GEL estimation;

(2) sufficient conditions for the gradient matrix $\bar{G}_n$ in Assumption 9 to have full rank in the limit;

(3) sufficient identification conditions for the spatial J test;

(4) proofs of lemmas;

(5) some expressions for the proof of Theorem 3.4.

(6) Monte Carlo results for the SARAR(1,1) and SE models.

A Sufficient identification conditions for the GMM and GEL estimation

In this section, we provide sufficient identification conditions for the GMM and hence GEL estimation of the SARAR(p,q) model. In the homoskedastic case, the conditions are similar to those in Liu et al. (2010) for the SARAR(1,1) model and in Lee and Liu (2010) for the SARAR(p,q) model. Our moment conditions include the variance parameter σ^2 but theirs do not, so we need to take into account this difference. In the heteroskedastic case, the conditions are modified from those in the homoskedastic case by involving different variances.

*Corresponding author. Tel.: +1 614 292 5508; fax: +1 614 292 4192. E-mail addresses: jin.fei@live.com (F. Jin), lee.1777@osu.edu (L.-F. Lee).

A.1 Homoskedastic case

The leading order term of the empirical linear moment $\frac{1}{n}Q_n'V_n(\theta)$ is the population moment $\frac{1}{n}Q_n'd_n(\theta)$, where $d_n(\theta) = \mathrm{E}[V_n(\theta)] = R_n(\tau)[\sum_{j=1}^p(\kappa_{0j}-\kappa_j)W_{jn}S_n^{-1}X_n\beta_0 + X_n(\beta_0-\beta)]$, and that of an empirical quadratic moment $\frac{1}{n}[V_n'(\theta)P_{ln}V_n(\theta) - \sigma^2\,\mathrm{tr}(P_{ln})]$ is

$$\begin{aligned}&\frac{1}{n}\,\mathrm{E}[V_n'(\theta)P_{ln}V_n(\theta) - \sigma^2\,\mathrm{tr}(P_{ln})]\\ &= \frac{1}{n}d_n'(\theta)P_{ln}d_n(\theta) + \frac{\sigma_0^2}{n}\,\mathrm{tr}\{[R_n(\tau)S_n(\kappa)S_n^{-1}R_n^{-1}]'P_{ln}[R_n(\tau)S_n(\kappa)S_n^{-1}R_n^{-1}]\} - \frac{\sigma^2}{n}\,\mathrm{tr}(P_{ln}).\end{aligned}$$

As β_0 only appears in $d_n(\theta)$ in these terms of population moment conditions, X_n needs to have full rank for large enough n for the identification and consistent estimation of β_0. Let $\Upsilon_n = [W_{1n}S_n^{-1}X_n\beta_0,\ldots,W_{pn}S_n^{-1}X_n\beta_0,X_n]$. Then $\frac{1}{n}Q_n'd_n(\theta) = \frac{1}{n}Q_n'R_n(\tau)\Upsilon_n[\kappa_{01}-\kappa_1,\ldots,\kappa_{0p}-\kappa_p,\beta_0'-\beta']'$, where $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n$ may or may not have full column rank for any τ in its parameter space.

If $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n$ has full column rank for any τ in its parameter space, then κ_0 and β_0 can be identified from the linear moments. As a result, only τ_0 and σ_0^2 need to be identified from the quadratic moments. With $\kappa=\kappa_0$ and $\beta=\beta_0$, $\lim_{n\to\infty}\frac{1}{n}\,\mathrm{E}[V_n'(\theta)P_{ln}V_n(\theta)-\sigma^2\,\mathrm{tr}(P_{ln})]=0$ for $l=1,\ldots,k_p$ become

$$\lim_{n\to\infty}\frac{1}{n}\{\sigma_0^2\,\mathrm{tr}[R_n'^{-1}R_n'(\tau)P_{ln}R_n(\tau)R_n^{-1}] - \sigma^2\,\mathrm{tr}(P_{ln})\} = 0 \text{ for } l=1,\ldots,k_p. \tag{A.1}$$

The identification of (τ_0',σ_0^2) via the quadratic moments is the same as that for the pure SAR process

$$U_n = \sum_{k=1}^q \tau_k M_{kn}U_n + V_n$$

via those moments, where $U_n = S_nY_n - X_n\beta_0$ is observable. Let Φ, Φ_{τ_k} and $\Phi_{\tau_{k_1k_2}}$ be $k_p\times 1$ vectors with the lth element being, respectively, $\Phi_l = \lim_{n\to\infty}\frac{1}{n}\,\mathrm{tr}(P_{ln})$, $\Phi_{\tau_k,l} = \lim_{n\to\infty}\frac{2}{n}\,\mathrm{tr}(P_{ln}M_{kn}R_n^{-1})$ and $\Phi_{\tau_{k_1k_2},l} = \lim_{n\to\infty}\frac{1}{n}\,\mathrm{tr}(R_n'^{-1}M_{k_1n}'P_{ln}M_{k_2n}R_n^{-1})$. Then, as $R_n(\tau) = R_n + \sum_{k=1}^q(\tau_{0k}-\tau_k)M_{kn}$, (A.1) is equivalent to $\frac{1}{\sigma_0^2}(\sigma_0^2-\sigma^2)\Phi + \sum_{k=1}^q(\tau_{0k}-\tau_k)\Phi_{\tau_k} + \sum_{k_1,k_2=1}^q(\tau_{0,k_1}-\tau_{k_1})(\tau_{0,k_2}-\tau_{k_2})\Phi_{\tau_{k_1k_2}} = 0$, which is a linear combination of Φ's with coefficients nonlinear in $(\tau_0'-\tau',\sigma_0^2-\sigma^2)$. It follows that (A.1) has a unique solution at (τ_0',σ_0^2) if and only if

$$c\Phi + \sum_{k=1}^q c_k\Phi_{\tau_k} + \sum_{k_1,k_2=1}^q c_{k_1}c_{k_2}\Phi_{\tau_{k_1k_2}} \neq 0 \text{ for any } (c,c_1,\ldots,c_q)\neq 0. \tag{A.2}$$

A sufficient condition is that all the Φ's with $k_2\le k_1$ are linearly independent.[1] Due to the presence of σ^2 in θ, it requires that $\Phi\neq 0$, which implies that at least one $\lim_{n\to\infty}\frac{1}{n}\,\mathrm{tr}(P_{ln})$ should be nonzero. In particular, at least one P_{ln} should not have a zero trace.

Some low level conditions for the full rank property of $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n$ and (A.2) can be derived. As $R_n(\tau) = I_n - \tau_1M_{1n} - \cdots - \tau_qM_{qn}$, $R_n(\tau)\Upsilon_n$ is a linear combination of Υ_n and $M_{jn}\Upsilon_n$ for $j=1,\ldots,q$. If M_{jn} is row-normalized, as X_n usually contains a column of ones corresponding to an intercept term, $M_{jn}X_n$ will also contain a column of ones. Let $X_n = [l_n, X_n^*]$, where X_n^* is the submatrix of X_n that excludes l_n, $X_{jn} = X_n^*$ if M_{jn} is row-normalized and $X_{jn} = X_n$ otherwise, and $\Upsilon_{jn} = [W_{1n}S_n^{-1}X_n\beta_0,\ldots,W_{pn}S_n^{-1}X_n\beta_0,X_{jn}]$, for $j=$

[1]Note that $\Phi_{\tau_{k_1k_2}} = \Phi_{\tau_{k_2k_1}}$ as P_{ln}'s are symmetric.

$1, \ldots, q$. A sufficient condition for $\lim_{n\to\infty} \frac{1}{n} Q_n' R_n(\tau)\Upsilon_n$ to have full rank for any τ in its parameter space is that $\lim_{n\to\infty} \frac{1}{n} Q_n'[\Upsilon_n, M_{1n}\Upsilon_{1n}, \ldots, M_{qn}\Upsilon_{qn}]$ has full column rank. This condition requires $[\Upsilon_n, M_{1n}\Upsilon_{1n}, \ldots, M_{qn}\Upsilon_{qn}]$ to have full column rank for large enough n, Q_n to have sufficient correlation with $[\Upsilon_n, M_{1n}\Upsilon_{1n}, \ldots, M_{qn}\Upsilon_{qn}]$, and the number of instruments in Q_n to be greater than or equal to the number of columns in $[\Upsilon_n, M_{1n}\Upsilon_{1n}, \ldots, M_{qn}\Upsilon_{qn}]$. As $S_n^{-1} = \sum_{j=0}^{\infty}(\kappa_{10}W_{1n} + \cdots + \kappa_{p0}W_{pn})^j$ if $\|\kappa_{10}W_{1n} + \cdots + \kappa_{p0}W_{pn}\| < 1$ for some matrix norm $\|\cdot\|$, Q_n may consist of independent columns of matrices with the form $f(W_{1n}, \ldots, W_{pn})X_n$ or $M_{jn}f(W_{1n}, \ldots, W_{pn})X_n$ for $j = 1, \ldots, q$, where $f(W_{1n}, \ldots, W_{pn})$ is a function of $W_{1n}, \ldots, W_{pn}$, e.g., a polynomial function, such that $\lim_{n\to\infty} \frac{1}{n} Q_n'[\Upsilon_n, M_{1n}\Upsilon_{1n}, \ldots, M_{qn}\Upsilon_{qn}]$ has full column rank.[2]

For (A.2), let $A^{(s)} = A + A'$ for any square matrix A, and φ_n be the matrix formed by the column vectors $\mathrm{vec}(I_n)$, $\mathrm{vec}((M_{kn}R_n^{-1})^{(s)})$ and $\mathrm{vec}((M_{k_2n}R_n^{-1}R_n'^{-1}M_{k_1n}')^{(s)})$, where $k, k_1 = 1, \ldots, q$ and $1 \le k_2 \le k_1$. Then $c\Phi + \sum_{k=1}^{q} c_k\Phi_{\tau_k} + \sum_{k_1,k_2=1}^{q} c_{k_1}c_{k_2}\Phi_{\tau_{k_1k_2}}$ is a linear combination of the columns of $\lim_{n\to\infty}\frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_pn})]'\varphi_n$, as $tr(AB) = \mathrm{vec}'(A')\,\mathrm{vec}(B)$ for two conformable matrices A and B. It follows that (A.2) holds if

$$\lim_{n\to\infty}\frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_pn})]'\varphi_n$$

has full column rank, which implies that φ_n has full column rank for large enough n, k_p is greater than or equal to the number of columns in φ_n, and P_{1n}, …, P_{k_pn} have sufficient correlation with I_n, $(M_{kn}R_n^{-1})^{(s)}$ and $(M_{k_2n}R_n^{-1}R_n'^{-1}M_{k_1n}')^{(s)}$ for $k, k_1 = 1, \ldots, q$ and $1 \le k_2 \le k_1$. Note that the number of columns in φ_n is $1 + q + \frac{1}{2}q(q+1)$, so (τ_0', σ_0^2) is over-identified under this sufficient condition for (A.2). For the rank of φ_n, as a linear combination of vectorizations of matrices is equal to the vectorization of a linear combination of matrices, φ_n having full rank is equivalent to linear independence of corresponding matrices. Since $R_n^{-1} = \sum_{j=0}^{\infty}(\tau_{10}M_{1n} + \cdots + \tau_{q0}M_{qn})^j$ if $\|\tau_{10}M_{1n} + \cdots + \tau_{q0}M_{qn}\| < 1$ for some matrix norm $\|\cdot\|$, P_{jn} can consist of matrices with the form $f(M_{1n}, \ldots, M_{qn})$, where $f(\cdot)$ is a function with function values being $n \times n$ symmetric matrices such that $\lim_{n\to\infty}\frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_pn})]'\varphi_n$ has full column rank.[3] As an example, we investigate the special case with $q = 1$. In this case,

$$\varphi_n = [\mathrm{vec}(I_n), \mathrm{vec}((M_{1n}R_n^{-1})^{(s)}), 2\,\mathrm{vec}(M_{1n}R_n^{-1}R_n'^{-1}M_{1n}')].$$

Then, for constants c_1, c_2 and c_3, $\frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_pn})]'\varphi_n[c_1, c_2, c_3]' = \frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_pn})]'\,\mathrm{vec}(c_1I_n + c_2(M_{1n}R_n^{-1})^{(s)} + 2c_3M_{1n}R_n^{-1}R_n'^{-1}M_{1n}')$. If I_n, $(M_{1n}R_n^{-1})^{(s)}$ and $M_{1n}R_n^{-1}R_n'^{-1}M_{1n}'$ are linearly independent and $\lim_{n\to\infty}\frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_pn})]'\varphi_n$ has full column rank, then (A.2) holds. The P_{ln}'s can be I_n, $M_{1n}^{(s)}$, $(M_{1n}^2)^{(s)}$, $M_{1n}M_{1n}'$, $M_{1n}^2M_{1n}'^2$, and so on.

[2] If we have initial consistent estimators $\tilde{\kappa}_n$ and $\tilde{\beta}_n$ of, respectively, κ_0 and β_0, then Υ_n can be estimated by $\hat{\Upsilon}_n = [W_{1n}S_n^{-1}(\tilde{\kappa}_n)X_n\tilde{\beta}_n, \ldots, W_{pn}S_n^{-1}(\tilde{\kappa}_n)X_n\tilde{\beta}_n, X_n]$, and Υ_{jn} by $\hat{\Upsilon}_{jn} = [W_{1n}S_n^{-1}(\tilde{\kappa}_n)X_n\tilde{\beta}_n, \ldots, W_{pn}S_n^{-1}(\tilde{\kappa}_n)X_n\tilde{\beta}_n, X_{jn}]$. Thus we may let $Q_n = [\hat{\Upsilon}_n, M_{1n}\hat{\Upsilon}_{1n}, \ldots, M_{qn}\hat{\Upsilon}_{qn}]$. Since this Q_n involves estimated parameters, the analysis for the higher order bias of the resulting estimator is more complicated and is omitted in the main text. A possible advantage of using this Q_n is that it may avoid the problem of many instruments and the possible high multicolinearity problem.

[3] If there is a consistent estimator $\tilde{\tau}$ of τ_0, then P_{jn}'s can be I_n, $(M_{kn}R_n^{-1}(\tilde{\tau}))^{(s)}$ and $(M_{k_2n}R_n^{-1}(\tilde{\tau})R_n'^{-1}(\tilde{\tau})M_{k_1n})^{(s)}$, where $k, k_1 = 1, \ldots, q$ and $1 \le k_2 \le k_1$.

If $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n$ does not have full column rank for some τ, it would likely be due to dependence of some $W_{jn}S_n^{-1}X_n\beta_0$ on X_n and other $W_{ln}S_n^{-1}X_n\beta_0$. In this case, assume that $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n$ has column rank k_x+p-p_0 for some $1\le p_0\le p$ for any τ in its parameter space. Without loss of generality, let $[W_{p_0+1,n}S_n^{-1}X_n\beta_0,\ldots,W_{pn}S_n^{-1}X_n\beta_0,X_n]$ have full column rank for large enough n. Then for $j=1,\ldots,p_0$, there exist α_{jl}'s and a_j's such that $W_{jn}S_n^{-1}X_n\beta_0=\sum_{l=p_0+1}^{p}W_{ln}S_n^{-1}X_n\beta_0\alpha_{jl}+X_na_j$ for large enough n. It follows that

$$Q_n'd_n(\theta)=Q_n'R_n(\tau)\Big\{\sum_{l=p_0+1}^{p}W_{ln}S_n^{-1}X_n\beta_0\Big[\sum_{j=1}^{p_0}\alpha_{jl}(\kappa_{0j}-\kappa_j)+\kappa_{0l}-\kappa_l\Big]+X_n\Big[\sum_{j=1}^{p_0}a_j(\kappa_{0j}-\kappa_j)+\beta_0-\beta\Big]\Big\}.$$

Thus, solutions to $\lim_{n\to\infty}\frac{1}{n}Q_n'd_n(\theta)=0$ are described by the relations

$$\beta=\sum_{j=1}^{p_0}a_j(\kappa_{0j}-\kappa_j)+\beta_0,\ \text{and}\ \kappa_l=\sum_{j=1}^{p_0}\alpha_{jl}(\kappa_{0j}-\kappa_j)+\kappa_{0l}\ \text{for}\ l=p_0+1,\ldots,p. \tag{A.3}$$

These relations imply that $d_n(\theta)=0$ and $(\beta_0',\kappa_{0,p_0+1},\ldots,\kappa_{0p})$ is identified as long as $(\kappa_{01},\ldots,\kappa_{0,p_0})$ is identified. With $d_n(\theta)=0$, $(\kappa_{01},\ldots,\kappa_{0,p_0},\tau_0',\sigma_0^2)$ can be identified from the quadratic moments, and the probability limits of quadratic moments $\lim_{n\to\infty}\frac{1}{n}\mathrm{E}[V_n'(\theta)P_{ln}V_n(\theta)-\sigma^2\operatorname{tr}(P_{ln})]=0$ for $l=1,\ldots,k_p$ reduce to

$$\lim_{n\to\infty}\frac{1}{n}\{\sigma_0^2\operatorname{tr}[R_n'^{-1}S_n'^{-1}S_n'(\kappa)R_n'(\tau)P_{ln}R_n(\tau)S_n(\kappa)S_n^{-1}R_n^{-1}]-\sigma^2\operatorname{tr}(P_{ln})\}=0\ \text{for}\ l=1,\ldots,k_p. \tag{A.4}$$

The identification of $(\kappa_0',\tau_0',\sigma_0^2)$ via the quadratic moments corresponds to that of the pure SARAR(p,q) model

$$\mathcal{Y}_n=\sum_{j=1}^{p}\kappa_jW_{jn}\mathcal{Y}_n+U_n,\quad U_n=\sum_{k=1}^{q}\tau_kM_{kn}U_n+V_n.$$

Under the additional condition that

$$\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n\ \text{has column rank}\ k_x+p-p_0\ \text{for some}\ 1\le p_0\le p\ \text{for any}\ \tau\ \text{in its parameter space}, \tag{A.5}$$

we shall use (A.3) to derive a condition equivalent to (A.4) holding only at $(\kappa',\tau',\sigma^2)=(\kappa_0',\tau_0',\sigma_0^2)$. With (A.3), $S_n(\kappa)=S_n+\sum_{j=1}^{p}(\kappa_{0j}-\kappa_j)W_{jn}=S_n+\sum_{j=1}^{p_0}(\kappa_{0j}-\kappa_j)(W_{jn}-\sum_{l=p_0+1}^{p}\alpha_{jl}W_{ln})$. Denote $\Delta_{jn}=(W_{jn}-\sum_{l=p_0+1}^{p}\alpha_{jl}W_{ln})S_n^{-1}R_n^{-1}$ for $j=1,\ldots,p_0$. Let Φ_{κ_j}, $\Phi_{\kappa_{j_1j_2}}$, $\Phi_{\kappa_j\tau_k}$, $\Phi_{\kappa_{j_1j_2}\tau_k}$, $\Phi_{\kappa_j\tau_{k_1k_2}}$ and $\Phi_{\kappa_{j_1j_2}\tau_{k_1k_2}}$ be $k_p\times1$ vectors with the lth element being, respectively, $\Phi_{\kappa_j,l}=\lim_{n\to\infty}\frac{2}{n}\operatorname{tr}(P_{ln}R_n\Delta_{jn})$, $\Phi_{\kappa_{j_1j_2},l}=\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}(\Delta_{j_1n}'R_n'P_{ln}R_n\Delta_{j_2n})$, $\Phi_{\kappa_j\tau_k,l}=\lim_{n\to\infty}\frac{2}{n}\operatorname{tr}(R_n'^{-1}M_{kn}'P_{ln}R_n\Delta_{jn}+P_{ln}M_{kn}\Delta_{jn})$, $\Phi_{\kappa_{j_1j_2}\tau_k,l}=\lim_{n\to\infty}\frac{2}{n}\operatorname{tr}(\Delta_{j_1n}'M_{kn}'P_{ln}R_n\Delta_{j_2n})$, $\Phi_{\kappa_j\tau_{k_1k_2},l}=\lim_{n\to\infty}\frac{2}{n}\operatorname{tr}(R_n'^{-1}M_{k_1n}'P_{ln}M_{k_2n}\Delta_{jn})$ and $\Phi_{\kappa_{j_1j_2}\tau_{k_1k_2},l}=\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}(\Delta_{j_1n}'M_{k_1n}'P_{ln}M_{k_2n}\Delta_{j_2n})$. With $d_n(\theta)=0$ and (A.3), (A.4) becomes

$$\begin{aligned}0=&\frac{1}{\sigma_0^2}(\sigma_0^2-\sigma^2)\Phi+\sum_{k=1}^{q}(\tau_{0k}-\tau_k)\Phi_{\tau_k}+\sum_{k_1,k_2=1}^{q}(\tau_{0,k_1}-\tau_{k_1})(\tau_{0,k_2}-\tau_{k_2})\Phi_{\tau_{k_1k_2}}+\sum_{j=1}^{p_0}(\kappa_{0j}-\kappa_j)\Phi_{\kappa_j}\\&+\sum_{j_1,j_2=1}^{p_0}(\kappa_{0,j_1}-\kappa_{j_1})(\kappa_{0,j_2}-\kappa_{j_2})\Phi_{\kappa_{j_1j_2}}+\sum_{j=1}^{p_0}\sum_{k=1}^{q}(\kappa_{0j}-\kappa_j)(\tau_{0k}-\tau_k)\Phi_{\kappa_j\tau_k}\\&+\sum_{j_1,j_2=1}^{p_0}\sum_{k=1}^{q}(\kappa_{0,j_1}-\kappa_{j_1})(\kappa_{0,j_2}-\kappa_{j_2})(\tau_{0k}-\tau_k)\Phi_{\kappa_{j_1j_2}\tau_k}+\sum_{j=1}^{p_0}\sum_{k_1,k_2=1}^{q}(\kappa_{0j}-\kappa_j)(\tau_{0,k_1}-\tau_{k_1})(\tau_{0,k_2}-\tau_{k_2})\Phi_{\kappa_j\tau_{k_1k_2}}\\&+\sum_{j_1,j_2=1}^{p_0}\sum_{k_1,k_2=1}^{q}(\kappa_{0,j_1}-\kappa_{j_1})(\kappa_{0,j_2}-\kappa_{j_2})(\tau_{0,k_1}-\tau_{k_1})(\tau_{0,k_2}-\tau_{k_2})\Phi_{\kappa_{j_1j_2}\tau_{k_1k_2}}.\end{aligned} \tag{A.6}$$

Thus (A.4) has a unique solution at $(\kappa', \tau', \sigma^2) = (\kappa'_0, \tau'_0, \sigma_0^2)$ if and only if the above linear combination of Φ's are nonzero when $(\kappa_1, \ldots, \kappa_{p_0}, \tau', \sigma^2) \neq (\kappa_{01}, \ldots, \kappa_{0,p_0}, \tau'_0, \sigma_0^2)$.

