Supplementary Appendix for

"A Robust Test for Network Generated Dependence"

Xiaodong Liu   Ingmar R. Prucha

April 25, 2018

B Supplementary Appendix

The supplementary appendix provides explicit proofs for some of the results claimed in the main part of the paper.

B.1 Proofs of Propositions

Proof of Proposition 1. Let $\mathbf{\Gamma} = (\mathbf{I}_q \otimes \widehat{\boldsymbol{\beta}}', \mathbf{I}_q)'$, then in light of (8) and (10) we have

$$\widehat{\mathbf{V}}_* = \mathrm{L}\widehat{\mathbf{V}} = \begin{bmatrix} \widehat{\mathbf{V}}^X \\ \widehat{\mathbf{V}}^U \end{bmatrix} \quad \text{and} \quad \widehat{\mathbf{\Phi}}_* = \mathrm{L}\widehat{\mathbf{\Phi}}\mathrm{L}' = \begin{bmatrix} \widehat{\mathbf{\Phi}}^{XX} & 0 \\ 0 & \widehat{\mathbf{\Phi}}^{UU} \end{bmatrix},$$

and

$$\widehat{\mathbf{V}} = \begin{bmatrix} \mathbf{\Gamma}'\widehat{\mathbf{V}}_* \\ \widehat{\mathbf{V}}_* \end{bmatrix} \quad \text{and} \quad \widehat{\mathbf{\Phi}} = \begin{bmatrix} \mathbf{\Gamma}'\widehat{\mathbf{\Phi}}_*\mathbf{\Gamma} & \mathbf{\Gamma}'\widehat{\mathbf{\Phi}}_* \\ \widehat{\mathbf{\Phi}}_*\mathbf{\Gamma} & \widehat{\mathbf{\Phi}}_* \end{bmatrix}.$$

Next observe that the Moore-Penrose generalized inverse of $\widehat{\mathbf{\Phi}}$ is, as is readily checked, given by

$$\widehat{\mathbf{\Phi}}^+ = \begin{bmatrix} \mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma} & \mathbf{\Gamma}'\mathbf{A} \\ \mathbf{A}\mathbf{\Gamma} & \mathbf{A} \end{bmatrix},$$

where $\mathbf{A} = (\mathbf{I} + \boldsymbol{\Gamma}\boldsymbol{\Gamma}')^{-1}\widehat{\boldsymbol{\Phi}}_*^{-1}(\mathbf{I} + \boldsymbol{\Gamma}\boldsymbol{\Gamma}')^{-1}$. Consequently,

$$
\begin{aligned}
\widehat{\mathbf{V}}'\widehat{\boldsymbol{\Phi}}^{+}\widehat{\mathbf{V}} &= \begin{bmatrix} \boldsymbol{\Gamma}'\widehat{\mathbf{V}}_* \\ \widehat{\mathbf{V}}_* \end{bmatrix}' \begin{bmatrix} \boldsymbol{\Gamma}'\mathbf{A}\boldsymbol{\Gamma} & \boldsymbol{\Gamma}'\mathbf{A} \\ \mathbf{A}\boldsymbol{\Gamma} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Gamma}'\widehat{\mathbf{V}}_* \\ \widehat{\mathbf{V}}_* \end{bmatrix} \\
&= \widehat{\mathbf{V}}_*'(\mathbf{I} + \boldsymbol{\Gamma}\boldsymbol{\Gamma}')\mathbf{A}(\mathbf{I} + \boldsymbol{\Gamma}\boldsymbol{\Gamma}')\widehat{\mathbf{V}}_* \\
&= \widehat{\mathbf{V}}_*'\widehat{\boldsymbol{\Phi}}_*^{-1}\widehat{\mathbf{V}}_*,
\end{aligned}
$$

which proves the claim. □

Proof of Proposition 2. Let $\mathbf{R} = \mathbf{I}_n - \rho_1\mathbf{W}_1 - \cdots - \rho_q\mathbf{W}_q$. If $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$, under the stated assumptions, $\mathbf{u} \sim N(\mathbf{0}, \boldsymbol{\Omega}_u)$ and $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Omega}_u)$, where $\boldsymbol{\Omega}_u = \sigma^2\mathbf{R}^{-1}\mathbf{R}'^{-1}$. The log-likelihood function is given by

$$
\ln L(\delta) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\det(\boldsymbol{\Omega}_u)| - \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Omega}_u^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \tag{B.1}
$$

where $\boldsymbol{\delta} = [\boldsymbol{\rho}', \boldsymbol{\beta}', \sigma^2]'$. The first-order derivatives of the log-likelihood function (B.1) are

$$
\begin{aligned}
\frac{\partial \ln L}{\partial \rho_r} &= \mathrm{tr}[\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{R}^{-1}] - \frac{1}{\sigma^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{R}'\frac{\partial \mathbf{R}}{\partial \rho_r}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \\
\frac{\partial \ln L}{\partial \boldsymbol{\beta}} &= \frac{1}{\sigma^2}\mathbf{X}'\mathbf{R}'\mathbf{R}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \\
\frac{\partial \ln L}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{R}'\mathbf{R}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).
\end{aligned}
$$

Evaluated at the true parameter value, the mean of the second-order derivatives of the log-likelihood function (B.1) are

$$
\begin{aligned}
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \rho_s} &= -\mathrm{tr}[\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \rho_s}\mathbf{R}^{-1}] - \mathrm{tr}[(\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{R}^{-1})'\frac{\partial \mathbf{R}}{\partial \rho_s}\mathbf{R}^{-1}], \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_s \partial \boldsymbol{\beta}} &= \mathbf{0}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_s \partial \sigma^2} &= \frac{1}{\sigma^2}\mathrm{tr}[\frac{\partial \mathbf{R}}{\partial \rho_s}\mathbf{R}^{-1}].
\end{aligned}
$$

Observe that when $\boldsymbol{\rho} = \mathbf{0}$, $\mathbf{R} = \mathbf{I}_n$ and $\frac{\partial}{\partial \rho_r}\mathbf{R} = -\mathbf{W}_r$. Evaluated at the restricted ML estimator $\widetilde{\boldsymbol{\delta}}$,

$$\frac{\partial \ln L}{\partial \rho_r}|_{\widetilde{\boldsymbol{\delta}}} = \frac{1}{\widetilde{\sigma}^2}\widetilde{\mathbf{u}}'\mathbf{W}_r\widetilde{\mathbf{u}}, \qquad \frac{\partial \ln L}{\partial \boldsymbol{\beta}}|_{\widetilde{\boldsymbol{\delta}}} = \mathbf{0}, \qquad \frac{\partial \ln L}{\partial \sigma^2}|_{\widetilde{\boldsymbol{\delta}}} = 0,$$

and

$$\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \rho_s}|_{\widetilde{\boldsymbol{\delta}}} = -2\mathrm{tr}(\overline{\mathbf{W}}_r\overline{\mathbf{W}}_s), \qquad \mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_s \partial \boldsymbol{\beta}}|_{\widetilde{\boldsymbol{\delta}}} = \mathbf{0}, \qquad \mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_s \partial \sigma^2}|_{\widetilde{\boldsymbol{\delta}}} = 0.$$

