Supplementary Appendix to

Semiparametric Estimation and Testing of Smooth Coefficient Spatial Autoregressive Models*

Emir Malikov[1] Yiguo Sun[2]

[1]Department of Agricultural Economics & Rur. Soc., Auburn University, Auburn, AL
[2]Department of Economics and Finance, University of Guelph, ON

First Draft: December 15, 2014
This Draft: February 25, 2017

Abstract

This paper considers a flexible semiparametric spatial autoregressive (mixed-regressive) model in which unknown coefficients are permitted to be nonparametric functions of some contextual variables to allow for potential nonlinearities and parameter heterogeneity in the spatial relationship. Unlike other semiparametric spatial dependence models, ours permits the spatial autoregressive parameter to meaningfully vary across units and thus allows the identification of a neighborhood-specific spatial dependence measure conditional on the vector of contextual variables. We propose several (locally) nonparametric GMM estimators for our model. The developed two-stage estimators incorporate both the linear and quadratic orthogonality conditions and are capable of accommodating a variety of data generating processes, including the instance of a pure spatially autoregressive semiparametric model with no relevant regressors as well as multiple partially linear specifications. All proposed estimators are shown to be consistent and asymptotically normal. We also contribute to the literature by putting forward two test statistics to test for parameter constancy in our model. Both tests are consistent.

Keywords: Consistent Test, Constrained Estimation, Local Linear Fitting, Nonparametric GMM, Partially Linear, Quadratic Moments, SAR, Spatial Lag

JEL Classification: C13, C14, C21

**Email*: emalikov@auburn.edu (Malikov), yisun@uoguelph.ca (Sun).

S1 Brief Mathematical Proofs of Theorems 5 and 6

Proof of Theorem 5. The proof of this theorem closely follows that of Theorem 1. Although the notation is the same, the represented variables can be different. Specifically, $\boldsymbol{\theta}_n = \xi_n \left[\begin{array}{cc} \widehat{\boldsymbol{\gamma}}(z) - \boldsymbol{\gamma}(z) & \left[\triangledown\widehat{\boldsymbol{\gamma}}(z) - \triangledown\boldsymbol{\gamma}(z)\right]\mathbf{H} \end{array} \right]$ is now a $(p_2+1)\times(q+1)$ matrix, $y_i^* = \dot{y}_i - \mathbf{m}_i'\boldsymbol{\gamma}(z) - \mathbf{m}_i'\triangledown\boldsymbol{\gamma}(z)(\mathbf{z}_i - z) = u_i\mathbf{m}_i'\boldsymbol{\Pi}(z_i^*)/2$ with $\mathbf{m}_i = \left[\left(\sum_{j\neq i} w_{ij}y_j\right), \mathbf{x}_{2i}'\right]'$, and $\boldsymbol{\Pi}(z_i^*)$ becomes a $(p_2+1)\times 1$ vector with its lth element being equal to $\Pi_l(\mathbf{z}_i^*) = (\mathbf{z}_i - z)'\triangledown^2\boldsymbol{\gamma}_l(\mathbf{z}_i^*)(\mathbf{z}_i - z)$ and $\mathbf{z}_i^*$ lying between $\mathbf{z}_i$ and z for $l = 1,\ldots,p_2+1$. Then, we define $\Lambda_n(\boldsymbol{\theta},\boldsymbol{\vartheta}) = \mathbf{g}_n(\boldsymbol{\theta},\boldsymbol{\vartheta})'\mathbf{g}_n(\boldsymbol{\theta},\boldsymbol{\vartheta})$, where $\boldsymbol{\vartheta}_n \equiv \xi_n\left[\widehat{\boldsymbol{\beta}}_1(z) - \boldsymbol{\beta}_1\right]$, $\mathbf{g}_n(\boldsymbol{\theta},\boldsymbol{\vartheta})$ is the same as $\mathbf{g}_n(\boldsymbol{\theta})$ defined in the proof of Theorem 1 with $\mathbf{u}(\boldsymbol{\theta})$ being replaced with the newly defined $\mathbf{u}(\boldsymbol{\theta},\boldsymbol{\vartheta})$, where the latter is an $n\times 1$ vector with a typical element being equal to $u_i(\boldsymbol{\theta},\boldsymbol{\vartheta}) = y_i^* - \xi_n^{-1}\mathbf{x}_{1i}'\boldsymbol{\vartheta} - \xi_n^{-1}\mathbf{m}_i'\boldsymbol{\theta}\mathcal{Z}_i(z)$. Again, $\{\xi_n\}$ is a sequence of positive constants such that $0 < M_1 < \|\boldsymbol{\theta}_n\|, \|\boldsymbol{\vartheta}_n\| < M_2 < \infty$ for all n.