Sufficient conditions for (A.5) and (A.6) to hold only at $(\kappa_1, \ldots, \kappa_{p_0}, \tau', \sigma^2) = (\kappa_{01}, \ldots, \kappa_{0,p_0}, \tau'_0, \sigma_0^2)$ can be derived similarly as for the case where $\lim_{n\to\infty} \frac{1}{n} Q'_n R_n(\tau)\Upsilon_n$ has full column rank for any τ in its parameter space. (A.5) holds if $\lim_{n\to\infty} \frac{1}{n} Q'_n[\Upsilon_n^*, M_{1n}\Upsilon_{1n}^*, \ldots, M_{qn}\Upsilon_{qn}^*]$ has full column rank, where

$$\Upsilon_n^* = [W_{p_0+1,n}S_n^{-1}X_n\beta_0, \ldots, W_{pn}S_n^{-1}X_n\beta_0, X_n]$$

and $\Upsilon_{jn}^* = [W_{p_0+1,n}S_n^{-1}X_n\beta_0, \ldots, W_{pn}S_n^{-1}X_n\beta_0, X_{jn}]$ for $j = 1, \ldots, q$. For (A.6), let φ_n^* be the matrix formed by the column vectors $\mathrm{vec}(I_n)$, $\mathrm{vec}((M_{kn}R_n^{-1})^{(s)})$, $\mathrm{vec}((M_{k_2n}R_n^{-1}R_n'^{-1}M'_{k_1n})^{(s)})$, $\mathrm{vec}((R_n\Delta_{jn})^{(s)})$, $\mathrm{vec}((R_n\Delta_{j_2n}\Delta'_{j_1n}R'_n)^{(s)})$, $\mathrm{vec}((R_n\Delta_{jn}R_n'^{-1}M'_{kn}+M_{kn}\Delta_{jn})^{(s)})$, $\mathrm{vec}((R_n\Delta_{jn}\Delta'_{j_1n}M'_{kn})^{(s)})$, $\mathrm{vec}((M_{kn}\Delta_{jn}R_n'^{-1}M'_{k_1n})^{(s)})$, $\mathrm{vec}(M_{kn}\Delta_{j_2n}\Delta'_{j_1n}M'_{kn})^{(s)})$, and $\mathrm{vec}(M_{kn}\Delta_{jn}\Delta'_{j_1n}M'_{k_3n})^{(s)})$, for $j, j_1 = 1, \ldots, p_0$; $k, k_1 = 1, \ldots, q$; $1 \le j_2 \le j_1$; $1 \le k_2 \le k_1$; and $1 \le k_3 < k$. Thus, (A.6) holds only at $(\kappa_1, \ldots, \kappa_{p_0}, \tau', \sigma^2) = (\kappa_{01}, \ldots, \kappa_{0,p_0}, \tau'_0, \sigma_0^2)$ if $\lim_{n\to\infty} \frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_p,n})]'\varphi_n^*$ has full column rank, which implies that φ_n^* has full column rank for large enough n, k_p is greater than or equal to the number of columns in φ_n^*, and P_{1n}, …, P_{k_pn} have sufficient correlation with matrices corresponding to the columns of φ_n^*. In the special case with $p = p_0 = 1$ and $q = 1$, $\Delta_{1n} = W_{1n}S_n^{-1}R_n^{-1}$. Then ϕ_n^* is the matrix formed by the column vectors $\mathrm{vec}(I_n)$, $\mathrm{vec}((M_{1n}R_n^{-1})^{(s)})$, $\mathrm{vec}((M_{1n}R_n^{-1}R_n'^{-1}M'_{1n})^{(s)})$, $\mathrm{vec}((R_n\Delta_{1n})^{(s)})$, $\mathrm{vec}((R_n\Delta_{1n}\Delta'_{1n}R'_n)^{(s)})$, $\mathrm{vec}((R_n\Delta_{1n}R_n'^{-1}M'_{1n}+M_{1n}\Delta_{1n})^{(s)})$, $\mathrm{vec}((R_n\Delta_{1n}\Delta'_{1n}M'_{1n})^{(s)})$, $\mathrm{vec}((M_{1n}\Delta_{1n}R_n'^{-1}M'_{1n})^{(s)})$, and $\mathrm{vec}(M_{1n}\Delta_{1n}\Delta'_{1n}M'_{1n})^{(s)})$. Then P_{ln}'s can be I_n, $M_{1n}^{(s)}$, $W_{1n}^{(s)}$, $(M_{1n}^2)^{(s)}$, $(W_{1n}^2)^{(s)}$, $M_{1n}M'_{1n}$, $W_{1n}W'_{1n}$, $M_{1n}^2M_{1n}'^2$, $W_{1n}^2W_{1n}'^2$, and so on. We summarize the identification conditions in the following assumption.

Assumption S1. *Either (i)* $\lim_{n\to\infty} \frac{1}{n} Q'_n R_n(\tau)\Upsilon_n$ *has full column rank for any* τ *in its parameter space, and* (A.1) *has a unique solution at* $(\tau', \sigma^2) = (\tau'_0, \sigma_0^2)$, *i.e.,* (A.2) *holds, or (ii)* $\lim_{n\to\infty} \frac{1}{n} Q'_n R_n(\tau)\Upsilon_n$ *has column rank* k_x+p-p_0 *for some* $1 \le p_0 \le p$ *for any* τ *in its parameter space, and* (A.4) *has a unique solution at* $(\kappa', \tau', \sigma^2) = (\kappa'_0, \tau'_0, \sigma_0^2)$, *i.e.,* (A.6) *holds only at* $(\kappa_1, \ldots, \kappa_{p_0}, \tau', \sigma^2) = (\kappa_{01}, \ldots, \kappa_{0,p_0}, \tau'_0, \sigma_0^2)$.

The following assumption contains sufficient conditions for Assumption S1.

Assumption S2. *Either (i)* $\lim_{n\to\infty} \frac{1}{n} Q'_n[\Upsilon_n, M_{1n}\Upsilon_{1n}, \ldots, M_{qn}\Upsilon_{qn}]$ *and* $\lim_{n\to\infty} \frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_pn})]'\varphi_n$ *have full column rank, or (ii)* $\lim_{n\to\infty} \frac{1}{n} Q'_n[\Upsilon_n^*, M_{1n}\Upsilon_{1n}^*, \ldots, M_{qn}\Upsilon_{qn}^*]$ *and* $\lim_{n\to\infty} \frac{1}{n}[\mathrm{vec}(P_{1n}), \ldots, \mathrm{vec}(P_{k_p,n})]'\varphi_n^*$ *have full column rank.*

A.2 Heteroskedastic case

Sufficient identification conditions in the heteroskedastic case can be similarly derived. Denote by Π_n the diagonal matrix with the ith diagonal element being σ_{ni}^2 for $i = 1, \ldots, n$. Let Ψ_{τ_k}, $\Psi_{\tau_{k_1k_2}}$, Ψ_{κ_j}, $\Psi_{\kappa_{j_1j_2}}$, $\Psi_{\kappa_j\tau_k}$, $\Psi_{\kappa_{j_1j_2}\tau_k}$, $\Psi_{\kappa_j\tau_{k_1k_2}}$ and $\Psi_{\kappa_{j_1j_2}\tau_{k_1k_2}}$ be $k_p\times 1$ vectors with the lth element being, respectively, $\Psi_{\tau_k,l} = \lim_{n\to\infty} \frac{2}{n}\,\mathrm{tr}(P_{ln}M_{kn}R_n^{-1}\Pi_n)$, $\Psi_{\tau_{k_1k_2},l} = \lim_{n\to\infty} \frac{1}{n}\,\mathrm{tr}(R_n'^{-1}M'_{k_1n}P_{ln}M_{k_2n}R_n^{-1}\Pi_n)$, $\Psi_{\kappa_j,l} = \lim_{n\to\infty} \frac{2}{n}\,\mathrm{tr}(P_{ln}R_n\Delta_{jn}\Pi_n)$,

$$\Psi_{\kappa_{j_1j_2},l} = \lim_{n\to\infty} \frac{1}{n}\,\mathrm{tr}(\Delta'_{j_1n}R'_nP_{ln}R_n\Delta_{j_2n}\Pi_n),$$

$\Psi_{\kappa_j\tau_k,l} = \lim_{n\to\infty} \frac{2}{n}\operatorname{tr}(R_n'^{-1}M_{kn}'P_{ln}R_n\Delta_{jn}\Pi_n + P_{ln}M_{kn}\Delta_{jn}\Pi_n)$, $\Psi_{\kappa_{j_1j_2}\tau_k,l} = \lim_{n\to\infty}\frac{2}{n}\operatorname{tr}(\Delta_{j_1n}'M_{kn}'P_{ln}R_n\Delta_{j_2n}\Pi_n)$, $\Psi_{\kappa_j\tau_{k_1k_2},l} = \lim_{n\to\infty}\frac{2}{n}\operatorname{tr}(R_n'^{-1}M_{k_1n}'P_{ln}M_{k_2n}\Delta_{jn}\Pi_n)$ and $\Psi_{\kappa_{j_1j_2}\tau_{k_1k_2},l} = \lim_{n\to\infty}\frac{1}{n}\operatorname{tr}(\Delta_{j_1n}'M_{k_1n}'P_{ln}M_{k_2n}\Delta_{j_2n}\Pi_n)$; let ϖ_n be the matrix formed by the column vectors $\operatorname{vec}((M_{kn}R_n^{-1}\Pi_n)^{(s)})$ and $\operatorname{vec}((M_{k_2n}R_n^{-1}\Pi_nR_n'^{-1}M_{k_1n}')^{(s)})$ for $k, k_1 = 1,\dots,q$ and $1 \le k_2 \le k_1$; and let ϖ_n^* be the matrix formed by the column vectors $\operatorname{vec}((M_{kn}R_n^{-1}\Pi_n)^{(s)})$, $\operatorname{vec}((M_{k_2n}R_n^{-1}\Pi_nR_n'^{-1}M_{k_1n}')^{(s)})$, $\operatorname{vec}((R_n\Delta_{jn}\Pi_n)^{(s)})$, $\operatorname{vec}((R_n\Delta_{j_2n}\Pi_n\Delta_{j_1n}'R_n')^{(s)})$,

$$\operatorname{vec}((R_n\Delta_{jn}\Pi_nR_n'^{-1}M_{kn}' + M_{kn}\Delta_{jn}\Pi_n)^{(s)}),$$

$\operatorname{vec}((R_n\Delta_{jn}\Pi_n\Delta_{j_1n}'M_{kn}')^{(s)})$, $\operatorname{vec}((M_{kn}\Delta_{jn}\Pi_nR_n'^{-1}M_{k_1n}')^{(s)})$, $\operatorname{vec}((M_{kn}\Delta_{j_2n}\Pi_n\Delta_{j_1n}'M_{kn}')^{(s)})$, and

$$\operatorname{vec}((M_{kn}\Delta_{jn}\Pi_n\Delta_{j_1n}'M_{k_3n}')^{(s)}),$$

for $j, j_1 = 1,\dots,p_0$; $k, k_1 = 1,\dots,q$; $1 \le j_2 \le j_1$; $1 \le k_2 \le k_1$; and $1 \le k_3 < k$. The following Assumption S3 is a sufficient identification assumption, and Assumption S4 provides low level sufficient conditions for Assumption S3.

Assumption S3. *Either (i)* $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n$ *has full column rank for any* τ *in its parameter space, and* $\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}[R_n'^{-1}R_n'(\tau)P_{ln}R_n(\tau)R_n^{-1}\Pi_n] = 0$ *for* $l = 1,\dots,k_p$ *have a unique solution at* $\tau = \tau_0$*, i.e.,* $\sum_{k=1}^q c_k\Psi_{\tau_k} + \sum_{k_1,k_2=1}^q c_{k_1}c_{k_2}\Psi_{\tau_{k_1k_2}} \neq 0$ *for any* $(c_1,\dots,c_q)\neq 0$*, or (ii)* $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n(\tau)\Upsilon_n$ *has column rank* $k_x + p - p_0$ *for some* $1 \le p_0 \le p$ *for any* τ *in its parameter space, and* $\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}[R_n'^{-1}S_n'^{-1}S_n'(\kappa)R_n'(\tau)P_{ln}R_n(\tau)S_n(\kappa)S_n^{-1}R_n^{-1}\Pi_n] = 0$ *for* $l = 1,\dots,k_p$ *have a unique solution at* $(\kappa',\tau') = (\kappa_0',\tau_0')$*, i.e.,*

$$\begin{aligned}
&\sum_{k=1}^q c_k\Psi_{\tau_k} + \sum_{k_1,k_2=1}^q c_{k_1}c_{k_2}\Psi_{\tau_{k_1k_2}} + \sum_{j=1}^{p_0} d_j\Psi_{\kappa_j} + \sum_{j_1,j_2=1}^{p_0} d_{j_1}d_{j_2}\Psi_{\kappa_{j_1j_2}} + \sum_{j=1}^{p_0}\sum_{k=1}^q d_jc_k\Psi_{\kappa_j\tau_k} \\
&+ \sum_{j_1,j_2=1}^{p_0}\sum_{k=1}^q d_{j_1}d_{j_2}c_k\Psi_{\kappa_{j_1j_2}\tau_k} + \sum_{j=1}^{p_0}\sum_{k_1,k_2=1}^q d_jc_{k_1}c_{k_2}\Psi_{\kappa_j\tau_{k_1k_2}} + \sum_{j_1,j_2=1}^{p_0}\sum_{k_1,k_2=1}^q d_{j_1}d_{j_2}c_{k_1}c_{k_2}\Psi_{\kappa_{j_1j_2}\tau_{k_1k_2}} \neq 0
\end{aligned} \tag{A.7}$$

for any $(c_1,\dots,c_q,d_1,\dots,d_{p_0}) \neq 0$.

Assumption S4. *Either (i)* $\lim_{n\to\infty}\frac{1}{n}Q_n'[\Upsilon_n, M_{1n}\Upsilon_{1n},\dots,M_{qn}\Upsilon_{qn}]$ *and* $\lim_{n\to\infty}\frac{1}{n}[\operatorname{vec}(P_{1n}),\dots,\operatorname{vec}(P_{k_pn})]'\varpi_n$ *have full column rank, or (ii)* $\lim_{n\to\infty}\frac{1}{n}Q_n'[\Upsilon_n^*, M_{1n}\Upsilon_{1n}^*,\dots,M_{qn}\Upsilon_{qn}^*]$ *and* $\lim_{n\to\infty}\frac{1}{n}[\operatorname{vec}(P_{1n}),\dots,\operatorname{vec}(P_{k_p,n})]'\varpi_n^*$ *have full column rank.*

B Sufficient conditions for $\lim_{n\to\infty}\bar{G}_n$ in Assumption 9 to have full rank

B.1 Homoskedastic case

In the homoskedastic case, the gradient matrix

$$\bar{G}_n = -\frac{1}{n}\begin{pmatrix}
2\sigma_0^2\operatorname{tr}(P_{1n}R_nW_{1n}S_n^{-1}R_n^{-1}) & \dots & 2\sigma_0^2\operatorname{tr}(P_{1n}R_nW_{pn}S_n^{-1}R_n^{-1}) & 2\sigma_0^2\operatorname{tr}(P_{1n}M_{1n}R_n^{-1}) & \dots & 2\sigma_0^2\operatorname{tr}(P_{1n}M_{qn}R_n^{-1}) & 0 & \operatorname{tr}(P_{1n}) \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
2\sigma_0^2\operatorname{tr}(P_{k_pn}R_nW_{1n}S_n^{-1}R_n^{-1}) & \dots & 2\sigma_0^2\operatorname{tr}(P_{k_pn}R_nW_{pn}S_n^{-1}R_n^{-1}) & 2\sigma_0^2\operatorname{tr}(P_{k_pn}M_{1n}R_n^{-1}) & \dots & 2\sigma_0^2\operatorname{tr}(P_{k_pn}M_{qn}R_n^{-1}) & 0 & \operatorname{tr}(P_{k_pn}) \\
Q_n'R_nW_{1n}S_n^{-1}X_n\beta_0 & \dots & Q_n'R_nW_{pn}S_n^{-1}X_n\beta_0 & 0 & \dots & 0 & Q_n'R_nX_n & 0
\end{pmatrix}.$$

Thus, by the last $(q+2)$ column blocks of $\bar{G}_n$, for $\bar{G}_n$ to have full column rank in the limit, $\lim_{n\to\infty}\frac{1}{n}Q_n'R_nX_n$ and

$$\lim_{n\to\infty}\frac{1}{n}[\mathrm{vec}(P_{1n}),\ldots,\mathrm{vec}(P_{k_pn})]'[\mathrm{vec}((M_{1n}R_n^{-1})^{(s)}),\ldots,\mathrm{vec}((M_{qn}R_n^{-1})^{(s)}),\mathrm{vec}(I_n)] \tag{B.1}$$

must have full column rank. By the last row blocks of $\bar{G}_n$, if we have the stronger conditions that $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ and (B.1) have full column rank, where $\Upsilon_n = [W_{1n}S_n^{-1}X_n\beta_0,\ldots,W_{pn}S_n^{-1}X_n\beta_0,X_n]$, then $\lim_{n\to\infty}\bar{G}_n$ has full column rank. If $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ does not have full column rank, by the first k_p row blocks of $\bar{G}_n$, sufficient conditions for $\lim_{n\to\infty}\bar{G}_n$ to have full column rank are that $\lim_{n\to\infty}\frac{1}{n}Q_n'R_nX_n$ and

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}[\mathrm{vec}(P_{1n}),\ldots,\mathrm{vec}(P_{k_pn})]'[&\mathrm{vec}((R_nW_{1n}S_n^{-1}R_n^{-1})^{(s)}),\ldots,\mathrm{vec}((R_nW_{pn}S_n^{-1}R_n^{-1})^{(s)}),\\ &\mathrm{vec}((M_{1n}R_n^{-1})^{(s)}),\ldots,\mathrm{vec}((M_{qn}R_n^{-1})^{(s)}),\mathrm{vec}(I_n)]\end{aligned} \tag{B.2}$$

have full column rank. Alternatively, we may derive some weaker sufficient conditions as in Section A.1. Assume that $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ has column rank k_x+p-p_0 for some $1\le p_0\le p$. Without loss of generality, let $[W_{p_0+1,n}S_n^{-1}X_n\beta_0,\ldots,W_{pn}S_n^{-1}X_n\beta_0,X_n]$ have full column rank for large enough n. Then for $j=1,\ldots,p_0$, there exist α_{jl}'s and a_j's such that $W_{jn}S_n^{-1}X_n\beta_0=\sum_{l=p_0+1}^{p}W_{ln}S_n^{-1}X_n\beta_0\alpha_{jl}+X_na_j$ for large enough n. Suppose that $\bar{G}_n[b_1,\ldots,b_{p_0},c_{p_0+1},\ldots,c_p,d_1,\ldots,d_q,c',e]'=0$, where b_j's, c_j's, d_j's and e are scalars and c is a $k_x\times 1$ vector. Using the last row blocks of $\bar{G}_n$, we have $\sum_{l=p_0+1}^{p}Q_n'R_nW_{ln}S_n^{-1}X_n\beta_0(\sum_{j=1}^{p_0}\alpha_{jl}b_j+c_l)+Q_n'R_nX_n(\sum_{j=1}^{p_0}a_jb_j+c)=0$. Hence,

$$\sum_{j=1}^{p_0}a_jb_j+c=0,\ \text{and}\ \sum_{j=1}^{p_0}\alpha_{jl}b_j+c_l=0,\ \text{for}\ l=p_0+1,\ldots,p.$$

By the above relations and the first k_p row blocks of $\bar{G}_n$, we have

$$\sum_{j=1}^{p_0}2\sigma_0^2\,\mathrm{tr}\Big[P_{mn}R_n\Big(W_{jn}-\sum_{l=p_0+1}^{p}W_{ln}\alpha_{jl}\Big)S_n^{-1}R_n^{-1}\Big]b_j+\sum_{k=1}^{q}2\sigma_0^2\,\mathrm{tr}(P_{mn}M_{kn}R_n^{-1})d_k+\mathrm{tr}(P_{mn})e=0,\ \text{for}\ m=1,\ldots,k_p.$$

Then sufficient conditions for $\lim_{n\to\infty}\bar{G}_n$ to have full column rank are that $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ has column rank k_x+p-p_0 for some $1\le p_0\le p$ and

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}[\mathrm{vec}(P_{1n}),\ldots,\mathrm{vec}(P_{k_pn})]'\Big[&\mathrm{vec}\Big(\Big(R_n\Big(W_{1n}-\sum_{l=p_0+1}^{p}W_{ln}\alpha_{1l}\Big)S_n^{-1}R_n^{-1}\Big)^{(s)}\Big),\ldots,\\ &\mathrm{vec}\Big(\Big(R_n\Big(W_{p_0n}-\sum_{l=p_0+1}^{p}W_{ln}\alpha_{p_0l}\Big)S_n^{-1}R_n^{-1}\Big)^{(s)}\Big),\mathrm{vec}((M_{1n}R_n^{-1})^{(s)}),\ldots,\mathrm{vec}((M_{qn}R_n^{-1})^{(s)}),\mathrm{vec}(I_n)\Big]\end{aligned} \tag{B.3}$$

has full column rank. We summarize the above sufficient conditions for $\lim_{n\to\infty}\bar{G}_n$ to have full column rank in the following assumption.

