Supplementary online appendix to 'Consistent specification testing under spatial dependence'

Abhimanyu Gupta and Xi Qu

August 29, 2022

S.A Additional simulation results: Unboundedly supported regressors and asymptotic critical values

This section provides additional simulation results using the same design as in Section 8 of the main body of the paper. Recall that the paper reports only bootstrap results for the compactly supported regressors case. Here we include results using asymptotic critical values for both the compactly and unbounded supported regressor cases, as well as bootstrap results for the latter, focusing on the SARARMA(0,1,0) model. The results are in Tables OT.1-OT.4 and our findings match those in the main text, with the bootstrap typically offering better size control.

S.B Proofs of Theorems 4.2 and 4.4

Proof of Theorem 4.2. From Corollary 4.1 and Lemma LS.2, $\|\Sigma(\widehat{\gamma})-\Sigma\| = O_p(\|\widehat{\gamma}-\gamma_0\|) = \sqrt{d_\gamma/n}$, so we have, from Assumption R.3,

$$\left\|\Sigma(\widehat{\gamma})^{-1}-\Sigma^{-1}\right\| \leq \left\|\Sigma(\widehat{\gamma})^{-1}\right\| \|\Sigma(\widehat{\gamma})-\Sigma\| \left\|\Sigma^{-1}\right\| = O_p(\|\widehat{\gamma}-\gamma_0\|) = \sqrt{d_\gamma/n}. \tag{S.B.1}$$

Similarly,

$$\begin{aligned} & \left\|\left(\frac{1}{n}\Psi'\Sigma(\widehat{\gamma})^{-1}\Psi\right)^{-1} - \left(\frac{1}{n}\Psi'\Sigma^{-1}\Psi\right)^{-1}\right\| \\ \leq & \left\|\left(\frac{1}{n}\Psi'\Sigma(\widehat{\gamma})^{-1}\Psi\right)^{-1}\right\| \left\|\frac{1}{n}\Psi'\left(\Sigma(\widehat{\gamma})^{-1}-\Sigma^{-1}\right)\Psi\right\| \left\|\left(\frac{1}{n}\Psi'\Sigma^{-1}\Psi\right)^{-1}\right\| \\ \leq & \sup_{\gamma\in\Gamma}\left\|\left(\frac{1}{n}\Psi'\Sigma(\gamma)^{-1}\Psi\right)^{-1}\right\| \left\|\Sigma(\widehat{\gamma})^{-1}-\Sigma^{-1}\right\| \left\|\frac{1}{\sqrt{n}}\Psi\right\|^2 = O_p(\|\widehat{\gamma}-\gamma_0\|) = \sqrt{d_\gamma/n}. \end{aligned}$$

By Assumption R.2, we have $\widehat{\alpha}-\alpha^* = O_p(1/\sqrt{n})$. Denote by $\theta^*(x) = \psi(x)'\beta^*$, where $\beta^* = \arg\min_\beta \mathcal{E}[y_i - \psi(x_i)'\beta)]^2$, and set $\theta_{ni} = \theta(x_i)$, $\theta_{0i} = \theta_0(x_i)$, $\widehat{\theta}_i = \psi_i'\widehat{\beta}$, $\widehat{f}_i = f(x_i,\widehat{\alpha})$, $f_i^* = f(x_i,\alpha^*)$. Then $\widehat{u}_i = y_i - f(x_i,\widehat{\alpha}) = u_i + \theta_{0i} - \widehat{f}_i$. Let $\theta_0 = (\theta_0(x_1),\ldots,\theta_0(x_n))'$ as before, with similar component-wise notation for the n-dimensional vectors θ^*, $\widehat{f}$, and u. As the approximation error is

$e = \theta_0 - \theta^* = \theta_0 - \Psi\beta^*$,

$$\begin{aligned}
\widehat{\theta} - \theta^* &= \Psi(\widehat{\beta} - \beta^*) = \Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}(u + \theta_0 - \Psi\beta^*) \\
&= \Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}(u + e),
\end{aligned}$$

so that

$$\begin{aligned}
n\widehat{m}_n &= \widehat{\sigma}^{-2}\widehat{v}'\Sigma\left(\widehat{\gamma}\right)^{-1}\widehat{u} = \widehat{\sigma}^{-2}\left(\widehat{\theta} - \widehat{f}\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(y - \widehat{f}\right) \\
&= \widehat{\sigma}^{-2}\left(\widehat{\theta} - \theta^* + \theta^* - \theta_0 + \theta_0 - \widehat{f}\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(u + \theta_0 - \widehat{f}\right) \\
&= \widehat{\sigma}^{-2}\left[\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}(u + e) - e + \theta_0 - \widehat{f}\right]'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(u + \theta_0 - \widehat{f}\right) \\
&= \widehat{\sigma}^{-2}u'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}u + \widehat{\sigma}^{-2}u'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(\theta_0 - \widehat{f}\right) \\
&- \widehat{\sigma}^{-2}\left(u + \theta_0 - \widehat{f}\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(I - \Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\right)e \\
&+ \widehat{\sigma}^{-2}\left(\theta_0 - \widehat{f}\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}u \\
&+ \widehat{\sigma}^{-2}(\theta_0 - \widehat{f})'\Sigma\left(\widehat{\gamma}\right)^{-1}(\theta_0 - \widehat{f}) \\
&= \widehat{\sigma}^{-2}u'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}u + \widehat{\sigma}^{-2}\left(A_1 + A_2 + A_3 + A_4\right),
\end{aligned}$$

say. First, for any vector g comprising of conditioned random variables,

$$\mathcal{E}\left[\left(u'\Sigma(\gamma)^{-1}g\right)^2\right] = g'\Sigma(\gamma)^{-1}\Sigma\Sigma(\gamma)^{-1}g \leq \sup_{\gamma\in\Gamma}\left\|\Sigma(\gamma)^{-1}\right\|^2\|\Sigma\|\,\|g\|^2 = O_p\left(\|g\|^2\right),$$

uniformly in $\gamma \in \Gamma$, where the expectation is taken conditional on g. Similarly,

$$\begin{aligned}
&\mathcal{E}\left[\left(u'\Sigma(\gamma)^{-1}\Psi\left(\Psi'\Sigma(\gamma)^{-1}\Psi\right)^{-1}\Psi'\Sigma(\gamma)^{-1}g\right)^2\right] \\
=\;& g'\Sigma(\gamma)^{-1}\Psi\left(\Psi'\Sigma(\gamma)^{-1}\Psi\right)^{-1}\Psi'\Sigma(\gamma)^{-1}\Sigma\Sigma(\gamma)^{-1}\Psi\left(\Psi'\Sigma(\gamma)^{-1}\Psi\right)^{-1}\Psi'\Sigma(\gamma)^{-1}g \\
\leq\;& \sup_{\gamma\in\Gamma}\left\|\Sigma(\gamma)^{-1}\right\|^4\|\Sigma\|\left\|\frac{1}{n}\Psi\left(\frac{1}{n}\Psi'\Sigma(\gamma)^{-1}\Psi\right)^{-1}\Psi'\right\|^2\|g\|^2 = O_p\left(\|g\|^2\right),
\end{aligned}$$

uniformly and, for any $j = 1, \ldots, d_\gamma$,

$$\begin{aligned}
\mathcal{E}\left[\left(u'\Sigma(\gamma)^{-1}\Sigma_j\left(\gamma\right)\Sigma\left(\gamma\right)^{-1}g\right)^2\right] &= g'\Sigma(\gamma)^{-1}\Sigma_j\left(\gamma\right)\Sigma\left(\gamma\right)^{-1}\Sigma\Sigma(\gamma)^{-1}\Sigma_j\left(\gamma\right)\Sigma\left(\gamma\right)^{-1}g \\
&\leq \sup_{\gamma\in\Gamma}\left\|\Sigma(\gamma)^{-1}\right\|^4\|\Sigma_j\left(\gamma\right)\|^2\|\Sigma\|\,\|g\|^2 = O_p\left(\|g\|^2\right).
\end{aligned}$$

Let Ψ_k be the k-th column of Ψ, $k = 1, \ldots, p$. Then, we have $\|\Psi_k/\sqrt{n}\| = O_p(1)$ and for any $\gamma \in \Gamma$,

$$\begin{aligned}
\mathcal{E}\left\|\frac{1}{\sqrt{n}}u'\Sigma\left(\gamma\right)^{-1}\Psi\right\|^2 &\leq \sum_{k=1}^{p}\mathcal{E}\left(u'\Sigma\left(\gamma\right)^{-1}\frac{1}{\sqrt{n}}\Psi_k\right)^2 = O_p\left(p\right), \\
\mathcal{E}\left\|\frac{1}{\sqrt{n}}u'\Sigma\left(\gamma\right)^{-1}\Sigma_j\left(\gamma\right)\Sigma\left(\gamma\right)^{-1}\Psi\right\|^2 &\leq \sum_{k=1}^{p}\mathcal{E}\left(u'\Sigma\left(\gamma\right)^{-1}\Sigma_j\left(\gamma\right)\Sigma\left(\gamma\right)^{-1}\frac{1}{\sqrt{n}}\Psi_k\right)^2 = O(p).
\end{aligned}$$

Therefore, by Chebyshev's inequality,

$$\sup_{\gamma\in\Gamma}\left\|\frac{1}{\sqrt{n}}u'\Sigma\left(\gamma\right)^{-1}\Psi\right\| = O_p(\sqrt{p}) \quad \text{and} \quad \sup_{\gamma\in\Gamma}\left\|\frac{1}{\sqrt{n}}u'\Sigma\left(\gamma\right)^{-1}\Sigma_j\left(\gamma\right)\Sigma\left(\gamma\right)^{-1}\Psi\right\| = O_p(\sqrt{p}).$$