Define $\widetilde{\mathbf{M}}(z) \equiv [\mathbf{X}_1, \mathbf{M}(z)]$ and $\boldsymbol{\alpha}_n \equiv \left[\mathrm{vec}(\boldsymbol{\vartheta}_n)', \mathrm{vec}(\boldsymbol{\theta}_n)'\right]'$. Then, following closely the proof of Theorem 1, we have $\boldsymbol{\alpha}_n = B_n^{-1}(z)A_n(z)'$, where $A_n(z) = A_{n1}(z) + A_{n2}(z) - A_{n3}(z)$ and $B_n(z)$ are as defined in the proof of Theorem 1 except for $\mathbf{M}(z)$ being replaced with $\widetilde{\mathbf{M}}(z)$, $\mathbf{C}(z)$ and $\Gamma_{j,l}$ $(j = 1,2,3)$ being calculated from the newly defined $\mathbf{m}_i$ and $\boldsymbol{\Pi}(z_i^*)$, $\Psi_{1,l}$, $\Psi_{2,l}$ and $\Psi_{3,l}$ being replaced with $\widetilde{\Psi}_{1,l} = \mathbf{u}'\left[\mathbf{P}_{n,l}\mathbf{K}_H(z)\right]^s\widetilde{\mathbf{M}}(z)$, $\widetilde{\Psi}_{2,l} = \mathbf{C}(z)'\left[\mathbf{P}_{n,l}\mathbf{K}_H(z)\right]^s\widetilde{\mathbf{M}}(z)$ and $\widetilde{\Psi}_{3,l} = \widetilde{\mathbf{M}}(z)'\left[\mathbf{P}_{n,l}\mathbf{K}_H(z)\right]^s\widetilde{\mathbf{M}}(z)$, respectively.

In Lemmas 8–10, we show that (A.2)–(A.3) continue to hold with the newly defined $\varkappa_A(\mathbf{H},z)$, $\varkappa_B(\mathbf{H},z)$ and $\boldsymbol{\Omega}(z)$ and that ξ_n is of order $\sqrt{n|\mathbf{H}|}$.

Next, letting $\mathbf{S}_{p_1,p_1+(q+1)(p_2+1)}$ be the first p_1 rows of the identity matrix $\mathbf{I}_{p_1+(q+1)(p_2+1)}$, under Assumption 5 we have

$$\begin{aligned}
\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1 &= \frac{1}{n\sqrt{n|\mathbf{H}|}}\sum_{i=1}^n \mathbf{S}_{p_1,p_1+(q+1)(p_2+1)}B_{n,-i}^{-1}(\mathbf{z}_i)A_{n,-i}(\mathbf{z}_i) \\
&\approx \frac{1}{n}\sum_{i=1}^n \mathbf{S}_{p_1,p_1+(q+1)(p_2+1)}\varkappa_B(\mathbf{H},z_i)^{-1}\varkappa_A(\mathbf{H},\mathbf{z}_i)' + \\
&\quad \frac{1}{n(n|\mathbf{H}|)^{3/2}}\sum_{i=1}^n f^{-2}(\mathbf{z}_i)\mathbf{S}_{p_1,p_1+(q+1)(p_2+1)}\varkappa_B(\mathbf{H},\mathbf{z}_i)^{-1}A_{n2,-i}(\mathbf{z}_i) \\
&= \frac{1}{n(n|\mathbf{H}|)^{3/2}}\sum_{i=1}^n f^{-2}(\mathbf{z}_i)\mathbf{S}_{p_1,p_1+(q+1)(p_2+1)}\varkappa_B(\mathbf{H},\mathbf{z}_i)^{-1}A_{n2,-i}(\mathbf{z}_i) + O_p\left(\|\mathbf{H}\|^2\right).
\end{aligned}$$

Note that the "$-i$" subscript indicates that the observation from the ith unit is not being used for calculation. By Lemma 11, we then obtain

$$\sqrt{n}\left(\widetilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1\right) \xrightarrow{d} \mathbb{N}\left(\mathbf{0}_{p_1}, \boldsymbol{\Omega}_0\right)$$

if $n\|\mathbf{H}\|^4 \to 0$. This completes the proof of this theorem. ■

Lemma 8 *Under Assumptions 1–3 and 9, we obtain* $(n|\mathbf{H}|)^{-2}\xi_n A_{n1}(z) = f^2(z)\varkappa_A(\mathbf{H},z)' + o_p\left(\|\mathbf{H}\|^2\right)$.
(i) If $\boldsymbol{\beta}_2(z) \neq \mathbf{0}_p$ holds over at least one non-empty subset,