Assumption S5. *Either (i)* $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ *and* (B.1) *have full column rank, or (ii)* $\lim_{n\to\infty}\frac{1}{n}Q_n'R_nX_n$ *and* (B.2) *have full column rank, or (iii)* $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ *has column rank* k_x+p-p_0 *for some* $1\le p_0\le p$ *and* (B.3) *has full column rank.*

The interpretation of the above assumption is similar to that of Assumption S1. For example, (B.1) having full column rank implies that $(M_{1n}R_n^{-1})^{(s)}$, …, $(M_{qn}R_n^{-1})^{(s)}$ and I_n are linearly independent, $k_p\ge q+1$, and P_{jn}'s have enough correlation with $(M_{1n}R_n^{-1})^{(s)}$, …, $(M_{qn}R_n^{-1})^{(s)}$ and I_n. We note that Assumption S2(i) implies Assumption S5(i), and Assumption S2(ii) implies Assumption S5(iii).

B.2 Heteroskedastic case

In the heteroskedastic case, the gradient matrix

$$\bar{G}_n = -\frac{1}{n}\begin{pmatrix} 2\operatorname{tr}(P_{1n}R_nW_{1n}S_n^{-1}R_n^{-1}\Pi_n) & \dots & 2\operatorname{tr}(P_{1n}R_nW_{pn}S_n^{-1}R_n^{-1}\Pi_n) & 2\operatorname{tr}(P_{1n}M_{1n}R_n^{-1}\Pi_n) & \dots & 2\operatorname{tr}(P_{1n}M_{qn}R_n^{-1}\Pi_n) & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 2\operatorname{tr}(P_{k_pn}R_nW_{1n}S_n^{-1}R_n^{-1}\Pi_n) & \dots & 2\operatorname{tr}(P_{k_pn}R_nW_{pn}S_n^{-1}R_n^{-1}\Pi_n) & 2\operatorname{tr}(P_{k_pn}M_{1n}R_n^{-1}\Pi_n) & \dots & 2\operatorname{tr}(P_{k_pn}M_{qn}R_n^{-1}\Pi_n) & 0 \\ Q_n'R_nW_{1n}S_n^{-1}X_n\beta_0 & \dots & Q_n'R_nW_{pn}S_n^{-1}X_n\beta_0 & 0 & \dots & 0 & Q_n'R_nX_n \end{pmatrix}.$$

Then, by an argument similar to that for the homoskedastic case, we can derive the following sufficient conditions for $\lim_{n\to\infty}\bar{G}_n$ to have full column rank.

Assumption S6. *Either (i)* $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ *and*

$$\lim_{n\to\infty}\frac{1}{n}[\operatorname{vec}(P_{1n}),\dots,\operatorname{vec}(P_{k_pn})]'[\operatorname{vec}((M_{1n}R_n^{-1}\Pi_n)^{(s)}),\dots,\operatorname{vec}((M_{qn}R_n^{-1}\Pi_n)^{(s)})]$$

have full column rank, or (ii) $\lim_{n\to\infty}\frac{1}{n}Q_n'R_nX_n$ *and*

$$\lim_{n\to\infty}\frac{1}{n}[\operatorname{vec}(P_{1n}),\dots,\operatorname{vec}(P_{k_pn})]'[\operatorname{vec}((R_nW_{1n}S_n^{-1}R_n^{-1}\Pi_n)^{(s)}),\dots,\operatorname{vec}((R_nW_{pn}S_n^{-1}R_n^{-1}\Pi_n)^{(s)}),$$
$$\operatorname{vec}((M_{1n}R_n^{-1}\Pi_n)^{(s)}),\dots,\operatorname{vec}((M_{qn}R_n^{-1}\Pi_n)^{(s)})]$$

have full column rank, or (iii) $\lim_{n\to\infty}\frac{1}{n}Q_n'R_n\Upsilon_n$ *has column rank* k_x+p-p_0 *for some* $1\le p_0\le p$ *and*

$$\lim_{n\to\infty}\frac{1}{n}[\operatorname{vec}(P_{1n}),\dots,\operatorname{vec}(P_{k_pn})]'\Big[\operatorname{vec}\Big(\Big(R_n\Big(W_{1n}-\sum_{l=p_0+1}^{p}W_{ln}\alpha_{1l}\Big)S_n^{-1}R_n^{-1}\Pi_n\Big)^{(s)}\Big),\dots,$$
$$\operatorname{vec}\Big(\Big(R_n\Big(W_{p_0n}-\sum_{l=p_0+1}^{p}W_{ln}\alpha_{p_0l}\Big)S_n^{-1}R_n^{-1}\Pi_n\Big)^{(s)}\Big),\operatorname{vec}((M_{1n}R_n^{-1}\Pi_n)^{(s)}),\dots,\operatorname{vec}((M_{qn}R_n^{-1}\Pi_n)^{(s)})\Big]$$

has full column rank.

C Sufficient identification conditions for the spatial J test

For the identification of the augmented model (11) for the spatial J test, compared with that of the original SARAR(p,q) model, we need to take into account additional terms from the predictor $\hat{Y}_n$, and possible inclusion of the generated regressor $\hat{Y}_n$ in the IV matrix Q_n if $\hat{Y}_n=\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$.

C.1 Homoskedastic case

With the pseudo-true value θ_n^* satisfying Assumption 12(i) and (ii), when $\hat{Y}_n=\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, let $\bar{Y}_n^*=\mathcal{S}_n^{-1}(\boldsymbol{\kappa}_n^*)\mathcal{X}_n\boldsymbol{\beta}_n^*$ and $\epsilon_n^*=0_{n\times1}$; when $\hat{Y}_n=\sum_{j=1}^{p_1}\hat{\boldsymbol{\kappa}}_{jn}\mathcal{W}_{jn}Y_n+\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, let $\bar{Y}_n^*=\sum_{j=1}^{p_1}\boldsymbol{\kappa}_{jn}^*\mathcal{W}_{jn}S_n^{-1}X_n\beta_0+\mathcal{X}_n\boldsymbol{\beta}_n^*$ and $\epsilon_n^*=\Gamma_nV_n$, where $\Gamma_n=\sum_{j=1}^{p_1}\boldsymbol{\kappa}_{jn}^*\mathcal{W}_{jn}S_n^{-1}R_n^{-1}$. The leading order terms for elements of $\hat{Y}_n$ are given by $\bar{Y}_n^*+\epsilon_n^*$. For $\hat{Y}_n=\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, note that $\frac{1}{n}V_n'A_n\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n=o_p(1)$ and $\frac{1}{n}b_n'A_n\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n=\frac{1}{n}b_n'A_n\mathcal{S}_n^{-1}(\boldsymbol{\kappa}_n^*)\mathcal{X}_n\boldsymbol{\beta}_n^*+o_p(1)$, where A_n is an $n\times n$ nonstochastic matrix bounded in both row and column sum norms and b_n is an $n\times1$

vector of uniformly bounded constants. Then $\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$ is asymptotically exogenous and it may be included in the IV matrix Q_n. To account for the possible randomness of Q_n, if $\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$ is in Q_n, let Q_n^* be the matrix obtained by replacing $\mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$ in Q_n with $\mathcal{S}_n^{-1}(\boldsymbol{\kappa}_n^*)\mathcal{X}_n\boldsymbol{\beta}_n^*$, and $Q_n^* = Q_n$ otherwise. Since $\mathcal{S}_n^{-1}(\boldsymbol{\kappa}_n^*)\mathcal{X}_n\boldsymbol{\beta}_n^*$ behaves like X_n, sufficient identification conditions are similar to those for the SARAR(p,q) model. For $\hat{Y}_n = \sum_{j=1}^{p_1} \hat{\boldsymbol{\kappa}}_{jn}\mathcal{W}_{jn}Y_n + \mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, due to the presence of the stochastic part ϵ_n^* and its correlation with vectors linear in V_n, this $\hat{Y}_n$ cannot be included in Q_n. On the other hand, unlike the case with the first $\hat{Y}_n$, η can be identified purely from the quadratic moments. Since the identification conditions of η from the linear moment with the second $\hat{Y}_n$ are essentially the same as those with the first $\hat{Y}_n$, we omit those. The following assumption gives some sufficient identification conditions.

Assumption S7. *(1) If $\hat{Y}_n = \mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, either (i) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)[\Upsilon_n, \bar{Y}_n^*]$ has full column rank for any τ in its parameter space, and* (A.1) *has a unique solution at $(\tau', \sigma^2) = (\tau_0', \sigma^2)$, or (ii) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)[\Upsilon_n, \bar{Y}_n^*]$ has column rank $k_x + 1 + p - p_0$ for some $1 \le p_0 \le p$ for any τ in its parameter space, and* (A.4) *has a unique solution at $(\kappa', \tau', \sigma^2) = (\kappa_0', \tau_0', \sigma_0^2)$.*

(2) If $\hat{Y}_n = \sum_{j=1}^{p_1}\hat{\boldsymbol{\kappa}}_{jn}\mathcal{W}_{jn}Y_n + \mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, either (i) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)\Upsilon_n$ has full column rank for any τ in its parameter space, and

$$\lim_{n\to\infty}\frac{1}{n}\{\sigma_0^2\operatorname{tr}[(R_n^{-1} - \eta\Gamma_n)'R_n'(\tau)P_{ln}R_n(\tau)(R_n^{-1} - \eta\Gamma_n)] - \sigma^2\operatorname{tr}(P_{ln})\} = 0 \text{ for } l = 1, \dots, k_p \tag{C.1}$$

have a unique solution at $(\tau', \sigma^2, \eta) = (\tau_0', \sigma_0^2, 0)$, or (ii) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)\Upsilon_n$ has column rank $k_x + p - p_0$ for some $1 \le p_0 \le p$ for any τ in its parameter space, and

$$\lim_{n\to\infty}\frac{1}{n}\{\sigma_0^2\operatorname{tr}[(S_n(\kappa)S_n^{-1}R_n^{-1} - \eta\Gamma_n)'R_n'(\tau)P_{ln}R_n(\tau)(S_n(\kappa)S_n^{-1}R_n^{-1} - \eta\Gamma_n)] - \sigma^2\operatorname{tr}(P_{ln})\} = 0 \text{ for } l = 1, \dots, k_p \tag{C.2}$$

have a unique solution at $(\kappa', \tau', \sigma^2, \eta) = (\kappa_0', \tau_0', \sigma_0^2, 0)$.

Equivalent conditions for (A.1) and (A.4) to have unique solutions at true parameter values are given in Assumption S1, and those for Assumption S7(2)(i) can be derived by expanding (C.1) as a polynomial linear in $(\sigma_0^2 - \sigma^2)$ and quadratic in $(\tau_0 - \tau)$ and η. For Assumption S7(2)(ii), again let $[W_{p_0+1,n}S_n^{-1}X_n\beta_0, \dots, W_{pn}S_n^{-1}X_n\beta_0, X_n]$ have full column rank for large enough n. Then for large enough n, there exist α_{jl} and a_j for $j = 1, \dots, p_0$ such that $W_{jn}S_n^{-1}X_n\beta_0 = \sum_{l=p_0+1}^p W_{ln}S_n^{-1}X_n\beta_0\alpha_{jl} + X_na_j$, and there exist α_l's and a such that $\bar{Y}_n^* = \sum_{l=p_0+1}^p W_{ln}S_n^{-1}X_n\beta_0\alpha_l + X_na$. It follows that $Q_n^{*'}[d_n(\theta) - \eta R_n(\tau)\bar{Y}_n^*] = Q_n^{*'}R_n(\tau)\{\sum_{l=p_0+1}^p W_{ln}S_n^{-1}X_n\beta_0[\sum_{j=1}^{p_0}\alpha_{jl}(\kappa_{0j} - \kappa_j) - \alpha_l\eta + \kappa_{0l} - \kappa_l] + X_n[\sum_{j=1}^{p_0}a_j(\kappa_{0j} - \kappa_j) - a\eta + \beta_0 - \beta]\}$. Thus, the solutions to $\operatorname{plim}_{n\to\infty}\frac{1}{n}Q_n'V_n(\vartheta) = \lim_{n\to\infty}\frac{1}{n}Q_n^{*'}[d_n(\theta) - \eta R_n(\tau)\bar{Y}_n^*] = 0$ are described by the relations

$$\beta = \sum_{j=1}^{p_0}a_j(\kappa_{0j} - \kappa_j) - a\eta + \beta_0, \text{ and } \kappa_l = \sum_{j=1}^{p_0}\alpha_{jl}(\kappa_{0j} - \kappa_j) - a_l\eta + \kappa_{0l} \text{ for } l = p_0 + 1, \dots, p. \tag{C.3}$$

Then $S_n(\kappa)S_n^{-1}R_n^{-1} - \eta\Gamma_n = R_n^{-1} + \sum_{j=1}^{p_0}(\kappa_{0j} - \kappa_j)\Delta_{jn} - \eta\Gamma_n^*$, where $\Gamma_n^* = \Gamma_n - \sum_{l=p_0+1}^p a_lW_{ln}S_n^{-1}R_n^{-1}$. In this expression, $(-\eta)$ is like an additional $(\kappa_{0j} - \kappa_j)$ and Γ_n^* is like a Δ_{jn}. Let $\Delta_{p_0+1,n} = \Gamma_n^*$ and define Φ's involving

$\Delta_{p_0+1,n}$ similarly to those in Section A.1. Then replacing p_0 in (A.6) by p_0+1 yields an equivalent condition for (C.2) to have a unique solution at the true parameter values. We omit that for simplicity.

C.2 Heteroskedastic case

In the heteroskedastic case, the following assumption is sufficient for the identification of the augmented model for the spatial J test.

Assumption S8. *(1) If $\hat{Y}_n = \mathcal{S}_n^{-1}(\hat{\boldsymbol{\kappa}}_n)\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, either (i) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)[\Upsilon_n, \bar{Y}_n^*]$ has full column rank for any τ in its parameter space, and $\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}[R_n'^{-1}R_n'(\tau)P_{ln}R_n(\tau)R_n^{-1}\Sigma_n]=0$ for $l=1,\ldots,k_p$ have a unique solution at $\tau=\tau_0$, or (ii) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)[\Upsilon_n, \bar{Y}_n^*]$ has column rank $k_x+1+p-p_0$ for some $1\le p_0\le p$ for any τ in its parameter space, and $\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}[R_n'^{-1}S_n'^{-1}S_n'(\kappa)R_n'(\tau)P_{ln}R_n(\tau)S_n(\kappa)S_n^{-1}R_n^{-1}\Sigma_n]=0$ for $l=1,\ldots,k_p$ have a unique solution at $(\kappa',\tau')=(\kappa_0',\tau_0')$.*

(2) If $\hat{Y}_n=\sum_{j=1}^{p_1}\hat{\boldsymbol{\kappa}}_{jn}\mathcal{W}_{jn}Y_n+\mathcal{X}_n\hat{\boldsymbol{\beta}}_n$, either (i) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)\Upsilon_n$ has full column rank for any τ in its parameter space, and $\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}[(R_n^{-1}-\eta\Gamma_n)'R_n'(\tau)P_{ln}R_n(\tau)(R_n^{-1}-\eta\Gamma_n)\Sigma_n]=0$ for $l=1,\ldots,k_p$ have a unique solution at $(\tau',\eta)=(\tau_0',0)$, or (ii) $\lim_{n\to\infty}\frac{1}{n}Q_n^{*'}R_n(\tau)\Upsilon_n$ has column rank k_x+p-p_0 for some $1\le p_0\le p$ for any τ in its parameter space, and $\lim_{n\to\infty}\frac{1}{n}\operatorname{tr}[(S_n(\kappa)S_n^{-1}R_n^{-1}-\eta\Gamma_n)'R_n'(\tau)P_{ln}R_n(\tau)(S_n(\kappa)S_n^{-1}R_n^{-1}-\eta\Gamma_n)\Sigma_n]=0$ for $l=1,\ldots,k_p$ have a unique solution at $(\kappa',\tau',\eta)=(\kappa_0',\tau_0',0)$.

D Proofs of lemmas

D.1 General lemmas

Lemma 1. *For $l=1,\ldots,s$, let $D_{ln}(\theta)=[d_{ln,ij}(\theta)]$ be $n\times n$ matrices which are bounded in row sum norm uniformly in $\theta\in\Theta$, and $u_{ln}=[u_{ln,i}]$ be $n\times 1$ vectors such that $\sup_{1\le l\le s}\sup_{1\le j\le n}\mathrm{E}\,|u_{ln,j}|^{a_l}=O(1)$ for $a_l>1$. Then,*

(i) $\sup_{\theta\in\Theta}\sup_{1\le i\le n}|\prod_{l=1}^s\sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}|=O_p(n^{\sum_{l=1}^s\frac{1}{a_l}})$;

(ii) $\sup_{\theta\in\Theta}\sup_{1\le i\le n}\mathrm{E}(|\prod_{l=1}^s\sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}|^{1/\sum_{l=1}^s\frac{1}{a_l}})=O(1)$ if $D_{ln}(\theta)$'s are nonstochastic;

(iii) $\frac{1}{n}\sum_{i=1}^n\prod_{l=1}^s\sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}=O_p(1)$ if $\sum_{l=1}^s\frac{1}{a_l}\le 1$, and $\sup_{\theta\in\Theta}\frac{1}{n}|\sum_{i=1}^n\prod_{l=1}^s\sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}|=O_p(1)$ if $\sum_{l=1}^s\frac{1}{a_l}\le 1$ and $D_{ln}(\theta)$'s are also bounded in column sum norm uniformly in $\theta\in\Theta$.

Proof. (i) There exists a finite $c_l>0$ such that $\frac{1}{c_l}+\frac{1}{a_l}=1$. By Hölder's inequality,

$$\Big|\sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}\Big|\le\sum_{j=1}^n|d_{ln,ij}(\theta)|^{\frac{1}{c_l}}|d_{ln,ij}(\theta)|^{\frac{1}{a_l}}|u_{ln,j}|\le\Big(\sum_{j=1}^n|d_{ln,ij}(\theta)|\Big)^{1/c_l}\Big(\sum_{j=1}^n|d_{ln,ij}(\theta)|\cdot|u_{ln,j}|^{a_l}\Big)^{1/a_l},$$

where $\sum_{j=1}^n|d_{ln,ij}(\theta)|\le c=\sup_n\sup_{1\le l\le s}\sup_{\theta\in\Theta}\|D_{ln}(\theta)\|_\infty<\infty$. Then,

$$\Big|\prod_{l=1}^s\sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}\Big|\le c^{\sum_{l=1}^s\frac{1}{c_l}}\Big(\sum_{l=1}^s\sum_{j=1}^n|d_{ln,ij}(\theta)|\cdot|u_{ln,j}|^{a_l}\Big)^{\sum_{l=1}^s\frac{1}{a_l}} \tag{D.1}$$

$$\leq c^{\sum_{l=1}^s \frac{1}{c_l}} \Big(\sum_{l=1}^s \sum_{j=1}^n c|u_{ln,j}|^{a_l}\Big)^{\sum_{l=1}^s \frac{1}{a_l}}, \tag{D.2}$$

where $\frac{1}{n}\sum_{l=1}^s \sum_{j=1}^n |u_{ln,j}|^{a_l} = O_p(1)$ by Markov's inequality. Hence the result holds.

(*ii*) By (D.1),

$$\sup_{\theta\in\Theta}\sup_{1\leq i\leq n} \mathrm{E}\Big(\Big|\prod_{l=1}^s \sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}\Big|^{1/\sum_{l=1}^s \frac{1}{a_l}}\Big) \leq c^{(\sum_{l=1}^s \frac{1}{c_l})/\sum_{l=1}^s \frac{1}{a_l}} \sup_{\theta\in\Theta}\sup_{1\leq i\leq n} \sum_{l=1}^s \sum_{j=1}^n |d_{ln,ij}(\theta)| \,\mathrm{E}\,|u_{ln,j}|^{a_l} = O(1).$$

Hence the result holds.