By the decomposition

$$\begin{aligned}
& u'\left(\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1} - \Sigma^{-1}\Psi[\Psi'\Sigma^{-1}\Psi]^{-1}\Psi'\Sigma^{-1}\right)u \\
= \; & u'\left(\Sigma\left(\widehat{\gamma}\right)^{-1} + \Sigma^{-1}\right)\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\left(\sum_{i=1}^{n}e_{in}e_{in}'\right)\left(\Sigma\left(\widehat{\gamma}\right)^{-1} - \Sigma^{-1}\right)u \\
& + u'\Sigma^{-1}\Psi\left([\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1} - [\Psi'\Sigma^{-1}\Psi]^{-1}\right)\Psi'\Sigma^{-1}u \\
= \; & u'\left(\Sigma\left(\widehat{\gamma}\right)^{-1} + \Sigma^{-1}\right)\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\left(\sum_{i=1}^{n}e_{in}e_{in}'\right)\sum_{j=1}^{d_\gamma}\left(\Sigma\left(\widetilde{\gamma}\right)^{-1}\Sigma_j\left(\widetilde{\gamma}\right)\Sigma\left(\widetilde{\gamma}\right)^{-1}\right) \\
\times \; & u(\widetilde{\gamma}_j - \gamma_{j0}) + u'\Sigma^{-1}\Psi\left([\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1} - [\Psi'\Sigma^{-1}\Psi]^{-1}\right)\Psi'\Sigma^{-1}u,
\end{aligned}$$

where e_{in} is an $n \times 1$ vector with i-th entry one and zeros elsewhere, so $\sum_{i=1}^{n} e_{in}e_{in}' = I_n$, and

$$\begin{aligned}
e_{in}'\left(\Sigma\left(\widehat{\gamma}\right)^{-1} - \Sigma^{-1}\right)u &= \sum_{j=1}^{d_\gamma}e_{in}'\left(\Sigma\left(\widetilde{\gamma}\right)^{-1}\Sigma_j\left(\widetilde{\gamma}\right)\Sigma\left(\widetilde{\gamma}\right)^{-1}\right)u(\widetilde{\gamma}_j - \gamma_{j0}) \\
&= e_{in}'\sum_{j=1}^{d_\gamma}\left(\Sigma\left(\widetilde{\gamma}\right)^{-1}\Sigma_j\left(\widetilde{\gamma}\right)\Sigma\left(\widetilde{\gamma}\right)^{-1}\right)u(\widetilde{\gamma}_j - \gamma_{j0})
\end{aligned}$$

where $\widetilde{\gamma}$ is a value between $\widehat{\gamma}$ and γ_0 due to the mean value theorem. We have

$$\begin{aligned}
& \left|u'\left(\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1} - \Sigma^{-1}\Psi[\Psi'\Sigma^{-1}\Psi]^{-1}\Psi'\Sigma^{-1}\right)u\right| \\
\leq \; & 2\sup_{\gamma\in\Gamma}\left\|\frac{1}{\sqrt{n}}u'\Sigma\left(\gamma\right)^{-1}\Psi\right\|\left\|\left(\frac{1}{n}\Psi'\Sigma\left(\gamma\right)^{-1}\Psi\right)^{-1}\right\|\sum_{j=1}^{d_\gamma}\left\|\frac{1}{\sqrt{n}}\Psi'\left(\Sigma\left(\gamma\right)^{-1}\Sigma_j\left(\gamma\right)\Sigma\left(\gamma\right)^{-1}\right)u\right\| \\
\times \; & |\widetilde{\gamma}_j - \gamma_{j0}| + \left\|\frac{1}{\sqrt{n}}u'\Sigma^{-1}\Psi\right\|^2\left\|\left(\frac{1}{n}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1} - \left(\frac{1}{n}\Psi'\Sigma^{-1}\Psi\right)^{-1}\right\| \\
= \; & O_p(\sqrt{p})O_p(d_\gamma\sqrt{p}/\sqrt{n}) + O_p(p)O_p(\sqrt{d_\gamma}/\sqrt{n}) = O_p(d_\gamma p/\sqrt{n}) = o_p(\sqrt{p}),
\end{aligned}$$

where the last equality holds under the conditions of the theorem.

It remains to show that

$$A_i = o_p\left(p^{1/2}\right), i = 1, \ldots, 4. \tag{S.B.2}$$

It is convenient to perform the calculations under H_ℓ, which covers H_0 as a particular case. Using the mean value theorem and either H_0 or H_ℓ, we can express

$$\theta_{0i} - \widehat{f}_i = f_i^* - \widehat{f}_i - (p^{1/4}/n^{1/2})h_i = \sum_{j=1}^{d_\alpha} \frac{\partial f(x_i, \widetilde{\alpha})}{\partial \alpha_j}(\alpha_j^* - \widetilde{\alpha}_j) - \frac{p^{1/4}}{n^{1/2}} h_i, \tag{S.B.3}$$

where $\widetilde{\alpha}_j$ is a value between α_j^* and $\widehat{\alpha}_j$. Then, for any $j = 1, \ldots, d_\alpha$, $\left|\alpha_j^* - \widetilde{\alpha}_j\right| = O_p(1/\sqrt{n})$. Based on

$$\sup_{\gamma \in \Gamma} \left| u'\Sigma(\gamma)^{-1}\Psi\left(\Psi'\Sigma(\gamma)^{-1}\Psi\right)^{-1}\Psi'\Sigma(\gamma)^{-1}g\right| = O_p\left(\|g\|\right) \text{ and } \sup_{\gamma\in\Gamma}\left|u'\Sigma(\gamma)^{-1}g\right| = O_p\left(\|g\|\right)$$

for any $\gamma \in \Gamma$ and any conditioned vector g, if we take $g = \partial f(x, \alpha)/\partial \alpha_j$ or $g = h$, then both satisfy $O_p\left(\|g\|\right) = O_p\left(\sqrt{n}\right)$ and it follows that

$$\begin{aligned} |A_1| &= \left|u'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(\theta_0 - \widehat{f}\right)\right| \leq \sup_{\gamma,\alpha}\sum_{j=1}^{d_\alpha}\left\|u'\Sigma(\gamma)^{-1}\frac{\partial f(x,\alpha)}{\partial \alpha_j}\right\|\left|\alpha_j^* - \widetilde{\alpha}_j\right| + \frac{p^{1/4}}{n^{1/2}}\sup_\gamma\left\|u'\Sigma(\gamma)^{-1}h\right\| \\ &= O_p(\sqrt{n})O_p\left(\frac{1}{\sqrt{n}}\right) + O\left(\frac{p^{1/4}}{n^{1/2}}\right)O_p(\sqrt{n}) = O_p(p^{1/4}) = o_p(p^{1/2}). \end{aligned}$$

Similarly,

$$\begin{aligned} |A_3| &= \left|u'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(\theta_0 - \widehat{f}\right)\right| \\ &\leq \sup_{\gamma,\alpha}\sum_{j=1}^{d_\alpha}\left\|u'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\frac{\partial f(x,\alpha)}{\partial\alpha_j}\right\|\left|\alpha_j^* - \widetilde{\alpha}_j\right| \\ &\quad + \frac{p^{1/4}}{n^{1/2}}\sup_\gamma\left\|u'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}h\right\| \\ &= O_p(1) + O_p(p^{1/4}) = O_p(p^{1/4}) = o_p(p^{1/2}). \end{aligned}$$

Also, by Assumptions R.2 and R.10, we have

$$\left\|\theta_0 - \widehat{f}\right\| \leq \sup_\alpha\sum_{j=1}^{d_\alpha}\left\|\frac{\partial f(x,\alpha)}{\partial\alpha_j}\right\|\left|\alpha_j^* - \widetilde{\alpha}_j\right| + \|h\|\frac{p^{1/4}}{n^{1/2}} = O_p(p^{1/4}). \tag{S.B.4}$$

By (3.2), we have $\|e\| = O(p^{-\mu}n^{1/2})$ and

$$\begin{aligned}
|A_2| &= \left|(u+\theta_0 - \widehat{f})' \left(\Sigma\left(\widehat{\gamma}\right)^{-1} - \Sigma\left(\widehat{\gamma}\right)^{-1}\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\right) e\right| \\
&\leq \sup_{\gamma} |u'\Sigma\left(\gamma\right)^{-1} e| + \sup_{\gamma} \left|u'\Sigma\left(\gamma\right)^{-1}\Psi[\Psi'\Sigma\left(\gamma\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\gamma\right)^{-1} e\right| \\
&\quad + \left\|\theta_0 - \widehat{f}\right\| \sup_{\gamma} \left(\left\|\Sigma\left(\gamma\right)^{-1}\right\| + \left\|\Sigma\left(\gamma\right)^{-1}\Psi[\Psi'\Sigma\left(\gamma\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\gamma\right)^{-1}\right\|\right) \|e\| \\
&= O_p(p^{-\mu}n^{1/2}) + O_p(p^{-\mu+1/4}n^{1/2}) = O_p(p^{-\mu+1/4}n^{1/2}) = o_p(\sqrt{p}).
\end{aligned}$$

where the last equality holds under the conditions of the theorem. Finally, under H_ℓ,

$$\begin{aligned}
A_4 &= \left(\theta_0 - \widehat{f}\right)' \Sigma\left(\widehat{\gamma}\right)^{-1} \left(\theta_0 - \widehat{f}\right) \\
&= \left(\theta_0 - \widehat{f}\right)' \Sigma^{-1} \left(\theta_0 - \widehat{f}\right) + \left(\theta_0 - \widehat{f}\right)' \left(\Sigma\left(\widehat{\gamma}\right)^{-1} - \Sigma^{-1}\right) \left(\theta_0 - \widehat{f}\right) \\
&= \frac{p^{1/2}}{n} h'\Sigma^{-1}h + o_p(1) + O_p\left(p^{1/2}d_\gamma^{1/2}/n^{1/2}\right) = \frac{p^{1/2}}{n} h'\Sigma^{-1}h + o_p(\sqrt{p}).
\end{aligned}$$

Combining these together, we have

$$n\widehat{m}_n = \widehat{\sigma}^{-2}\widehat{v}'\Sigma\left(\widehat{\gamma}\right)^{-1}\widehat{u} = \frac{1}{\sigma_0^2}\varepsilon'\mathscr{V}\varepsilon + \frac{p^{1/2}}{n} h'\Sigma^{-1}h + o_p(\sqrt{p}),$$

under H_ℓ and the same expression holds with $h = 0$ under H_0.

□

Proof of Theorem 4.4. (1) Follows from Theorems 4.2 and 4.3. (2) Following reasoning analogous to the proofs of Theorems 4.2 and 4.3, it can be shown that under H_1, $\widehat{m}_n = n^{-1}\sigma^{*-2}(\theta_0 - f^*)'\Sigma\left(\gamma^*\right)^{-1}(\theta_0 - f^*) + o_p(1)$. Then,

$$\mathscr{T}_n = \left(n\widehat{m}_n - p\right)/\sqrt{2p} = \left(n/\sqrt{p}\right)(\theta_0 - f^*)'\Sigma\left(\gamma^*\right)^{-1}(\theta_0 - f^*)/\left(\sqrt{2}n\sigma^{*2}\right) + o_p\left(n/\sqrt{p}\right)$$

and for any nonstochastic sequence $\{C_n\}$, $C_n = o(n/p^{1/2})$, $P(\mathscr{T}_n > C_n) \to 1$, so that consistency follows. (3) Follows from Theorems 4.2 and 4.3. □

S.C Proof of Theorem 5.1

Proof. We prove the result under H_1, which is the more challenging case as it involves nonparametric estimation. The proof under H_0 is similar. We will show $\widehat{\phi} \xrightarrow{p} \phi_0$, whence $\widehat{\beta} \xrightarrow{p} \beta_0$ and $\widehat{\sigma}^2 \xrightarrow{p} \sigma_0^2$ follow from (5.3) and (5.4) respectively. First note that

$$\mathcal{L}(\phi)-\mathcal{L}=\log\overline{\sigma}^2(\phi)/\overline{\sigma}^2-n^{-1}\log\left|T'(\lambda)\Sigma(\gamma)^{-1}T(\lambda)\Sigma\right|=\log\overline{\sigma}^2(\phi)/\sigma^2(\phi)-\log\overline{\sigma}^2/\sigma_0^2+\log r(\phi), \tag{S.C.1}$$

where recall that $\sigma^2(\phi)=n^{-1}\sigma_0^2 tr\left(T'(\lambda)\Sigma(\gamma)^{-1}T(\lambda)\Sigma\right),\ \overline{\sigma}^2=\overline{\sigma}^2(\phi_0)=n^{-1}u'E'MEu$, using (5.4) and also $r(\phi)=n^{-1}tr\left(T'(\lambda)\Sigma(\gamma)^{-1}T(\lambda)\Sigma\right)/\left|T'(\lambda)\Sigma(\gamma)^{-1}T(\lambda)\Sigma\right|^{1/n}$.

