A Supplement to "A spatial extension of generalized autoregressive conditional heteroscedasticity models"

1 Proofs of Theorems 1-3

Here we show the detailed proofs for Theorems 1-3. The asymptotic properties of the first step estimator are proofed in the similar manner to (Lee (2004), Yu et al (2008) and Yang (2018)).

1.1 Proof of Theorem1

The consistency of $\hat{\theta}$ will follow from the uniform convergence of $\frac{1}{n}(\log L_n(\theta) - Q_n(\theta))$ to zero on Θ and the uniqueness identification condition that, for any $\epsilon > 0, \limsup_{n\to\infty} \max_{\theta\in N_\epsilon^c(\theta_0)} \frac{1}{n}(Q_n(\theta) - Q_n(\theta_0)) < 0$, where $N_\epsilon^c(\theta_0)$ is the complement of an open neighborhood of θ_0 in Θ of diameter ϵ (Theorem 3.4 of white (1994)).

First, we shall prove the uniform convergence of $\frac{1}{n}(\log L_n(\theta) - Q_n(\theta))$ to zero on Θ. It is proved in the same way as (Yang (2018)) that $\inf_{\theta\in\Theta} \sigma_n^{*2}(\theta)$ is bounded away from zero and thus $\sup_{\theta\in\Theta} |\hat{\sigma}_n^2(\theta) - \sigma_n^{*2}(\theta)| = o_p(1)$. Following (Lee (2004) and Yu et al (2008)), We can show that $|\log \hat{\sigma}_n^2(\theta) - \log \sigma_n^{*2}(\theta)| = o_p(1)$ uniformly on Θ and thus $\sup_{\theta\in\Theta} \left|\frac{1}{n}(\log L_n(\theta) - Q_n(\theta))\right| = o_p(1)$.

Secondly, we shall prove the identification uniqueness condition. Because the partial derivatives of each term are uniformly bounded, $\frac{1}{n}Q_n(\theta) = \frac{1}{2}(\log 2\pi + 1) - \frac{1}{2}\log \sigma_n^{*2}(\theta) - \frac{1}{n}\log |R_n(\lambda)| + \frac{1}{n}\log |S_n(\theta)|$ is uniformly equicontinuous on Θ. Let an auxiliary process be $Y_n = \lambda W_n Y_n + \rho W_n Y_n + R_n(\lambda) V_n$ where $V_n \sim N(0, \sigma_0^2 I_n)$. The log-likelihood function of the above auxiliary process is given by

$$\log L_{p,n}(\theta, \sigma^2) \quad = \quad -\frac{n}{2}\log(2\pi\sigma^2(\theta)) - \log |R_n(\lambda)| + \log |S_n(\theta)| - \frac{1}{2\sigma^2} Y_n' S_n'(\theta) R_n'^{-1}(\lambda) R_n^{-1}(\lambda) S_n(\theta) Y_n.$$

Let E_p be the expectation under this auxiliary process and $Q_{p,n}(\theta) = \max_{\sigma^2} E_p(\log L_{p,n}(\theta))$. By information Inequality (Ferguson (1996)), $\frac{1}{n}(Q_{p,n}(\theta) - Q_{p,n}(\theta_0) \leq 0$ for all $\theta \in \Theta$. Thus, the identification uniqueness condition holds by contradiction in the same way as (Lee (2004)). The consistency of $\hat{\theta}$ follow form uniform

convergence and the identification uniqueness condition. This completes the proof of the theorem. □

1.2 Proof of Theorem 2

By the Taylor expansion, we have

$$0 = \frac{1}{\sqrt{n}}\frac{\partial \log L_n(\psi_0)}{\partial \psi} + \left(\frac{1}{n}\frac{\partial^2 \log L_n(\bar{\psi}_n)}{\partial \psi \partial \psi'}\right)\sqrt{n}(\hat{\psi}_n - \psi_0),$$

where $\bar{\psi}_n$ lies between $\hat{\psi}_n$ and ψ_0. Thus, the asymptotic normality of $\hat{\psi}_n$ follows if

The asymptotic normality of $\frac{1}{\sqrt{n}}\frac{\partial \log L_n(\psi_0)}{\partial \psi}$ follows from the central limit theorems for linear-quadratic forms in Kelejian and Prucha (2001). Each score function holds the assumptions and the asymptotic normality of each score function follows. Finally, the Cramér-Wold devise (Proposition 6.3.1 of Brockwell and Davis (1991)) leads to the joint asymptotic normality.