Then, the LM test statistic is given by

$$\begin{aligned}
\mathrm{LM}_u &= [\frac{\partial \ln L}{\partial \boldsymbol{\delta}'}]_{\widetilde{\boldsymbol{\delta}}}[-\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\delta}\partial \boldsymbol{\delta}'}]_{\widetilde{\boldsymbol{\delta}}}^{-1}[\frac{\partial \ln L}{\partial \boldsymbol{\delta}}]_{\widetilde{\boldsymbol{\delta}}} \\
&= \begin{bmatrix} \frac{\partial \ln L}{\partial \boldsymbol{\rho}}|_{\widetilde{\boldsymbol{\delta}}} \\ \mathbf{0}_{K_x \times 1} \\ 0 \end{bmatrix}' \begin{bmatrix} -\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\rho}\partial \boldsymbol{\rho}'}|_{\widetilde{\boldsymbol{\delta}}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'}|_{\widetilde{\boldsymbol{\delta}}} & -\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta}\partial \sigma^2}|_{\widetilde{\boldsymbol{\delta}}} \\ \mathbf{0} & -\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta}'\partial \sigma^2}|_{\widetilde{\boldsymbol{\delta}}} & -\mathrm{E}\frac{\partial^2 \ln L}{\partial (\sigma^2)^2}|_{\widetilde{\boldsymbol{\delta}}} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial \ln L}{\partial \boldsymbol{\rho}}|_{\widetilde{\boldsymbol{\delta}}} \\ \mathbf{0}_{K_x \times 1} \\ 0 \end{bmatrix} \\
&= [\frac{\partial \ln L}{\partial \boldsymbol{\rho}}]'_{\widetilde{\boldsymbol{\delta}}}[-\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\rho}\partial \boldsymbol{\rho}'}]_{\widetilde{\boldsymbol{\delta}}}^{-1}[\frac{\partial \ln L}{\partial \boldsymbol{\rho}}]_{\widetilde{\boldsymbol{\delta}}} \\
&= \widetilde{\mathbf{V}}^{U\prime}(\widetilde{\boldsymbol{\Phi}}^{UU})^{-1}\widetilde{\mathbf{V}}^{U},
\end{aligned}$$

which proves the claim. □

Proof of Proposition 3. Let $\mathbf{R} = \mathbf{I}_n - \rho_1\mathbf{W}_1 - \cdots - \rho_q\mathbf{W}_q$ and $\mathbf{S} = \mathbf{I}_n - \lambda_1\mathbf{W}_1 - \cdots - \lambda_q\mathbf{W}_q$. If $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$, under the stated assumptions, $\mathbf{y} \sim N(\mathbf{S}^{-1}\boldsymbol{\mu}_y, \boldsymbol{\Omega}_y)$, where $\boldsymbol{\mu}_y = \mathbf{X}\boldsymbol{\beta} + \sum_{r=1}^{q}\mathbf{W}_r\mathbf{X}\boldsymbol{\gamma}_r$ and $\boldsymbol{\Omega}_y = \mathbf{S}^{-1}\boldsymbol{\Omega}_u\mathbf{S}'^{-1}$. The log-likelihood function is given by

$$\ln L(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\det(\boldsymbol{\Omega}_y)| - \frac{1}{2}(\mathbf{y} - \mathbf{S}^{-1}\boldsymbol{\mu}_y)'\boldsymbol{\Omega}_y^{-1}(\mathbf{y} - \mathbf{S}^{-1}\boldsymbol{\mu}_y), \tag{B.2}$$

where $\boldsymbol{\theta} = [\boldsymbol{\rho}', \boldsymbol{\lambda}', \boldsymbol{\gamma}_1', \cdots, \boldsymbol{\gamma}_q', \boldsymbol{\beta}', \sigma^2]'$. The first-order derivatives of the log-likelihood function (B.2) are

$$
\begin{aligned}
\frac{\partial \ln L}{\partial \rho_r} &= \operatorname{tr}[\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{R}^{-1}] - \frac{1}{\sigma^2}(\mathbf{S}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \textstyle\sum_{s=1}^q \mathbf{W}_s\mathbf{X}\boldsymbol{\gamma}_s)'\mathbf{R}'\frac{\partial \mathbf{R}}{\partial \rho_r}(\mathbf{S}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \textstyle\sum_{s=1}^q \mathbf{W}_s\mathbf{X}\boldsymbol{\gamma}_s) \\
\frac{\partial \ln L}{\partial \lambda_r} &= -\operatorname{tr}[\mathbf{W}_r\mathbf{S}^{-1}] + \frac{1}{\sigma^2}(\mathbf{S}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \textstyle\sum_{s=1}^q \mathbf{W}_s\mathbf{X}\boldsymbol{\gamma}_s)'\mathbf{R}'\mathbf{R}\mathbf{W}_r\mathbf{y} \\
\frac{\partial \ln L}{\partial \boldsymbol{\gamma}_r} &= \frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}_r'\mathbf{R}'\mathbf{R}(\mathbf{S}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \textstyle\sum_{s=1}^q \mathbf{W}_s\mathbf{X}\boldsymbol{\gamma}_s) \\
\frac{\partial \ln L}{\partial \boldsymbol{\beta}} &= \frac{1}{\sigma^2}\mathbf{X}'\mathbf{R}'\mathbf{R}(\mathbf{S}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \textstyle\sum_{s=1}^q \mathbf{W}_s\mathbf{X}\boldsymbol{\gamma}_s) \\
\frac{\partial \ln L}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(\mathbf{S}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \textstyle\sum_{s=1}^q \mathbf{W}_s\mathbf{X}\boldsymbol{\gamma}_s)'\mathbf{R}'\mathbf{R}(\mathbf{S}\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \textstyle\sum_{s=1}^q \mathbf{W}_s\mathbf{X}\boldsymbol{\gamma}_s).
\end{aligned}
$$

Evaluated at the true parameter value, the mean of the second-order derivatives of the log-likelihood function (B.2) are