We have $\overline{\sigma}^2(\phi)=n^{-1}\left\{S^{-1\prime}(\Psi\beta_0+u)\right\}'S'(\lambda)E(\gamma)'M(\gamma)E(\gamma)S(\lambda)S^{-1}(\Psi\beta_0+u)=c_1(\phi)+c_2(\phi)+c_3(\phi)$, where

$$\begin{aligned}
c_1(\phi) &= n^{-1}\beta_0'\Psi'T'(\lambda)E(\gamma)'M(\gamma)E(\gamma)T(\lambda)\Psi\beta_0,\\
c_2(\phi) &= n^{-1}\sigma_0^2 tr\left(T'(\lambda)E(\gamma)'M(\gamma)E(\gamma)T(\lambda)\Sigma\right),\\
c_3(\phi) &= n^{-1}tr\left(T'(\lambda)E(\gamma)'M(\gamma)E(\gamma)T(\lambda)\left(uu'-\sigma_0^2\Sigma\right)\right)\\
&+ 2n^{-1}\beta_0'\Psi'T'(\lambda)E(\gamma)'M(\gamma)E(\gamma)T(\lambda)u.
\end{aligned}$$

Note that in the particular cases of Theorems 4.1 and 6.1, where $T(\lambda)=I_n$, the c_1 term vanishes because $M(\gamma)E(\gamma)\Psi=0$ and $M(\tau)E(\tau)\Psi=0$. Proceeding with the current, more general proof

$$\begin{aligned}
\log\frac{\overline{\sigma}^2(\phi)}{\sigma^2(\phi)} &= \log\frac{\overline{\sigma}^2(\phi)}{(c_1(\phi)+c_2(\phi))}+\log\frac{c_1(\phi)+c_2(\phi)}{\sigma^2(\phi)}\\
&= \log\left(1+\frac{c_3(\phi)}{c_1(\phi)+c_2(\phi)}\right)+\log\left(1+\frac{c_1(\phi)-f(\phi)}{\sigma^2(\phi)}\right),
\end{aligned}$$

where $f(\phi)=n^{-1}\sigma_0^2 tr\left(E'^{-1}T'(\lambda)E(\gamma)'(I_n-M(\gamma))E(\gamma)T(\lambda)E^{-1}\right)$. Then (S.C.1) implies

$$\begin{aligned}
P\left(\left\|\widehat{\phi}-\phi_0\right\|\in\overline{\mathcal{N}}^{\phi}(\eta)\right) &= P\left(\inf_{\phi\in\overline{\mathcal{N}}^{\phi}(\eta)}\mathcal{L}(\phi)-\mathcal{L}\le 0\right)\\
&\le P\left(\log\left(1+\sup_{\phi\in\overline{\mathcal{N}}^{\phi}(\eta)}\left|\frac{c_3(\phi)}{c_1(\phi)+c_2(\phi)}\right|\right)+\left|\log\left(\overline{\sigma}^2/\sigma_0^2\right)\right|\right.\\
&\qquad\left.\ge\inf_{\phi\in\overline{\mathcal{N}}^{\phi}(\eta)}\left(\log\left(1+\frac{c_1(\phi)-f(\phi)}{\sigma^2(\phi)}\right)+\log r(\phi)\right)\right),
\end{aligned}$$

where recall that $\overline{\mathcal{N}}^{\phi}(\eta)=\Phi\backslash\mathcal{N}^{\phi}(\eta)$, $\mathcal{N}^{\phi}(\eta)=\{\phi:\|\phi-\phi_0\|<\eta\}\cap\Phi$. Because $\overline{\sigma}^2/\sigma_0^2\xrightarrow{p}1$, the property $\log(1+x)=x+o(x)$ as $x\to 0$ implies that it is sufficient to show that

$$\sup_{\phi\in\overline{\mathcal{N}}^{\phi}(\eta)}\left|\frac{c_3(\phi)}{c_1(\phi)+c_2(\phi)}\right|\xrightarrow{p}0, \tag{S.C.2}$$

$$\sup_{\phi\in\overline{\mathcal{N}}^{\phi}(\eta)}\left|\frac{f(\phi)}{\sigma^2(\phi)}\right|\xrightarrow{p}0, \tag{S.C.3}$$

$$P\left(\inf_{\phi\in\overline{\mathcal{N}}^{\phi}(\eta)}\left\{\frac{c_1(\phi)}{\sigma^2(\phi)}+\log r(\phi)\right\}>0\right)\quad\longrightarrow\quad 1. \tag{S.C.4}$$

Because $\overline{\mathcal{N}}^{\phi}(\eta)\subseteq\left\{\Lambda\times\overline{\mathcal{N}}^{\gamma}(\eta/2)\right\}\cup\left\{\overline{\mathcal{N}}^{\lambda}(\eta/2)\times\Gamma\right\}$, we have

$$\begin{aligned}P\left(\inf_{\phi\in\overline{\mathcal{N}}^{\phi}(\eta)}\left\{\frac{c_1(\phi)}{\sigma^2(\phi)}+\log r(\phi)\right\}>0\right) &\geq P\left(\min\left\{\inf_{\Lambda\times\overline{\mathcal{N}}^{\gamma}(\eta/2)}\frac{c_1(\phi)}{\sigma^2(\phi)},\inf_{\overline{\mathcal{N}}^{\lambda}(\eta/2)}\log r(\phi)\right\}>0\right)\\ &\geq P\left(\min\left\{\inf_{\Lambda\times\overline{\mathcal{N}}^{\gamma}(\eta/2)}\frac{c_1(\phi)}{C},\inf_{\overline{\mathcal{N}}^{\lambda}(\eta/2)}\log r(\phi)\right\}>0\right),\end{aligned}$$

from Assumption SAR.2, whence Assumptions SAR.3 and SAR.4 imply (S.C.4). Again using Assumption SAR.2, uniformly in ϕ, $\left|f(\phi)/\sigma^2(\phi)\right|=O_p(|f(\phi)|)$ and

$$\begin{aligned}|f(\phi)| &= O_p\left(tr\left(E'^{-1}T'(\lambda)\Sigma(\gamma)^{-1}\Psi\left(\Psi'\Sigma(\gamma)^{-1}\Psi\right)^{-1}\Psi'\Sigma(\gamma)^{-1}T(\lambda)E^{-1}\right)/n\right)\\ &= O_p\left(tr\left(E'^{-1}T'(\lambda)\Sigma(\gamma)^{-1}\Psi\Psi'\Sigma(\gamma)^{-1}T(\lambda)E^{-1}\right)/n^2\right)=O_p\left(\left\|\Psi'\Sigma(\gamma)^{-1}T(\lambda)E^{-1}/n\right\|_F^2\right)\\ &= O_p\left(\|\Psi/n\|_F^2\,\overline{\varphi}^2\left(\Sigma(\gamma)^{-1}\right)\|T(\lambda)\|^2\left\|E^{-1}\right\|^2\right)=O_p\left(\|\Psi/n\|_F^2\,\|T(\lambda)\|^2\,\overline{\varphi}(\Sigma)/\underline{\varphi}^2(\Sigma(\gamma))\right)\\ &= O_p\left(\|T(\lambda)\|^2/n\right),\end{aligned} \tag{S.C.5}$$

where we have twice made use of the inequality

$$\|AB\|_F\leq\|A\|_F\|B\| \tag{S.C.6}$$

for generic multiplication compatible matrices A and B. (S.C.3) now follows by Assumption SAR.1 and compactness of Λ because $T(\lambda)=I_n+\sum_{j=1}^{d_\lambda}(\lambda_{0j}-\lambda_j)G_j$. Finally consider (S.C.2). We first prove pointwise convergence. For any fixed $\phi\in\overline{\mathcal{N}}^{\phi}(\eta)$ and large enough n, Assumptions SAR.2 and SAR.4 imply

$$\{c_1(\phi)\}^{-1} = O_p\left(\|\beta_0\|^{-2}\right)=O_p(1) \tag{S.C.7}$$

$$\{c_2(\phi)\}^{-1} = O_p(1), \tag{S.C.8}$$

because $\left\{n^{-1}\sigma_0^2 tr\left(T'(\lambda)\Sigma(\gamma)^{-1}T(\lambda)E^{-1}\right)\right\}^{-1}=O_p(1)$ and, proceeding like in the bound for $|f(\phi)|$, $tE'^{-1}r\left(E'^{-1}T'(\lambda)E(\gamma)'\left(I-M(\gamma)\right)E(\gamma)T(\lambda)E^{-1}\right)=O_p\left(\|T(\lambda)\|^2/n\right)=O_p(1/n)$. In fact it is worth noting for the equicontinuity argument presented later that Assumptions SAR.2 and SAR.4 actually imply that (S.C.7) and (S.C.8) hold uniformly over $\overline{\mathcal{N}}^{\phi}(\eta)$, a property not needed for the present pointwise arguments. Thus $c_3(\phi)/(c_1(\phi)+c_2(\phi))=O_p(|c_3(\phi)|)$ where, writing $\mathfrak{B}(\phi)=$

$T'(\lambda)E(\gamma)'M(\gamma)E(\gamma)T(\lambda)$ with typical element $\mathfrak{b}_{rs}(\phi)$, $r,s=1,\ldots,n$, $c_3(\phi)$ has mean 0 and variance

$$O_p\left(\frac{\|\mathfrak{B}(\phi)\Sigma\|_F^2}{n^2}+\frac{\sum_{r,s,t,v=1}^{n}\mathfrak{b}_{rs}(\phi)\mathfrak{b}_{tv}(\phi)\kappa_{rstv}}{n^2}+\frac{\left\|\beta_0'\Psi'\mathfrak{B}(\phi)E^{-1}\right\|^2}{n^2}\right), \tag{S.C.9}$$

with κ_{rstv} denoting the fourth cumulant of u_r,u_s,u_t,u_v, $r,s,t,v=1,\ldots,n$. Under the linear process assumed in Assumption R.4 it is known that

$$\sum_{r,s,t,v=1}^{n}\kappa_{rstv}^2=O(n). \tag{S.C.10}$$

Using (S.C.6) and Assumptions SAR.1 and R.3, the first term in parentheses in (S.C.9) is

$$\begin{aligned}O_p\left(\|\mathfrak{B}(\phi)\|_F^2\,\overline{\varphi}^2(\Sigma)/n^2\right) &= O_p\left(\|T(\lambda)\|_F^2\,\|E(\gamma)\|^4\,\|M(\gamma)\|^2\,\|T(\lambda)\|^2/n^2\right)\\ &= O_p\left(\|T(\lambda)\|^4/n\underline{\varphi}^2(\Sigma(\gamma))\right)=O_p\left(\|T(\lambda)\|^4/n\right),\end{aligned} \tag{S.C.11}$$

while the second is similarly

$$O_p\left\{\left(\|\mathfrak{B}(\phi)\|_F^2/n\right)\left(\sum_{r,s,t,v=1}^{n}\kappa_{rstv}^2/n^2\right)^{\frac{1}{2}}\right\}=o_p\left(\|T(\lambda)\|^4\right), \tag{S.C.12}$$

using (S.C.10). Finally, the third term in parentheses in (S.C.9) is

$$O_p\left(\|\mathfrak{B}(\phi)\|^2/n\right)=O_p\left(\|T(\lambda)\|^4/n\right). \tag{S.C.13}$$

By compactness of Λ and Assumption SAR.1, (S.C.11), (S.C.12) and (S.C.13) are negligible, thus pointwise convergence is established.