Let $D_{\psi\psi}$ be $\frac{1}{n}\frac{\partial^2 \log L_n(\psi_0)}{\partial \psi \partial \psi'} - E\big(\frac{1}{n}\frac{\partial^2 \log L_n(\psi_0)}{\partial \psi \partial \psi'}\big)$. Then, $D_{\psi\psi}$ has the elements:

$$\begin{aligned}
D_{\beta\beta'} &= 0, \\
D_{\beta\sigma^2} &= -\frac{1}{n\sigma_0^4}X_n' R_n'^{-1} V_n, \\
D_{\beta\rho} &= -\frac{1}{n\sigma_0^2}X_n' R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n V_n, \\
D_{\beta\lambda} &= \frac{1}{n\sigma_0^2}X_n'(R_n'^{-1} W_n' R_n'^{-1} + R_n'^{-1} R_n^{-1} W - R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n) V_n, \\
D_{\sigma^2\sigma^2} &= \frac{1}{\sigma_0^4} - \frac{1}{n\sigma_0^6} V_n' V_n, \\
D_{\sigma^2\rho} &= -\frac{1}{n\sigma_0^4}\beta_0' X_n' S_n'^{-1} W_n' R_n'^{-1} V_n - \frac{1}{n\sigma_0^4}(V_n' R_n' S_n'^{-1} W_n' R_n'^{-1} V_n - \sigma_0^2 tr(S_n'^{-1} W_n')), \\
D_{\sigma^2\lambda} &= -\frac{1}{n\sigma_0^4}\beta_0' X_n' S_n'^{-1} W_n' R_n'^{-1} V_n + \frac{1}{n\sigma_0^4}(V_n' W_n' R_n'^{-1} V_n - \sigma_0^2 tr(W_n' R_n'^{-1})) \\
&\quad -\frac{1}{n\sigma_0^4}(V_n' R_n' S_n'^{-1} W_n' R_n'^{-1} V_n - \sigma_0^2 tr(S_n'^{-1} W_n')), \\
D_{\rho\rho} &= -\frac{2}{n\sigma_0}\beta_0' X_n' S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n V_n \\
&\quad -\frac{1}{n\sigma_0^2}(V_n' R_n' S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n V_n - \sigma_0^2 tr(R_n' S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n)), \\
D_{\rho\lambda} &= \frac{1}{n\sigma_0^2}\beta_0' X_n'(S_n'^{-1} W_n' R_n'^{-1} W_n' R_n'^{-1} + S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n - 2 S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n) V_n \\
&\quad +\frac{1}{n\sigma_0^2}(V_n' R_n' S_n'^{-1} W_n' R_n'^{-1} W_n' R_n'^{-1} V_n - \sigma_0^2 tr(S_n'^{-1} W_n' R_n'^{-1} W_n')) \\
&\quad +\frac{1}{n\sigma_0^2}(V_n' R_n' S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n V_n - \sigma_0^2 tr(R_n' S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n)) \\
&\quad -\frac{1}{n\sigma_0^2}(V_n' R_n' S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n V_n - \sigma_0^2 tr(R_n' S_n'^{-1} W_n' R_n'^{-1} R_n^{-1} W_n S_n^{-1} R_n)),
\end{aligned}$$