$$
\begin{aligned}
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \rho_s} &= -\operatorname{tr}[\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \rho_s}\mathbf{R}^{-1}] - \operatorname{tr}[(\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{R}^{-1})'\frac{\partial \mathbf{R}}{\partial \rho_s}\mathbf{R}^{-1}], \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \lambda_s} &= \operatorname{tr}[\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{W}_s\mathbf{S}^{-1}\mathbf{R}^{-1}] + \operatorname{tr}[\frac{\partial \mathbf{R}'}{\partial \rho_r}\mathbf{R}\mathbf{W}_s\mathbf{S}^{-1}(\mathbf{R}'\mathbf{R})^{-1}], \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \boldsymbol{\gamma}_s} &= \mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \boldsymbol{\beta}} = \mathbf{0}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \sigma^2} &= \frac{1}{\sigma^2}\operatorname{tr}[\frac{\partial \mathbf{R}}{\partial \rho_r}\mathbf{R}^{-1}],
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \lambda_s} &= -\operatorname{tr}[\mathbf{W}_r\mathbf{S}^{-1}\mathbf{W}_s\mathbf{S}^{-1}] - \operatorname{tr}[(\mathbf{R}\mathbf{W}_r\mathbf{S}^{-1}\mathbf{R}^{-1})'\mathbf{R}\mathbf{W}_s\mathbf{S}^{-1}\mathbf{R}^{-1}] \\
&\quad -\frac{1}{\sigma^2}[\mathbf{R}\mathbf{W}_r\mathbf{S}^{-1}(\mathbf{X}\boldsymbol{\beta} + \textstyle\sum_{r=1}^q \mathbf{W}_r\mathbf{X}\boldsymbol{\gamma}_r)]'[\mathbf{R}\mathbf{W}_s\mathbf{S}^{-1}(\mathbf{X}\boldsymbol{\beta} + \textstyle\sum_{r=1}^q \mathbf{W}_r\mathbf{X}\boldsymbol{\gamma}_r)], \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \boldsymbol{\gamma}_s} &= -\frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}_s'\mathbf{R}'\mathbf{R}\mathbf{W}_r\mathbf{S}^{-1}(\mathbf{X}\boldsymbol{\beta} + \textstyle\sum_{r=1}^q \mathbf{W}_r\mathbf{X}\boldsymbol{\gamma}_r), \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \boldsymbol{\beta}} &= -\frac{1}{\sigma^2}\mathbf{X}'\mathbf{R}'\mathbf{R}\mathbf{W}_r\mathbf{S}^{-1}(\mathbf{X}\boldsymbol{\beta} + \textstyle\sum_{r=1}^q \mathbf{W}_r\mathbf{X}\boldsymbol{\gamma}_r), \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \sigma^2} &= -\frac{1}{\sigma^2}\operatorname{tr}[\mathbf{R}\mathbf{W}_r\mathbf{S}^{-1}\mathbf{R}^{-1}],
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\gamma}_r \partial \boldsymbol{\gamma}_s'} &= -\frac{1}{\sigma^2}\mathbf{X}'\mathbf{W}_r'\mathbf{R}'\mathbf{R}\mathbf{W}_s\mathbf{X}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta} \partial \boldsymbol{\gamma}_r'} &= -\frac{1}{\sigma^2}\mathbf{X}'\mathbf{R}'\mathbf{R}\mathbf{W}_r\mathbf{X}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} &= -\frac{1}{\sigma^2}\mathbf{X}'\mathbf{R}'\mathbf{R}\mathbf{X}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\gamma}_r \partial \sigma^2} &= \mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta} \partial \sigma^2} = \mathbf{0}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial (\sigma^2)^2} &= -\frac{n}{2\sigma^4}.
\end{aligned}
$$

Observe that when $\boldsymbol{\lambda} = \boldsymbol{\rho} = \mathbf{0}$, $\mathbf{S} = \mathbf{R} = \mathbf{I}_n$ and $\frac{\partial}{\partial \rho_r}\mathbf{R} = -\mathbf{W}_r$ for $r = 1, \cdots, q$. Evaluated at the restricted ML estimator $\widehat{\boldsymbol{\theta}}$, we have $\frac{\partial \ln L}{\partial \rho_r}|_{\widehat{\boldsymbol{\theta}}} = \frac{1}{\widehat{\sigma}^2}\widehat{\mathbf{u}}' W_r \widehat{\mathbf{u}}$, $\frac{\partial \ln L}{\partial \lambda_r}|_{\widehat{\boldsymbol{\theta}}} = \frac{1}{\widehat{\sigma}^2}\mathbf{y}' W_r' \widehat{\mathbf{u}}$, $\frac{\partial \ln L}{\partial \boldsymbol{\gamma}_r}|_{\widehat{\boldsymbol{\theta}}} = \frac{1}{\widehat{\sigma}^2}\mathbf{X}' W_r' \widehat{\mathbf{u}}$, $\frac{\partial \ln L}{\partial \boldsymbol{\beta}}|_{\widehat{\boldsymbol{\theta}}} = \mathbf{0}$, $\frac{\partial \ln L}{\partial \sigma^2}|_{\widehat{\boldsymbol{\theta}}} = 0$, and

$$
\begin{aligned}
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \rho_s}|_{\widehat{\boldsymbol{\theta}}} &= \mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \lambda_s}|_{\widehat{\boldsymbol{\theta}}} = -2\mathrm{tr}(\overline{\mathbf{W}}_r \overline{\mathbf{W}}_s), \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \lambda_s}|_{\widehat{\boldsymbol{\theta}}} &= -2\mathrm{tr}(\overline{\mathbf{W}}_r \overline{\mathbf{W}}_s) - \frac{1}{\widehat{\sigma}^2}\widehat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{W}_r'\mathbf{W}_s\mathbf{X}\widehat{\boldsymbol{\beta}}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \boldsymbol{\gamma}_s}|_{\widehat{\boldsymbol{\theta}}} &= -\frac{1}{\widehat{\sigma}^2}\mathbf{X}'\mathbf{W}_s'\mathbf{W}_r\mathbf{X}\widehat{\boldsymbol{\beta}}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \boldsymbol{\beta}}|_{\widehat{\boldsymbol{\theta}}} &= -\frac{1}{\widehat{\sigma}^2}\mathbf{X}'\mathbf{W}_r\mathbf{X}\widehat{\boldsymbol{\beta}}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\gamma}_r \partial \boldsymbol{\gamma}_s'}|_{\widehat{\boldsymbol{\theta}}} &= -\frac{1}{\widehat{\sigma}^2}\mathbf{X}'\mathbf{W}_r'\mathbf{W}_s\mathbf{X}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta} \partial \boldsymbol{\gamma}_r'}|_{\widehat{\boldsymbol{\theta}}} &= -\frac{1}{\widehat{\sigma}^2}\mathbf{X}'\mathbf{W}_r\mathbf{X}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}|_{\widehat{\boldsymbol{\theta}}} &= -\frac{1}{\widehat{\sigma}^2}\mathbf{X}'\mathbf{X}, \\
\mathrm{E}\frac{\partial^2 \ln L}{\partial (\sigma^2)^2}|_{\widehat{\boldsymbol{\theta}}} &= -\frac{n}{2(\widehat{\sigma}^2)^2},
\end{aligned}
$$

$$
\mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \gamma_{k,s}}|_{\widehat{\boldsymbol{\theta}}} = \mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \beta_k}|_{\widehat{\boldsymbol{\theta}}} = \mathrm{E}\frac{\partial^2 \ln L}{\partial \rho_r \partial \sigma^2}|_{\widehat{\boldsymbol{\theta}}} = \mathrm{E}\frac{\partial^2 \ln L}{\partial \lambda_r \partial \sigma^2}|_{\widehat{\boldsymbol{\theta}}} = \mathrm{E}\frac{\partial^2 \ln L}{\partial \gamma_{k,r} \partial \sigma^2}|_{\widehat{\boldsymbol{\theta}}} = \mathrm{E}\frac{\partial^2 \ln L}{\partial \beta_k \partial \sigma^2}|_{\widehat{\boldsymbol{\theta}}} = 0.
$$