Uniform convergence will follow from an equicontinuity argument. First, for arbitrary $\varepsilon>0$ we can find points $\phi_*=(\lambda_*',\gamma_*')'$, possibly infinitely many, such that the neighborhoods $\|\phi-\phi^*\|<\varepsilon$ form an open cover of $\overline{\mathcal{N}}^{\phi}(\eta)$. Since Φ is compact any open cover has a finite subcover and thus we may in fact choose finitely many $\phi_*=(\lambda_*',\gamma_*')'$, whence it suffices to prove

$$\sup_{\|\phi-\phi_*\|<\varepsilon}\left|\frac{c_3(\phi)}{c_1(\phi)+c_2(\phi)}-\frac{c_3(\phi_*)}{c_1(\phi_*)+c_2(\phi_*)}\right|\xrightarrow{p}0.$$

Proceeding as in Gupta and Robinson (2018), we denote the two components of $c_3(\phi)$ by $c_{31}(\phi)$, $c_{32}(\phi)$, and are left with establishing the negligibility of

$$\frac{|c_{31}(\phi)-c_{31}(\phi_*)|}{c_2(\phi)}+\frac{|c_{32}(\phi)-c_{32}(\phi_*)|}{c_1(\phi)}+\frac{|c_3(\phi_*)|}{c_1(\phi)c_1(\phi_*)}|c_1(\phi_*)-c_1(\phi)|$$

$$+ \quad \frac{|c_3(\phi_*)|}{c_2(\phi)\, c_2(\phi_*)} |c_2(\phi_*) - c_2(\phi)|, \tag{S.C.14}$$

uniformly on $\|\phi - \phi_*\| < \varepsilon$. By the fact that (S.C.7) and (S.C.8) hold uniformly over Φ, we first consider only the numerators in the first two terms in (S.C.14). As in the proof of Theorem 1 of Delgado and Robinson (2015), (S.C.6) implies that $\mathcal{E}\left(\sup_{\|\phi-\phi_*\|<\varepsilon} |c_{31}(\phi) - c_{31}(\phi_*)|\right)$ is bounded by

$$n^{-1}\left(\mathcal{E}\|u\|^2 + \sigma_0^2 tr\Sigma\right) \sup_{\|\phi-\phi_*\|<\varepsilon} \|\mathfrak{B}(\phi) - \mathfrak{B}(\phi_*)\| = O_p\left(\sup_{\|\phi-\phi_*\|<\varepsilon} \|\mathfrak{B}(\phi) - \mathfrak{B}(\phi_*)\|\right),$$

because $\mathcal{E}\|u\|^2 = O(n)$ and $tr\Sigma = O(n)$. $\mathfrak{B}(\phi) - \mathfrak{B}(\phi_*)$ can be written as

$$\begin{aligned} & (T(\lambda) - T(\lambda_*))' E(\gamma)' M(\gamma) E(\gamma) T(\lambda) + T(\lambda_*)' \Sigma'(\gamma_*) M(\gamma_*) E(\gamma_*) (T(\lambda) - T(\lambda_*)) \\ + \quad & T'(\lambda_*) \left(E(\gamma)' M(\gamma) E(\gamma) - E(\gamma_*)' M(\gamma_*) E(\gamma_*)\right) T(\lambda), \end{aligned} \tag{S.C.15}$$

which, by the triangle inequality, has spectral norm bounded by

$$\begin{aligned} & \|T(\lambda) - T(\lambda_*)\| \left(\|E(\gamma)\|^2 \|T(\lambda)\| + \|E(\gamma_*)\|^2 \|T(\lambda_*)\|\right) \\ + \quad & \|T(\lambda_*)\| \left\|E(\gamma)' M(\gamma) E(\gamma) - E(\gamma_*)' M(\gamma_*) E(\gamma_*)\right\| \|T(\lambda)\| \\ = \quad & O_p\left(\|T(\lambda) - T(\lambda_*)\| + \left\|E(\gamma)' M(\gamma) E(\gamma) - E(\gamma_*)' M(\gamma_*) E(\gamma_*)\right\|\right). \end{aligned} \tag{S.C.16}$$

By Assumption SAR.1 the first term in parentheses on the right side of (S.C.16) is bounded uniformly on $\|\phi - \phi_*\| < \varepsilon$ by

$$\sum_{j=1}^{d_\lambda} |\lambda_j - \lambda_{*j}| \|G_j\| \le \max_{j=1,\dots,d_\lambda} \|G_j\| \|\lambda - \lambda_*\| = O_p(\varepsilon), \tag{S.C.17}$$

while because $E(\gamma)' M(\gamma) E(\gamma) = n^{-1} \Sigma(\gamma)^{-1} \Psi \left(n^{-1} \Psi' \Sigma(\gamma)^{-1} \Psi\right)^{-1} \Psi' \Sigma(\gamma)^{-1}$ for any $\gamma \in \Gamma$, the second one can be decomposed into terms with bounds typified by

$$\begin{aligned} & n^{-1} \left\|\Sigma(\gamma)^{-1} - \Sigma(\gamma_*)^{-1}\right\| \|\Psi\|^2 \left\|\left(n^{-1} \Psi' \Sigma(\gamma)^{-1} \Psi\right)^{-1}\right\| \left\|\Sigma(\gamma)^{-1}\right\|^2 \\ \le \quad & n^{-1} \|\Sigma(\gamma) - \Sigma(\gamma_*)\| \|\Psi\|^2 \left\|\left(n^{-1} \Psi' \Sigma(\gamma)^{-1} \Psi\right)^{-1}\right\| \left\|\Sigma(\gamma)^{-1}\right\|^3 \left\|\Sigma(\gamma_*)^{-1}\right\| \\ = \quad & O_p(\|\Sigma(\gamma) - \Sigma(\gamma_*)\|) = O_p(\varepsilon), \end{aligned}$$

uniformly on $\|\phi - \phi_*\| < \varepsilon$, by Assumptions R.3 and R.8, Proposition 4.1 and the inequality $\|A\| \le \|A\|_F$ for a generic matrix A, so that

$$\sup_{\|\phi-\phi_*\|<\varepsilon} \|\mathfrak{B}(\phi) - \mathfrak{B}(\phi_*)\| = O_p(\varepsilon). \tag{S.C.18}$$

Thus equicontinuity of the first term in (S.C.14) follows because ε is arbitrary. The equicontinuity of the second term in (S.C.14) follows in much the same way. Indeed $\sup_{\|\phi-\phi_*\|<\varepsilon} c_{32}(\phi) - c_{32}(\phi_*) = 2n^{-1}\beta_0'\Psi' \sup_{\|\phi-\phi_*\|<\varepsilon} (\mathfrak{B}(\phi) - \mathfrak{B}(\phi_*))\, u = O_p\left(\sup_{\|\phi-\phi_*\|<\varepsilon} \|\mathfrak{B}(\phi) - \mathfrak{B}(\phi_*)\|\right) = O_p(\varepsilon)$, using earlier arguments and (S.C.18). Because $c_1(\phi)$ is bounded and bounded away from zero in probability (see S.C.7) for sufficiently large n and all $\phi \in \overline{\mathcal{N}}^{\phi}(\eta)$, the third term in (S.C.14) may be bounded by $|c_3(\phi_*)|/c_1(\phi_*)\,(1 + c_1(\phi_*)/c_1(\phi)) \xrightarrow{p} 0$, convergence being uniform on $\|\phi - \phi_*\| < \varepsilon$ by pointwise convergence of $c_3(\phi)/\left(c_1(\phi) + c_2(\phi)\right)$, cf. Gupta and Robinson (2018). The uniform convergence to zero of the fourth term in (S.C.14) follows in identical fashion, because $c_2(\phi)$ is bounded and bounded away from zero (see (S.C.8)) in probability for sufficiently large n and all $\phi \in \overline{\mathcal{N}}^{\phi}(\eta)$. This concludes the proof. □

S.D Lemmas

Lemma LS.1. *Under the conditions of Theorem 4.1,* $c_1(\gamma) = n^{-1}\beta'\Psi'E'(\gamma)M(\gamma)E(\gamma)\Psi\beta + o_p(1)$.

Proof. First,

$$c_1(\gamma) = n^{-1}\beta'\Psi'E'(\gamma)M(\gamma)E(\gamma)\Psi\beta + c_{12}(\gamma) + c_{13}(\gamma),$$

with $c_{12}(\gamma) = 2n^{-1}e'E'(\gamma)M(\gamma)E(\gamma)\Psi\beta$ and $c_{13}(\gamma) = n^{-1}e'E'(\gamma)M(\gamma)E(\gamma)e$. It is readily seen that $c_{12}(\gamma)$ and $c_{13}(\gamma)$ are negligible. □

Lemma LS.2. *Under the conditions of Theorem 4.2 or Theorem 5.2,* $\|\widehat{\gamma} - \gamma_0\| = O_p\left(\sqrt{d_\gamma/n}\right)$.