$$
\begin{aligned}
D_{\lambda\lambda} \;=\; & \frac{1}{n\sigma_0^2}\beta_0' X_n'(2S_n'^{-1}W_n'R_n'^{-1}W_n'R_n'^{-1} + S_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_n - 2S_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_nS_n'^{-1}R_n \\
& -2R_n'^{-1}W_n'R_n'^{-1}W_n'R_n'^{-1} - R_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_n + 2R_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_nS_n^{-1}R_n \\
& +2R_n'^{-1}W_n'R_n'^{-1}W_n'R_n^{-1} + R_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_n - R_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_nS_n^{-1}R_n)V_n \\
& +\frac{2}{n\sigma_0^2}(V_n'R_n'S_n'^{-1}W_n'R_n'^{-1}W_n'R_n'^{-1}V_n - \sigma_0^2 tr(S_n'^{-1}W_n'R_n'^{-1}W_n')) \\
& +\frac{1}{n\sigma_0^2}(V_n'R_n'S_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_nV_n - \sigma_0^2 tr(R_n'S_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_n)) \\
& -\frac{1}{n\sigma_0^2}(V_n'R_n'S_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_nS_n^{-1}R_nV_n - \sigma_0^2 tr(R_n'S_n'^{-1}W_n'R_n'^{-1}R_n^{-1}W_nS_n^{-1}R_n)) \\
& -\frac{2}{n\sigma_0^2}(V_n'W_n'R_n'^{-1}W_n'R_n'^{-1}V_n - \sigma_0^2 tr(W_n'R_n'^{-1}W_n'R_n'^{-1})) \\
& +\frac{1}{n\sigma_0^2}(V_n'W_n'R_n'^{-1}R_n^{-1}W_nV_n - \sigma_0^2 tr(W_n'R_n'^{-1}R_n^{-1}W_n)) \\
& +\frac{1}{n\sigma_0^2}(V_n'W_n'R_n'^{-1}R_n^{-1}W_nS_n^{-1}R_nV_n - \sigma_0^2 tr(W_n'R_n'^{-1}R_n^{-1}W_nS_n^{-1}R_n)).
\end{aligned}
$$

Thus, the elements of $D_{\psi\psi}$ are decomposed into sums of the forms: $\frac{1}{n}X_n'A_n(\theta)V_n$, $\frac{1}{n}\beta_0'X_n'A_n(\theta)V_n$, $\frac{1}{n}(V_n'A_n(\theta)V_n - E(V_n'A_n(\theta)V_n))$ and $\frac{1}{\sigma_0^4} - \frac{1}{n\sigma_0^6}V_n'V_n$, where a matrix $A_n(\theta)$ is uniformly bounded in both row and column sums. Each matrix converges to zero. Therefore, it follow that $\frac{1}{n}\frac{\partial^2 \log L_n(\psi_0)}{\partial\psi\partial\psi'} - E\big(\frac{1}{n}\frac{\partial^2 \log L_n(\psi_0)}{\partial\psi\partial\psi'}\big) \xrightarrow{p} 0$.

Here, $\bar{\sigma}^{-r} = \sigma_0^{-r} + o_p(1), r = 2, 4, 6$ because $\bar{\sigma}^2 \xrightarrow{p} \sigma_0^2$ and σ^r appears in $H_n(\psi) \equiv \frac{\partial^2}{\partial\psi\partial\psi'}\log L_n(\psi)$ multiplicatively, thus it results in an asymptotically negligible error to replace $\bar{\sigma}^2$ by σ_0^2. The elements of the Hessian matrix, $H_n(\psi) \equiv \frac{\partial^2}{\partial\psi\partial\psi'}\log L_n(\psi)$, are decomposed into sums of terms of the forms: $X_n'A_n(\theta)X_n$, $X_n'A_n(\theta)Y_n$, $X_n'A_n(\theta)V(\theta)$, $Y_n'A_n(\theta)Y_n$, $\frac{n}{2\sigma^4} - \frac{1}{\sigma^6}V_n'(\theta)V_n(\theta), Y_n'A_n(\theta)V_n(\theta), V_n'(\theta)A_n(\theta)V_n(\theta)$ and $tr(A_n(\theta))$, where a matrix $A_n(\theta)$ is uniformly bounded in both row and column sums. The differences between each term at $\bar{\psi}$ and ψ_0 converge to zero in probability. Hence, $\frac{1}{n}\frac{\partial^2 \log L_n(\bar{\psi}_n)}{\partial\psi\partial\psi'} - \frac{1}{n}\frac{\partial^2 \log L_n(\psi_0)}{\partial\psi\partial\psi'} \xrightarrow{p} 0$. This completes the proof of the theorem. □