Let

$$\mathbf{A}_{11} = -\begin{bmatrix} \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\lambda}\partial\boldsymbol{\lambda}'}|_{\widehat{\boldsymbol{\theta}}} & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\lambda}\partial\partial\boldsymbol{\gamma}_1'}|_{\widehat{\boldsymbol{\theta}}} & \cdots & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\lambda}\partial\boldsymbol{\gamma}_q'}|_{\widehat{\boldsymbol{\theta}}} & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\lambda}\partial\boldsymbol{\rho}'}|_{\widehat{\boldsymbol{\theta}}} \\ \mathrm{E}\frac{\partial^2 \ln L}{\partial\partial\boldsymbol{\gamma}_1\partial\boldsymbol{\lambda}'}|_{\widehat{\boldsymbol{\theta}}} & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\gamma}_1\partial\boldsymbol{\gamma}_1'}|_{\widehat{\boldsymbol{\theta}}} & \cdots & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\gamma}_1\partial\boldsymbol{\gamma}_q'}|_{\widehat{\boldsymbol{\theta}}} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\gamma}_q\partial\boldsymbol{\lambda}'}|_{\widehat{\boldsymbol{\theta}}} & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\gamma}_q\partial\boldsymbol{\gamma}_1'}|_{\widehat{\boldsymbol{\theta}}} & \cdots & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\gamma}_q\partial\boldsymbol{\gamma}_q'}|_{\widehat{\boldsymbol{\theta}}} & \mathbf{0} \\ \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\rho}\partial\boldsymbol{\lambda}'}|_{\widehat{\boldsymbol{\theta}}} & \mathbf{0} & \cdots & \mathbf{0} & \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\rho}\partial\boldsymbol{\rho}'}|_{\widehat{\boldsymbol{\theta}}} \end{bmatrix}$$

$$\mathbf{A}_{21} = -[\mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\beta}\partial\boldsymbol{\lambda}'}|_{\widehat{\boldsymbol{\theta}}}, \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\beta}\partial\boldsymbol{\gamma}_1'}|_{\widehat{\boldsymbol{\theta}}}, \cdots, \mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\beta}\partial\boldsymbol{\gamma}_q'}|_{\widehat{\boldsymbol{\theta}}}, \mathbf{0}]$$

$$\mathbf{A}_{22} = -\mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}|_{\widehat{\boldsymbol{\theta}}},$$

then, the LM test statistic is given by

$$\begin{aligned} \mathrm{LM}_y &= [\frac{\partial \ln L}{\partial\boldsymbol{\theta}'}]_{\widehat{\boldsymbol{\theta}}}[-\mathrm{E}\frac{\partial^2 \ln L}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}]^{+}_{\widehat{\boldsymbol{\theta}}}[\frac{\partial \ln L}{\partial\boldsymbol{\theta}}]_{\widehat{\boldsymbol{\theta}}} \\ &= \begin{bmatrix} \frac{\partial \ln L}{\partial\lambda}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial\boldsymbol{\gamma}_1}|_{\widehat{\boldsymbol{\theta}}} \\ \vdots \\ \frac{\partial \ln L}{\partial\boldsymbol{\gamma}_q}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial\boldsymbol{\rho}}|_{\widehat{\boldsymbol{\theta}}} \\ \mathbf{0}_{K_x\times 1} \\ 0 \end{bmatrix}' \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{21}' & 0 \\ \mathbf{A}_{21} & \mathbf{A}_{22} & 0 \\ 0 & 0 & -\mathrm{E}\frac{\partial^2 \ln L}{\partial(\sigma^2)^2}|_{\widehat{\boldsymbol{\theta}}} \end{bmatrix}^{+} \begin{bmatrix} \frac{\partial \ln L}{\partial\lambda}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial\boldsymbol{\gamma}_1}|_{\widehat{\boldsymbol{\theta}}} \\ \vdots \\ \frac{\partial \ln L}{\partial\boldsymbol{\gamma}_q}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial\boldsymbol{\rho}}|_{\widehat{\boldsymbol{\theta}}} \\ \mathbf{0}_{K_x\times 1} \\ 0 \end{bmatrix}. \end{aligned}$$

Let $\widehat{\mathbf{V}}$ and $\widehat{\boldsymbol{\Phi}}$ be defined as in (8), then using results on the generalized inverse of partitioned matrices given in, e.g., Trenkler and Schipp (1993), we have

$$\mathrm{LM}_y = \begin{bmatrix} \frac{\partial \ln L}{\partial \lambda}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial \gamma_1}|_{\widehat{\boldsymbol{\theta}}} \\ \vdots \\ \frac{\partial \ln L}{\partial \gamma_q}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial \boldsymbol{\rho}}|_{\widehat{\boldsymbol{\theta}}} \end{bmatrix}' \left[\ (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{+} \ \right] \begin{bmatrix} \frac{\partial \ln L}{\partial \lambda}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial \gamma_1}|_{\widehat{\boldsymbol{\theta}}} \\ \vdots \\ \frac{\partial \ln L}{\partial \gamma_q}|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial \ln L}{\partial \boldsymbol{\rho}}|_{\widehat{\boldsymbol{\theta}}} \end{bmatrix} = \widehat{\mathbf{V}}'\widehat{\boldsymbol{\Phi}}^{+}\widehat{\mathbf{V}},$$

observing that $\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} = \widehat{\sigma}^{-4}\widehat{\boldsymbol{\Phi}}$. By Proposition 1 we have $\widehat{\mathbf{V}}'\widehat{\boldsymbol{\Phi}}^{+}\widehat{\mathbf{V}} = (\mathrm{L}\widehat{\mathbf{V}})'(\mathrm{L}\widehat{\boldsymbol{\Phi}}\mathrm{L}')^{-1}(\mathrm{L}\widehat{\mathbf{V}})$, which proves the claim. □

Proof of Proposition 4. Under the maintained assumptions, $\mathrm{E}[n^{-1/2}\mathbf{X}'\mathbf{u}] = \mathbf{0}$, $\mathrm{cov}[n^{-1/2}\mathbf{X}'\mathbf{u}] = n^{-1}\mathbf{X}'\boldsymbol{\Omega}_u\mathbf{X} = O(1)$, which implies $n^{-1/2}\mathbf{X}'\mathbf{u} = O_p(1)$. Furthermore, as $(n^{-1}\mathbf{X}'\mathbf{X})^{-1} = O(1)$, we have $n^{1/2}(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = (n^{-1}\mathbf{X}'\mathbf{X})^{-1}n^{-1/2}\mathbf{X}'\mathbf{u} = O_p(1)$. Let $\mathbf{A}$ be some $n \times n$ matrix whose row and column sums are uniformly bounded in absolute value by some finite constant. Then,

$$\begin{aligned} n^{-1}\widetilde{\mathbf{u}}'\mathbf{A}\widetilde{\mathbf{u}} &= n^{-1}\mathbf{u}'\mathbf{A}\mathbf{u} - n^{-1}\mathbf{u}'\mathbf{A}\mathbf{X}(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) - (\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'n^{-1}\mathbf{X}'\mathbf{A}\mathbf{u} + (\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'n^{-1}\mathbf{X}'\mathbf{A}\mathbf{X}(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\ &= n^{-1}\mathbf{u}'\mathbf{A}\mathbf{u} + o_p(1) \end{aligned}$$