$$\varkappa_{A,Q}(\mathbf{H},z) = \frac{\mu_{1,2}(k)}{2}\left[\begin{array}{cc} \mathbf{E}_2(\mathbf{H},z)\mathbf{E}_1(z) & \mathbf{0}_{q(p_2+1)}' \end{array}\right] \tag{S1.1}$$

$$\varkappa_{A,P}\left(\mathbf{H},z\right) = \sum_{l=1}^{m}\left(\widetilde{\mathbf{F}}_{3,l}\left(\mathbf{H},z\right)\left[\widetilde{\mathbf{F}}_{1,l}\left(\mathbf{H},z\right)+\frac{1}{4}\widetilde{\mathbf{F}}_{2,l}\left(z\right)\right]+2\mathbf{F}_{4,l}\left(\mathbf{H},z\right)\widetilde{\mathbf{F}}_{1,l}\left(\mathbf{H},z\right)\right); \quad \text{(S1.2)}$$

(ii) If $\boldsymbol{\beta}_1 = \mathbf{0}_{p_1}$ *and* $\boldsymbol{\beta}_2\left(z\right) = \mathbf{0}_{p_2}$ *hold over their domains but* $\rho\left(z\right)$ *is not constant over at least one non-empty subset, we have* $\varkappa_{A,Q}\left(\mathbf{H},z\right) = O_p\left(\left\|\mathbf{H}\right\|^2\left(n\left|\mathbf{H}\right|\right)^{-1/2}\right)$ *and* $\varkappa_{A,P}\left(\mathbf{H},z\right) = 2\sum_{l=1}^{m}\mathbf{F}_{4,l}\left(\mathbf{H},z\right)\widetilde{\mathbf{F}}_{1,l}\left(\mathbf{H},z\right)$*;*
(iii) If $\boldsymbol{\beta}_1 = \mathbf{0}_{p_1}$*,* $\boldsymbol{\beta}_2\left(z\right) = \mathbf{0}_{p_2}$ *and* $\rho\left(z\right) \equiv \rho_0$ *hold over their domains, we have* $\varkappa_{A,P}\left(\mathbf{H},z\right) = O_p\left(\left(n\sqrt{\left|\mathbf{H}\right|}\right)^{-1}\right)$.

Proof. Note that $\Gamma_{1,l}$ and $\Gamma_{3,l}$ have the same limit results as in the proof of Lemma 1.

First, we consider case *(i)*. Applying straightforward calculation, we obtain

$$\begin{aligned}\frac{1}{n\left|\mathbf{H}\right|}\mathbf{Q}(z)'\mathbf{K}_H(z)\widetilde{\mathbf{M}}(z) &= \frac{1}{n\left|\mathbf{H}\right|}\sum_{i=1}^{n}\pi_i\left[\mathcal{Z}_i(z)\otimes\mathbf{Q}_{n,i}\right]\left[\mathbf{x}_{1i}',\mathcal{Z}_i(z)'\otimes\mathbf{m}_i'\right] \\ &= f\left(z\right)\begin{bmatrix}\mathbf{E}_{1x}\left(z\right) & \mathbf{E}_{1m}\left(z\right) & \mathbf{0}_{d\times(q(p_2+1))} \\ \mathbf{0}_{(dq)\times p_1} & \mathbf{0}_{dq\times(p_2+1)} & \mu_{1,2}\left(k\right)\mathbf{I}_q\otimes\mathbf{E}_{1m}\left(z\right)\end{bmatrix}+ \\ &\quad O_p\left(\left\|\mathbf{H}\right\|^2+\left(n\left|\mathbf{H}\right|\right)^{-1/2}\right), \end{aligned} \quad \text{(S1.3)}$$

where $\mathbf{E}_{1x}\left(z\right)\equiv n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[\mathbf{Q}_{n,i}\mathbf{x}_{1i}'|\mathbf{z}_i=z\right]$, $\mathbf{E}_{1m}\left(z\right)\equiv n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[\mathbf{Q}_{n,i}\mathbf{m}_i'|\mathbf{z}_i=z\right]$ and

$$\mathbf{E}_1\left(z\right)\equiv\begin{bmatrix}\mathbf{E}_{1x}\left(z\right) & \mathbf{E}_{1m}\left(z\right)\end{bmatrix}. \quad \text{(S1.4)}$$