1.3 Proof of Theorem 3

The estimator for α is

$$
\hat{\alpha}_n = (1-\hat{\lambda})\log\left(\frac{1}{n}\sum_{i=1}^{n}\exp\{(R_n^{-1}(\hat{\lambda})[S(\hat{\theta})Y_n - Z_n\hat{\delta}])_i\}\right).
$$

Here,

$$\begin{aligned} S(\hat{\theta})Y_n - Z_n\hat{\delta} &= Y_n - \hat{\lambda}W_nY_n - \hat{\rho}W_nY_n - Z_n\hat{\delta}, \\ &= (\lambda_0 - \hat{\lambda})W_nY_n + (\rho_0 - \hat{\rho})W_nY_n + Z_n(\delta_0 - \hat{\delta}) + \alpha_0\mathbf{1}_n + R_n \log \boldsymbol{\varepsilon}^2, \\ &= D + \alpha_0\mathbf{1}_n + R_n \log \boldsymbol{\varepsilon}^2, \end{aligned}$$

where $D = (\lambda_0 - \hat{\lambda})W_nY_n + (\rho_0 - \hat{\rho})W_nY_n + Z_n(\delta_0 - \hat{\delta})$.

Because $R_n^{-1}(\hat{\lambda})(S(\hat{\theta})Y_n - Z_n\hat{\delta}) = \frac{\alpha_0}{1-\hat{\lambda}}\mathbf{1}_n + R_n^{-1}(\hat{\lambda})D + R_n^{-1}(\hat{\lambda})R_n \log \boldsymbol{\varepsilon}^2$,

$$\frac{1}{n}\sum_{i=1}^{n} \exp\{(R_n^{-1}(\hat{\lambda})[S(\hat{\theta})Y_n - Z_n\hat{\delta}])_i\} = \exp\left(\frac{\alpha_0}{1-\lambda}\right)\frac{1}{n}\sum_{i=1}^{n} \exp\{(R_n^{-1}(\hat{\lambda})D + R_n^{-1}(\hat{\lambda})R_n \log \boldsymbol{\varepsilon}^2)_i\}.$$

Thus,

$$\hat{\alpha} - \alpha_0 = (1-\hat{\lambda}) \log\left(\frac{1}{n}\sum_{i=1}^{n} \exp\{(R_n^{-1}(\hat{\lambda})D + R_n^{-1}(\hat{\lambda})R_n \log \boldsymbol{\varepsilon}^2)_i\}\right). \tag{1}$$

To prove consistency, it is sufficient that the right side of (1) converges to zero in probability.

By the consistency of the first step estimator, $R_n^{-1}(\hat{\lambda})D = o_p(1)$. Similarly, $R_n^{-1}(\hat{\lambda})R_n \log \boldsymbol{\varepsilon}^2 = \log \boldsymbol{\varepsilon}^2 + o_p(1)$. Thus, $R_n^{-1}(\hat{\lambda})D + R_n^{-1}(\hat{\lambda})R_n \log \boldsymbol{\varepsilon}^2 = \log \boldsymbol{\varepsilon}^2 + o_p(1)$.

The variance of ε_i is 1 and the fourth moment of ε_i exists. By the law of large number and the continuous mapping theorem, $\log(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2) = o_p(1)$.

Therefore, $\hat{\alpha} - \alpha_0 = o_p(1)$ and the consistency of the second step estimator is validated. □