Proof. We show the details for the setting of Theorem 4.2 and omit the details for the setting of Theorem 5.2. Write $l = \partial L(\beta_0, \gamma_0)/\partial\gamma$. By Robinson (1988), we have $\|\widehat{\gamma} - \gamma_0\| = O_p(\|l\|)$. Now $l = \left(l_1, \ldots, l_{d_\gamma}\right)'$, with $l_j = n^{-1}tr\left(\Sigma^{-1}\Sigma_j\right) - n^{-1}\sigma_0^{-2}u'\Sigma^{-1}\Sigma_j\Sigma^{-1}u$. Next, $\mathcal{E}\|l\|^2 = \sum_{j=1}^{d_\gamma}\mathcal{E}\left(l_j^2\right)$ and

$$\mathcal{E}\left(l_j^2\right) = \frac{1}{n^2\sigma_0^4}var\left(u'\Sigma^{-1}\Sigma_j\Sigma^{-1}u\right) = \frac{1}{n^2\sigma_0^4}var\left(\varepsilon'B'\Sigma^{-1}\Sigma_j\Sigma^{-1}B\varepsilon\right) = \frac{1}{n^2\sigma_0^4}var\left(\varepsilon'D_j\varepsilon\right), \qquad \text{(S.D.1)}$$

say. But, writing $d_{j,st}$ for a typical element of the infinite dimensional matrix D_j, we have

$$var\left(\varepsilon'D_j\varepsilon\right) = \left(\mu_4 - 3\sigma_0^4\right)\sum_{s=1}^{\infty} d_{j,ss}^2 + 2\sigma_0^4 tr\left(D_j^2\right) = \left(\mu_4 - 3\sigma_0^4\right)\sum_{s=1}^{\infty} d_{j,ss}^2 + 2\sigma_0^4\sum_{s,t=1}^{\infty} d_{j,st}^2. \qquad \text{(S.D.2)}$$

Next, by Assumptions R.4, R.3 and R.9

$$\sum_{s=1}^{\infty} d_{j,ss}^2 = \sum_{s=1}^{\infty}\left(b_s'\Sigma^{-1}\Sigma_j\Sigma^{-1}b_s\right)^2 \le \left(\sum_{s=1}^{\infty}\|b_s\|^2\right)\left\|\Sigma^{-1}\right\|^2\|\Sigma_j\| = O\left(\sum_{j=1}^{n}\sum_{s=1}^{\infty} b_{js}^{*2}\right) = O(n). \qquad \text{(S.D.3)}$$

Similarly,

$$\sum_{s,t=1}^{\infty} d_{j,st}^2 = \sum_{s=1}^{\infty} b_s' \Sigma^{-1}\Sigma_j \Sigma^{-1} \left(\sum_{t=1}^{\infty} b_t b_t' \right) \Sigma^{-1}\Sigma_j \Sigma^{-1} b_s = \sum_{s=1}^{\infty} b_s' \Sigma^{-1}\Sigma_j \Sigma^{-1}\Sigma_j \Sigma^{-1} b_s = O(n). \quad \text{(S.D.4)}$$

Using (S.D.3) and (S.D.4) in (S.D.2) implies that $\mathcal{E}\left(l_j^2\right) = O\left(n^{-1}\right)$, by (S.D.1). Thus we have $\mathcal{E}\left\|l\right\|^2 = O\left(d_\gamma/n\right)$, and thus $\|l\| = O_p\left(\sqrt{d_\gamma/n}\right)$, by Markov's inequality, proving the lemma. □

Lemma LS.3. *Under the conditions of Theorem 4.3,* $\mathcal{E}\left(\sigma_0^{-2}\varepsilon'\mathscr{V}\varepsilon\right) = p$ *and* $Var\left(\sigma_0^{-2}\varepsilon'\mathscr{V}\varepsilon\right)/2p \to 1$.

Proof. As $\mathcal{E}\left(\sigma_0^{-2}\varepsilon'\mathscr{V}\varepsilon\right) = tr\left(\mathcal{E}[B'\Sigma^{-1}\Psi(\Psi'\Sigma^{-1}\Psi)^{-1}\Psi'\Sigma^{-1}B]\right) = p$, and

$$Var\left(\frac{1}{\sigma_0^2}\varepsilon'\mathscr{V}\varepsilon\right) = \left(\frac{\mu_4}{\sigma_0^4} - 3\right)\sum_{s=1}^{\infty}\mathcal{E}(v_{ss}^2) + \mathcal{E}[tr(\mathscr{V}\mathscr{V}') + tr(\mathscr{V}^2)] = \left(\frac{\mu_4}{\sigma_0^4} - 3\right)\sum_{s=1}^{\infty} v_{ss}^2 + 2p, \quad \text{(S.D.5)}$$

it suffices to show that

$$(2p)^{-1}\sum_{s=1}^{\infty} v_{ss}^2 \xrightarrow{p} 0. \quad \text{(S.D.6)}$$

Because $v_{ss} = b_s'\mathscr{M}b_s$, we have $v_{ss}^2 = \left(\sum_{i,j=1}^{n} b_{is}b_{js}m_{ij}\right)^2$. Thus, using Assumption R.4 and (A.5), we have

$$\begin{aligned}\sum_{s=1}^{\infty} v_{ss}^2 &\leq \left(\sup_{i,j}|m_{ij}|\right)^2 \sum_{s=1}^{\infty}\left(\sum_{i,j=1}^{n}\left|b_{is}^*\right|\left|b_{js}^*\right|\right)^2 = O_p\left(p^2 n^{-2}\left(\sup_s \sum_{i=1}^{n}\left|b_{is}^*\right|\right)^3 \sum_{i=1}^{n}\sum_{s=1}^{\infty}\left|b_{is}^*\right|\right) \\ &= O_p\left(p^2 n^{-1}\right), \end{aligned} \quad \text{(S.D.7)}$$

establishing (S.D.6) because $p^2/n \to 0$. □

Lemma LS.4. *Under the conditions of Theorem 6.2,* $\|\widehat{\tau} - \tau_0\| = O_p\left(\sqrt{d_\tau/n}\right)$.

Proof. The proof is similar to that of Lemma LS.2 and is omitted. □

Denote $H(\gamma) = I_n + \sum_{j=m_1+1}^{m_1+m_2}\gamma_j W_j$ and $K(\gamma) = I_n - \sum_{j=1}^{m_1}\gamma_j W_j$. Let $G_j(\gamma) = W_j K^{-1}(\gamma)$, $j = 1,\ldots,m_1$, $T_j = H^{-1}(\gamma)W_j$, $j = m_1+1,\ldots,m_1+m_2$ and, for a generic matrix A, denote $\overline{A} = A + A'$. Our final conditions may differ according to whether the W_j are of general form or have 'single nonzero diagonal block structure', see e.g Gupta and Robinson (2015). To define these, denote by V an $n\times n$ block diagonal matrix with i-th block V_i, a $s_i \times s_i$ matrix, where $\sum_{i=1}^{m_1+m_2} s_i = n$, and for $i = 1,...,m_1+m_2$ obtain W_j from V by replacing each V_j, $j \neq i$, by a matrix of zeros. Thus $V = \sum_{i=1}^{m_1+m_2} W_j$.

Lemma LS.5. *For the spatial error model with SARMA(p, q) errors, if*

$$\sup_{\gamma\in\Gamma^o}\left(\left\|K^{-1}(\gamma)\right\|+\left\|K'^{-1}(\gamma)\right\|+\left\|H^{-1}(\gamma)\right\|+\left\|H'^{-1}(\gamma)\right\|\right)+\max_{j=1,\ldots,m_1+m_2}\|W_j\|<C, \tag{S.D.8}$$

then

$$(D\Sigma(\gamma))\left(\gamma^\dagger\right)=A^{-1}(\gamma)\left(\sum_{j=1}^{m_1}\gamma_j^\dagger\overline{H^{-1}(\gamma)G_j(\gamma)}+\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^\dagger\overline{T_j(\gamma)}\right)A'^{-1}(\gamma).$$

Proof. We first show that $D\Sigma\in\mathscr{L}\left(\Gamma^o,\mathcal{M}^{n\times n}\right)$. Clearly, $D\Sigma$ is a linear map and (S.D.8)

$$\left\|(D\Sigma(\gamma))\left(\gamma^\dagger\right)\right\|\le C\left\|\gamma^\dagger\right\|_1,$$

in the general case and

$$\left\|(D\Sigma(\gamma))\left(\gamma^\dagger\right)\right\|\le C\max_{j=1,\ldots,m_1+m_2}\left|\gamma_j^\dagger\right|,$$

in the 'single nonzero diagonal block' case. Thus $D\Sigma$ is a bounded linear operator between two normed linear spaces, i.e. it is a continuous linear operator.

With $A(\gamma)=H^{-1}(\gamma)K(\gamma)$, we now show that

$$\frac{\left\|A^{-1}\left(\gamma+\gamma^\dagger\right)A'^{-1}\left(\gamma+\gamma^\dagger\right)-A^{-1}\left(\gamma\right)A'^{-1}\left(\gamma\right)-\left(D\Sigma(\gamma)\right)\left(\gamma^\dagger\right)\right\|}{\left\|\gamma^\dagger\right\|_g}\to 0,\ \text{as}\ \left\|\gamma^\dagger\right\|_g\to 0, \tag{S.D.9}$$

where $\|\cdot\|_g$ is either the 1-norm or the max norm on Γ. First, note that

$$\begin{aligned}
& A^{-1}\left(\gamma+\gamma^\dagger\right)A'^{-1}\left(\gamma+\gamma^\dagger\right)-A^{-1}(\gamma)A'^{-1}(\gamma)\\
= & A^{-1}\left(\gamma+\gamma^\dagger\right)\left(A^{-1}\left(\gamma+\gamma^\dagger\right)-A^{-1}(\gamma)\right)'+\left(A^{-1}\left(\gamma+\gamma^\dagger\right)-A^{-1}(\gamma)\right)A^{-1}(\gamma)\\
= & -A^{-1}\left(\gamma+\gamma^\dagger\right)A'^{-1}\left(\gamma+\gamma^\dagger\right)\left(A\left(\gamma+\gamma^\dagger\right)-A(\gamma)\right)'A'^{-1}(\gamma)\\
- & A^{-1}\left(\gamma+\gamma^\dagger\right)\left(A\left(\gamma+\gamma^\dagger\right)-A(\gamma)\right)A^{-1}(\gamma)A'^{-1}(\gamma).
\end{aligned} \tag{S.D.10}$$

Next,

$$\begin{aligned}
A\left(\gamma+\gamma^\dagger\right)-A(\gamma) &= H^{-1}\left(\gamma+\gamma^\dagger\right)K\left(\gamma+\gamma^\dagger\right)-H^{-1}\left(\gamma\right)K\left(\gamma\right)\\
&= H^{-1}\left(\gamma+\gamma^\dagger\right)\left(K\left(\gamma+\gamma^\dagger\right)-K(\gamma)\right)\\
&+ H^{-1}\left(\gamma+\gamma^\dagger\right)\left(H\left(\gamma\right)-H\left(\gamma+\gamma^\dagger\right)\right)H^{-1}\left(\gamma\right)K\left(\gamma\right)\\
&= -H^{-1}\left(\gamma+\gamma^\dagger\right)\left(\sum_{j=1}^{m_1}\gamma_j^\dagger W_j+\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^\dagger W_jH^{-1}(\gamma)K(\gamma)\right).
\end{aligned} \tag{S.D.11}$$