(*iii*) If $\sum_{l=1}^s \frac{1}{a_l} \leq 1$, by (D.1) and Jensen's inequality for a concave function,

$$\Big|\frac{1}{n}\sum_{i=1}^n \prod_{l=1}^s \sum_{j=1}^n d_{ln,ij}(\theta)u_{ln,j}\Big| \leq c^{\sum_{l=1}^s \frac{1}{c_l}} \big(\frac{1}{n}\sum_{i=1}^n \sum_{l=1}^s \sum_{j=1}^n |d_{ln,ij}(\theta)|\cdot|u_{ln,j}|^{a_l}\big)^{\sum_{l=1}^s \frac{1}{a_l}}$$

$$\leq c^{\sum_{l=1}^s \frac{1}{c_l}} \big(\frac{1}{n}\sum_{l=1}^s \sum_{j=1}^n |u_{ln,j}|^{a_l} \sup_n \sup_{1\leq l\leq s}\sup_{\theta\in\Theta} \|D_{ln}(\theta)\|_1\big)^{\sum_{l=1}^s \frac{1}{a_l}}.$$

Since $\frac{1}{n}\sum_{i=1}^n\sum_{l=1}^s\sum_{j=1}^n |d_{ln,ij}(\theta)|\cdot|u_{ln,j}|^{a_l} = O_p(1)$ for any $\theta\in\Theta$ and $\frac{1}{n}\sum_{l=1}^s\sum_{j=1}^n |u_{ln,j}|^{a_l} = O_p(1)$ by Markov's inequality, the two results follow by, respectively, the first and second inequalities. □

Lemma 2. *Suppose that v_{ni}'s are independent with zero mean and $\mathrm{E}(v_{ni}^2) = \sigma_{ni}^2$ for $i = 1,\ldots,n$, and $[a_{ln,ij}]$, $[b_{ln,ij}]$, $[c_{ln,ij}]$, $[d_{ln,ij}]$, $[e_{ln,ij}]$, $[f_{ln,ij}]$, $[g_{ln,ij}]$ and $[h_{ln,ij}]$ for $l = 1,2$ are $n\times n$ nonstochastic matrices with bounded row sum norms. Then,*

(i) for $r_{ni}^{(l)} = a_{ln,ii}(v_{ni}^2-\sigma_{ni}^2) + b_{ln,ii}v_{ni} + (c_{ln,ii}+d_{ln,ii}v_{ni})\sum_{j=1}^{i-1} e_{ln,ij}v_{nj} + \sum_{j=1}^{i-1}(g_{ln,ij}v_{nj}\sum_{k=1}^{j-1} h_{ln,ik}v_{nk})$ with $l = 1$ and 2, if $\sup_n\sup_{1\leq i\leq n}\mathrm{E}(v_{ni}^4) < \infty$, then $\frac{1}{n}\sum_{i=1}^n \mathrm{E}(r_{ni}^{(1)}r_{ni}^{(2)}) = O(1)$ and $\frac{1}{n}\sum_{i=1}^n[r_{ni}^{(1)}r_{ni}^{(2)} - \mathrm{E}(r_{ni}^{(1)}r_{ni}^{(2)})] = o_p(1)$;

(ii) for

$$r_{ni}^{(l)} = a_{ln,ii}(v_{ni}^2-\sigma_{ni}^2)+b_{ln,ii}v_{ni}+(c_{ln,ii}+d_{ln,ii}v_{ni})\sum_{j=1}^{i-1} e_{ln,ij}v_{nj}+\sum_{j=1}^{i-1} f_{ln,ij}(v_{nj}^2-\sigma_{nj}^2)+\sum_{j=1}^{i-1}\Big(g_{ln,ij}v_{nj}\sum_{k=1}^{j-1} h_{ln,ik}v_{nk}\Big)$$

with $l = 1$ and 2, if $\sup_n\sup_{1\leq i\leq n}\mathrm{E}(v_{ni}^8) < \infty$, then $\frac{1}{n}\sum_{i=1}^n[r_{ni}^{(1)}r_{ni}^{(2)} - \mathrm{E}(r_{ni}^{(1)}r_{ni}^{(2)})] = O_p(n^{-1/2})$.

Proof. (*i*) We shall prove the results for the simplified $r_{ni}^{(l)} = b_{ln,ii}v_{ni} + (c_{ln,ii}+d_{ln,ii}v_{ni})\sum_{j=1}^{i-1} e_{ln,ij}v_{nj} + \sum_{j=1}^{i-1}(g_{ln,ij}v_{nj}\sum_{k=1}^{j-1} h_{ln,ik}v_{nk})$ for $l = 1$ and 2, and point out that the results with the original $r_{ni}^{(l)}$'s hold similarly. For notational simplicity, we omit the subscript n in all terms though they are understood to depend on n. Since $\mathrm{E}(r_i^{(1)}r_i^{(2)}) = \sigma_i^2 b_{1,ii}b_{2,ii} + (c_{1,ii}c_{2,ii}+\sigma_i^2 d_{1,ii}d_{2,ii})\sum_{j=1}^{i-1} e_{1,ij}e_{2,ij}\sigma_j^2 + \sum_{j=1}^{i-1}(g_{1,ij}g_{2,ij}\sigma_j^2\sum_{k=1}^{j-1} h_{1,ik}h_{2,ik}\sigma_k^2)$,

$$\sup_n\sup_{1\leq i\leq n}|\,\mathrm{E}(r_i^{(1)}r_i^{(2)})| \leq \sup_n\sup_{1\leq i\leq n}\Big[\sigma_i^2|b_{1,ii}b_{2,ii}| + (|c_{1,ii}c_{2,ii}|+\sigma_i^2|d_{1,ii}d_{2,ii}|)\sum_{j=1}^{i-1}|e_{1,ij}e_{2,ij}\sigma_j^2|$$

$$+\sum_{j=1}^{i-1}\Big(|g_{1,ij}g_{2,ij}\sigma_j^2|\sum_{k=1}^{j-1}|h_{1,ik}h_{2,ik}\sigma_k^2|\Big)\Big] < c,$$

for some constant c. Thus, $\frac{1}{n}\sum_{i=1}^n \mathrm{E}(r_i^{(1)}r_i^{(2)}) = O(1)$. To prove the convergence of $\frac{1}{n}\sum_{i=1}^n [r_i^{(1)}r_i^{(2)} - \mathrm{E}(r_i^{(1)}r_i^{(2)})]$, rewrite $r_i^{(1)}r_i^{(2)} - \mathrm{E}(r_i^{(1)}r_i^{(2)}) = \Delta_{1i} + \Delta_{2i}$, where

$$\begin{aligned}
\Delta_{1i} &= b_{1,ii}b_{2,ii}(v_i^2 - \sigma_i^2) + [b_{1,ii}c_{2,ii}v_i + b_{1,ii}d_{2,ii}(v_i^2 - \sigma_i^2)]\sum_{j=1}^{i-1} e_{2,ij}v_j + [b_{2,ii}c_{1,ii}v_i + b_{2,ii}d_{1,ii}(v_i^2 - \sigma_i^2)]\sum_{j=1}^{i-1} e_{1,ij}v_j \\
&\quad + b_{1,ii}v_i\sum_{j=1}^{i-1}\Big(g_{2,ij}v_j\sum_{k=1}^{j-1} h_{2,ik}v_k\Big) + b_{2,ii}v_i\sum_{j=1}^{i-1}\Big(g_{1,ij}v_j\sum_{k=1}^{j-1} h_{1,ik}v_k\Big) \\
&\quad + [(c_{1,ii}d_{2,ii} + c_{2,ii}d_{1,ii})v_i + d_{1,ii}d_{2,ii}(v_i^2 - \sigma_i^2)]\Big[\sum_{j=1}^{i-1} e_{1,ij}e_{2,ij}(v_j^2 - \sigma_j^2) + \sum_{j=1}^{i-1}\sum_{k=1}^{j-1}(e_{1,ij}e_{2,ik} + e_{2,ij}e_{1,ik})v_jv_k\Big] \\
&\quad + [(c_{1,ii}d_{2,ii} + c_{2,ii}d_{1,ii})v_i + d_{1,ii}d_{2,ii}(v_i^2 - \sigma_i^2)]\sum_{j=1}^{i-1} e_{1,ij}e_{2,ij}\sigma_j^2 \\
&\quad + d_{1,ii}v_i\sum_{j=1}^{i-1}\Big(e_{1,ij}g_{2,ij}(v_j^2 - \sigma_j^2)\sum_{k=1}^{j-1} h_{2,ik}v_k\Big) + d_{2,ii}v_i\sum_{j=1}^{i-1}\Big(e_{2,ij}g_{1,ij}(v_j^2 - \sigma_j^2)\sum_{k=1}^{j-1} h_{1,ik}v_k\Big) \\
&\quad + d_{1,ii}v_i\sum_{j=1}^{i-1}\Big(g_{2,ij}v_j\sum_{k=1}^{j-1} e_{1,ik}h_{2,ik}(v_k^2 - \sigma_k^2)\Big) + d_{2,ii}v_i\sum_{j=1}^{i-1}\Big(g_{1,ij}v_j\sum_{k=1}^{j-1} e_{2,ik}h_{1,ik}(v_k^2 - \sigma_k^2)\Big) \\
&\quad + d_{1,ii}v_i\sum_{j=1}^{i-1}\Big(e_{1,ij}g_{2,ij}\sigma_j^2\sum_{k=1}^{j-1} h_{2,ik}v_k\Big) + d_{2,ii}v_i\sum_{j=1}^{i-1}\Big(e_{2,ij}g_{1,ij}\sigma_j^2\sum_{k=1}^{j-1} h_{1,ik}v_k\Big) \\
&\quad + d_{1,ii}v_i\sum_{j=1}^{i-1}\Big(g_{2,ij}v_j\sum_{k=1}^{j-1} e_{1,ik}h_{2,ik}\sigma_k^2\Big) + d_{2,ii}v_i\sum_{j=1}^{i-1}\Big(g_{1,ij}v_j\sum_{k=1}^{j-1} e_{2,ik}h_{1,ik}\sigma_k^2\Big) \\
&\quad + d_{1,ii}v_i\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1} v_jv_kv_l(e_{1,ij}g_{2,ik}h_{2,il} + g_{2,ij}e_{1,ik}h_{2,il} + g_{2,ij}h_{2,ik}e_{1,il}) \\
&\quad + d_{2,ii}v_i\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1} v_jv_kv_l(e_{2,ij}g_{1,ik}h_{1,il} + g_{1,ij}e_{2,ik}h_{1,il} + g_{1,ij}h_{1,ik}e_{2,il}),
\end{aligned}$$

and

$$\begin{aligned}
\Delta_{2i} &= b_{1,ii}d_{2,ii}\sigma_i^2\sum_{j=1}^{i-1} e_{2,ij}v_j + b_{2,ii}d_{1,ii}\sigma_i^2\sum_{j=1}^{i-1} e_{1,ij}v_j \\
&\quad + (c_{1,ii}c_{2,ii} + \sigma_i^2 d_{1,ii}d_{2,ii})\Big[\sum_{j=1}^{i-1} e_{1,ij}e_{2,ij}(v_j^2 - \sigma_j^2) + \sum_{j=1}^{i-1}\sum_{k=1}^{j-1}(e_{1,ij}e_{2,ik} + e_{2,ij}e_{1,ik})v_jv_k\Big] \\
&\quad + c_{1,ii}\sum_{j=1}^{i-1}\Big(e_{1,ij}g_{2,ij}(v_j^2 - \sigma_j^2)\sum_{k=1}^{j-1} h_{2,ik}v_k\Big) + c_{2,ii}\sum_{j=1}^{i-1}\Big(e_{2,ij}g_{1,ij}(v_j^2 - \sigma_j^2)\sum_{k=1}^{j-1} h_{1,ik}v_k\Big) \\
&\quad + c_{1,ii}\sum_{j=1}^{i-1}\Big(g_{2,ij}v_j\sum_{k=1}^{j-1} e_{1,ik}h_{2,ik}(v_k^2 - \sigma_k^2)\Big) + c_{2,ii}\sum_{j=1}^{i-1}\Big(g_{1,ij}v_j\sum_{k=1}^{j-1} e_{2,ik}h_{1,ik}(v_k^2 - \sigma_k^2)\Big) \\
&\quad + c_{1,ii}\sum_{j=1}^{i-1}\Big(e_{1,ij}g_{2,ij}\sigma_j^2\sum_{k=1}^{j-1} h_{2,ik}v_k\Big) + c_{2,ii}\sum_{j=1}^{i-1}\Big(e_{2,ij}g_{1,ij}\sigma_j^2\sum_{k=1}^{j-1} h_{1,ik}v_k\Big) \\
&\quad + c_{1,ii}\sum_{j=1}^{i-1}\Big(g_{2,ij}v_j\sum_{k=1}^{j-1} e_{1,ik}h_{2,ik}\sigma_k^2\Big) + c_{2,ii}\sum_{j=1}^{i-1}\Big(g_{1,ij}v_j\sum_{k=1}^{j-1} e_{2,ik}h_{1,ik}\sigma_k^2\Big) \\
&\quad + c_{1,ii}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{l=1}^{k-1} v_jv_kv_l(e_{1,ij}g_{2,ik}h_{2,il} + g_{2,ij}e_{1,ik}h_{2,il} + g_{2,ij}h_{2,ik}e_{1,il})
\end{aligned}$$

$$
\begin{aligned}
&+ c_{2,ii} \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} \sum_{l=1}^{k-1} v_j v_k v_l (e_{2,ij} g_{1,ik} h_{1,il} + g_{1,ij} e_{2,ik} h_{1,il} + g_{1,ij} h_{1,ik} e_{2,il}) \\
&+ \sum_{j=1}^{i-1} g_{1,ij} g_{2,ij} (v_j^2 - \sigma_j^2) \sum_{k=1}^{j-1} h_{1,ik} h_{2,ik} (v_k^2 - \sigma_k^2) \\
&+ \sum_{j=1}^{i-1} g_{1,ij} g_{2,ij} \sigma_j^2 \sum_{k=1}^{j-1} h_{1,ik} h_{2,ik} (v_k^2 - \sigma_k^2) + \sum_{j=1}^{i-1} g_{1,ij} g_{2,ij} (v_j^2 - \sigma_j^2) \sum_{k=1}^{j-1} h_{1,ik} h_{2,ik} \sigma_k^2 \\
&+ \sum_{j=1}^{i-1} g_{1,ij} g_{2,ij} (v_j^2 - \sigma_j^2) \sum_{k=1}^{j-1} \sum_{l=1}^{k-1} (h_{1,ik} h_{2,il} + h_{1,il} h_{2,ik}) v_k v_l + \sum_{j=1}^{i-1} g_{1,ij} g_{2,ij} \sigma_j^2 \sum_{k=1}^{j-1} \sum_{l=1}^{k-1} (h_{1,ik} h_{2,il} + h_{1,il} h_{2,ik}) v_k v_l \\
&+ \sum_{j=1}^{i-1} g_{2,ij} v_j \sum_{k=1}^{j-1} g_{1,ik} h_{2,ik} (v_k^2 - \sigma_k^2) \sum_{l=1}^{k-1} h_{1,il} v_l + \sum_{j=1}^{i-1} g_{1,ij} v_j \sum_{k=1}^{j-1} g_{2,ik} h_{1,ik} (v_k^2 - \sigma_k^2) \sum_{l=1}^{k-1} h_{2,il} v_l \\
&+ \sum_{j=1}^{i-1} g_{2,ij} v_j \sum_{k=1}^{j-1} g_{1,ik} h_{2,ik} \sigma_k^2 \sum_{l=1}^{k-1} h_{1,il} v_l + \sum_{j=1}^{i-1} g_{1,ij} v_j \sum_{k=1}^{j-1} g_{2,ik} h_{1,ik} \sigma_k^2 \sum_{l=1}^{k-1} h_{2,il} v_l \\
&+ \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} (g_{1,ij} g_{2,ik} + g_{2,ij} g_{1,ik}) v_j v_k \sum_{l=1}^{k-1} h_{1,il} h_{2,il} (v_l^2 - \sigma_l^2) + \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} (g_{1,ij} g_{2,ik} + g_{2,ij} g_{1,ik}) v_j v_k \sum_{l=1}^{k-1} h_{1,il} h_{2,il} \sigma_l^2 \\
&+ \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} \sum_{l=1}^{k-1} \sum_{m=1}^{l-1} v_j v_k v_l v_m (g_{1,ij} h_{1,ik} g_{2,il} h_{2,im} + g_{1,ij} g_{2,ik} h_{1,il} h_{2,im} \\
&\quad + g_{1,ij} g_{2,ik} h_{2,il} h_{1,im} + g_{2,ij} g_{1,ik} h_{1,il} h_{2,im} + g_{2,ij} g_{1,ik} h_{2,il} h_{1,im} + g_{2,ij} h_{2,ik} g_{1,il} h_{1,im}).
\end{aligned}
$$

Note that Δ_{1i}'s are martingale differences, and Δ_{2i} only involves $v_1, \ldots, v_{i-1}$. Each term in Δ_{1i} has the form $\prod_{l=1}^{s} \sum_{j=1}^{n} d_{ln,ij}(\theta) u_{ln,j}$ in Lemma 1(*ii*). Under the assumption that $\sup_n \sup_{1 \le i \le n} \mathrm{E}\,|v_i^4| < \infty$, by Lemma 1(*ii*), Δ_{1i}'s are uniformly integrable. Thus, by the martingale law of large numbers in Davidson (1994, p. 299, Theorem 19.7), $\frac{1}{n}\sum_{i=1}^{n} \Delta_{1i} = o_p(1)$. This argument still holds for the original $r_i^{(1)}$ and $r_i^{(2)}$ with the assumption $\sup_n \sup_{1 \le i \le n} \mathrm{E}\,|v_i^4| < \infty$. For Δ_{2i}, because each term in its expression, e.g., $v_j v_k v_l v_m g_{1,ij} h_{1,ik} g_{2,il} h_{2,im}$, has mean zero and is only correlated with any other similar term with the same subscripts j, k, l, m as j, k, l, m are different, the sum over all subscripts i, j, k, l, m of each term divided by n has a variance of order $O(n^{-1})$ under the assumption $\sup_n \sup_{1 \le i \le n} \mathrm{E}(v_i^4) < \infty$. For example,

$$
\operatorname{var}\Big(\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} \sum_{l=1}^{k-1} \sum_{m=1}^{l-1} v_j v_k v_l v_m g_{1,ij} h_{1,ik} g_{2,il} h_{2,im}\Big) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{i-1} g_{1,ij}^2 \sigma_j^2 \sum_{k=1}^{j-1} h_{1,ik}^2 \sigma_k^2 \sum_{l=1}^{k-1} g_{2,il}^2 \sigma_l^2 \sum_{m=1}^{l-1} h_{2,im}^2 \sigma_m^2 = O(n^{-1}).
$$

Thus, $\frac{1}{n}\sum_{i=1}^{n} \Delta_{2i} = O_p(n^{-1/2})$. With the original $r_i^{(1)}$ and $r_i^{(2)}$, the argument still applies under the assumption $\sup_n \sup_{1 \le i \le n} \mathrm{E}(v_i^4) < \infty$. Hence, $\frac{1}{n}\sum_{i=1}^{n} [r_i^{(1)} r_i^{(2)} - \mathrm{E}(r_i^{(1)} r_i^{(2)})] = o_p(1)$.

(*ii*) We can decompose $r_i^{(1)} r_i^{(2)} - \mathrm{E}(r_i^{(1)} r_i^{(2)}) = \Delta_{1i} + \Delta_{2i}$ in a way similar to that in (*i*). In (*i*), $\frac{1}{n}\sum_{i=1}^{n} [r_i^{(1)} r_i^{(2)} - \mathrm{E}(r_i^{(1)} r_i^{(2)})]$ is only shown to be $o_p(1)$ but may not be $O_p(n^{-1/2})$, because a martingale CLT is used to show that $\frac{1}{n}\sum_{i=1}^{n} \Delta_{1i} = o_p(1)$. With the assumption $\sup_n \sup_{1 \le i \le n} \mathrm{E}(v_i^8) < \infty$, we can show directly that the sample average of each term in the expression of $\frac{1}{n}\sum_{i=1}^{n} \Delta_{1i}$ has a variance of order $O(n^{-1})$. Then $\frac{1}{n}\sum_{i=1}^{n} \Delta_{1i} = O_p(n^{-1/2})$. Similar to (*i*), $\frac{1}{n}\sum_{i=1}^{n} \Delta_{2i} = O_p(n^{-1/2})$. Hence the result holds. □

D.2 Lemmas for the SARAR(p,q) model

Lemma 3. *Suppose that $A_{rn}(\theta) = [a_{rn,ij}(\theta)]$ for $r = 1, 2$ are square matrices of dimension n, $b_{rn}(\theta) = [b_{rn,i}(\theta)]$ for $r = 1, 2$ are column vectors of dimension n, and their elements are nonstochastic functions of $\theta \in \Theta$. Assume that elements of $A_{rn}(\theta)$ and $b_{rn}(\theta)$ are differentiable with respect to θ, the sequences $\{A_{rn}(\theta)\}$ and $\{\frac{\partial A_{rn}(\theta)}{\partial \theta_j}\}$ for $r = 1, 2$ and $j = 1, \ldots, k_\theta$ are bounded in both row and column sum norms, and $\{b_{rn}(\theta)\}$ and $\{\frac{\partial b_{rn}(\theta)}{\partial \theta_j}\}$ for $r = 1, 2$ and $j = 1, \ldots, k_\theta$ are bounded in row sum norm, uniformly in a neighborhood of θ_0.*

Let $\xi_{rn,i}(\theta) = a_{rn,ii}(\theta)[v_{ni}^2(\theta) - \sigma^2] + 2v_{ni}(\theta)\sum_{j=1}^{i-1} a_{rn,ij}(\theta)v_{nj}(\theta) + b_{rn,i}(\theta)v_{ni}(\theta)$ for $r = 1, 2$ if the disturbances v_{ni}'s are homoskedastic, and $\xi_{rn,i}(\theta) = 2v_{ni}(\theta)\sum_{j=1}^{i-1} a_{rn,ij}(\theta)v_{nj}(\theta) + b_{rn,i}(\theta)v_{ni}(\theta)$ for $r = 1, 2$ if v_{ni}'s are heteroskedastic. Assume that $\hat\theta_n = \theta_0 + o_p(1)$. Then, under Assumptions 1–4, $\frac{1}{n}\sum_{i=1}^n \xi_{1n,i}(\hat\theta_n)\xi_{2n,i}(\hat\theta_n) = \frac{1}{n}\sum_{i=1}^n \mathrm{E}[\xi_{1n,i}(\theta_0)\xi_{2n,i}(\theta_0)] + o_p(1)$.