since by standard argumentation $n^{-1}\mathbf{X}'\mathbf{A}\mathbf{u} = o_p(1)$ and $n^{-1}\mathbf{X}'\mathbf{A}\mathbf{X} = O(1)$. Next observe that $n^{-1}\mathrm{E}[\mathbf{u}'\mathbf{A}\mathbf{u}] = n^{-1}\mathrm{tr}[\mathbf{A}\boldsymbol{\Omega}_u]$ is bounded by a finite constant under the maintained assumptions, and

$$\mathrm{cov}[n^{-1}\mathbf{u}'\mathbf{A}\mathbf{u}] = n^{-2}2\mathrm{tr}[\mathbf{C}^2] + n^{-2}\sum_{i=1}^{n} c_{ii}^2[\mathrm{E}(\varepsilon_i/\sigma)^4 - 3]$$

with $\mathbf{C} = [c_{ij}] = \sigma^2\mathbf{R}'^{-1}\overline{\mathbf{A}}\mathbf{R}^{-1}$. Since the elements of $\mathbf{C}$ and $\mathbf{C}^2$ are uniformly bounded in absolute value under the maintained assumptions we have $\mathrm{cov}[n^{-1}\mathbf{u}'\mathbf{A}\mathbf{u}] = o(1)$, and hence by Chebychev's inequality $n^{-1}\widetilde{\mathbf{u}}'\mathbf{A}\widetilde{\mathbf{u}} = n^{-1}\mathrm{E}[\mathbf{u}'\mathbf{A}\mathbf{u}] + o_p(1)$. Using this result with $\mathbf{A} = \mathbf{W}_r$ shows that $n^{-1}\widetilde{\mathbf{V}}^U = n^{-1}\boldsymbol{\mu}^U + o_p(1)$, and using this result with $\mathbf{A} = \mathbf{I}$ shows that $\widetilde{\sigma}^2 = \overline{\sigma}^2 + o_p(1)$. Hence, under the maintained assumptions, $(n^{-1}\widetilde{\mathbf{\Phi}}^{UU})^{-1} - (n^{-1}\mathbf{\Phi}^{UU})^{-1} = o_p(1)$ and $(n^{-1}\widetilde{\mathbf{\Phi}}^{UU})^{-1} = O_p(1)$. Consequently, $n^{-1}\mathcal{I}_u^2(q) = n^{-1}\boldsymbol{\mu}^{U\prime}(\mathbf{\Phi}^{UU})^{-1}\boldsymbol{\mu}^U + \nu_n$ where $\nu_n = o_p(1)$. Let $a = c_\mu^2/C_\phi$ and observe that $n^{-1}\boldsymbol{\mu}^{U\prime}(\mathbf{\Phi}^{UU})^{-1}\boldsymbol{\mu}^U \geq a > 0$. Since $\nu_n = o_p(1)$, there exists an n_ε such that $P(|\nu_n| \geq a/2) \leq \varepsilon$ for all $n \geq n_\varepsilon$. Consequently for all $n \geq n_\varepsilon$,

$$
\begin{aligned}
\Pr(\mathcal{I}_u^2(q) \leq \gamma) &= \Pr(\boldsymbol{\mu}^{U\prime}(\mathbf{\Phi}^{UU})^{-1}\boldsymbol{\mu}^U + n\nu_n \leq \gamma) \\
&\leq \Pr(na + n\nu_n \leq \gamma) \\
&= \Pr(na + n\nu_n \leq \gamma, |\nu_n| < a/2) + \Pr(na + n\nu_n \leq \gamma, |\nu_n| \geq a/2) \\
&\leq \Pr(na/2 \leq \gamma) + \Pr(|\nu_n| \geq a/2) \\
&\leq \Pr(na/2 \leq \gamma) + \varepsilon
\end{aligned}
$$

Now let n_γ be such that $n_\gamma a/2 > \gamma$, then, for all $n \geq \max(n_\varepsilon, n_\gamma)$, $\Pr(na/2 \leq \gamma) = 0$ and thus $\Pr(\mathcal{I}_u^2(q) \leq \gamma) \leq \varepsilon$, which proves that $\lim_{n\to\infty} \Pr(\mathcal{I}_u^2(q) \leq \gamma) = 0$ for any $\gamma > 0$ as claimed. □

Proof of Proposition 5. Under the alternative hypothesis, $\widehat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{d} + \mathbf{M}_X\mathbf{S}^{-1}\mathbf{u}$. Under the maintained assumptions, $n^{-1}\widehat{\mathbf{V}}^X = n^{-1}\underline{\boldsymbol{\mu}}^X + o_p(1)$ and $n^{-1}\underline{\boldsymbol{\mu}}^X = O(1)$ by argumentation as used in the proof of Proposition 4. Let $\mathbf{A}$ be some $n \times n$ matrix whose row and column sums are uniformly bounded in absolute value by some finite constant. Then,

$$
\begin{aligned}
n^{-1}\widehat{\mathbf{u}}'\mathbf{A}\widehat{\mathbf{u}} &= n^{-1}\mathbf{d}'\mathbf{A}\mathbf{d} + 2n^{-1}\mathbf{d}'\overline{\mathbf{A}}\mathbf{M}_X\mathbf{S}^{-1}\mathbf{u} + n^{-1}\mathbf{u}'\mathbf{S}'^{-1}\mathbf{M}_X\mathbf{A}\mathbf{M}_X\mathbf{S}^{-1}\mathbf{u} \\
&= n^{-1}\mathbf{d}'\mathbf{A}\mathbf{d} + 2n^{-1}\mathbf{d}'\overline{\mathbf{A}}\mathbf{M}_X\mathbf{S}^{-1}\mathbf{u} \\
&\quad + n^{-1}\mathbf{u}'\mathbf{S}'^{-1}\mathbf{A}\mathbf{S}^{-1}\mathbf{u} + (n^{-1}\mathbf{u}'\mathbf{S}'^{-1}\mathbf{X})(n^{-1}\mathbf{X}'\mathbf{X})^{-1}(n^{-1}\mathbf{X}'\mathbf{A}\mathbf{X})(n^{-1}\mathbf{X}'\mathbf{X})^{-\mathbf{1}}(n^{-1}\mathbf{X}'\mathbf{S}^{-1}\mathbf{u}) \\
&\quad + (n^{-1}\mathbf{u}'\mathbf{S}'^{-1}\mathbf{X})(n^{-1}\mathbf{X}'\mathbf{X})^{-1}(n^{-1}\mathbf{X}'\mathbf{A}\mathbf{S}^{-1}\mathbf{u}) + (n^{-1}\mathbf{u}'\mathbf{S}'^{-1}\mathbf{A}\mathbf{X})(n^{-1}\mathbf{X}'\mathbf{X})^{-\mathbf{1}}(n^{-1}\mathbf{X}'\mathbf{S}^{-1}\mathbf{u}) \\
&= n^{-1}\mathbf{d}'\mathbf{A}\mathbf{d} + n^{-1}\mathbf{u}'\mathbf{S}'^{-1}\mathbf{A}\mathbf{S}^{-1}\mathbf{u} + o_p(1).
\end{aligned}
$$

since $n^{-1}\mathbf{d}'\overline{\mathbf{A}}\mathbf{M}_X\mathbf{S}^{-1}\mathbf{u} = o_p(1)$, $n^{-1}\mathbf{X}'\mathbf{S}^{-1}\mathbf{u} = o_p(1)$, $n^{-1}\mathbf{X}'\mathbf{A}\mathbf{S}^{-1}\mathbf{u} = o_p(1)$ by argumentation as used in the proof of Proposition 4, and $(n^{-1}\mathbf{X}'\mathbf{X})^{-1} = O(1)$. Under the maintained assumptions,