Also, we have

$$\begin{aligned}\mathbf{E}_2\left(\mathbf{H},z\right) \equiv\; & n^{-1}\sum_{i=1}^{n}\left(\operatorname{tr}\left\{\mathbf{H}\frac{\partial^2\rho(z)}{\partial z\partial z'}\mathbf{H}\right\}\mathbb{E}\left[\left(\mathbf{W}\mathbf{y}\right)_i\mathbf{Q}_{n,i}'|\mathbf{z}_i=z\right]+\right. \\ & \left.\mathbb{E}\left[\operatorname{tr}\left\{\mathbf{H}\frac{\partial^2\left[\boldsymbol{\beta}_2(z)'\mathbf{x}_{2,i}\right]}{\partial z\partial z'}\mathbf{H}\right\}\mathbf{Q}_{n,i}'\middle|\mathbf{z}_i=z\right]\right). \end{aligned} \quad \text{(S1.5)}$$

Further, for any given l, we have

$$\begin{aligned}\frac{\widetilde{\Psi}_{1,l}}{n\left|\mathbf{H}\right|} &= \frac{1}{n\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(p_{l,ij}\pi_j+p_{l,ji}\pi_i\right)u_i\left[\mathbf{x}_{1i}',\left(\mathcal{Z}_j'(z)\otimes\mathbf{m}_j'\right)\right] \\ &= 2f\left(z\right)\widetilde{\mathbf{F}}_{1,l}\left(\mathbf{H},z\right)+o_p\left(1\right) \end{aligned} \quad \text{(S1.6)}$$

and

$$\begin{aligned}\frac{\widetilde{\Psi}_{2,l}}{n\left|\mathbf{H}\right|} &= \frac{1}{2n\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(p_{l,ij}\pi_j+p_{l,ji}\pi_i\right)\mathbf{m}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\left[\mathbf{x}_{1i}',\left(\mathcal{Z}_j'(z)\otimes\mathbf{m}_j'\right)\right] \\ &= \frac{1}{2n\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j=1}^{n}p_{l,ij}\pi_j\mathbf{m}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\left[\mathbf{x}_{1i}',\left(\mathcal{Z}_j'(z)\otimes\mathbf{m}_j'\right)\right]+o_p\left(1\right) \\ &= \frac{1}{2}f\left(z\right)\widetilde{\mathbf{F}}_{2,l}\left(z\right)+o_p\left(1\right), \end{aligned} \quad \text{(S1.7)}$$

where $\widetilde{\mathbf{F}}_{1,l}(\mathbf{H},z) = \left[\mathbf{0}'_{p_1}, \left[1,\mathbf{0}'_q\right]\otimes\mathbf{F}_{1,l}(\mathbf{H},z)\right]$ and $\widetilde{\mathbf{F}}_{2,l}(\mathbf{H},z) = \left[\mathbf{F}_{21,l}(z), \left[1,\mathbf{0}'_q\right]\otimes\mathbf{F}_{22,l}(z)\right]$ with $\mathbf{F}_{21,l}(z) = n^{-1}\sum_{i\neq j}\mathbb{E}\left[p_{l,ij}\pi_j\mathbf{m}'_i\mathbf{\Pi}(\mathbf{z}^*_i)\mathbf{x}'_{1j}|\mathbf{z}_j = z\right]$ and $\mathbf{F}_{22,l}(z) = n^{-1}\sum_{i\neq j}\mathbb{E}\left[p_{l,ij}\mathbf{m}'_i\mathbf{\Pi}(\mathbf{z}^*_i)\mathbf{m}'_j|\mathbf{z}_j = z\right]$. Also,

$$\widetilde{\mathbf{F}}_{3,l}(\mathbf{H},z) = \frac{1}{n}\sum_{i\neq j}\operatorname{tr}\left\{\mathbf{H}\mathbb{E}\left[p_{l,ij}\mathbf{m}'_i\mathbf{\Pi}(\mathbf{z}^*_i)\left((\mathbf{W}\mathbf{y})_j\frac{\partial^2\rho(z)}{\partial z\partial z'} + \frac{\partial^2\left(\mathbf{x}'_{2j}\boldsymbol{\beta}_2(z)\right)}{\partial z\partial z'}\right)\Bigg|\mathbf{z}_j = z\right]\mathbf{H}\right\}. \tag{S1.8}$$

Again, we have $\Gamma_{1,l}(\mathbf{H},z)/(n|\mathbf{H}|) = O_p(1)$, $\widetilde{\Psi}_{2,l}/(n|\mathbf{H}|) = O_p(1)$ and $\Gamma_{2,l}/(n|\mathbf{H}|) = O_p\left(\|\mathbf{H}\|^2\right)$. We also obtain

$$\frac{\bar{\Gamma}_{1,l}\widetilde{\Psi}_{1,l}}{(n|\mathbf{H}|)^2} = O_p\left(\frac{\|\mathbf{H}\|^2}{\sqrt{n|\mathbf{H}|}}\right) \quad\text{and}\quad \frac{\Gamma_{3,l}\widetilde{\Psi}_{1,l}}{(n|\mathbf{H}|)^2} = O_p\left(\frac{1}{n\sqrt{|\mathbf{H}|}}\right).$$