Proof. By the mean value theorem,

$$\frac{1}{n}\sum_{i=1}^n \xi_{1n,i}(\hat\theta_n)\xi_{2n,i}(\hat\theta_n) = \frac{1}{n}\sum_{i=1}^n \xi_{1n,i}(\theta_0)\xi_{2n,i}(\theta_0) + \sum_{l=1}^{k_\theta}\frac{1}{n}\sum_{i=1}^n\Big[\frac{\partial \xi_{1n,i}(\check\theta_n)}{\partial\theta_l}\xi_{2n,i}(\check\theta_n) + \xi_{1n,i}(\check\theta_n)\frac{\partial \xi_{2n,i}(\check\theta_n)}{\partial\theta_l}\Big](\hat\theta_{nl} - \theta_{0l}),$$

where $\check\theta_n$ lies between θ_0 and $\hat\theta_n$. We shall show that the second term on the r.h.s. of the above equation goes to zero in probability. Note that $V_{ni}(\theta) = e_{ni}'V_n(\theta)$ and $\sum_{j=1}^{i-1} a_{rn,ij}(\theta)v_{nj}(\theta) = \mathrm{tril}[A_{rn}(\theta)]V_n(\theta)$, where e_{ni} is the ith unit column vector of dimension n, and $\mathrm{tril}(A)$ for a square matrix A denotes the strictly lower triangular matrix formed by the elements below the diagonal of A. Under the assumptions in the lemma, $\mathrm{tril}[A_{rn}(\theta)]$ and $\frac{\partial\,\mathrm{tril}[A_{rn}(\theta)]}{\partial\theta_l}$ for $l = 1, \ldots, k_\theta$ are bounded in both row and column sum norms uniformly in a neighborhood of θ_0. Since $Y_n = S_n^{-1}(X_n\beta_0 + R_n^{-1}V_n)$,

$$\begin{aligned}
V_n(\theta) &= \Big[R_n + \sum_{k=1}^q(\tau_{0k} - \tau_k)M_{kn}\Big]\Big\{\Big[S_n + \sum_{j=1}^p(\kappa_{0j} - \kappa_j)W_{jn}\Big]S_n^{-1}(X_n\beta_0 + R_n^{-1}V_n) - X_n\beta_0 + X_n(\beta_0 - \beta)\Big\}\\
&= R_nX_n(\beta_0 - \beta) + \sum_{j=1}^p(\kappa_{0j} - \kappa_j)R_nW_{jn}S_n^{-1}X_n\beta_0 + \sum_{k=1}^q M_{kn}X_n(\beta_0 - \beta)(\tau_{0k} - \tau_k)\\
&\quad + \sum_{j=1}^p\sum_{k=1}^q(\kappa_{0j} - \kappa_j)(\tau_{0k} - \tau_k)M_{kn}W_{jn}S_n^{-1}X_n\beta_0 + V_n + \sum_{j=1}^p(\kappa_{0j} - \kappa_j)R_nW_{jn}S_n^{-1}R_n^{-1}V_n\\
&\quad + \sum_{k=1}^q(\tau_{0k} - \tau_k)M_{kn}R_n^{-1}V_n + \sum_{j=1}^p\sum_{k=1}^q(\kappa_{0j} - \kappa_j)(\tau_{0k} - \tau_k)M_{kn}W_{jn}S_n^{-1}R_n^{-1}V_n,
\end{aligned} \tag{D.3}$$

which is linear in V_n and quadratic in $(\theta_0 - \theta)$. In (D.3), terms that do not involve V_n have uniformly bounded elements, and terms that involve V_n have matrices in front of V_n bounded in both row and column sum norms. We can expand $\frac{1}{n}\sum_{i=1}^n[\frac{\partial\xi_{1n,i}(\theta)}{\partial\theta_l}\xi_{2n,i}(\theta) + \xi_{1n,i}(\theta)\frac{\partial\xi_{2n,i}(\theta)}{\partial\theta_l}]$ by using (D.3) such that it is a sum of terms that have the form in Lemma 1(iii) with $u_{ln,i} = v_{ni}$. Then $\frac{1}{n}\sum_{i=1}^n[\frac{\partial\xi_{1n,i}(\theta)}{\partial\theta_l}\xi_{2n,i}(\theta) + \xi_{1n,i}(\theta)\frac{\partial\xi_{2n,i}(\theta)}{\partial\theta_l}] = O_p(1)$ uniformly in a neighborhood of θ_0. Since $\hat\theta_n = \theta_0 + o_p(1)$, $\frac{1}{n}\sum_{i=1}^n[\frac{\partial\xi_{1n,i}(\check\theta_n)}{\partial\theta_l}\xi_{2n,i}(\check\theta_n) + \xi_{1n,i}(\check\theta_n)\frac{\partial\xi_{2n,i}(\check\theta_n)}{\partial\theta_l}] = O_p(1)$. Thus,

$$\frac{1}{n}\sum_{i=1}^n \xi_{1n,i}(\hat\theta_n)\xi_{2n,i}(\hat\theta_n) = \frac{1}{n}\sum_{i=1}^n \xi_{1n,i}(\theta_0)\xi_{2n,i}(\theta_0) + o_p(1).$$

By Lemma 2(i), $\frac{1}{n}\sum_{i=1}^n \xi_{1n,i}(\theta_0)\xi_{2n,i}(\theta_0) = \frac{1}{n}\sum_{i=1}^n \mathrm{E}[\xi_{1n,i}(\theta_0)\xi_{2n,i}(\theta_0)] + o_p(1)$. Hence, the result in the lemma follows. □

Lemma 4. *Under Assumptions 1–4,* $\sup_{\theta\in\Theta}\sup_{1\le i\le n}\|g_{ni}(\theta)\| = O_p(n^{2/(4+\iota)})$*, where* $\|\cdot\|$ *denotes the Euclidean norm.*

Proof. The expression for $g_{ni}(\theta)$ is given in (4), in which $\omega_{ln,i}(\theta) = p_{ln,ii}[(e'_{ni}V_n(\theta))^2 - \sigma^2] + 2e'_{ni}V_n(\theta)e'_{ni}\text{tril}(P_{ln})V_n(\theta)$ in the homoskedastic case, and $\omega_{ln,i}(\theta) = 2e'_{ni}V_n(\theta)e'_{ni}\text{tril}(P_{ln})V_n(\theta)$, where $V_n(\theta)$ is linear in V_n and quadratic in θ in (D.3), and $\text{tril}(P_{ln})$ is bounded in both row and column sum norms. Using the expression of $V_n(\theta)$ in (D.3), each element of $g_{ni}(\theta)$ can be expanded as a polynomial of θ whose coefficients have the form $\prod_{l=1}^{s}\sum_{j=1}^{n} d_{ln,ij}(\theta)u_{ln,j}$ with $u_{ln,j} = v_{nj}$ and $s = 1$ or 2 in Lemma 1(i). Thus, the result follows by Lemma 1(i). □

Let $G_n^{(j)}(\theta) = \frac{\partial G_n(\theta)}{\partial\theta_j}$, $G_n^{(jk)}(\theta) = \frac{\partial^2 G_n(\theta)}{\partial\theta_j\partial\theta_k}$, $G_n^{(jkl)}(\theta) = \frac{\partial^3 G_n(\theta)}{\partial\theta_j\partial\theta_k\partial\theta_l}$ and $G_{ni}^{(j)}(\theta) = \frac{\partial G_{ni}(\theta)}{\partial\theta_j}$, where $G_{ni}(\theta) = \frac{\partial g_{ni}(\theta)}{\partial\theta'}$.

Lemma 5. *Under Assumptions 1–4,* $\sup_{\theta\in\Theta}\frac{1}{n}\sum_{i=1}^{n}\|g_{ni}(\theta)\|^2$, $\sup_{\theta\in\Theta}\frac{1}{n}\sum_{i=1}^{n}\|G_{ni}(\theta)\|$, $\sup_{\theta\in\Theta}\frac{1}{n}\sum_{i=1}^{n}\|G_{ni}^{(j)}(\theta)\|$, $\sup_{\theta\in\Theta}\|g_n(\theta)\|$, $\sup_{\theta\in\Theta}\|G_n(\theta)\|$, $\sup_{\theta\in\Theta}\|G_n^{(j)}(\theta)\|$, $\sup_{\theta\in\Theta}\|G_n^{(jk)}(\theta)\|$, *and* $\sup_{\theta\in\Theta}\|G_n^{(jkl)}(\theta)\|$ *are all of order* $O_p(1)$.

Proof. By (D.3) and the proof of Lemma 4, we can expand $\frac{1}{n}\sum_{i=1}^{n}|\omega_{ln,i}(\theta)|^2$ and $\frac{1}{n}\sum_{i=1}^{n}|v_{ni}(\theta)|^2$ as polynomials of θ. Since $g_n(\theta)$ is quadratic in $V_n(\theta)$, each element of $g_n(\theta)$ can be expanded as a polynomial of θ. Each coefficient of those polynomials is $O_p(1)$ by Lemma 1(iii). Hence the results hold. □

Let $g_{ni}^{(j)}(\theta) = \frac{\partial g_{ni}(\theta)}{\partial\theta_j}$, $g_{ni}^{(jk)}(\theta) = \frac{\partial^2 g_{ni}(\theta)}{\partial\theta_j\partial\theta_k}$, $g_{ni}^{(jkl)}(\theta) = \frac{\partial^3 g_{ni}(\theta)}{\partial\theta_j\partial\theta_k\partial\theta_l}$, and $g_{ni}^{(jklr)}(\theta) = \frac{\partial^4 g_{ni}(\theta)}{\partial\theta_j\partial\theta_k\partial\theta_l\partial\theta_r}$.

Lemma 6. *Under Assumptions 1–4,* $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}(\theta)g'_{ni}(\theta)\|$, $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(j)}(\theta)g'_{ni}(\theta)\|$, $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(jk)}(\theta)g'_{ni}(\theta)\|$, $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(j)}(\theta)g_{ni}^{(k)'}(\theta)\|$, $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(jkl)}(\theta)g'_{ni}(\theta)\|$, $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(jk)}(\theta)g_{ni}^{(l)'}(\theta)\|$, $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(jklr)}(\theta)g'_{ni}(\theta)\|$, $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(jkl)}(\theta)g_{ni}^{(r)'}(\theta)\|$ *and* $\sup_{\theta\in\Theta}\|\frac{1}{n}\sum_{i=1}^{n}g_{ni}^{(jk)}(\theta)g_{ni}^{(lr)'}(\theta)\|$ *have order* $O_p(1)$.

Proof. As in the proof of Lemma 5, $\frac{1}{n}\sum_{i=1}^{n}g_{ni}(\theta)g'_{ni}(\theta)$ can be expanded as a polynomial of θ with coefficients being $O_p(1)$ by Lemma 1(iii). Then the results in the lemma follow. □

Lemma 7. *Under Assumptions 1–4, (i)* $\frac{1}{n}\sum_{i=1}^{n}g_{ni}(\theta_0)g'_{ni}(\theta_0) = \bar{\Omega}_n + o_p(1)$*, (ii)* $\frac{1}{n}\sum_{i=1}^{n}[\mathrm{E}\, g_{ni}^{(j)}(\theta_0)]g'_{ni}(\theta_0) = O_p(n^{-1/2})$*, and (iii)* $\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[g_{ni}^{(j)}(\theta_0)g'_{ni}(\theta_0)] = O(1)$*; under Assumptions 1–4 and 10, (iv)* $\frac{1}{n}\sum_{i=1}^{n}g_{ni}(\theta_0)g'_{ni}(\theta_0) = \bar{\Omega}_n + O_p(n^{-1/2})$*, and (v)* $\frac{1}{n}\sum_{i=1}^{n}\{g_{ni}^{(j)}(\theta_0)g'_{ni}(\theta_0) - \mathrm{E}[g_{ni}^{(j)}(\theta_0)g'_{ni}(\theta_0)]\} = O_p(n^{-1/2})$*, for* $j = 1, \ldots, k_\theta$.

Proof. We omit the subscript n in relevant terms for notational simplicity. Since P_j's are symmetric, each element of $g_i(\theta_0)$ has the linear-quadratic form $a_{ii}(v_i^2 - \sigma_0^2) + 2v_i\sum_{j=1}^{i-1}a_{ij}v_j + b_iv_i$ or $2v_i\sum_{j=1}^{i-1}a_{ij}v_j + b_iv_i$, where a_{ij} is the (i,j)th element of an $n\times n$ nonstochastic matrix with bounded row and column sum norms, and b_i for $i = 1, \ldots, n$ are constants bounded uniformly in i. Then (i) and (iv) follow, respectively, by Lemma 2(i) and Lemma 2(ii). It remains to show (ii), (iii) and (v).

The lth element of $g_i^{(k)}(\theta)$ for $1 \le l \le k_p$ is $\frac{\partial\omega_{l,i}(\theta)}{\partial\theta_k} = 2p_{l,ii}v_i(\theta)\frac{\partial v_i(\theta)}{\partial\theta_k} - p_{l,ii}\frac{\partial\sigma^2}{\partial\theta_k} + 2\sum_{j=1}^{i-1}p_{l,ij}\big[v_i(\theta)\frac{\partial v_j(\theta)}{\partial\theta_k} + v_j(\theta)\frac{\partial v_i(\theta)}{\partial\theta_k}\big]$ in the homoskedastic case, and $\frac{\partial\omega_{l,i}(\theta)}{\partial\theta_k} = 2\sum_{j=1}^{i-1}p_{l,ij}\big[v_i(\theta)\frac{\partial v_j(\theta)}{\partial\theta_k} + v_j(\theta)\frac{\partial v_i(\theta)}{\partial\theta_k}\big]$ in the heteroskedastic case; and the last k_q elements are $Q_i\frac{\partial v_i(\theta)}{\partial\theta_k}$. By (D.3), $\frac{\partial v_i(\theta_0)}{\partial\theta_k}$ has the form $a_{ki} + \sum_{r=1}^{n}b_{k,ir}v_r$, where a_{ki} is bounded uniformly in k and i, and $b_{k,ir}$ is the (i,r)th element of an $n\times n$ nonstochastic matrix bounded in both row and column sum norms. Hence, every element of $g_i^{(k)}(\theta_0)$ has the form

$$\Xi_{ki}=2p_{l,ii}v_i\Big(a_{ki}+\sum_{s=1}^n b_{k,is}v_s\Big)-p_{l,ii}\frac{\partial\sigma^2}{\partial\theta_k}+2\sum_{j=1}^{i-1}p_{l,ij}\Big[v_i\Big(a_{kj}+\sum_{s=1}^n b_{k,js}v_s\Big)+v_j\Big(a_{ki}+\sum_{s=1}^n b_{k,is}v_s\Big)\Big].$$

Thus, $\mathrm{E}(\Xi_{ki})=2p_{l,ii}b_{k,ii}\sigma_i^2-p_{l,ii}\frac{\partial\sigma^2}{\partial\theta_k}+2\sum_{j=1}^{i-1}p_{l,ij}(b_{k,ji}\sigma_i^2+b_{k,ij}\sigma_j^2)$ is bounded uniformly in i and k. Note that the variance of the term in (ii) only involves the first fourth moments of v_j, then the result (ii) follows by Lemma 2(ii).

Below we shall prove (iii) in the lemma and (vi) $\frac{1}{n}\sum_{i=1}^n\big\{[g_i^{(k)}(\theta_0)-\mathrm{E}(g_i^{(k)}(\theta_0))]g_i'(\theta_0)-\mathrm{E}[g_i^{(k)}(\theta_0)g_i'(\theta_0)]\big\}=O_p(n^{-1/2})$. The result (v) in the lemma follows by (iii) and (vi). Write $\Xi_{ki}-\mathrm{E}(\Xi_{ki})=\Xi_{1k,i}+\Xi_{2k,i}$, where

$$\begin{aligned}\Xi_{1k,i}&=2p_{l,ii}a_{ki}v_i+2p_{l,ii}v_i\sum_{s=1}^{i-1}b_{k,is}v_s+2p_{l,ii}b_{k,ii}(v_i^2-\sigma_i^2)+2v_i\sum_{j=1}^{i-1}p_{l,ij}a_{kj}\\&\quad+2v_i\sum_{j=1}^{i-1}p_{l,ij}\sum_{s=1}^{i-1}b_{k,js}v_s+2(v_i^2-\sigma_i^2)\sum_{j=1}^{i-1}p_{l,ij}b_{k,ji}+2a_{ki}\sum_{j=1}^{i-1}p_{l,ij}v_j+2\sum_{j=1}^{i-1}p_{l,ij}v_j\sum_{s=1}^{j-1}b_{k,is}v_s\\&\quad+2\sum_{j=1}^{i-1}p_{l,ij}b_{k,ij}(v_j^2-\sigma_j^2)+2\sum_{j=1}^{i-1}p_{l,ij}v_j\sum_{s=j+1}^{i-1}b_{k,is}v_s+2b_{k,ii}v_i\sum_{j=1}^{i-1}p_{l,ij}v_j,\end{aligned}$$

and

$$\begin{aligned}\Xi_{2k,i}&=2p_{l,ii}v_i\Big(\sum_{s=i+1}^n b_{k,is}v_s\Big)+2v_i\Big(\sum_{j=1}^{i-1}p_{l,ij}\sum_{s=i+1}^n b_{k,js}v_s\Big)+2\sum_{j=1}^{i-1}p_{l,ij}v_j\sum_{s=i+1}^n b_{k,is}v_s\\&=2p_{l,ii}v_i\Big(\sum_{s=i+1}^n b_{k,is}v_s\Big)+2v_i\Big(\sum_{s=i+1}^n v_s\sum_{j=1}^{i-1}p_{l,ij}b_{k,js}\Big)+2\sum_{s=i+1}^n b_{k,is}v_s\sum_{j=1}^{i-1}p_{l,ij}v_j.\end{aligned}$$

Note that $\Xi_{1k,i}$ has the form of $r_i^{(1)}$ in Lemma 2(ii) because $2\sum_{j=1}^{i-1}p_{l,ij}v_j\sum_{s=j+1}^{i-1}b_{k,is}v_s=2\sum_{s=1}^{i-1}b_{k,is}v_s\sum_{j=1}^{s-1}p_{l,ij}v_j$. Thus $\frac{1}{n}\sum_{i=1}^n\mathrm{E}[\Xi_{1k,i}g_i'(\theta_0)]=O(1)$ and $\frac{1}{n}\sum_{i=1}^n\Xi_{1k,i}g_i'(\theta_0)-\frac{1}{n}\sum_{i=1}^n\mathrm{E}[\Xi_{1k,i}g_i'(\theta_0)]=O_p(n^{-1/2})$. For $\Xi_{2k,i}$, we shall show that $\frac{1}{n}\sum_{i=1}^n\mathrm{E}[\Xi_{2k,i}g_i'(\theta_0)]=0$ and $\frac{1}{n}\sum_{i=1}^n\Xi_{2k,i}g_i'(\theta_0)-\frac{1}{n}\sum_{i=1}^n\mathrm{E}[\Xi_{2k,i}g_i'(\theta_0)]=O_p(n^{-1/2})$. Each term in $\Xi_{2k,i}$ has the form $(\sum_{s=i+1}^n b^*_{k,is}v_s)(\sum_{j=1}^i p^*_{l,ij}v_j)$, where $b^*_{k,is}$ is the (i,s)th element of a general $n\times n$ nonstochastic matrix with bounded row and column sum norms uniformly in k. The term $\sum_{j=1}^i p^*_{l,ij}v_j$ is a special form of $r_i^{(1)}$ in Lemma 2(ii). Compared with the form of $r_i^{(1)}r_i^{(2)}$ in Lemma 2(ii), each element of $\Xi_{2k,i}g_i'(\theta_0)$ has the additional term $\sum_{s=i+1}^n b^*_{k,is}v_s$. In the proof of Lemma 2(ii), If we multiply each term in $r_i^{(1)}r_i^{(2)}-\mathrm{E}(r_i^{(1)}r_i^{(2)})$ by $\sum_{s=i+1}^n b^*_{k,is}v_s$, then the obtained terms have zero expected values and the sum over all subscripts i,j,k,l of those terms divided by n still has the order $O_p(n^{-1/2})$, because the summation $\sum_{s=i+1}^n b^*_{k,is}v_s$ starts from $s=i+1$. Hence, $\mathrm{E}(\Xi_{2k,i}r_i^{(2)})=0$ and $\frac{1}{n}\sum_{i=1}^n[\Xi_{2k,i}g_i'(\theta_0)-\mathrm{E}(\Xi_{2k,i}g_i'(\theta_0))]=O_p(n^{-1/2})$. Then (iii) and (vi) follow. □

The first order condition for the initial GMM can be written as

$$0=-\begin{pmatrix}G_n'(\tilde\theta_n)\tilde\lambda_n\\ g_n(\tilde\theta_n)+\hat J_n\tilde\lambda_n\end{pmatrix},\tag{D.4}$$

where $\tilde\lambda_n=-\hat J_n^{-1}g_n(\tilde\theta_n)$. Let $\tilde\gamma_n=(\tilde\theta_n',\tilde\lambda_n')'$ and $\gamma_0=(\theta_0',0_{1\times k_g})'$.