Substituting (S.D.11) in (S.D.10) implies that

$$A^{-1}\left(\gamma+\gamma^{\dagger}\right)A'^{-1}\left(\gamma+\gamma^{\dagger}\right)-A^{-1}(\gamma)A'^{-1}(\gamma)=\Delta_1\left(\gamma,\gamma^{\dagger}\right)+\Delta_2\left(\gamma,\gamma^{\dagger}\right)=\Delta\left(\gamma,\gamma^{\dagger}\right), \quad \text{(S.D.12)}$$

say, where

$$\begin{aligned}\Delta_1\left(\gamma,\gamma^{\dagger}\right) &= A^{-1}\left(\gamma+\gamma^{\dagger}\right)A'^{-1}\left(\gamma+\gamma^{\dagger}\right)\left(\sum_{j=1}^{m_1}\gamma_j^{\dagger}W_j'+K'(\gamma)H'^{-1}(\gamma)\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^{\dagger}W_j'\right)\\ &\times H'^{-1}\left(\gamma+\gamma^{\dagger}\right)A'^{-1}(\gamma),\\ \Delta_2\left(\gamma,\gamma^{\dagger}\right) &= A^{-1}\left(\gamma+\gamma^{\dagger}\right)H^{-1}\left(\gamma+\gamma^{\dagger}\right)\left(\sum_{j=1}^{m_1}\gamma_j^{\dagger}W_j+\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^{\dagger}W_jH^{-1}(\gamma)K(\gamma)\right)\\ &\times A^{-1}(\gamma)A'^{-1}(\gamma).\end{aligned}$$

From the definitions above and recalling that $A(\gamma)=H^{-1}(\gamma)K(\gamma)$, we can write

$$\Delta\left(\gamma,\gamma^{\dagger}\right)=A^{-1}\left(\gamma+\gamma^{\dagger}\right)\Upsilon\left(\gamma,\gamma^{\dagger}\right)A'^{-1}(\gamma), \quad \text{(S.D.13)}$$

with

$$\begin{aligned}\Upsilon\left(\gamma,\gamma^{\dagger}\right) &= \sum_{j=1}^{m_1}\gamma_j^{\dagger}G_j'\left(\gamma+\gamma^{\dagger}\right)H'^{-1}\left(\gamma+\gamma^{\dagger}\right)+A'^{-1}\left(\gamma+\gamma^{\dagger}\right)A'(\gamma)\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^{\dagger}T_j'\left(\gamma+\gamma^{\dagger}\right)\\ &+ \sum_{j=1}^{m_1}\gamma_j^{\dagger}H^{-1}\left(\gamma+\gamma^{\dagger}\right)G_j(\gamma)+\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^{\dagger}T_j\left(\gamma+\gamma^{\dagger}\right).\end{aligned}$$

Then (S.D.12) implies that

$$\begin{aligned}&A^{-1}\left(\gamma+\gamma^{\dagger}\right)A'^{-1}\left(\gamma+\gamma^{\dagger}\right)-A^{-1}(\gamma)A'^{-1}(\gamma)-(D\Sigma(\gamma))\left(\gamma^{\dagger}\right)\\ =\ &A^{-1}\left(\gamma+\gamma^{\dagger}\right)A'^{-1}\left(\gamma+\gamma^{\dagger}\right)-A^{-1}(\gamma)A'^{-1}(\gamma)-\Delta\left(\gamma,\gamma^{\dagger}\right)-(D\Sigma(\gamma))\left(\gamma^{\dagger}\right)+\Delta\left(\gamma,\gamma^{\dagger}\right)\\ =\ &\Delta\left(\gamma,\gamma^{\dagger}\right)-(D\Sigma(\gamma))\left(\gamma^{\dagger}\right), \quad \text{(S.D.14)}\end{aligned}$$

so to prove (S.D.9) it is sufficient to show that

$$\frac{\left\|\Delta\left(\gamma,\gamma^{\dagger}\right)-(D\Sigma(\gamma))\left(\gamma^{\dagger}\right)\right\|}{\left\|\gamma^{\dagger}\right\|_g}\to 0 \text{ as } \left\|\gamma^{\dagger}\right\|_g\to 0. \quad \text{(S.D.15)}$$

The numerator in (S.D.15) can be written as $\sum_{i=1}^{7}\Pi_i\left(\gamma,\gamma^{\dagger}\right)A'^{-1}(\gamma)$ by adding, subtracting and grouping terms, where (omitting the argument $\left(\gamma,\gamma^{\dagger}\right)$)

$$\begin{aligned}
\Pi_1 &= A^{-1}\left(\gamma+\gamma^{\dagger}\right)\sum_{j=1}^{m_1}\gamma_j^{\dagger}G_j'\left(\gamma+\gamma^{\dagger}\right)H'^{-1}(\gamma)\left(H(\gamma)-H\left(\gamma+\gamma^{\dagger}\right)\right)'H'^{-1}\left(\gamma+\gamma^{\dagger}\right),\\
\Pi_2 &= A^{-1}\left(\gamma+\gamma^{\dagger}\right)\sum_{j=1}^{m_1}\gamma_j^{\dagger}H^{-1}\left(\gamma+\gamma^{\dagger}\right)\left(H(\gamma)-H\left(\gamma+\gamma^{\dagger}\right)\right)H^{-1}(\gamma)G_j\left(\gamma\right),\\
\Pi_3 &= A^{-1}\left(\gamma+\gamma^{\dagger}\right)\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^{\dagger}\left(A^{-1}\left(\gamma+\gamma^{\dagger}\right)-A^{-1}\left(\gamma\right)\right)T_j'\left(\gamma+\gamma^{\dagger}\right),\\
\Pi_4 &= \left(A^{-1}\left(\gamma+\gamma^{\dagger}\right)-A^{-1}\left(\gamma\right)\right)\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^{\dagger}\overline{T_j\left(\gamma+\gamma^{\dagger}\right)},\\
\Pi_5 &= A^{-1}(\gamma)\sum_{j=m_1+1}^{m_1+m_2}\gamma_j^{\dagger}\overline{H^{-1}\left(\gamma+\gamma^{\dagger}\right)\left(H(\gamma)-H\left(\gamma+\gamma^{\dagger}\right)\right)H^{-1}(\gamma)W_j},\\
\Pi_6 &= \Delta\left(\gamma,\gamma^{\dagger}\right)\sum_{j=1}^{m_1}\gamma_j^{\dagger}W_j'H'^{-1}(\gamma),\\
\Pi_7 &= \left(A^{-1}\left(\gamma+\gamma^{\dagger}\right)-A^{-1}\left(\gamma\right)\right)\sum_{j=1}^{m_1}\gamma_j^{\dagger}H^{-1}(\gamma)G_j(\gamma).
\end{aligned}$$

By (S.D.8), (S.D.13) and replication of earlier techniques, we have

$$\max_{i=1,\dots,7}\sup_{\gamma\in\Gamma^o}\left\|\Pi_i\left(\gamma,\gamma^{\dagger}\right)A^{-1}(\gamma)\right\|\le C\left\|\gamma^{\dagger}\right\|_g^2, \tag{S.D.16}$$

where the norm used on the RHS of (S.D.16) depends on whether we are considering the general case or the 'single nonzero diagonal block' case. Thus

$$\frac{\left\|\Delta\left(\gamma,\gamma^{\dagger}\right)-\left(D\Sigma(\gamma)\right)\left(\gamma^{\dagger}\right)\right\|}{\left\|\gamma^{\dagger}\right\|_g}\le C\left\|\gamma^{\dagger}\right\|_g\to 0 \text{ as } \left\|\gamma^{\dagger}\right\|_g\to 0,$$

proving (S.D.15) and thus (S.D.9). □

Corollary CS.1. *For the spatial error model with SAR(m_1) errors,*

$$\left(D\Sigma(\gamma)\right)\left(\gamma^{\dagger}\right)=K^{-1}(\gamma)\sum_{j=1}^{m_1}\gamma_j^{\dagger}\overline{G_j(\gamma)}K'^{-1}(\gamma).$$

Proof. Taking $m_2=0$ in Lemma LS.5, the elements involving sums from m_1+1 to m_1+m_2 do not arise and $H(\gamma)=I_n$, proving the claim. □

Corollary CS.2. *For the spatial error model with SMA(m_2) errors,*

$$(D\Sigma(\gamma))\left(\gamma^{\dagger}\right) = H(\gamma)\sum_{j=1}^{m_2}\gamma_j^{\dagger}\overline{T_j(\gamma)}H'(\gamma).$$

Proof. Taking $m_1 = 0$ in Lemma LS.5, the elements involving sums from 1 to m_1 do not arise and $K(\gamma) = I_n$, proving the claim. □

Lemma LS.6. *For the spatial error model with MESS(m_1) errors, if*

$$\max_{j=1,\ldots,m_1}\left(\|W_j\| + \left\|W_j'\right\|\right) < 1, \tag{S.D.17}$$

then

$$(D\Sigma(\gamma))\left(\gamma^{\dagger}\right) = \exp\left(\sum_{j=1}^{m_1}\gamma_j\left(W_j + W_j'\right)\right)\sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j + W_j'\right).$$

Proof. Clearly $D\Sigma \in \mathscr{L}\left(\Gamma^o, \mathcal{M}^{n\times n}\right)$. Next,

$$\begin{aligned}
&\left\|A^{-1}\left(\gamma+\gamma^{\dagger}\right)A'^{-1}\left(\gamma+\gamma^{\dagger}\right) - A^{-1}(\gamma)A'^{-1}(\gamma) - (D\Sigma(\gamma))\left(\gamma^{\dagger}\right)\right\| \\
=\;& \left\|\exp\left(\sum_{j=1}^{m_1}\left(\gamma_j+\gamma_j^{\dagger}\right)\left(W_j+W_j'\right)\right) - \exp\left(\sum_{j=1}^{m_1}\gamma_j\left(W_j+W_j'\right)\right) - (D\Sigma(\gamma))\left(\gamma^{\dagger}\right)\right\| \\
=\;& \left\|\exp\left(\sum_{j=1}^{m_1}\gamma_j\left(W_j+W_j'\right)\right)\left(\exp\left(\sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j+W_j'\right)\right) - I_n - \sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j+W_j'\right)\right)\right\| \\
\le\;& \left\|\exp\left(\sum_{j=1}^{m_1}\gamma_j\left(W_j+W_j'\right)\right)\right\|\left\|\exp\left(\sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j+W_j'\right)\right) - I_n - \sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j+W_j'\right)\right\| \\
\le\;& C\left\|I_n + \sum_{j=1}^{p}\gamma_j^{\dagger}\left(W_j+W_j'\right) + \sum_{k=2}^{\infty}\left\{\sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j+W_j'\right)\right\}^k - I_n - \sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j+W_j'\right)\right\| \\
\le\;& C\left\|\sum_{k=2}^{\infty}\left\{\sum_{j=1}^{m_1}\gamma_j^{\dagger}\left(W_j+W_j'\right)\right\}^k\right\| \le C\sum_{k=2}^{\infty}\sum_{j=1}^{m_1}\left|\gamma_j^{\dagger}\right|\left\|\left(W_j+W_j'\right)\right\|^k \\
\le\;& C\sum_{k=2}^{\infty}\left\|\gamma^{\dagger}\right\|_g^k,
\end{aligned} \tag{S.D.18}$$

by (S.D.17), without loss of generality, and again the norm used in (S.D.18) depending on whether we are in the general or the 'single nonzero diagonal block' case. Thus

$$\frac{\left\|A^{-1}\left(\gamma+\gamma^{\dagger}\right)A'^{-1}\left(\gamma+\gamma^{\dagger}\right)-A^{-1}(\gamma)A'^{-1}(\gamma)-(D\Sigma(\gamma))\left(\gamma^{\dagger}\right)\right\|}{\|\gamma^{\dagger}\|_g}\leq C\sum_{k=2}^{\infty}\left\|\gamma^{\dagger}\right\|_g^{k-1}\to 0,$$

as $\left\|\gamma^{\dagger}\right\|_g\to 0$, proving the claim. □

Theorem TS.1. *Under the conditions of Theorem 4.4 or 5.3, $\mathscr{T}_n-\mathscr{T}_n^a=o_p(1)$ as $n\to\infty$.*