Second, we consider case *(ii)*. For this case, $\mathbf{m}_i \equiv [\bar{u}_i, \mathbf{x}'_{2i}]'$ and $\mathbf{E}_{1m}(z) \equiv \left[\mathbf{0}_d, n^{-1}\sum_{i=1}^n\mathbb{E}\left[\mathbf{Q}_{n,i}\mathbf{x}'_{2i}|\mathbf{z}_i = z\right]\right]$. In addition, we have

$$\frac{\widetilde{\Psi}_{2,l}}{n|\mathbf{H}|} = \frac{1}{2n|\mathbf{H}|}\sum_{i=1}^n\sum_{j=1}^n(p_{l,ij}\pi_j + p_{l,ji}\pi_i)\eta_i(z)\left[\bar{u}_i\mathbf{x}'_{1j}, \mathcal{Z}'_j(z)\otimes\left[\bar{u}_i\bar{u}_j, \bar{u}_i\mathbf{x}'_{2j}\right]\right] = O_p\left(\|\mathbf{H}\|^2\right).$$

Combining all the above results gives $(n|\mathbf{H}|)^{-2}\xi_nA_{n1}(z) \approx (n|\mathbf{H}|)^{-2}\sum_{l=1}^m\widetilde{\Gamma}_{1,l}\widetilde{\Psi}_{1,l}$, which verifies case *(ii)*.

Next, consider case *(iii)* such that $\boldsymbol{\beta}_1 = \mathbf{0}_{p_1}$, $\boldsymbol{\beta}_2(z) = \mathbf{0}_{p_2}$ and $\rho(z) = \rho_0$ hold over their domains. Then, we have $\mathbf{C}(z) = \mathbf{0}_n$, which implies $\Gamma_{1,l} = 0$, $\Gamma_{2,l} = 0$ and $\widetilde{\Psi}_{2,l} = \mathbf{0}'_{p_1+(p_2+1)(q+1)}$. Hence, $(n|\mathbf{H}|)^{-2}\xi_nA_{n1}(z) = (2n|\mathbf{H}|)^{-2}\sum_{l=1}^m\Gamma_{3,l}\widetilde{\Psi}'_{1,l} = O_p\left(\left(n\sqrt{|\mathbf{H}|}\right)^{-1}\right)$. This completes the proof of this lemma. ■

Lemma 9 *Under Assumptions **1–3** and **9**, we obtain*

$$\frac{\xi_n^2}{(n|\mathbf{H}|)^2}B_n(z) = f^2(z)\varkappa_B(\mathbf{H},z) + O_p\left((\xi_n\|\mathbf{H}\|)^{-1} + (\xi_n\|\mathbf{H}\|)^{-2}\right) + o_p(1), \tag{S1.9}$$

where

$$\varkappa_{B,Q}(\mathbf{H},z) = \begin{bmatrix} \mathbf{E}_{1x}(z)'\mathbf{E}_{1x}(z) & \mathbf{0}_{p_1\times(p_2+1)} & \mathbf{0}_{p_1\times(q(p_2+1))} \\ \mathbf{0}_{d\times p_1} & \mathbf{E}_{1m}(z)\mathbf{E}_{1m}(z)' & \mathbf{0}_{d\times(q(p_2+1))} \\ \mathbf{0}_{(dq)\times p_1} & \mathbf{0}_{(dq)\times(p_2+1)} & \mu^2_{1,2}(k)\mathbf{I}_q\otimes\left[\mathbf{E}_{1m}(z)\mathbf{E}_{1m}(z)'\right] \end{bmatrix}$$

$$\varkappa_{B,P}(\mathbf{H},z) = \sum_{l=1}^m\left[2\widetilde{\mathbf{F}}_{1,l}(\mathbf{H},z) + \frac{1}{2}\widetilde{\mathbf{F}}_{2,l}(z)\right]'\left[2\widetilde{\mathbf{F}}_{1,l}(\mathbf{H},z) + \frac{1}{2}\widetilde{\mathbf{F}}_{2,l}(z)\right].$$

Proof. By (S1.3), we have

$$\frac{1}{(n|\mathbf{H}|)^2}\widetilde{\mathbf{M}}(z)'\mathbf{\Xi}_H(z)\widetilde{\mathbf{M}}(z) \approx f^2(z)\varkappa_{B,Q}(\mathbf{H},z). \tag{S1.10}$$