Lemma 8. *Under Assumptions 1–5 and 8–9,* $\sqrt{n}(\tilde{\gamma}_n-\gamma_0) = \tilde{\xi}_n+n^{-1/2}\tilde{\psi}_n+O_p(n^{-1})$*, where* $\tilde{\xi}_n = -(\bar{K}_n^J)^{-1}\binom{0}{\sqrt{n}g_n(\theta_0)} = O_p(1)$ *and* $\tilde{\psi}_n = -(\bar{K}_n^J)^{-1}\begin{pmatrix} 0 & \sqrt{n}(G_n' - \bar{G}_n') \\ \sqrt{n}(G_n - \bar{G}_n) & \xi_n^J \end{pmatrix}\tilde{\xi}_n - \frac{1}{2}(\bar{K}_n^J)^{-1}\sum_{j=1}^{k_\theta+k_g}\tilde{\xi}_{nj}\bar{K}_{nj}\tilde{\xi}_n = O_p(1)$*, where* $\bar{K}_n^J = \begin{pmatrix} 0 & \bar{G}_n' \\ \bar{G}_n & \bar{J}_n \end{pmatrix}$, $\bar{K}_{nj} = \begin{pmatrix} 0 & \bar{G}_n^{(j)'} \\ \bar{G}_n^{(j)} & 0 \end{pmatrix}$ *for* $1 \le j \le k_\theta$*, and* $\bar{K}_{nj} = \begin{pmatrix} [\bar{G}_n^{(1)'}e_{k_g,j-k_\theta},\ldots,\bar{G}_n^{(k_\theta)'}e_{k_g,j-k_\theta}] & 0 \\ 0 & 0 \end{pmatrix}$ *for* $k_\theta + 1 \le j \le k_\theta + k_g$.

Proof. As shown in Lee and Liu (2010), $\tilde{\theta}_n = \theta_0 + O_p(n^{-1/2})$. Then by the mean value theorem and Lemma 5, $g_n(\tilde{\theta}_n) = O_p(n^{-1/2})$. Thus $\tilde{\lambda}_n = -\hat{J}_n^{-1}g_n(\tilde{\theta}_n) = O_p(n^{-1/2})$. It follows that $\tilde{\gamma}_n - \gamma_0 = O_p(n^{-1/2})$. Together with Assumption 8, the first order condition (D.4) for the initial GMM is equal to

$$0 = -\begin{pmatrix} G_n'(\tilde{\theta}_n)\tilde{\lambda}_n \\ g_n(\tilde{\theta}_n) + (\bar{J}_n + n^{-1/2}\xi_n^J)\tilde{\lambda}_n \end{pmatrix} + O_p(n^{-3/2}).$$

By a second order Taylor expansion of the first vector on the right hand side at γ_0, and using Lemma 5,

$$0 = -\begin{pmatrix} 0 \\ g_n(\theta_0) \end{pmatrix} - K_n^J(\tilde{\gamma}_n - \gamma_0) - \frac{1}{2}\sum_{j=1}^{k_\theta+k_g}(\tilde{\gamma}_{nj} - \gamma_{0j})K_{nj}(\tilde{\gamma}_n - \gamma_0) + O_p(n^{-3/2}),$$

where $K_n^J = \begin{pmatrix} 0 & G_n'(\theta_0) \\ G_n(\theta_0) & \bar{J}_n + n^{-1/2}\xi_n^J \end{pmatrix}$, $K_{nj} = \begin{pmatrix} 0 & G_n^{(j)'}(\theta_0) \\ G_n^{(j)}(\theta_0) & 0 \end{pmatrix}$ for $1 \le j \le k_\theta$, and $K_{nj} = \begin{pmatrix} [G_n^{(1)'}(\theta_0)e_{k_g,j-k_\theta},\ldots,G_n^{(k_\theta)'}(\theta_0)e_{k_g,j-k_\theta}] & 0 \\ 0 & 0 \end{pmatrix}$ for $k_\theta + 1 \le j \le k_\theta + k_g$. As $\bar{K}_n^J = \mathrm{E}(K_n^J) = O(1)$ and $\bar{K}_{nj} = \mathrm{E}(K_{nj}) = O(1)$ for all $j = 1,\ldots,k_\theta + k_g$,

$$\begin{aligned} \sqrt{n}(\tilde{\gamma}_n - \gamma_0) &= -(\bar{K}_n^J)^{-1}\begin{pmatrix} 0 \\ \sqrt{n}g_n(\theta_0) \end{pmatrix} - (\bar{K}_n^J)^{-1}(K_n^J - \bar{K}_n^J)\sqrt{n}(\tilde{\gamma}_n - \gamma_0) \\ &\quad - \frac{\sqrt{n}}{2}(\bar{K}_n^J)^{-1}\sum_{j=1}^{k_\theta+k_g}(\tilde{\gamma}_{nj} - \gamma_{0j})K_{nj}(\tilde{\gamma}_n - \gamma_0) + O_p(n^{-1}). \end{aligned} \tag{D.5}$$

As every element of $g_n(\theta)$ is a linear-quadratic form of $V_n(\theta)$, which is linear in V_n by (D.3), $G_n(\theta_0) - \bar{G}_n = O_p(n^{-1/2})$ and $G_n^{(j)}(\theta_0) - \bar{G}_n^{(j)}(\theta_0) = O_p(n^{-1/2})$ by Lemma 1(*iii*). It follows that $K_n^J - \bar{K}_n^J = O_p(n^{-1/2})$ and $K_{nj} - \bar{K}_{nj} = O_p(n^{-1/2})$. Hence, $\sqrt{n}(\tilde{\gamma}_n - \gamma_0) = \tilde{\xi}_n + O_p(n^{-1/2})$, where $\tilde{\xi}_n = -(\bar{K}_n^J)^{-1}\binom{0}{\sqrt{n}g_n(\theta_0)} = O_p(1)$. Substituting $K_{nj} - \bar{K}_{nj} = O_p(n^{-1/2})$ and $\sqrt{n}(\tilde{\gamma}_n - \gamma_0) = \tilde{\xi}_n + O_p(n^{-1/2})$ into (D.5) yields $\sqrt{n}(\tilde{\gamma}_n - \gamma_0) = \tilde{\xi}_n + \tilde{\psi}_n + O_p(n^{-1})$. ☐

Lemma 9. *Under Assumptions 1–5 and 8–9,* $\Omega_n(\tilde{\theta}_n) = \bar{\Omega}_n + o_p(1)$*; under the additional Assumption 10,*

$$\sqrt{n}[\Omega_n(\tilde{\theta}_n) - \bar{\Omega}_n] = \xi_n^\Omega + O_p(n^{-1/2}),$$

where $\xi_n^\Omega = \sqrt{n}[\frac{1}{n}\sum_{i=1}^n g_{ni}(\theta_0)g_{ni}'(\theta_0) - \bar{\Omega}_n] + \sum_{j=1}^{k_\theta}\{\frac{1}{n}\sum_{i=1}^n \mathrm{E}[g_{ni}(\theta_0)g_{ni}^{(j)'}(\theta_0) + g_{ni}^{(j)}(\theta_0)g_{ni}'(\theta_0)]\}\tilde{\xi}_{nj} = O_p(1)$.

Proof. By a first order Taylor expansion and Lemma 6,

$$\Omega_n(\tilde{\theta}_n) = \bar{\Omega}_n + \Big(\frac{1}{n}\sum_{i=1}^n g_{ni}(\theta_0)g_{ni}'(\theta_0) - \bar{\Omega}_n\Big) + \sum_{j=1}^{k_\theta}\Big\{\frac{1}{n}\sum_{i=1}^n[g_{ni}(\theta_0)g_{ni}^{(j)'}(\theta_0) + g_{ni}^{(j)}(\theta_0)g_{ni}'(\theta_0)]\Big\}(\tilde{\theta}_{nj} - \theta_{0j}) + O_p(n^{-1}).$$

Under Assumptions 1–5 and 8–9, by Lemma 7(i), the second term on the r.h.s. of the above equation is $o_p(1)$; by Lemma 6, the third term is $o_p(1)$. Thus the first result follows. The second result requires the existence of higher order moments of disturbances, which is maintained in Assumption 10. Substituting the expression for $(\tilde{\theta}_n - \theta_0)$ in Lemma 8 into the above equation and keeping only terms with order $O_p(n^{-1/2})$ by using Lemma 7, we obtain the result. $\square$

Lemma 10. *Under Assumptions 1–4, for any* ζ *with* $\zeta > \frac{2}{4+\iota}$ *and* $\Lambda_n = \{\lambda : \|\lambda\| \le n^{-\zeta}\}$, $\sup_{\theta\in\Theta,\lambda\in\Lambda_n,1\le i\le n}|\lambda' g_{ni}(\theta)| \xrightarrow{p} 0$, *and* $\Lambda_n \subset \Lambda_n(\theta)$ *for all* $\theta \in \Theta$ *w.p.a.1.*

Proof. Let $b_n = \sup_{1\le i\le n}\sup_{\theta\in\Theta}\|g_{ni}(\theta)\|$. By Lemma 4, $b_n = O_p(n^{2/(4+\iota)})$. Then by the Cauchy-Schwarz inequality, $\sup_{\theta\in\Theta,\lambda\in\Lambda_n,1\le i\le n}|\lambda' g_{ni}(\theta)| \le n^{-\zeta} b_n = O_p(n^{2/(4+\iota)-\zeta}) = o_p(1)$. Given the first conclusion, w.p.a.1. $\lambda' g_{ni}(\theta) \in \mathcal{V}$ for all $1 \le i \le n$, $\theta \in \Theta$ and $\|\lambda\| \le n^{-\zeta}$. $\square$

With the above lemmas, Lemmas 11–12 for the GEL estimation follow by arguments similar to those for Lemmas A2–A3 in Newey and Smith (2004), thus their proofs are omitted.

E Some expressions for the proof of Theorem 3.4

$$\frac{\partial h_{ni}(\gamma)}{\partial \gamma'} = \begin{pmatrix} [G_{ni}^{(1)'}(\theta)\lambda, \ldots, G_{ni}^{(k_\theta)'}(\theta)\lambda] & G'_{ni}(\theta) \\ G_{ni}(\theta) & 0 \end{pmatrix},$$

$$\frac{\partial^2 h_{ni}(\gamma)}{\partial \gamma_j \partial \gamma'} = \begin{pmatrix} [G_{ni}^{(1j)'}(\theta)\lambda, \ldots, G_{ni}^{(k_\theta,j)'}(\theta)\lambda] & G_{ni}^{(j)'}(\theta) \\ G_{ni}^{(j)}(\theta) & 0 \end{pmatrix} \text{ for } 1 \le j \le k_\theta,$$

$$\frac{\partial^2 h_{ni}(\gamma)}{\partial \gamma_j \partial \gamma'} = \begin{pmatrix} [G_{ni}^{(1)'}(\theta)e_{k_g,j-k_\theta}, \ldots, G_{ni}^{(k_\theta)'}(\theta)e_{k_g,j-k_\theta}] & 0 \\ 0 & 0 \end{pmatrix} \text{ for } k_\theta + 1 \le j \le k_\theta + k_g,$$

$$\frac{\partial^3 h_{ni}(\gamma)}{\partial \gamma_k \partial \gamma_j \partial \gamma'} = \begin{pmatrix} [G_{ni}^{(1jk)'}(\theta)\lambda, \ldots, G_{ni}^{(k_\theta,jk)'}(\theta)\lambda] & G_{ni}^{(jk)'}(\theta) \\ G_{ni}^{(jk)}(\theta) & 0 \end{pmatrix} \text{ for } 1 \le j \le k_\theta, \text{ and } 1 \le k \le k_\theta,$$

$$\frac{\partial^3 h_{ni}(\gamma)}{\partial \gamma_k \partial \gamma_j \partial \gamma'} = \begin{pmatrix} [G_{ni}^{(1j)'}(\theta)e_{k_g,k-k_\theta}, \ldots, G_{ni}^{(k_\theta,j)'}(\theta)e_{k_g,k-k_\theta}] & 0 \\ 0 & 0 \end{pmatrix} \text{ for } 1 \le j \le k_\theta \text{ and } k_\theta + 1 \le k \le k_\theta + k_g,$$

$$\frac{\partial^3 h_{ni}(\gamma)}{\partial \gamma_k \partial \gamma_j \partial \gamma'} = \begin{pmatrix} [G_{ni}^{(1k)'}(\theta)e_{k_g,j-k_\theta}, \ldots, G_{ni}^{(k_\theta,k)'}(\theta)e_{k_g,j-k_\theta}] & 0 \\ 0 & 0 \end{pmatrix} \text{ for } k_\theta + 1 \le j \le k_\theta + k_g \text{ and } 1 \le k \le k_\theta,$$

and $\frac{\partial^3 h_{ni}(\gamma)}{\partial \gamma_k \partial \gamma_j \partial \gamma'} = 0$ for $k_\theta + 1 \le j \le k_\theta + k_g$ and $k_\theta + 1 \le k \le k_\theta + k_g$. Let g_{nit} be the tth element of g_{ni}, and $g_{nit}^{(k)}$ be the tth element of $g_{ni}^{(k)}$. With the above derivatives, by the chain rule of differentiation,

$$\frac{\partial m_{ni}(\gamma_0)}{\partial \gamma'} = -\begin{pmatrix} 0 & G'_{ni}(\theta_0) \\ G_{ni}(\theta_0) & g_{ni}(\theta_0)g'_{ni}(\theta_0) \end{pmatrix},$$

$$\frac{\partial^2 m_{ni}(\gamma_0)}{\partial \gamma_j \partial \gamma'} = -\begin{pmatrix} 0 & G_{ni}^{(j)'} \\ G_{ni}^{(j)} & g_{ni}^{(j)}g'_{ni} + g_{ni}g_{ni}^{(j)'} \end{pmatrix} \text{ for } 1 \le j \le k_\theta,$$

$$\frac{\partial^2 m_{ni}(\gamma_0)}{\partial\gamma_j\partial\gamma'} = -\begin{pmatrix} [G_{ni}^{(1)'}e_{k_g,s},\ldots,G_{ni}^{(k_\theta)'}e_{k_g,s}] & G'_{ni}e_{k_g,s}g'_{ni}+g_{nis}G'_{ni} \\ g_{ni}e'_{k_g,s}G_{ni}+g_{nis}G_{ni} & -\rho_3 g_{nis}g_{ni}g'_{ni} \end{pmatrix} \text{ for } k_\theta+1\le j\le k_\theta+k_g \text{ and } s=j-k_\theta,$$

$$\frac{\partial^3 m_{ni}(\gamma_0)}{\partial\gamma_k\partial\gamma_j\partial\gamma'} = -\begin{pmatrix} 0 & G_{ni}^{(jk)'} \\ G_{ni}^{(jk)} & g_{ni}^{(jk)}g'_{ni}+g_{ni}g_{ni}^{(jk)'}+g_{ni}^{(j)}g_{ni}^{(k)'}+g_{ni}^{(k)}g_{ni}^{(j)'} \end{pmatrix} \text{ for } 1\le j\le k_\theta \text{ and } 1\le k\le k_\theta,$$

$$\frac{\partial^3 m_{ni}(\gamma_0)}{\partial\gamma_k\partial\gamma_j\partial\gamma'} = -\begin{pmatrix} [G_{ni}^{(1j)'}e_{k_g,k-k_\theta},\ldots,G_{ni}^{(k_\theta,j)'}e_{k_g,k-k_\theta}] & G_{ni}^{(j)'}e_{k_g,t}g'_{ni}+G'_{ni}e_{k_g,t}g_{ni}^{(j)'}+g_{nit}^{(j)}G'_{ni}+g_{nit}G_{ni}^{(j)'} \\ g_{ni}e'_{k_g,t}G_{ni}^{(j)}+g_{ni}^{(j)}e'_{k_g,t}G_{ni}+g_{nit}^{(j)}G_{ni}+g_{nit}G_{ni}^{(j)} & -\rho_3 g_{nit}^{(j)}g_{ni}g'_{ni}-\rho_3 g_{nit}(g_{ni}^{(j)}g'_{ni}+g_{ni}g_{ni}^{(j)'}) \end{pmatrix}$$

for $1\le j\le k_\theta$, $k_\theta+1\le k\le k_\theta+k_g$, and $t=k-k_\theta$,

$$\frac{\partial^3 m_{ni}(\gamma_0)}{\partial\gamma_k\partial\gamma_j\partial\gamma'} = -\begin{pmatrix} [G_{ni}^{(1k)'}e_{k_g,j-k_\theta},\ldots,G_{ni}^{(k_\theta,k)'}e_{k_g,j-k_\theta}] & G_{ni}^{(k)'}e_{k_g,t}g'_{ni}+G'_{ni}e_{k_g,t}g_{ni}^{(k)'}+g_{nit}^{(k)}G'_{ni}+g_{nit}G_{ni}^{(k)'} \\ g_{ni}e'_{k_g,t}G_{ni}^{(k)}+g_{ni}^{(k)}e'_{k_g,t}G_{ni}+g_{nit}^{(k)}G_{ni}+g_{nit}G_{ni}^{(k)} & -\rho_3 g_{nit}^{(k)}g_{ni}g'_{ni}-\rho_3 g_{nit}(g_{ni}^{(k)}g'_{ni}+g_{ni}g_{ni}^{(k)'}) \end{pmatrix}$$

for $k_\theta+1\le j\le k_\theta+k_g$, $1\le k\le k_\theta$ and $t=j-k_\theta$, and

$$\frac{\partial^3 m_{ni}(\gamma_0)}{\partial\gamma_k\partial\gamma_j\partial\gamma'} = \begin{pmatrix} -G'_{ni}e_{k_g,s}e'_{k_g,t}G_{ni}-G'_{ni}e_{k_g,t}e'_{k_g,s}G_{ni} & \rho_3 g_{nis}G'_{ni}e_{k_g,t}g'_{ni}+\rho_3 g_{nit}G'_{ni}e_{k_g,s}g'_{ni}+\rho_3 g_{nis}g_{nit}G'_{ni} \\ \rho_3 g_{nis}g_{ni}e'_{k_g,t}G_{ni}+\rho_3 g_{nit}g_{ni}e'_{k_g,s}G_{ni}+\rho_3 g_{nis}g_{nit}G_{ni} & \rho_4 g_{nis}g_{nit}g_{ni}g'_{ni} \end{pmatrix}$$
$$-\begin{pmatrix} g_{nis}[G_{ni}^{(1)'}e_{k_g,t},\ldots,G_{ni}^{(k_\theta)'}e_{k_g,t}]+g_{nit}[G_{ni}^{(1)'}e_{k_g,s},\ldots,G_{ni}^{(k_\theta)'}e_{k_g,s}] & 0 \\ 0 & 0 \end{pmatrix},$$

for $k_\theta+1\le j\le k_\theta+k_g$, $k_\theta+1\le k\le k_\theta+k_g$, $s=j-k_\theta$ and $t=k-k_\theta$.

F Monte Carlo results for the SARAR(1,1) and SE models

In this section, we report Monte Carlo results for the SARAR(1,1) model (14), and the SE model, which is a special case of (14).

For the estimation of model (14), in the homoskedastic case, we use the moment vector $\frac{1}{n}[V_n'V_n-n\sigma_0^2, V_n'W_nV_n, V_n'W_n^2V_n-\sigma_0^2\operatorname{tr}(W_n^2), V_n'(X_n, W_nX_n^*, W_n^2X_n^*)]'$; in the heteroskedastic case, we use the moment vector $\frac{1}{n}[V_n'W_nV_n, V_n'(W_n^2-\operatorname{diag}(W_n^2))V_n, V_n'(X_n, W_nX_n^*, W_n^2X_n^*)]'$. The Monte Carlo results reported in Tables F.1–F.3 display patterns similar to those for the SARAR(2,0) model in the main text.

We consider tests of spatial error dependence in both the SE model and the SARAR(1,1) model. For the SE model, in the homoskedastic case, the parameter restriction tests "PT_{GMM}", "PT_{EL}" and "PT_{ET}" are based on the moment vector $\frac{1}{n}[V_n'V_n-n\sigma_0^2, V_n'W_nV_n, V_n'W_n^2V_n-\sigma_0^2\operatorname{tr}(W_n^2), V_n'(X_n, W_nX_n^*, W_n^2X_n^*)]'$; the overidentification tests "OT_{GMM}", "OT_{EL}" and "OT_{ET}" are based on the moment vector $\frac{1}{n}[V_n'W_nV_n, V_n'X_n]'$; the robust, EL and ET Moran's I tests are based on the moment $\frac{1}{n}V_n'W_nV_n$. For the latter three tests, OLS residuals are used to formulate test statistics. In the heteroskedastic case, the above tests are also considered, among which parameter restriction tests are based on the moment vector $\frac{1}{n}[V_n'W_nV_n, V_n'(W_n^2-\operatorname{diag}(W_n^2))V_n, V_n'(X_n, W_nX_n^*, W_n^2X_n^*)]'$ robust to unknown heteroskedasticity. In addition, we investigate the GMM parameter restriction test "PT^*_{GMM}" based on the moment vector $\frac{1}{n}[V_n'V_n-n\sigma_0^2, V_n'W_nV_n, V_n'W_n^2V_n-\sigma_0^2\operatorname{tr}(W_n^2), V_n'(X_n, W_nX_n^*, W_n^2X_n^*)]'$, and the conventional Moran's I test "Moran*", which do not take into account unknown heteroskedasticity.

Table F.4 presents empirical sizes of tests for $\tau_0 = 0$ in the SE model. $\mathrm{PT}_{\mathrm{EL}}$ and $\mathrm{PT}_{\mathrm{ET}}$ have relatively large sizes for small sample cases and sizes for the larger sample size $n = 400$ have improved. In the heteroskedastic case, $\mathrm{PT}^*_{\mathrm{GMM}}$ and Moran* have large size distortions and the distortions do not improve with the larger sample size $n = 400$. Other tests have relatively small size distortions. Powers of these tests except $\mathrm{PT}^*_{\mathrm{GMM}}$ and Moran* are presented in Table F.5. Their powers are generally similar for different valid tests, but are higher for the homoskedastic model than those of the heteroskedastic model. R^2 does not have much impact on powers. These tests are powerful in cases with a larger τ_0 and a larger sample size in the data generating process (DGP).