Proof. It suffices to show that $n\widetilde{m}_n=n\widehat{m}_n+o_p(\sqrt{p})$. As $\widehat{\eta}=y-\widehat{\theta}$, $\widehat{u}=y-\widehat{f}$, and $\widehat{v}=\widehat{\theta}-\widehat{f}$, we have $\widehat{u}=\widehat{\eta}+\widehat{v}$ and

$$\begin{aligned}
n\widetilde{m}_n &= \widehat{\sigma}^{-2}\left(\widehat{u}'\Sigma\left(\widehat{\gamma}\right)^{-1}\widehat{u}-\widehat{\eta}'\Sigma\left(\widehat{\gamma}\right)^{-1}\widehat{\eta}\right)=\widehat{\sigma}^{-2}\left(2\widehat{u}'\Sigma\left(\widehat{\gamma}\right)^{-1}\widehat{v}-\widehat{v}'\Sigma\left(\widehat{\gamma}\right)^{-1}\widehat{v}\right)\\
&= 2n\widehat{m}_n-\widehat{\sigma}^{-2}\left[\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}(u+e)-e+\theta_0-\widehat{f}\right]'\\
&\quad \Sigma\left(\widehat{\gamma}\right)^{-1}\left[\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}(u+e)-e+\theta_0-\widehat{f}\right]\\
&= 2n\widehat{m}_n-\widehat{\sigma}^{-2}u'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}u-\widehat{\sigma}^{-2}\left(\theta_0-\widehat{f}\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(\theta_0-\widehat{f}\right)\\
&\quad +\widehat{\sigma}^{-2}\left(2(\theta_0-\widehat{f})-e\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(I-\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\right)e\\
&\quad -2\widehat{\sigma}^{-2}\left(\theta_0-\widehat{f}\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}u\\
&= 2n\widehat{m}_n-\left(n\widehat{m}_n-\widehat{\sigma}^{-2}\left(A_1+A_2+A_3+A_4\right)\right)-\widehat{\sigma}^{-2}A_4\\
&\quad +\widehat{\sigma}^{-2}\left(2(\theta_0-\widehat{f})-e\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(I-\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\right)e-2\widehat{\sigma}^{-2}A_3\\
&= n\widehat{m}_n+\widehat{\sigma}^{-2}\left(A_1+A_2-A_3\right)\\
&\quad +\widehat{\sigma}^{-2}\left(2(\theta_0-\widehat{f})-e\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(I-\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\right)e. \qquad \text{(S.D.19)}
\end{aligned}$$

In the proof of Theorem 4.2, we have shown that

$$\left|\left(\theta_0-\widehat{f}\right)'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(I-\Psi[\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi]^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\right)e\right|=o_p(\sqrt{p})$$

in the process of proving $|A_2|=o_p(\sqrt{p})$. Along with

$$\begin{aligned}
&\left|e'\Sigma\left(\widehat{\gamma}\right)^{-1}\left(I-\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\right)e\right|\\
\leq\ &\left|e'\Sigma\left(\widehat{\gamma}\right)^{-1}e\right|+\left|e'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\left(\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}\Psi\right)^{-1}\Psi'\Sigma\left(\widehat{\gamma}\right)^{-1}e\right|\\
\leq\ &\|e\|^2\sup_{\gamma\in\Gamma}\left\|\Sigma\left(\gamma\right)^{-1}\right\|+\|e\|^2\sup_{\gamma\in\Gamma}\left\|\Sigma\left(\gamma\right)^{-1}\right\|^2\left\|\frac{1}{n}\Psi\left(\frac{1}{n}\Psi'\Sigma\left(\gamma\right)^{-1}\Psi\right)^{-1}\Psi'\right\|
\end{aligned}$$

$$= O_p\left(\|e\|^2\right) = O_p\left(p^{-2\mu}n\right) = o_p(\sqrt{p}),$$

we complete the proof that $n\widetilde{m}_n = n\widehat{m}_n + o_p(\sqrt{p})$. In the SAR setting of Section 5,

$$\begin{aligned} n\widetilde{m}_n &= \widehat{\sigma}^{-2}\left(\widehat{u}'\Sigma(\widehat{\gamma})^{-1}\widehat{u} - \widehat{\eta}'\Sigma(\widehat{\gamma})^{-1}\widehat{\eta}\right) = \widehat{\sigma}^{-2}\left(2\widehat{u}'\Sigma(\widehat{\gamma})^{-1}\widehat{v} - \widehat{v}'\Sigma(\widehat{\gamma})^{-1}\widehat{v}\right) \\ &= 2n\widehat{m}_n - \widehat{\sigma}^{-2}\left[\Psi\left(\Psi'\Sigma(\widehat{\gamma})^{-1}\Psi\right)^{-1}\Psi'\Sigma(\widehat{\gamma})^{-1}\left(u + e + \sum_{j=1}^{d_\lambda}(\lambda_{j_0} - \widehat{\lambda}_j)W_j y\right) - e + \theta_0 - \widehat{f}\right]' \\ &\quad \Sigma(\widehat{\gamma})^{-1}\left[\Psi\left(\Psi'\Sigma(\widehat{\gamma})^{-1}\Psi\right)^{-1}\Psi'\Sigma(\widehat{\gamma})^{-1}\left(u + e + \sum_{j=1}^{d_\lambda}(\lambda_{j_0} - \widehat{\lambda}_j)W_j y\right) - e + \theta_0 - \widehat{f}\right]. \end{aligned}$$

Compared to the expression in (S.D.19), we have the additional terms

$$-\widehat{\sigma}^{-2}\sum_{j=1}^{d_\lambda}(\lambda_{j_0} - \widehat{\lambda}_j)W_j y'\Sigma(\widehat{\gamma})^{-1}\Psi\left(\Psi'\Sigma(\widehat{\gamma})^{-1}\Psi\right)^{-1}\Psi'\Sigma(\widehat{\gamma})^{-1}\sum_{j=1}^{d_\lambda}(\lambda_{j_0} - \widehat{\lambda}_j)W_j y$$

and

$$-2\widehat{\sigma}^{-2}\sum_{j=1}^{d_\lambda}(\lambda_{j_0} - \widehat{\lambda}_j)W_j y'\Sigma(\widehat{\gamma})^{-1}\Psi\left(\Psi'\Sigma(\widehat{\gamma})^{-1}\Psi\right)^{-1}\Psi'\Sigma(\widehat{\gamma})^{-1}\left(u + \theta_0 - \widehat{f}\right).$$

Both terms are $o_p(\sqrt{p})$ from the orders of A_5 and A_6 in the proof of Theorem 5.2. Hence, in the SAR setting, $n\widetilde{m}_n = n\widehat{m}_n + o_p(\sqrt{p})$ also holds.

We now present similar calculations that justify the validity of our bootstrap test for the SARARMA(m_1,m_2, m_3) model. The bootstrapped test statistic is constructed with

$$n\widehat{m}_n^* = \widehat{v}^{*\prime}\Sigma(\widehat{\gamma}^*)^{-1}\widehat{u}^* = (\widehat{\theta}_n^* - f(x, \widehat{\alpha}_n^*))'\Sigma(\widehat{\gamma}^*)^{-1}\left((I_n - \sum_{k=1}^{m_1}\widehat{\lambda}_k^* W_{1k})y^* - f(x, \widehat{\alpha}_n^*)\right).$$

Let $J_n = (I_n - \frac{1}{n}l_n l_n')$. As $y = S(\lambda)^{-1}(\theta(x) + R(\gamma)\xi)$, we have

$$\begin{aligned} \widetilde{\xi} &= J_n\widehat{\xi} \\ &= J_n\left(\left(\sum_{l=1}^{m_3}\gamma_{3l}W_{3l} + I_n\right)^{-1} + \left(\sum_{l=1}^{m_3}\gamma_{3l}W_{3l} + I_n\right)^{-1}\sum_{l=1}^{m_3}(\gamma_{3l} - \widehat{\gamma}_{3l})W_{3l}\left(\sum_{l=1}^{m_3}\widehat{\gamma}_{3l}W_{3l} + I_n\right)^{-1}\right) \\ &\quad \times\left(I_n - \sum_{l=1}^{m_2}\gamma_{2l}W_{2l} + \sum_{l=1}^{m_2}(\gamma_{2l} - \widehat{\gamma}_{2l})W_{2l}\right)\left(S(\lambda)y - \theta(x) + \sum_{k=1}^{m_1}(\lambda_k - \widehat{\lambda}_k)W_{1k}y + \theta(x) - \psi'\widehat{\beta}\right) \\ &= \xi - \frac{1}{n}l_n l_n'\xi + J_n\left(\sum_{l=1}^{m_3}\gamma_{3l}W_{3l} + I_n\right)^{-1}\left(I_n - \sum_{l=1}^{m_2}\gamma_{2l}W_{2l}\right)\left(\sum_{k=1}^{m_1}(\lambda_k - \widehat{\lambda}_k)W_{1k}y + \theta(x) - \psi'\widehat{\beta}\right) \end{aligned}$$

$$+J_n\left(\sum_{l=1}^{m_3}\gamma_{3l}W_{3l}+I_n\right)^{-1}\sum_{l=1}^{m_2}(\gamma_{2l}-\widehat{\gamma}_{2l})W_{2l}\left(S(\lambda)y-\theta(x)+\sum_{k=1}^{m_1}(\lambda_k-\widehat{\lambda}_k)W_{1k}y+\theta(x)-\psi'\widehat{\beta}\right)$$

$$+J_n\left(\sum_{l=1}^{m_3}\gamma_{3l}W_{3l}+I_n\right)^{-1}\sum_{l=1}^{m_3}(\gamma_{3l}-\widehat{\gamma}_{3l})W_{3l}\left(\sum_{l=1}^{m_3}\widehat{\gamma}_{3l}W_{3l}+I_n\right)^{-1}$$

$$\times\left(I_n-\sum_{l=1}^{m_2}\gamma_{2l}W_{2l}+\sum_{l=1}^{m_2}(\gamma_{2l}-\widehat{\gamma}_{2l})W_{2l}\right)\left(S(\lambda)y-\theta(x)+\sum_{k=1}^{m_1}(\lambda_k-\widehat{\lambda}_k)W_{1k}y+\theta(x)-\psi'\widehat{\beta}\right),$$

which can be written as

$$\widetilde{\xi}=\xi+\sum_{j=1}^{r}\zeta_{1n,j}p_{nj}+\sum_{j=1}^{s}\zeta_{2n,j}Q_{nj}\xi,$$

where p_{nj} is an n-dimensional vector with bounded elements, $Q_{nj}=[q_{nj,i}]$ is an $n\times n$ matrix with bounded row and column sum norms, and $\zeta_{1n,j}$ and $\zeta_{2n,j}$'s are equal to $l_n'\xi/n$, elements of $\lambda_k-\widehat{\lambda}_k$, $\gamma_{2l}-\widehat{\gamma}_{2l}$, $\theta(x)-\psi'\widehat{\beta}$ or their products. This differs from the proof of Lemma 2 in Jin and Lee (2015) in the term $\theta(x)-\psi'\widehat{\beta}$ and potentially increasing order of d_γ. Then, $\zeta_{1n,j}=O_p(\sqrt{p^{1/2}/n}\vee\sqrt{d_\gamma/n})$ and $\zeta_{2n,j}=O_p(\sqrt{p^{1/2}/n}\vee\sqrt{d_\gamma/n})$, instead of $O_p(\sqrt{1/n})$ as in Jin and Lee (2015). Based on this result, the assumptions in Theorem 4 of Su and Qu (2017) hold, so the validility of our bootstrap test directly follows.