In addition, we have

$$\frac{\widetilde{\Psi}_{3,l}}{n|\mathbf{H}|} = \frac{1}{n|\mathbf{H}|}\sum_{i=1}^n\sum_{j=1}^n(p_{l,ij}\pi_j + p_{l,ji}\pi_i)\begin{bmatrix} \mathbf{x}_{1i}\mathbf{x}'_{1i} & \mathbf{x}_{1i}\left[\mathcal{Z}'_j(z)\otimes\mathbf{m}'_j\right] \\ \left[\mathcal{Z}_i(z)\otimes\mathbf{m}_i\right]\mathbf{x}'_{1j} & \left[\mathcal{Z}_i(z)\mathcal{Z}'_j(z)\right]\otimes\left(\mathbf{m}_i\mathbf{m}'_j\right) \end{bmatrix}$$

$$= \quad O_p\left(\|\mathbf{H}\|^{-1}\right).$$

Therefore, combining all the above results yields

$$\frac{\xi_n^2}{(n\,|\mathbf{H}|)^2} B_n(z) \approx \frac{1}{(n\,|\mathbf{H}|)^2}\sum_{l=1}^{m}\left[\left(\widetilde{\Psi}_{1,l}+\widetilde{\Psi}_{2,l}\right)'\left(\widetilde{\Psi}_{1,l}+\widetilde{\Psi}_{2,l}\right)\right]+\frac{1}{n\,|\mathbf{H}|}\widetilde{\mathbf{M}}(z)'\boldsymbol{\Xi}_H(z)\widetilde{\mathbf{M}}(z)+ O_p\left(\xi_n^{-1}\|\mathbf{H}\|^{-1}+\xi_n^{-2}\|\mathbf{H}\|^{-2}\right)=\varkappa_B(\mathbf{H},z)+o_p(1).$$

This completes the proof of this lemma. ■

Remark If $\text{diag}\{\mathbf{P}_{n,l}\}=0$ for all l, we have $\varkappa_{B,P}(\mathbf{H},z)=4^{-1}\sum_{l=1}^{m}\widetilde{\mathbf{F}}_{2,l}(z)'\,\widetilde{\mathbf{F}}_{2,l}(z)$. In addition, if $\boldsymbol{\beta}_1=\mathbf{0}_{p_1}$ and $\boldsymbol{\beta}_2(z)=\mathbf{0}_{p_2}$ hold over their domains, $\varkappa_B(\mathbf{H},z)$ is again singular because $\widetilde{\mathbf{F}}_{2,l}(z)=\mathbf{0}'_{p_1+(q+1)(p_2+1)}$ and $\mathbf{E}_{1m}(z)\,\mathbf{E}_{1m}(z)'$ is singular.

Lemma 10 *Under Assumptions **1**–**3** and **9**, we obtain*

$$\frac{\xi_n A_{n,2}(z)}{(n\,|\mathbf{H}|)^{3/2}} \xrightarrow{d} \mathbb{N}\Big(\mathbf{0},\ f^3(z)\,R_2(K)\,\boldsymbol{\Omega}(z)\Big), \tag{S1.11}$$

where $\boldsymbol{\Omega}(z)$ is defined in the proof.

Proof. Since the proof closely follows that of Lemma 3, we omit the details to conserve space. In this lemma, we have

$$\Delta_n(z)=\begin{bmatrix}\mathbf{E}_{1x}(z) & \mathbf{E}_{1m}(z) & \mathbf{0}_{d\times(q(p_2+1))}\\ \mathbf{0}_{(dq)\times p_1} & \mathbf{0}_{dq\times(p_2+1)} & \mu_{1,2}(k)\,\mathbf{I}_q\otimes\mathbf{E}_{1m}(z)\end{bmatrix},$$

so that

$$\boldsymbol{\Omega}_Q(z)=\begin{bmatrix}\mathbf{E}_1(z)'\,\mathbf{E}_3(z)\,\mathbf{E}_1(z) & \mathbf{0}_{(p+1)\times[q(p_2+1)]}\\ \mathbf{0}_{[q(p_2+1)]\times(p+1)} & R_2^{-1}(K)\,\mu_{1,2}^2(k)\,\mu_{2,2}(k)\,\mathbf{I}_q\otimes\left[\mathbf{E}_{1m}(z)'\,\mathbf{E}_3(z)\,\mathbf{E}_{1m}(z)\right]\end{bmatrix}. \tag{S1.12}$$