For the SARAR(1,1) model, parameter restriction tests are based on moment conditions similar to those for the SE model. Overidentification tests are based on the moment vector $\frac{1}{n}[V_n' W_n V_n, V_n'(W_n^2 - \mathrm{diag}(W_n^2))V_n, V_n'(X_n, W_n(I_n - \hat{\kappa}_n W_n)^{-1} X_n \hat{\beta}_n)]'$, where $\hat{\kappa}_n$ and $\hat{\beta}_n$ are the FOGMM estimator of the SAR model as described above. To compute Moran's I tests, we use the 2SLS estimator $\hat{\phi}_n$ of $\phi = (\kappa, \beta')'$ with the IV matrix $Q_n = [X_n, W_n X_n^*, W_n^2 X_n^*]$ for the SAR model. The test statistics employ the moment condition $g_n(\theta) = g_{1n}(\theta) - \frac{\partial g_{1n}(\theta)}{\partial \phi'}(\frac{\partial g_{2n}(\theta)}{\partial \phi'})^{-1} g_{2n}(\theta)$, where $g_{1n}(\theta) = \frac{1}{n} V_n'(\theta) W_n V_n(\theta)$ and $g_{2n}(\theta) = \frac{1}{n} Z_n' Q_n (Q_n' Q_n)^{-1} Q_n' V_n(\theta)$ with $Z_n = [W_n Y_n, X_n]$. Thus $g_{2n}(\hat{\theta}_n) = 0$, where $\hat{\theta}_n = (0, \hat{\phi}_n')'$.

Test results on $\tau_0 = 0$ in the SARAR(1,1) model (14) are reported in Tables F.6 and F.7. When $n = 144$, the size distortions of parameter restriction tests are larger than those of overidentification tests, and those of Moran's I tests are smallest; when $n = 400$, all sizes are generally close to the nominal 5%. Different versions of parameter restriction tests have similar powers. So are different versions of overidentification tests and those of Moran's I tests. Parameter restriction tests are more powerful than overidentification tests, and the latter ones are generally more powerful than Moran's I tests. With larger R^2, sample sizes, and τ_0 in the DGP, all tests tend to be more powerful.

References

Davidson, J., 1994. Stochastic limit theory: An introduction for econometricians. Oxford University Press, USA.

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

Liu, X., Lee, L.F., Bollinger, C.R., 2010. An efficient GMM estimator of spatial autoregressive models. Journal of Econometrics 159, 303–319.

Newey, W.K., Smith, R.J., 2004. Higher order properties of GMM and generalized empirical likelihood estimators. Econometrica 72, 219–255.

Table F.1: Biases, standard errors and RMSEs of estimators for the SARAR model (14) in the homoskedastic case

R^2, κ_0, τ_0		κ	τ	β_1	β_2	β_3	σ^2
	$n = 144$						
0.8, 0.2, 0.2	GMM	-0.003[0.080]0.080	-0.010[0.131]0.131	0.009[0.157]0.157	-0.003[0.032]0.032	-0.001[0.031]0.031	-0.011[0.015]0.019
	EL	-0.003[0.079]0.079	-0.013[0.127]0.127	0.010[0.154]0.154	-0.003[0.032]0.032	-0.001[0.031]0.031	-0.007[0.015]0.016
	ET	-0.003[0.078]0.078	-0.013[0.126]0.127	0.009[0.153]0.154	-0.003[0.032]0.032	-0.001[0.031]0.031	-0.009[0.015]0.018
0.8, 0.2, 0.4	GMM	-0.002[0.081]0.082	-0.010[0.119]0.120	0.008[0.169]0.170	-0.002[0.031]0.031	-0.002[0.031]0.031	-0.012[0.015]0.019
	EL	0.000[0.079]0.079	-0.019[0.115]0.117	0.003[0.165]0.165	-0.001[0.031]0.031	-0.002[0.031]0.031	-0.007[0.015]0.016
	ET	0.000[0.079]0.079	-0.018[0.115]0.117	0.003[0.165]0.165	-0.001[0.031]0.031	-0.002[0.031]0.031	-0.009[0.015]0.018
0.8, 0.4, 0.2	GMM	-0.005[0.069]0.069	-0.009[0.127]0.127	0.013[0.172]0.173	-0.001[0.031]0.031	-0.000[0.030]0.030	-0.012[0.015]0.019
	EL	-0.006[0.067]0.068	-0.011[0.122]0.122	0.016[0.169]0.169	-0.001[0.031]0.031	-0.001[0.030]0.030	-0.007[0.015]0.017
	ET	-0.006[0.067]0.067	-0.010[0.122]0.122	0.016[0.168]0.169	-0.001[0.031]0.031	-0.001[0.030]0.030	-0.010[0.015]0.018
0.8, 0.4, 0.4	GMM	-0.006[0.076]0.076	-0.011[0.121]0.122	0.016[0.198]0.199	-0.002[0.030]0.030	0.000[0.032]0.032	-0.011[0.015]0.019
	EL	-0.005[0.075]0.075	-0.017[0.118]0.119	0.013[0.195]0.196	-0.002[0.030]0.030	0.000[0.032]0.032	-0.007[0.015]0.016
	ET	-0.005[0.075]0.075	-0.016[0.118]0.119	0.014[0.195]0.196	-0.002[0.030]0.030	0.000[0.032]0.032	-0.009[0.015]0.018
0.4, 0.2, 0.2	GMM	-0.005[0.177]0.177	-0.021[0.211]0.212	0.018[0.357]0.358	-0.006[0.077]0.078	-0.004[0.076]0.076	-0.081[0.089]0.120
	EL	-0.002[0.160]0.160	-0.025[0.191]0.193	0.008[0.327]0.327	-0.004[0.077]0.077	-0.002[0.076]0.076	-0.050[0.089]0.102
	ET	-0.002[0.160]0.160	-0.024[0.191]0.192	0.008[0.328]0.328	-0.004[0.078]0.078	-0.001[0.076]0.076	-0.066[0.088]0.110
0.4, 0.2, 0.4	GMM	0.000[0.185]0.185	-0.026[0.200]0.201	0.015[0.394]0.394	-0.001[0.079]0.079	-0.005[0.077]0.077	-0.081[0.093]0.123
	EL	0.008[0.169]0.169	-0.035[0.184]0.187	-0.004[0.365]0.365	0.001[0.079]0.079	-0.003[0.077]0.077	-0.050[0.092]0.105
	ET	0.009[0.169]0.169	-0.035[0.183]0.187	-0.004[0.366]0.366	0.002[0.079]0.079	-0.003[0.077]0.077	-0.066[0.092]0.113
0.4, 0.4, 0.2	GMM	-0.017[0.175]0.176	-0.019[0.208]0.209	0.048[0.446]0.448	-0.005[0.081]0.081	-0.003[0.078]0.078	-0.079[0.094]0.123
	EL	-0.023[0.164]0.165	-0.012[0.193]0.193	0.058[0.419]0.423	-0.003[0.080]0.080	-0.001[0.077]0.077	-0.049[0.094]0.106
	ET	-0.023[0.161]0.163	-0.011[0.191]0.191	0.057[0.415]0.419	-0.003[0.080]0.080	-0.001[0.077]0.077	-0.065[0.093]0.114
0.4, 0.4, 0.4	GMM	-0.021[0.184]0.186	-0.021[0.197]0.198	0.054[0.478]0.481	-0.005[0.076]0.076	-0.007[0.076]0.077	-0.082[0.093]0.124
	EL	-0.019[0.169]0.170	-0.023[0.180]0.181	0.048[0.443]0.446	-0.003[0.076]0.076	-0.005[0.077]0.077	-0.052[0.093]0.107
	ET	-0.019[0.169]0.170	-0.023[0.179]0.181	0.047[0.442]0.444	-0.003[0.076]0.076	-0.005[0.076]0.077	-0.068[0.092]0.114
	$n = 400$						
0.8, 0.2, 0.2	GMM	-0.000[0.043]0.043	-0.004[0.069]0.069	0.001[0.086]0.086	-0.000[0.017]0.017	-0.001[0.018]0.018	-0.004[0.009]0.010
	EL	-0.000[0.043]0.043	-0.004[0.068]0.069	0.001[0.086]0.086	-0.000[0.017]0.017	-0.001[0.018]0.018	-0.002[0.009]0.009
	ET	-0.000[0.043]0.043	-0.004[0.068]0.068	0.001[0.086]0.086	-0.000[0.017]0.017	-0.001[0.018]0.018	-0.003[0.009]0.010
0.8, 0.2, 0.4	GMM	-0.002[0.048]0.048	-0.002[0.065]0.065	0.004[0.098]0.098	-0.001[0.019]0.019	-0.001[0.018]0.018	-0.004[0.009]0.010
	EL	-0.001[0.048]0.048	-0.004[0.065]0.065	0.003[0.097]0.098	-0.001[0.019]0.019	-0.000[0.018]0.018	-0.002[0.009]0.009
	ET	-0.001[0.048]0.048	-0.004[0.065]0.065	0.003[0.097]0.097	-0.001[0.019]0.019	-0.000[0.018]0.018	-0.003[0.009]0.010
0.8, 0.4, 0.2	GMM	-0.002[0.040]0.040	-0.003[0.073]0.073	0.006[0.099]0.099	-0.000[0.019]0.019	-0.001[0.018]0.018	-0.004[0.009]0.010
	EL	-0.003[0.039]0.040	-0.003[0.072]0.072	0.007[0.098]0.098	-0.000[0.019]0.019	-0.001[0.018]0.018	-0.002[0.009]0.009
	ET	-0.003[0.039]0.039	-0.003[0.072]0.072	0.007[0.098]0.098	-0.000[0.019]0.019	-0.001[0.018]0.018	-0.003[0.009]0.009
0.8, 0.4, 0.4	GMM	-0.003[0.045]0.045	-0.005[0.070]0.070	0.008[0.116]0.117	-0.000[0.017]0.017	-0.001[0.018]0.018	-0.004[0.009]0.010
	EL	-0.003[0.045]0.045	-0.006[0.069]0.070	0.008[0.116]0.116	-0.000[0.017]0.017	-0.001[0.018]0.018	-0.002[0.009]0.009
	ET	-0.003[0.045]0.045	-0.006[0.069]0.070	0.008[0.116]0.116	-0.000[0.017]0.017	-0.001[0.018]0.018	-0.003[0.009]0.009
0.4, 0.2, 0.2	GMM	-0.001[0.106]0.106	-0.008[0.118]0.118	0.006[0.209]0.209	-0.002[0.043]0.044	-0.002[0.047]0.047	-0.031[0.054]0.063
	EL	-0.000[0.101]0.101	-0.009[0.113]0.113	0.005[0.202]0.202	-0.001[0.043]0.043	-0.001[0.047]0.047	-0.017[0.054]0.057
	ET	-0.000[0.101]0.101	-0.009[0.113]0.113	0.005[0.202]0.202	-0.001[0.043]0.043	-0.001[0.046]0.046	-0.024[0.054]0.059
0.4, 0.2, 0.4	GMM	-0.004[0.108]0.108	-0.007[0.110]0.110	0.011[0.232]0.233	-0.002[0.044]0.044	-0.002[0.043]0.043	-0.032[0.055]0.064
	EL	-0.000[0.105]0.105	-0.012[0.107]0.108	0.004[0.227]0.227	-0.001[0.044]0.044	-0.002[0.043]0.043	-0.017[0.055]0.058
	ET	0.000[0.104]0.104	-0.012[0.106]0.107	0.003[0.226]0.226	-0.001[0.044]0.044	-0.002[0.043]0.043	-0.024[0.055]0.060
0.4, 0.4, 0.2	GMM	-0.009[0.094]0.094	-0.001[0.117]0.117	0.016[0.234]0.234	0.001[0.045]0.045	0.001[0.045]0.045	-0.029[0.054]0.062
	EL	-0.011[0.091]0.092	0.002[0.113]0.113	0.020[0.228]0.229	0.001[0.045]0.045	0.001[0.045]0.045	-0.015[0.054]0.056
	ET	-0.011[0.091]0.091	0.002[0.113]0.113	0.020[0.227]0.228	0.001[0.045]0.045	0.001[0.045]0.045	-0.022[0.054]0.058
0.4, 0.4, 0.4	GMM	-0.003[0.106]0.106	-0.011[0.121]0.121	0.013[0.277]0.278	-0.002[0.043]0.043	-0.003[0.045]0.045	-0.033[0.055]0.064
	EL	-0.003[0.102]0.102	-0.012[0.116]0.117	0.013[0.267]0.268	-0.001[0.043]0.043	-0.003[0.045]0.045	-0.019[0.055]0.058
	ET	-0.003[0.102]0.102	-0.012[0.116]0.116	0.012[0.267]0.267	-0.001[0.043]0.043	-0.002[0.045]0.045	-0.026[0.054]0.060

$\beta_0 = [0.5, 0.5, 0.5]'$.

Table F.2: Biases, standard errors and RMSEs of estimators for the SARAR model (14) in the heteroskedastic case

R^2, κ_0, τ_0		κ	τ	β_1	β_2	β_3
	$n = 144$					
0.8, 0.2, 0.2	GMM	-0.001[0.048]0.048	0.001[0.229]0.229	0.005[0.108]0.108	-0.001[0.022]0.022	-0.001[0.022]0.022
	EL	-0.001[0.048]0.048	-0.025[0.235]0.237	0.006[0.110]0.110	-0.002[0.023]0.023	-0.001[0.022]0.022
	ET	-0.001[0.046]0.046	-0.022[0.229]0.230	0.005[0.106]0.106	-0.001[0.022]0.022	-0.001[0.021]0.021
0.8, 0.2, 0.4	GMM	-0.001[0.051]0.051	0.013[0.232]0.232	0.004[0.131]0.131	-0.000[0.023]0.023	-0.001[0.022]0.022
	EL	0.001[0.051]0.051	-0.021[0.258]0.259	0.001[0.127]0.127	-0.000[0.023]0.023	-0.000[0.023]0.023
	ET	0.000[0.050]0.050	-0.018[0.245]0.246	0.001[0.123]0.123	-0.000[0.022]0.022	-0.000[0.022]0.022
0.8, 0.4, 0.2	GMM	-0.001[0.047]0.047	0.000[0.240]0.240	0.005[0.130]0.130	-0.001[0.022]0.022	-0.001[0.021]0.021
	EL	-0.001[0.041]0.041	-0.032[0.234]0.236	0.005[0.116]0.116	-0.001[0.021]0.021	-0.000[0.021]0.021
	ET	-0.001[0.041]0.041	-0.029[0.229]0.230	0.004[0.114]0.114	-0.001[0.020]0.020	-0.000[0.020]0.020
0.8, 0.4, 0.4	GMM	-0.003[0.050]0.050	-0.004[0.230]0.230	-0.001[0.270]0.270	-0.001[0.021]0.021	-0.001[0.021]0.021
	EL	-0.003[0.049]0.049	-0.043[0.238]0.242	0.008[0.146]0.146	-0.001[0.021]0.021	-0.001[0.021]0.021
	ET	-0.003[0.047]0.047	-0.036[0.224]0.227	0.010[0.178]0.179	-0.001[0.020]0.020	-0.001[0.020]0.020
0.4, 0.2, 0.2	GMM	-0.001[0.121]0.121	-0.002[0.277]0.277	0.009[0.370]0.370	-0.005[0.054]0.054	-0.005[0.055]0.055
	EL	0.001[0.114]0.114	-0.035[0.261]0.263	0.010[0.263]0.263	-0.004[0.053]0.053	-0.004[0.053]0.053
	ET	0.001[0.109]0.109	-0.030[0.255]0.256	0.010[0.253]0.253	-0.004[0.051]0.051	-0.004[0.051]0.051
0.4, 0.2, 0.4	GMM	-0.005[0.134]0.134	0.003[0.269]0.269	0.031[0.702]0.703	-0.005[0.059]0.059	-0.005[0.058]0.059
	EL	0.000[0.124]0.124	-0.041[0.248]0.251	0.006[0.297]0.297	-0.003[0.057]0.057	-0.003[0.057]0.057
	ET	0.001[0.121]0.121	-0.033[0.258]0.260	0.005[0.288]0.288	-0.003[0.056]0.056	-0.003[0.055]0.055
0.4, 0.4, 0.2	GMM	-0.011[0.122]0.122	0.017[0.279]0.279	0.034[0.357]0.358	-0.004[0.055]0.055	-0.003[0.056]0.056
	EL	-0.011[0.111]0.112	-0.012[0.263]0.263	0.035[0.333]0.335	-0.002[0.054]0.054	-0.002[0.055]0.055
	ET	-0.010[0.106]0.106	-0.009[0.254]0.254	0.031[0.310]0.312	-0.002[0.051]0.051	-0.001[0.053]0.053
0.4, 0.4, 0.4	GMM	-0.009[0.132]0.133	-0.001[0.255]0.255	0.028[0.460]0.461	-0.006[0.056]0.056	-0.004[0.054]0.054
	EL	-0.003[0.119]0.119	-0.040[0.244]0.247	0.012[0.351]0.351	-0.004[0.054]0.054	-0.002[0.052]0.052
	ET	-0.003[0.116]0.116	-0.035[0.238]0.240	0.011[0.344]0.345	-0.004[0.053]0.053	-0.002[0.051]0.051
	$n = 400$					
0.8, 0.2, 0.2	GMM	-0.000[0.025]0.025	0.002[0.127]0.127	0.001[0.060]0.060	-0.000[0.012]0.012	-0.000[0.012]0.012
	EL	-0.000[0.025]0.025	-0.012[0.127]0.128	0.002[0.060]0.060	-0.000[0.012]0.012	-0.000[0.012]0.012
	ET	-0.000[0.025]0.025	-0.011[0.125]0.126	0.001[0.059]0.059	-0.000[0.012]0.012	-0.000[0.012]0.012
0.8, 0.2, 0.4	GMM	0.002[0.029]0.029	-0.004[0.114]0.114	-0.004[0.071]0.072	0.001[0.013]0.013	0.000[0.013]0.013
	EL	0.002[0.029]0.029	-0.018[0.115]0.116	-0.005[0.072]0.072	0.001[0.013]0.013	0.000[0.013]0.013
	ET	0.002[0.029]0.029	-0.017[0.113]0.114	-0.005[0.071]0.071	0.001[0.013]0.013	0.000[0.013]0.013
0.8, 0.4, 0.2	GMM	-0.001[0.024]0.024	0.004[0.120]0.120	0.005[0.066]0.066	-0.001[0.012]0.012	-0.001[0.011]0.011
	EL	-0.001[0.024]0.024	-0.009[0.119]0.119	0.005[0.067]0.067	-0.001[0.012]0.012	-0.001[0.011]0.011
	ET	-0.001[0.024]0.024	-0.008[0.117]0.118	0.005[0.065]0.065	-0.001[0.012]0.012	-0.001[0.011]0.011
0.8, 0.4, 0.4	GMM	-0.001[0.027]0.027	-0.001[0.115]0.115	0.003[0.082]0.082	0.000[0.012]0.012	-0.000[0.012]0.012
	EL	-0.001[0.027]0.027	-0.015[0.116]0.116	0.004[0.080]0.080	-0.000[0.012]0.012	-0.000[0.012]0.012
	ET	-0.001[0.026]0.026	-0.014[0.113]0.114	0.004[0.078]0.078	-0.000[0.012]0.012	-0.000[0.012]0.012
0.4, 0.2, 0.2	GMM	0.002[0.062]0.062	-0.002[0.140]0.140	-0.003[0.145]0.145	-0.001[0.029]0.029	-0.002[0.029]0.029
	EL	0.002[0.063]0.063	-0.015[0.141]0.142	-0.003[0.147]0.147	-0.001[0.029]0.029	-0.002[0.029]0.029
	ET	0.003[0.062]0.062	-0.014[0.139]0.139	-0.003[0.144]0.144	-0.001[0.029]0.029	-0.001[0.029]0.029
0.4, 0.2, 0.4	GMM	-0.002[0.069]0.069	0.001[0.129]0.129	0.007[0.174]0.174	-0.001[0.031]0.031	-0.002[0.030]0.030
	EL	-0.001[0.069]0.069	-0.015[0.123]0.124	0.004[0.175]0.175	-0.000[0.031]0.031	-0.001[0.031]0.031
	ET	-0.001[0.068]0.068	-0.014[0.121]0.122	0.004[0.172]0.172	-0.000[0.031]0.031	-0.001[0.030]0.030
0.4, 0.4, 0.2	GMM	-0.005[0.060]0.060	0.014[0.138]0.139	0.015[0.175]0.176	-0.001[0.028]0.029	-0.002[0.029]0.029
	EL	-0.005[0.059]0.059	0.001[0.136]0.136	0.016[0.166]0.167	-0.001[0.028]0.028	-0.002[0.029]0.029
	ET	-0.005[0.058]0.058	0.002[0.134]0.134	0.015[0.162]0.163	-0.001[0.028]0.028	-0.002[0.028]0.028
0.4, 0.4, 0.4	GMM	-0.000[0.067]0.067	-0.003[0.129]0.129	0.006[0.200]0.200	-0.000[0.030]0.030	-0.001[0.029]0.029
	EL	-0.001[0.068]0.068	-0.017[0.129]0.130	0.007[0.202]0.202	-0.000[0.030]0.030	-0.001[0.030]0.030
	ET	-0.001[0.067]0.067	-0.016[0.127]0.128	0.006[0.197]0.197	-0.000[0.029]0.029	-0.001[0.029]0.029

$\beta_0 = [0.5, 0.5, 0.5]'$.