□

References

Delgado, M. and P. M. Robinson (2015). Non-nested testing of spatial correlation. *Journal of Econometrics 187*, 385–401.

Gupta, A. and P. M. Robinson (2015). Inference on higher-order spatial autoregressive models with increasingly many parameters. *Journal of Econometrics 186*, 19–31.

Gupta, A. and P. M. Robinson (2018). Pseudo maximum likelihood estimation of spatial autoregressive models with increasing dimension. *Journal of Econometrics 202*, 92–107.

Jin, F. and L. F. Lee (2015). On the bootstrap for Moran's i test for spatial dependence. *Journal of Econometrics 184*, 295–314.

Robinson, P. M. (1988). The stochastic difference between econometric statistics. *Econometrica 56*, 531–548.

	PS			Trig			B-s		
	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
$n = 60$									
$c = 0$	0.01	0.032	0.05	0.01	0.028	0.054	0.02	0.042	0.064
	0.036	0.084	0.122	0.02	0.056	0.084	0.044	0.008	0.11
$c = 3$	0.07	0.156	0.194	0.166	0.248	0.296	0.208	0.302	0.372
	0.454	0.58	0.658	0.172	0.29	0.358	0.166	0.274	0.346
$c = 6$	0.37	0.532	0.644	0.688	0.806	0.854	0.688	0.82	0.884
	0.998	1	1	0.676	0.822	0.866	0.576	0.726	0.81
$n = 100$									
$c = 0$	0.008	0.03	0.044	0.006	0.012	0.028	0.016	0.028	0.042
	0.022	0.052	0.068	0.004	0.028	0.05	0.018	0.048	0.062
$c = 3$	0.352	0.478	0.574	0.27	0.39	0.484	0.376	0.518	0.614
	0.54	0.666	0.744	0.288	0.412	0.508	0.316	0.462	0.544
$c = 6$	0.984	0.99	0.99	0.956	0.986	0.992	0.98	0.992	0.994
	0.998	0.998	0.998	0.948	0.99	0.992	0.956	0.99	0.996
$n = 200$									
$c = 0$	0.002	0.016	0.034	0.002	0.014	0.034	0.038	0.074	0.102
	0.008	0.026	0.048	0.012	0.028	0.036	0.01	0.036	0.074
$c = 3$	0.176	0.29	0.356	0.164	0.256	0.312	0.388	0.354	0.606
	0.34	0.496	0.582	0.144	0.274	0.356	0.168	0.282	0.376
$c = 6$	0.888	0.942	0.96	0.818	0.898	0.934	0.944	0.974	0.986
	0.99	0.998	1	0.816	0.904	0.944	0.862	0.932	0.954

Table OT.1: Rejection probabilities of SARARMA(0,1,0) using asymptotic test $\mathscr{T}_n$ at 1, 5, 10% levels, power series (**PS**), trigonometric (**Trig**) and B-spline (**B-s**) bases. Compactly supported regressors.

	PS			Trig			B-s		
	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
$n = 60$									
$c = 0$	0.01	0.032	0.05	0.01	0.028	0.054	0.06	0.01	0.016
	0.036	0.084	0.122	0.02	0.056	0.084	0.044	0.008	0.116
$c = 3$	0.07	0.156	0.194	0.16	0.252	0.292	0.09	0.138	0.186
	0.454	0.58	0.658	0.174	0.29	0.358	0.166	0.272	0.34
$c = 6$	0.37	0.532	0.644	0.682	0.798	0.85	0.514	0.644	0.714
	0.998	1	1	0.676	0.822	0.866	0.572	0.714	0.8
$n = 100$									
$c = 0$	0.008	0.03	0.044	0.006	0.012	0.026	0	0.004	0.006
	0.022	0.052	0.068	0.006	0.028	0.05	0.018	0.05	0.062
$c = 3$	0.352	0.478	0.574	0.268	0.396	0.486	0.158	0.23	0.288
	0.54	0.666	0.744	0.288	0.412	0.508	0.322	0.466	0.55
$c = 6$	0.984	0.99	0.99	0.958	0.986	0.992	0.918	0.97	0.98
	0.998	0.998	0.998	0.952	0.99	0.992	0.96	0.99	0.998
$n = 200$									
$c = 0$	0.002	0.016	0.034	0.002	0.018	0.038	0	0	0
	0.008	0.026	0.048	0.012	0.028	0.032	0.01	0.036	0.064
$c = 3$	0.176	0.29	0.356	0.156	0.258	0.312	0.022	0.03	0.044
	0.34	0.496	0.582	0.144	0.272	0.352	0.154	0.266	0.352
$c = 6$	0.888	0.942	0.96	0.816	0.908	0.936	0.43	0.522	0.554
	0.99	0.998	1	0.816	0.904	0.944	0.856	0.924	0.944

Table OT.2: Rejection probabilities of SARARMA(0,1,0) using asymptotic test $\mathscr{T}_n{}^a$ at 1, 5, 10% levels, power series (**PS**), trigonometric (**Trig**) and B-spline (**B-s**) bases. Compactly supported regressors.

		PS	$\mathscr{T}_n = \mathscr{T}_n^a$		**Trig**	$\mathscr{T}_n$		**Trig**	$\mathscr{T}_n^a$
	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
$n = 60$									
$c = 0$	0.02	0.05	0.072	0.016	0.038	0.052	0.016	0.038	0.052
	0.038	0.082	0.11	0.038	0.06	0.08	0.038	0.06	0.08
$c = 3$	0.106	0.158	0.224	0.062	0.11	0.146	0.062	0.11	0.146
	0.152	0.25	0.31	0.09	0.158	0.204	0.09	0.158	0.204
$c = 6$	0.552	0.686	0.73	0.234	0.352	0.482	0.236	0.354	0.43
	0.634	0.774	0.82	0.404	0.542	0.642	0.404	0.542	0.642
$n = 100$									
$c = 0$	0.008	0.024	0.036	0.002	0.018	0.036	0.002	0.018	0.036
	0.024	0.05	0.068	0.012	0.026	0.052	0.012	0.026	0.052
$c = 3$	0.162	0.262	0.342	0.142	0.22	0.286	0.142	0.22	0.286
	0.216	0.332	0.408	0.164	0.274	0.35	0.164	0.274	0.35
$c = 6$	0.824	0.894	0.926	0.79	0.868	0.892	0.79	0.866	0.894
	0.888	0.944	0.952	0.862	0.896	0.928	0.862	0.896	0.928
$n = 200$									
$c = 0$	0.006	0.018	0.032	0.008	0.022	0.032	0.008	0.022	0.032
	0.012	0.032	0.068	0.01	0.026	0.046	0.01	0.026	0.046
$c = 3$	0.096	0.182	0.258	0.076	0.152	0.212	0.078	0.15	0.208
	0.126	0.24	0.33	0.098	0.184	0.26	0.098	0.184	0.26
$c = 6$	0.754	0.858	0.892	0.596	0.728	0.794	0.596	0.724	0.79
	0.84	0.918	0.944	0.684	0.794	0.866	0.684	0.792	0.866

Table OT.3: Rejection probabilities of SARARMA(0,1,0) using asymptotic tests $\mathscr{T}_n, \mathscr{T}_n{}^a$ at 1, 5, 10% levels, power series (**PS**) and trigonometric (**Trig**) bases. Unboundedly supported regressors.

		PS	$\mathscr{T}_n{}^* = \mathscr{T}_n^{a*}$		Trig	$\mathscr{T}_n^*$		Trig	$\mathscr{T}_n^{a*}$
	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
$n = 60$									
$c = 0$	0.008	0.058	0.108	0.01	0.046	0.124	0.01	0.046	0.124
	0.008	0.042	0.094	0.006	0.044	0.102	0.006	0.044	0.102
$c = 3$	0.052	0.17	0.318	0.036	0.14	0.21	0.036	0.14	0.21
	0.034	0.16	0.184	0.034	0.132	0.234	0.034	0.132	0.234
$c = 6$	0.35	0.67	0.808	0.16	0.392	0.556	0.16	0.392	0.558
	0.262	0.656	0.794	0.204	0.468	0.66	0.204	0.468	0.66
$n = 100$									
$c = 0$	0.006	0.05	0.102	0.006	0.05	0.11	0.004	0.05	0.112
	0.012	0.054	0.128	0.004	0.044	0.112	0.004	0.044	0.112
$c = 3$	0.13	0.342	0.516	0.128	0.324	0.488	0.126	0.32	0.488
	0.122	0.326	0.498	0.114	0.298	0.474	0.114	0.298	0.474
$c = 6$	0.766	0.932	0.974	0.728	0.92	0.974	0.728	0.92	0.972
	0.774	0.934	0.968	0.732	0.898	0.952	0.732	0.898	0.952
$n = 200$									
$c = 0$	0.03	0.056	0.088	0.028	0.06	0.098	0.028	0.06	0.098
	0.028	0.084	0.128	0.022	0.068	0.118	0.022	0.068	0.118
$c = 3$	0.17	0.346	0.49	0.132	0.286	0.384	0.13	0.288	0.38
	0.178	0.34	0.488	0.128	0.274	0.416	0.128	0.274	0.416
$c = 6$	0.794	0.92	0.966	0.682	0.866	0.93	0.678	0.864	0.93
	0.84	0.936	0.976	0.698	0.888	0.93	0.698	0.888	0.93

Table OT.4: Rejection probabilities of SARARMA(0,1,0) using bootstrap tests $\mathscr{T}_n{}^*, \mathscr{T}_n{}^{a*}$ at 1, 5, 10% levels, power series (**PS**) and trigonometric (**Trig**) bases. Unboundedly supported regressors.