In addition, we obtain

$$\boldsymbol{\Omega}_P(z) = \sum_{l=1}^{m}\sum_{l'=1}^{m}\widetilde{\mathbf{F}}_{2,l}(z)'\,\widetilde{\mathbf{F}}_{2,l'}(z)\left(\mathbb{V}ar\left(u_1^2\right)\lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[p_{l,ii}p_{l',ii}|\mathbf{z}_i=z\right]+\frac{1}{4}\lim_{n\to\infty}\frac{1}{n}\sum_{i'\neq i\neq j}\mathbb{E}\left[p_{l,ij}p_{l',i'j}\bar{\mathbf{m}}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\bar{\mathbf{m}}_{i'}'\boldsymbol{\Pi}(\mathbf{z}_{i'}^*)|\mathbf{z}_j=z\right]\right) \tag{S1.13}$$

and

$$\boldsymbol{\Omega}_{Q,P}(z) = 2^{-1}\sum_{l=1}^{m}\widetilde{\mathbf{F}}_{2,l}(z)'\left[\left(\sum_{l'=1}^{m}\lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[p_{l',ii}\mathbf{Q}_{n,i}'|\mathbf{z}_i=z\right]+\lim_{n\to\infty}\frac{1}{n}\sum_{i\neq j}\mathbb{E}\left[p_{l,ij}\bar{\mathbf{m}}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\mathbf{Q}_{n,j}'|\mathbf{z}_j=z\right]\right)\lim_{n\to\infty}\mathbf{E}_1(z),\ \mathbf{0}'_{q(p_2+1)}\right]. \tag{S1.14}$$

This completes the proof of this lemma. ■

Lemma 11 *Under Assumptions **1–3**, **5** and **9**, we obtain*

$$\frac{1}{n^2\left|\mathbf{H}\right|^{3/2}}\sum_{i=1}^{n} f^{-2}\left(\mathbf{z}_i\right)\mathbf{S}_{p_1,p_1+(q+1)(p_2+1)}\varkappa_B\left(\mathbf{H},z_i\right)^{-1}A_{n2,-i}\left(\mathbf{z}_i\right)\ \overset{d}{\to}\ \mathbb{N}\Big(\mathbf{0}_{p_1},\ \mathbf{\Omega}_0\Big), \tag{S1.15}$$

where $\mathbf{\Omega}_0$ is defined in the proof.

Proof. Let $\boldsymbol{\alpha}\neq\mathbf{0}$ be a column vector of dimension p_1. Under Assumption 5, the results given in Lemma 9 hold uniformly over the domain of $B_n\left(z\right)$. Denoting $\psi\left(\mathbf{H},z_i\right)=\mathbf{S}_{p_1,p_1+(q+1)(p_2+1)}\varkappa_B\left(\mathbf{H},z_i\right)^{-1}$ and building on the proofs of Lemmas 3 and 10, we then have

$$\frac{1}{n\left(n\left|\mathbf{H}\right|\right)^{3/2}}\sum_{i=1}^{n} f^{-2}\left(\mathbf{z}_i\right)\boldsymbol{\alpha}'\psi\left(\mathbf{H},z_i\right)A_{n2,-i}\left(\mathbf{z}_i\right)'\approx\mathbf{u}'\mathbf{A}_n\mathbf{u}+\mathbf{b}_n\mathbf{u}\equiv\Lambda_n,$$

where $\mathbf{A}_n=\sum_{i=1}^{n}\mathbf{A}_n\left(\mathbf{z}_i\right)$ and $\mathbf{b}_n=\sum_{i=1}^{n}\mathbf{b}_n\left(\mathbf{z}_i\right)$ with

$$\begin{aligned}
\mathbf{A}_n\left(\mathbf{z}_i\right) &= \frac{f^{-1}\left(\mathbf{z}_i\right)}{4n\sqrt{n\left|\mathbf{H}\right|}}\boldsymbol{\alpha}'\psi\left(\mathbf{H},z_i\right)\sum_{l=1}^{m}\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}\right)'\left[\mathbf{P}_{n,l}\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\right]^s\\
\mathbf{b}_n\left(\mathbf{z}_i\right) &= \frac{f^{-1}\left(\mathbf{z}_i\right)}{n\sqrt{n\left|\mathbf{H}\right|}}\boldsymbol{\alpha}'\psi\left(\mathbf{H},z_i\right)\left(\frac{1}{2}\sum_{l=1}^{m}\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}\right)'\bar{\mathbf{C}}(\mathbf{z}_i)'\left[\mathbf{P}_{n,l}\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\right]^s+\Delta_n\left(\mathbf{z}_i\right)'\mathbf{Q}\left(\mathbf{z}_i\right)'\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\right),
\end{aligned}$$

and $\widetilde{\mathbf{K}}_H(\mathbf{z}_i)$ equals $\mathbf{K}_H(\mathbf{z}_i)$ with its (i,i)th element being equal to zero due to the leave-one-out procedure. Since $\{\mathbf{A}_n\left(\mathbf{z}_i\right)\}$ and $\{\mathbf{b}_n\left(\mathbf{z}_i\right)\}$ are uncorrelated with $\{u_i\}$, in order to apply Kelejian & Prucha's (2001) central limit theorem for linear-quadratic forms with minor modification, we need to check on the following two conditions:

$$n^{-1}\sum_{k=1}^{n}\mathbb{E}\left[\left|b_{n,k}\right|^{2+\delta_1}\right]\leq M\qquad\text{for some }\delta_1>0 \tag{S1.16}$$

$$0<c_0/n<\mathbb{V}ar\left[\Lambda_n\right]<c_1/n<\infty, \tag{S1.17}$$

where $b_{n,k}$ is the kth element of $\mathbf{b}_n$.