Table F.3: Coverage probabilities of 95% confidence intervals for the SARAR model (14)

		Homoskedastic case						Heteroskedastic case				
R^2, κ_0, τ_0		κ	τ	β_1	β_2	β_3	σ^2	κ	τ	β_1	β_2	β_3
	$n = 144$											
0.8, 0.2, 0.2	GMM	0.893	0.893	0.902	0.910	0.926	0.769	0.921	0.880	0.921	0.906	0.924
	EL	0.912	0.917	0.915	0.926	0.935	0.873	0.929	0.922	0.929	0.923	0.940
	ET	0.923	0.925	0.927	0.935	0.947	0.828	0.933	0.931	0.935	0.939	0.949
0.8, 0.2, 0.4	GMM	0.920	0.917	0.925	0.917	0.920	0.787	0.933	0.856	0.932	0.926	0.920
	EL	0.932	0.941	0.928	0.921	0.932	0.886	0.938	0.907	0.930	0.937	0.932
	ET	0.938	0.952	0.939	0.936	0.938	0.848	0.947	0.919	0.940	0.949	0.938
0.8, 0.4, 0.2	GMM	0.908	0.907	0.899	0.926	0.923	0.781	0.925	0.873	0.927	0.932	0.937
	EL	0.922	0.920	0.922	0.938	0.935	0.882	0.941	0.904	0.946	0.946	0.957
	ET	0.932	0.930	0.926	0.946	0.945	0.838	0.947	0.921	0.947	0.954	0.963
0.8, 0.4, 0.4	GMM	0.916	0.893	0.923	0.919	0.914	0.793	0.934	0.848	0.934	0.951	0.931
	EL	0.930	0.920	0.934	0.932	0.930	0.885	0.935	0.911	0.935	0.958	0.938
	ET	0.937	0.929	0.941	0.946	0.936	0.847	0.949	0.924	0.945	0.964	0.949
0.4, 0.2, 0.2	GMM	0.848	0.833	0.861	0.917	0.923	0.750	0.919	0.872	0.902	0.925	0.927
	EL	0.878	0.880	0.896	0.928	0.939	0.866	0.934	0.907	0.920	0.945	0.940
	ET	0.885	0.889	0.907	0.935	0.944	0.817	0.948	0.919	0.929	0.950	0.949
0.4, 0.2, 0.4	GMM	0.841	0.840	0.856	0.896	0.919	0.751	0.911	0.841	0.929	0.916	0.927
	EL	0.875	0.878	0.893	0.915	0.934	0.854	0.928	0.897	0.931	0.935	0.943
	ET	0.880	0.892	0.902	0.929	0.939	0.825	0.936	0.905	0.944	0.947	0.953
0.4, 0.4, 0.2	GMM	0.864	0.850	0.870	0.912	0.912	0.776	0.909	0.851	0.914	0.913	0.913
	EL	0.897	0.867	0.891	0.932	0.925	0.890	0.928	0.901	0.920	0.946	0.941
	ET	0.911	0.887	0.910	0.940	0.930	0.849	0.936	0.912	0.929	0.952	0.950
0.4, 0.4, 0.4	GMM	0.836	0.838	0.851	0.916	0.918	0.765	0.912	0.869	0.922	0.902	0.913
	EL	0.872	0.866	0.874	0.929	0.928	0.856	0.925	0.906	0.924	0.922	0.927
	ET	0.881	0.880	0.887	0.937	0.937	0.821	0.937	0.923	0.938	0.935	0.938
	$n = 400$											
0.8, 0.2, 0.2	GMM	0.945	0.932	0.941	0.920	0.951	0.886	0.942	0.916	0.924	0.937	0.948
	EL	0.943	0.934	0.944	0.921	0.955	0.932	0.937	0.922	0.929	0.934	0.946
	ET	0.948	0.942	0.948	0.927	0.958	0.917	0.941	0.929	0.934	0.940	0.950
0.8, 0.2, 0.4	GMM	0.928	0.945	0.934	0.958	0.936	0.887	0.943	0.929	0.940	0.950	0.942
	EL	0.933	0.947	0.936	0.957	0.941	0.938	0.931	0.938	0.941	0.949	0.938
	ET	0.934	0.949	0.941	0.964	0.944	0.911	0.939	0.946	0.944	0.954	0.942
0.8, 0.4, 0.2	GMM	0.948	0.942	0.950	0.940	0.947	0.868	0.942	0.914	0.957	0.949	0.951
	EL	0.953	0.955	0.946	0.940	0.950	0.921	0.942	0.933	0.954	0.949	0.952
	ET	0.955	0.956	0.949	0.946	0.956	0.893	0.943	0.936	0.957	0.951	0.958
0.8, 0.4, 0.4	GMM	0.937	0.920	0.940	0.934	0.950	0.888	0.946	0.931	0.939	0.955	0.956
	EL	0.943	0.929	0.939	0.938	0.951	0.937	0.939	0.942	0.928	0.952	0.950
	ET	0.943	0.933	0.943	0.941	0.953	0.916	0.945	0.945	0.935	0.958	0.957
0.4, 0.2, 0.2	GMM	0.922	0.928	0.922	0.932	0.937	0.880	0.951	0.923	0.943	0.937	0.952
	EL	0.931	0.934	0.929	0.934	0.942	0.924	0.945	0.932	0.937	0.940	0.950
	ET	0.933	0.936	0.934	0.939	0.941	0.909	0.951	0.940	0.944	0.942	0.956
0.4, 0.2, 0.4	GMM	0.911	0.908	0.904	0.937	0.941	0.886	0.948	0.929	0.946	0.935	0.947
	EL	0.922	0.920	0.923	0.938	0.944	0.923	0.947	0.941	0.945	0.933	0.944
	ET	0.925	0.922	0.927	0.944	0.948	0.904	0.952	0.945	0.948	0.935	0.947
0.4, 0.4, 0.2	GMM	0.917	0.909	0.934	0.933	0.939	0.876	0.950	0.934	0.958	0.940	0.961
	EL	0.926	0.917	0.942	0.936	0.942	0.927	0.942	0.944	0.950	0.939	0.966
	ET	0.930	0.920	0.945	0.937	0.945	0.902	0.944	0.949	0.960	0.940	0.967
0.4, 0.4, 0.4	GMM	0.906	0.898	0.907	0.932	0.944	0.868	0.940	0.938	0.943	0.944	0.951
	EL	0.910	0.917	0.915	0.936	0.947	0.917	0.934	0.943	0.935	0.946	0.945
	ET	0.914	0.922	0.920	0.940	0.950	0.897	0.946	0.948	0.941	0.952	0.949

The variance matrix of a GMM estimator $\hat{\theta}_n$ is computed as $\frac{1}{n}[G_n'(\hat{\theta}_n)\Omega_n^{-1}(\hat{\theta}_n)G_n(\hat{\theta}_n)]^{-1}$, and that of a GEL estimator $\dot{\gamma}_n = (\dot{\theta}_n', \dot{\lambda}_n')'$ is computed as $\frac{1}{n}\Delta_n^{-1}(\dot{\gamma}_n)\begin{pmatrix}0 & 0\\ 0 & \Omega_n(\dot{\theta}_n)\end{pmatrix}\Delta_n^{-1}(\dot{\gamma}_n)$, where $\Delta_n(\gamma)$ is the second order derivative matrix of the GEL objective function given in the proof of Theorem 3.2.

Table F.4: Empirical sizes of tests for $\tau_0 = 0$ in an SE model

	Homoskedastic case				Heteroskedastic case			
	$n = 144$		$n = 400$		$n = 144$		$n = 400$	
	$R^2 = 0.8$	$R^2 = 0.4$	$R^2 = 0.8$	$R^2 = 0.4$	$R^2 = 0.8$	$R^2 = 0.4$	$R^2 = 0.8$	$R^2 = 0.4$
PT_{GMM}	0.034	0.061	0.051	0.050	0.057	0.068	0.053	0.058
PT_{EL}	0.072	0.107	0.061	0.061	0.113	0.144	0.082	0.078
PT_{ET}	0.065	0.092	0.063	0.060	0.094	0.122	0.080	0.074
OT_{GMM}	0.043	0.050	0.053	0.042	0.040	0.050	0.052	0.050
OT_{EL}	0.049	0.063	0.054	0.044	0.058	0.073	0.057	0.059
OT_{ET}	0.050	0.059	0.054	0.044	0.050	0.067	0.058	0.057
Moran	0.043	0.050	0.054	0.043	0.041	0.052	0.055	0.050
Moran_{EL}	0.049	0.066	0.054	0.044	0.055	0.063	0.058	0.055
Moran_{ET}	0.052	0.060	0.054	0.044	0.047	0.064	0.058	0.058
PT^*_{GMM}					0.105	0.107	0.156	0.175
Moran^*					0.011	0.013	0.020	0.018

"PT_{GMM}", "PT_{EL}" and "PT_{ET}" denote, respectively, the GMM, EL and ET parameter restriction tests; "OT_{GMM}", "OT_{EL}" and "OT_{ET}" denote, respectively, the GMM, EL and ET overidentification tests; "Moran", "Moran_{EL} " and "Moran_{ET} " denote, respectively, the robust, EL and ET Moran's I tests; "PT^*_{GMM}" denotes the GMM parameter restriction test without taking into account unknown heteroskedasticity; and "Moran^*" denotes the conventional Moran's I test that does not take into account unknown heteroskedasticity. The nominal size is 5%.

Table F.5: Powers of tests for $\tau_0 = 0$ in an SE model

		$n = 144$			$n = 400$		
		$\tau_0 = 0.2$	$\tau_0 = 0.4$	$\tau_0 = 0.6$	$\tau_0 = 0.2$	$\tau_0 = 0.4$	$\tau_0 = 0.6$
Homoskedastic case							
$R^2 = 0.8$	PT_{GMM}	0.441	0.952	1.000	0.901	1.000	1.000
	PT_{EL}	0.551	0.982	1.000	0.915	1.000	1.000
	PT_{ET}	0.537	0.980	1.000	0.910	1.000	1.000
	OT_{GMM}	0.474	0.977	1.000	0.913	1.000	1.000
	OT_{EL}	0.500	0.984	1.000	0.913	1.000	1.000
	OT_{ET}	0.503	0.985	1.000	0.915	1.000	1.000
	Moran	0.471	0.978	1.000	0.912	1.000	1.000
	Moran_{EL}	0.497	0.982	1.000	0.914	1.000	1.000
	Moran_{ET}	0.505	0.983	1.000	0.915	1.000	1.000
$R^2 = 0.4$	PT_{GMM}	0.429	0.959	0.999	0.911	1.000	1.000
	PT_{EL}	0.546	0.985	1.000	0.921	1.000	1.000
	PT_{ET}	0.517	0.982	1.000	0.926	1.000	1.000
	OT_{GMM}	0.461	0.973	1.000	0.918	1.000	1.000
	OT_{EL}	0.493	0.977	1.000	0.922	1.000	1.000
	OT_{ET}	0.488	0.975	1.000	0.922	1.000	1.000
	Moran	0.461	0.970	1.000	0.918	1.000	1.000
	Moran_{EL}	0.492	0.977	1.000	0.922	1.000	1.000
	Moran_{ET}	0.488	0.976	1.000	0.923	1.000	1.000
Heteroskedastic case							
$R^2 = 0.8$	PT_{GMM}	0.196	0.550	0.915	0.348	0.932	1.000
	PT_{EL}	0.270	0.631	0.927	0.384	0.932	1.000
	PT_{ET}	0.245	0.625	0.933	0.385	0.939	1.000
	OT_{GMM}	0.127	0.480	0.881	0.296	0.904	1.000
	OT_{EL}	0.133	0.477	0.874	0.294	0.887	1.000
	OT_{ET}	0.133	0.498	0.886	0.304	0.904	1.000
	Moran	0.122	0.457	0.873	0.285	0.898	1.000
	Moran_{EL}	0.131	0.469	0.872	0.287	0.887	1.000
	Moran_{ET}	0.136	0.489	0.885	0.302	0.901	1.000
$R^2 = 0.4$	PT_{GMM}	0.186	0.586	0.930	0.371	0.916	1.000
	PT_{EL}	0.267	0.654	0.944	0.394	0.926	1.000
	PT_{ET}	0.249	0.651	0.952	0.399	0.927	1.000
	OT_{GMM}	0.119	0.496	0.903	0.294	0.869	0.999
	OT_{EL}	0.134	0.489	0.887	0.297	0.864	0.999
	OT_{ET}	0.134	0.512	0.901	0.300	0.873	0.999
	Moran	0.114	0.476	0.900	0.289	0.865	0.999
	Moran_{EL}	0.131	0.479	0.885	0.295	0.861	0.999
	Moran_{ET}	0.131	0.502	0.902	0.299	0.871	0.999

"PT_{GMM}", "PT_{EL}" and "PT_{ET}" denote, respectively, the GMM, EL and ET parameter restriction tests; "OT_{GMM}", "OT_{EL}" and "OT_{ET}" denote, respectively, the GMM, EL and ET overidentification tests; and "Moran", "Moran_{EL} " and "Moran_{ET} " denote, respectively, the robust, EL and ET Moran's I tests.

Table F.6: Empirical sizes of tests for $\tau_0 = 0$ in the SARAR model (14)

	$n = 144$				$n = 400$			
	$R^2 = 0.8$		$R^2 = 0.4$		$R^2 = 0.8$		$R^2 = 0.4$	
	$\kappa_0 = 0.2$	$\kappa_0 = 0.4$	$\kappa_0 = 0.2$	$\kappa_0 = 0.4$	$\kappa_0 = 0.2$	$\kappa_0 = 0.4$	$\kappa_0 = 0.2$	$\kappa_0 = 0.4$
Homoskedastic case								
PT_{GMM}	0.049	0.049	0.043	0.048	0.055	0.047	0.048	0.045
PT_{EL}	0.083	0.089	0.067	0.066	0.065	0.050	0.058	0.053
PT_{ET}	0.072	0.079	0.057	0.064	0.068	0.054	0.057	0.056
OT_{GMM}	0.047	0.048	0.048	0.052	0.052	0.050	0.050	0.053
OT_{EL}	0.071	0.070	0.066	0.075	0.065	0.055	0.056	0.061
OT_{ET}	0.068	0.070	0.063	0.066	0.061	0.055	0.057	0.061
Moran	0.054	0.047	0.041	0.052	0.051	0.048	0.059	0.043
$Moran_{EL}$	0.060	0.059	0.054	0.060	0.056	0.052	0.061	0.046
$Moran_{ET}$	0.059	0.059	0.053	0.060	0.056	0.053	0.062	0.046
Heteroskedastic case								
PT_{GMM}	0.054	0.069	0.092	0.067	0.052	0.042	0.063	0.051
PT_{EL}	0.105	0.123	0.127	0.125	0.068	0.060	0.076	0.073
PT_{ET}	0.093	0.103	0.117	0.104	0.062	0.056	0.071	0.069
OT_{GMM}	0.032	0.049	0.040	0.035	0.056	0.045	0.055	0.049
OT_{EL}	0.075	0.093	0.073	0.086	0.065	0.055	0.078	0.064
OT_{ET}	0.064	0.081	0.070	0.075	0.064	0.056	0.072	0.066
Moran	0.043	0.058	0.040	0.048	0.047	0.037	0.048	0.053
$Moran_{EL}$	0.061	0.079	0.063	0.060	0.056	0.044	0.061	0.061
$Moran_{ET}$	0.056	0.071	0.056	0.058	0.055	0.041	0.059	0.062

"PT_{GMM}", "PT_{EL}" and "PT_{ET}" denote, respectively, the GMM, EL and ET parameter restriction tests; "OT_{GMM}", "OT_{EL}" and "OT_{ET}" denote, respectively, the GMM, EL and ET overidentification tests; and "Moran", "$Moran_{EL}$ " and "$Moran_{ET}$ " denote, respectively, the robust, EL and ET Moran's I tests. The nominal size is 5%.

Table F.7: Powers of tests for $\tau_0 = 0$ in the SARAR model (14)

		$n = 144$, $\kappa_0 = 0.2$			$n = 144$, $\kappa_0 = 0.4$			$n = 400$, $\kappa_0 = 0.2$			$n = 400$, $\kappa_0 = 0.4$		
		$\tau_0 = 0.2$	$\tau_0 = 0.4$	$\tau_0 = 0.6$	$\tau_0 = 0.2$	$\tau_0 = 0.4$	$\tau_0 = 0.6$	$\tau_0 = 0.2$	$\tau_0 = 0.4$	$\tau_0 = 0.6$	$\tau_0 = 0.2$	$\tau_0 = 0.4$	$\tau_0 = 0.6$
Homoskedastic case													
$R^2 = 0.8$	PT_{GMM}	0.301	0.825	0.986	0.297	0.819	0.985	0.758	0.998	1.000	0.735	1.000	1.000
	PT_{EL}	0.399	0.895	0.997	0.390	0.881	0.998	0.790	0.999	1.000	0.777	1.000	1.000
	PT_{ET}	0.392	0.883	0.996	0.381	0.874	0.998	0.791	0.999	1.000	0.778	1.000	1.000
	OT_{GMM}	0.207	0.762	0.979	0.238	0.755	0.989	0.668	0.998	1.000	0.681	0.999	1.000
	OT_{EL}	0.276	0.811	0.986	0.299	0.809	0.991	0.682	0.999	1.000	0.702	0.999	1.000
	OT_{ET}	0.272	0.807	0.986	0.287	0.805	0.991	0.682	0.998	1.000	0.705	0.999	1.000
	Moran	0.282	0.828	0.988	0.253	0.808	0.992	0.746	0.998	1.000	0.737	1.000	1.000
	$Moran_{EL}$	0.308	0.841	0.991	0.280	0.818	0.990	0.753	0.998	1.000	0.744	1.000	1.000
	$Moran_{ET}$	0.315	0.846	0.993	0.279	0.829	0.994	0.752	0.998	1.000	0.745	1.000	1.000
$R^2 = 0.4$	PT_{GMM}	0.147	0.488	0.833	0.176	0.506	0.820	0.362	0.912	0.997	0.432	0.928	1.000
	PT_{EL}	0.207	0.572	0.868	0.236	0.574	0.860	0.382	0.916	0.996	0.438	0.936	1.000
	PT_{ET}	0.188	0.552	0.861	0.232	0.561	0.858	0.394	0.921	0.997	0.442	0.939	1.000
	OT_{GMM}	0.105	0.350	0.679	0.145	0.380	0.672	0.265	0.840	0.991	0.328	0.880	0.999
	OT_{EL}	0.137	0.412	0.728	0.173	0.422	0.731	0.292	0.852	0.990	0.339	0.881	0.999
	OT_{ET}	0.132	0.404	0.718	0.174	0.417	0.725	0.287	0.853	0.991	0.342	0.887	0.999
	Moran	0.048	0.125	0.097	0.025	0.033	0.045	0.259	0.828	0.978	0.214	0.761	0.920
	$Moran_{EL}$	0.060	0.159	0.138	0.040	0.046	0.056	0.271	0.831	0.979	0.227	0.778	0.929
	$Moran_{ET}$	0.061	0.154	0.137	0.038	0.048	0.056	0.269	0.834	0.982	0.228	0.778	0.927
Heteroskedastic case													
$R^2 = 0.8$	PT_{GMM}	0.202	0.545	0.919	0.188	0.553	0.926	0.369	0.908	1.000	0.366	0.914	0.999
	PT_{EL}	0.278	0.590	0.921	0.249	0.617	0.938	0.390	0.910	1.000	0.378	0.911	0.999
	PT_{ET}	0.255	0.595	0.926	0.239	0.610	0.946	0.403	0.919	1.000	0.387	0.922	0.999
	OT_{GMM}	0.136	0.439	0.862	0.137	0.451	0.872	0.307	0.857	0.999	0.273	0.863	0.999
	OT_{EL}	0.174	0.478	0.856	0.172	0.468	0.875	0.310	0.852	0.999	0.277	0.856	0.999
	OT_{ET}	0.171	0.490	0.872	0.166	0.477	0.890	0.317	0.864	0.999	0.288	0.865	0.999
	Moran	0.116	0.382	0.842	0.087	0.393	0.838	0.283	0.846	0.999	0.250	0.843	0.998
	$Moran_{EL}$	0.129	0.400	0.822	0.105	0.404	0.820	0.286	0.837	0.999	0.246	0.834	0.997
	$Moran_{ET}$	0.133	0.414	0.849	0.104	0.422	0.848	0.296	0.850	0.999	0.257	0.847	0.998
$R^2 = 0.4$	PT_{GMM}	0.202	0.516	0.871	0.186	0.495	0.845	0.339	0.891	1.000	0.308	0.892	0.999
	PT_{EL}	0.242	0.515	0.868	0.224	0.500	0.839	0.322	0.878	0.996	0.291	0.872	0.995
	PT_{ET}	0.224	0.509	0.878	0.210	0.501	0.849	0.336	0.889	0.998	0.302	0.882	0.997
	OT_{GMM}	0.116	0.341	0.695	0.111	0.323	0.633	0.244	0.789	0.987	0.221	0.739	0.983
	OT_{EL}	0.145	0.362	0.709	0.141	0.351	0.656	0.226	0.763	0.982	0.222	0.729	0.976
	OT_{ET}	0.142	0.377	0.723	0.129	0.356	0.666	0.233	0.783	0.985	0.223	0.749	0.982
	Moran	0.052	0.232	0.589	0.037	0.165	0.411	0.190	0.753	0.989	0.158	0.690	0.981
	$Moran_{EL}$	0.073	0.276	0.604	0.052	0.192	0.432	0.204	0.747	0.981	0.167	0.687	0.974
	$Moran_{ET}$	0.070	0.276	0.621	0.051	0.196	0.448	0.206	0.759	0.988	0.171	0.699	0.982

"PT_{GMM}", "PT_{EL}" and "PT_{ET}" denote, respectively, the GMM, EL and ET parameter restriction tests; "OT_{GMM}", "OT_{EL}" and "OT_{ET}" denote, respectively, the GMM, EL and ET overidentification tests; and "Moran", "$Moran_{EL}$ " and "$Moran_{ET}$ " denote, respectively, the robust, EL and ET Moran's I tests.