$n^{-1}\mathbf{d}'\mathbf{A}\mathbf{d} = O(1)$. Let $\mathbf{B} = \mathbf{S}'^{-1}\mathbf{A}\mathbf{S}^{-1}$, and observe that $n^{-1}\mathrm{E}[\mathbf{u}'\mathbf{B}\mathbf{u}] = n^{-1}\mathrm{tr}(\mathbf{B}\boldsymbol{\Omega}_u) = O(1)$ under the maintained assumptions, and

$$\mathrm{cov}[n^{-1}\mathbf{u}'\mathbf{B}\mathbf{u}] = n^{-2}2\mathrm{tr}[\mathbf{C}^2] + n^{-2}\sum_{i=1}^{n} c_{ii}^2[\mathrm{E}(\varepsilon_i/\sigma)^4 - 3]$$

with $\mathbf{C} = [c_{ij}] = \sigma^2\mathbf{R}'^{-1}\overline{\mathbf{B}}\mathbf{R}^{-1}$. Since the elements of $\mathbf{C}$ and $\mathbf{C}^2$ are uniformly bounded in absolute value under the maintained assumptions we have $\mathrm{cov}[n^{-1}\mathbf{u}'\mathbf{B}\mathbf{u}] = o(1)$, and hence by Chebychev's inequality $n^{-1}\widehat{\mathbf{u}}'\mathbf{A}\widehat{\mathbf{u}} = n^{-1}\mathbf{d}'\mathbf{A}\mathbf{d} + n^{-1}\mathrm{tr}(\mathbf{B}\boldsymbol{\Omega}_u) + o_p(1)$. Using this result with $\mathbf{A} = \mathbf{W}_r$ shows that $n^{-1}\widehat{\mathbf{V}}^U = n^{-1}\underline{\boldsymbol{\mu}}^U + o_p(1)$, and using this result with $\mathbf{A} = \mathbf{I}_n$ shows that $\widehat{\sigma}^2 = \underline{\sigma}^2 + o_p(1)$. Hence, under the maintained assumptions, $n^{-1}\mathrm{L}\widehat{\mathbf{V}} = n^{-1}\underline{\boldsymbol{\mu}} + o_p(1)$, $(n^{-1}\mathrm{L}\widehat{\boldsymbol{\Phi}}\mathrm{L}')^{-1} - (n^{-1}\underline{\boldsymbol{\Phi}})^{-1} = o_p(1)$, and $(n^{-1}\mathrm{L}\widehat{\boldsymbol{\Phi}}\mathrm{L}')^{-1} = O_p(1)$. Consequently, $n^{-1}\mathcal{I}_y^2(q) = n^{-1}\underline{\boldsymbol{\mu}}'(\underline{\boldsymbol{\Phi}})^{-1}\underline{\boldsymbol{\mu}} + \nu_n$ where $\nu_n = o_p(1)$. The rest of the proof follows by the same argument as that in the proof of Proposition 4. □

Proof of Proposition 6. The proposition is proven by adapting the argumentation of Lieberman (1994). In the following let $k = p + q$ and $t = t_1 + \cdots + t_k$. It is readily checked that

$$\mathrm{E}\left[\left(\frac{\mathbf{u}'\mathbf{A}\mathbf{u} + \mathbf{a}'\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}}\right)^p \left(\frac{\mathbf{u}'\mathbf{B}\mathbf{u} + \mathbf{b}'\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}}\right)^q\right] = \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} \left.\frac{\partial^k M(t_a, t_b, t_1 + \cdots + t_k)}{\partial t_a^p \partial t_b^q}\right|_{t_a, t_b = 0} dt_1 \cdots dt_k$$

Observe that

$$\left.\frac{\partial^k M(t_a, t_b, t_1 + \cdots + t_k)}{\partial t_a^p \partial t_b^q}\right|_{t_a, t_b = 0} = M^{(k)}(0, 0, t_1 + \cdots + t_k)\exp[h(0, 0, t_1 + \cdots + t_k)],$$

where

$$\begin{aligned} M^{(k)}(0, 0, t_1 + \cdots + t_k) &= \left[\left.\frac{\partial^k M(t_a, t_b, t_1 + \cdots + t_k)}{\partial t_a^p \partial t_b^q}\right|_{t_a, t_b = 0}\right] \Big/ M(0, 0, t_1 + \cdots + t_k), \\ h(0, 0, t_1 + \cdots + t_k) &= \log M(0, 0, t_1 + \cdots + t_k). \end{aligned}$$

Observe that

$$h_l(0,0,t_1+\cdots+t_k) = \frac{\partial h(0,0,t_1+\cdots+t_k)}{\partial t_l} = \frac{\mathrm{E}\left\{\mathbf{u}'\mathbf{S}\mathbf{u}\exp\left[(t_1+\cdots+t_k)\mathbf{u}'\mathbf{S}\mathbf{u}\right]\right\}}{\mathrm{E}\left\{\exp\left[(t_1+\cdots+t_k)\mathbf{u}'\mathbf{S}\mathbf{u}\right]\right\}} > 0$$

since $\mathbf{S}$ is positive definite. Thus, over the range of integration $(-\infty,0]^k$, the function $h(0,0,t_1+\cdots+t_k)$ attains its maximum at the boundary point $t_1=\cdots=t_k=0$. Using the Laplace approximation of integrals (see, e.g., Olver, 1997) yields

$$\begin{aligned}
& \int_{-\infty}^{0}\cdots\int_{-\infty}^{0} \left.\frac{\partial^k M(t_a,t_b,t_1+\cdots+t_k)}{\partial t_a^p \partial t_b^q}\right|_{t_a,t_b=0} dt_1\cdots dt_k \\
= & \int_{-\infty}^{0}\cdots\int_{-\infty}^{0} M^{(k)}(0,0,t_1+\cdots+t_k)\exp[h(0,0,t_1+\cdots+t_k)]dt_1\cdots dt_k \\
\simeq & \; M^{(k)}(0,0,0)\exp[h(0,0,0)]/\Pi_{l=1}^{k} h_l(0,0,0)
\end{aligned}$$

with $h_l(0,0,t_1+\cdots+t_k) = \partial h(0,0,t_1+\cdots+t_k)/\partial t_l$. Next observe that

$$\begin{aligned}
M^{(k)}(0,0,0) &= \left[\left.\frac{\partial^k M(0,0,t_1+\cdots+t_k)}{\partial t_a^p \partial t_b^q}\right|_{t_a,t_b=0}\right] \Big/ M(0,0,0) = \mathrm{E}\left[(\mathbf{u}'\mathbf{A}\mathbf{u}+\mathbf{a}'\mathbf{u})^p(\mathbf{u}'\mathbf{B}\mathbf{u}+\mathbf{b}'\mathbf{u})^q\right], \\
h(0,0,0) &= 0, \\
h_l(0,0,0) &= \left[\left.\frac{\partial M(0,0,t_1+\cdots+t_k)}{\partial t_l}\right|_{t_1=\cdots=t_k=0}\right] \Big/ M(0,0,0) = \mathrm{E}\mathbf{u}'\mathbf{S}\mathbf{u},
\end{aligned}$$

and thus

$$\mathrm{E}\left[\left(\frac{\mathbf{u}'\mathbf{A}\mathbf{u}+\mathbf{a}'\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}}\right)^p\left(\frac{\mathbf{u}'\mathbf{B}\mathbf{u}+\mathbf{b}'\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}}\right)^q\right] \simeq \frac{\mathrm{E}\left[(\mathbf{u}'\mathbf{A}\mathbf{u}+\mathbf{a}'\mathbf{u})^p(\mathbf{u}'\mathbf{B}\mathbf{u}+\mathbf{b}'\mathbf{u})^q\right]}{\left[\mathrm{E}(\mathbf{u}'\mathbf{S}\mathbf{u})\right]^{p+q}}.$$

$\square$

B.2 Derivation of Laplace Approximated Moments

Corresponding to the partitioning of $\widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}} = [\widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}}^{Y\prime}, \widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}}^{X\prime}, \widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}}^{U\prime}]'$ consider the following partitioning of $\boldsymbol{\mu}_L = \mathrm{E}_L[\widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}}]$ and $\boldsymbol{\Phi}_L = \mathrm{E}_L[\widehat{\sigma}_{\mathrm{u}}^{-4}\widehat{\mathbf{V}}\widehat{\mathbf{V}}']$:

$$\boldsymbol{\mu}_L = \begin{bmatrix} \boldsymbol{\mu}_L^Y \\ \boldsymbol{\mu}_L^X \\ \boldsymbol{\mu}_L^U \end{bmatrix}, \qquad \boldsymbol{\Phi}_L = \begin{bmatrix} \boldsymbol{\Phi}_L^{YY} & \boldsymbol{\Phi}_L^{YX} & \boldsymbol{\Phi}_L^{YU} \\ \boldsymbol{\Phi}_L^{XY} & \boldsymbol{\Phi}_L^{XX} & \boldsymbol{\Phi}_L^{XU} \\ \boldsymbol{\Phi}_L^{UY} & \boldsymbol{\Phi}_L^{UX} & \boldsymbol{\Phi}_L^{UU} \end{bmatrix}.$$

The elements of $\widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}}^Y, \widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}}^X$, and $\widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{V}}^U$ corresponding to $\mathbf{W}_r$ are given by

$$\widehat{\sigma}_{\mathrm{u}}^{-2}\mathbf{y}'\mathbf{W}_r'\widehat{\mathbf{u}} = \frac{\mathbf{u}'\mathbf{A}_{y,r}\mathbf{u} + \mathbf{a}_{y,r}'\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}}, \quad \widehat{\sigma}_{\mathrm{u}}^{-2}\mathbf{x}_k'\mathbf{W}_r'\widehat{\mathbf{u}} = \frac{\mathbf{a}_{k,r}'\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}}, \quad \widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{u}}'\mathbf{W}_r'\widehat{\mathbf{u}} = \frac{\mathbf{u}'\mathbf{A}_{u,r}\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}},$$

with $\mathbf{A}_{y,r} = (\mathbf{M}_X\mathbf{W}_r + \mathbf{W}_r'\mathbf{M}_X)/2$, $\mathbf{a}_{y,r} = \mathbf{M}_X\mathbf{W}_r\mathbf{X}\boldsymbol{\beta}$, $\mathbf{a}_{k,r} = \mathbf{M}_X\mathbf{W}_r\mathbf{x}_k$, $\mathbf{A}_{u,r} = \mathbf{M}_X\overline{W}_r\mathbf{M}_X$, and $\mathbf{S} = (n - K_x)^{-1}\mathbf{M}_X$. By Lemma A.1 in Kelejian and Prucha (2010) we have for any conformably symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ and vectors $\mathbf{a}$ and $\mathbf{b}$:

$$\begin{aligned}
&\mathrm{E}(\mathbf{u}'\mathbf{A}\mathbf{u} + \mathbf{a}'\mathbf{u}) = \sigma^2\mathrm{tr}(\mathbf{A}) \\
&\mathrm{E}\left[(\mathbf{u}'\mathbf{A}\mathbf{u} + \mathbf{a}'\mathbf{u})(\mathbf{u}'\mathbf{B}\mathbf{u} + \mathbf{b}'\mathbf{u})\right] = 2\sigma^4\mathrm{tr}(\mathbf{A}\mathbf{B}) + \sigma^2\mathbf{a}'\mathbf{b} \\
&+ \left[\mu^{(4)} - 3\sigma^4\right]\mathrm{vec}_D(\mathbf{A})'\mathrm{vec}_D(\mathbf{B}) + \mu^{(3)}\left[\mathbf{a}'\mathrm{vec}_D(\mathbf{B}) + \mathbf{b}'\mathrm{vec}_D(\mathbf{A})\right] + \sigma^4\mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B})
\end{aligned}$$

Using Proposition 6 and observing furthermore that $\mathrm{E}(\mathbf{u}'\mathbf{S}\mathbf{u}) = \sigma^2$ it is then readily seen that

$$\boldsymbol{\mu}_L^Y = \left[\mathrm{tr}(\overline{\mathbf{W}}_r\mathbf{M}_X)\right]_{r=1,\cdots,q}, \quad \boldsymbol{\mu}_L^X = \mathbf{0}, \quad \boldsymbol{\mu}_L^U = \left[\mathrm{tr}(\overline{\mathbf{W}}_r\mathbf{M}_X)\right]_{r=1,\cdots,q},$$

and

$$
\begin{aligned}
\boldsymbol{\Phi}_L^{YY} &= \Big[\mathrm{tr}(\mathbf{W}_r\mathbf{M}_X\mathbf{W}_s\mathbf{M}_X + \mathbf{W}_r'\mathbf{M}_X\mathbf{W}_s\mathbf{M}_X) + \sigma^{-2}\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}_r'\mathbf{M}_X\mathbf{W}_s\mathbf{X}\boldsymbol{\beta} \\
&\quad +(\sigma^{-4}\mu^{(4)}-3)\mathrm{vec}_D(\mathbf{W}_r'\mathbf{M}_X)'\mathrm{vec}_D(\mathbf{W}_s'\mathbf{M}_X) + \sigma^{-4}\mu^{(3)}[\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}_r'\mathbf{M}_X\mathrm{vec}_D(\mathbf{W}_s'\mathbf{M}_X) \\
&\quad +\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}_s'\mathbf{M}_X\mathrm{vec}_D(\mathbf{W}_r'\mathbf{M}_X)] + \mathrm{tr}(\overline{\mathbf{W}}_r\mathbf{M}_X)\mathrm{tr}(\overline{\mathbf{W}}_s\mathbf{M}_X)\Big]_{r,s=1,\cdots,q}, \\
\boldsymbol{\Phi}_L^{YX} &= \Big[\sigma^{-2}\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}_r'\mathbf{M}_X\mathbf{W}_s\mathbf{X} + \sigma^{-4}\mu^{(3)}\mathrm{vec}_D(\mathbf{W}_r'\mathbf{M}_X)\mathbf{M}_X\mathbf{W}_s\mathbf{X}\Big]_{r,s=1,\cdots,q}, \\
\boldsymbol{\Phi}_L^{YU} &= \Big[2\mathrm{tr}(\overline{\mathbf{W}}_r\mathbf{M}_X\overline{\mathbf{W}}_s\mathbf{M}_X) + (\sigma^{-4}\mu^{(4)}-3)\mathrm{vec}_D(\mathbf{W}_r'\mathbf{M}_X)\mathrm{vec}_D(\mathbf{M}_X\overline{\mathbf{W}}_s\mathbf{M}_X) \\
&\quad +\sigma^{-4}\mu^{(3)}\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}_r'\mathbf{M}_X\mathrm{vec}_D(\mathbf{M}_X\overline{\mathbf{W}}_s\mathbf{M}_X) + \mathrm{tr}(\overline{\mathbf{W}}_r\mathbf{M}_X)\mathrm{tr}(\overline{\mathbf{W}}_s\mathbf{M}_X)\Big]_{r,s=1,\cdots,q},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\Phi}_L^{XX} &= \big[\sigma^{-2}\mathbf{X}'\mathbf{W}_r'\mathbf{M}_X\mathbf{W}_s\mathbf{X}\big]_{r,s=1,\cdots,q}, \\
\boldsymbol{\Phi}_L^{XU} &= \Big[\sigma^{-4}\mu^{(3)}\mathbf{X}'\mathbf{W}_r'\mathbf{M}_X\mathrm{vec}_D(\mathbf{M}_X\overline{\mathbf{W}}_s\mathbf{M}_X)\Big]_{r,s=1,\cdots,q}, \\
\boldsymbol{\Phi}_L^{UU} &= [2\mathrm{tr}(\overline{\mathbf{W}}_r\mathbf{M}_X\overline{\mathbf{W}}_s\mathbf{M}_X) + (\sigma^{-4}\mu^{(4)}-3)\mathrm{vec}_D(\mathbf{M}_X\overline{\mathbf{W}}_r\mathbf{M}_X)'\mathrm{vec}_D(\mathbf{M}_X\overline{\mathbf{W}}_s\mathbf{M}_X) \\
&\quad +\mathrm{tr}(\overline{\mathbf{W}}_r\mathbf{M}_X)\mathrm{tr}(\overline{\mathbf{W}}_s\mathbf{M}_X)]_{r,s=1,\cdots,q}.
\end{aligned}
$$

Note that $\boldsymbol{\mu}_L$ does not depend on any unknown parameters, and hence $\widehat{\boldsymbol{\mu}}_L = \boldsymbol{\mu}_L$. The estimator for $\widehat{\boldsymbol{\Phi}}_L$ is obtained by replacing $\boldsymbol{\beta}$ by the OLS estimator and σ^2 by $\widehat{\sigma}^2$ or $\widehat{\sigma}_{\mathrm{u}}^2$, and $\mu^{(3)}$ and $\mu^{(4)}$ by $n^{-1}\sum_{i=1}^n \widehat{u}_i^3$ and $n^{-1}\sum_{i=1}^n \widehat{u}_i^3$, respectively.

Remark B.1. Cliff and Ord (1981) gives results on the exact mean and variance (and higher moments) of ratios of quadratic forms under normality. Drukker and Prucha (2013) also give results on the covariance of ratios of quadratic forms under normality. We now can use those results to check on the approximation error for the Laplace approximation when the disturbances are normally distributed. In particular, for spatial weight matrices $\mathbf{W}_r$ and $\mathbf{W}_s$, let

$$
Q_r = \widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{u}}'\mathbf{W}_r\widehat{\mathbf{u}} = \frac{\mathbf{u}'\mathbf{A}_r\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}}, \qquad \text{and} \qquad Q_s = \widehat{\sigma}_{\mathrm{u}}^{-2}\widehat{\mathbf{u}}'\mathbf{W}_s\widehat{\mathbf{u}} = \frac{\mathbf{u}'\mathbf{A}_s\mathbf{u}}{\mathbf{u}'\mathbf{S}\mathbf{u}},
$$

where $\mathbf{A}_r = \mathbf{M}_X\overline{\mathbf{W}}_r\mathbf{M}_X$, $\mathbf{A}_s = \mathbf{M}_X\overline{\mathbf{W}}_s\mathbf{M}_X$ and $\mathbf{S} = (n-K_x)^{-1}\mathbf{M}_X$. Provided $\mathbf{u} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$ it follows from Drukker and Prucha (2013) that

$$\mathrm{E}\left[Q_r^p Q_s^q\right] = \frac{\mathrm{E}\left[(\mathbf{u}'\mathbf{A}_r\mathbf{u})^p(\mathbf{u}'\mathbf{A}_s\mathbf{u})^q\right]}{\mathrm{E}\left[(\mathbf{u}'\mathbf{S}\mathbf{u})^{p+q}\right]},$$

for $p, q = 0, 1$. Let E_L denote the Laplace approximation of the expected value, then in light of Proposition 6,

$$\begin{aligned}
\frac{\mathrm{E}_L\left[Q_r\right]}{\mathrm{E}\left[Q_r\right]} &= \frac{\mathrm{E}(\mathbf{u}'\mathbf{S}\mathbf{u})}{\mathrm{E}(\mathbf{u}'\mathbf{S}\mathbf{u})} = 1, \\
\frac{\mathrm{E}_L\left[Q_r Q_s\right]}{\mathrm{E}\left[Q_r Q_s\right]} &= \frac{\mathrm{E}\left[(\mathbf{u}'\mathbf{S}\mathbf{u})^2\right]}{\left[\mathrm{E}(\mathbf{u}'\mathbf{S}\mathbf{u})\right]^2} = 1 + 2(n-K_x)^{-1} = 1 + O(n^{-1})
\end{aligned}$$

observing that $\mathrm{E}(\mathbf{u}'\mathbf{S}\mathbf{u}) = \sigma^2$ and $\mathrm{E}\left[(\mathbf{u}'\mathbf{S}\mathbf{u})^2\right] = 2\sigma^4\mathrm{tr}(\mathbf{S}^2) + \sigma^4\mathrm{tr}(\mathbf{S}) = [2(n-K_x)^{-1} + 1]\sigma^4$.

References

Cliff, A. and Ord, J. (1981). *Spatial Processes, Models and Applications*, Pion, London.

Drukker, D. M. and Prucha, I. R. (2013). Small sample properties of the $I^2(q)$ test statistic for spatial dependence, *Spatial Economic Analysis* **8**: 271–292.

Kelejian, H. H. and Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics* **157**: 53–67.

Lieberman, O. (1994). A Laplace approximation to the moments of a ratio of quadratic forms, *Biometrika* **81**: 681–690.

Olver, F. W. J. (1997). *Asymptotics and Special Functions*, CRC Press, New York.

Trenkler, G. and Schipp, B. (1993). Generalized inverses of partitioned matrices, *Econometric Theory* **9**: 530–533.