First, we verify (S1.16). Applying direct calculations gives

$$\begin{aligned}
b_{n,k} &= \frac{1}{4n\sqrt{n\left|\mathbf{H}\right|}}\boldsymbol{\alpha}'\sum_{i=1}^{n} f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\sum_{l=1}^{m}\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}\right)'\sum_{j=1}^{n}\left(p_{l,kj}\pi_{ji}+p_{l,jk}\pi_{ki}\right)\bar{\mathbf{m}}_j'\mathbf{\Pi}(\mathbf{z}_j^*,\mathbf{z}_i)+\\
&\quad\frac{1}{n\sqrt{n\left|\mathbf{H}\right|}}\boldsymbol{\alpha}'\sum_{i=1}^{n} f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\Delta_n\left(\mathbf{z}_i\right)'\pi_{ik}\mathcal{Z}_k(\mathbf{z}_i)\otimes\mathbf{Q}_{n,k},
\end{aligned}$$

where we have defined $\pi_{ij}\equiv\mathcal{K}_H(\mathbf{z}_i,z_j)=\mathcal{K}(\mathbf{H}^{-1}(\mathbf{z}_i-z_j))$ and $\mathcal{Z}_k(\mathbf{z}_i)\equiv\left[1,\left(\mathbf{z}_k-\mathbf{z}_i\right)'\mathbf{H}^{-1}\right]'$, and $\mathbf{\Pi}(\mathbf{z}_j^*,\mathbf{z}_k)$ denotes a $(p_2+1)\times 1$ vector whose lth element equals $\Pi_l(\mathbf{z}_j^*,\mathbf{z}_i)=(\mathbf{z}_j-\mathbf{z}_i)'\bigtriangledown^2\gamma_l(\mathbf{z}_j^*)(\mathbf{z}_j-\mathbf{z}_i)$. Note that the leave-one-out procedure results in $\pi_{ii}=0$ all i. For each given l, applying Minkoski's inequality yields

$$\mathbb{E}\left(\left|\boldsymbol{\alpha}'\sum_{i=1}^{n} f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}\right)'\sum_{j=1}^{n}\left(p_{l,kj}\pi_{ji}+p_{l,jk}\pi_{ki}\right)\bar{\mathbf{m}}_j'\mathbf{\Pi}(\mathbf{z}_j^*,\mathbf{z}_i)\right|^{2+\delta_1}\right)$$

$$
\begin{aligned}
&\leq M\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathbb{E}\left|\left(p_{l,kj}\pi_{ji}+p_{l,jk}\pi_{ki}\right)\bar{\mathbf{m}}_j'\mathbf{\Pi}(\mathbf{z}_j^*,\mathbf{z}_i)\right|^{2+\delta_1}\right)^{1/(2+\delta_1)}\right]^{2+\delta_1}\\
&\leq Mn^{2+\delta_1}\left|\mathbf{H}\right|
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left|\boldsymbol{\alpha}'\sum_{i=1}^{n}f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\Delta_n\left(\mathbf{z}_i\right)'\pi_{ik}\mathcal{Z}_k(\mathbf{z}_i)\otimes\mathbf{Q}_{n,k}\right|^{2+\delta_1}\\
\leq\ & M\left[\sum_{i=1}^{n}\left(\mathbb{E}\left|\pi_{ik}\left\|\mathcal{Z}_k(\mathbf{z}_i)\right\|\left\|\mathbf{Q}_{n,k}\right\|\right|^{2+\delta_1}\right)^{1/(2+\delta_1)}\right]^{2+\delta_1}\\
\leq\ & Mn^{2+\delta_1}\left|\mathbf{H}\right|.
\end{aligned}
$$

Hence, we obtain $\mathbb{E}\left[\left|b_{n,k}\right|^{2+\delta_1}\right]\leq M\left|\mathbf{H}\right|\left(n\left|\mathbf{H}\right|\right)^{-(1+\delta_1/2)}$. This gives (S1.16).

Second, we verify (S1.17). Because $\pi_{ii}=0$ for all i and $\mathrm{tr}\left\{\mathbf{P}_{n,l}\right\}=0$ for all l, we have

$$
\begin{aligned}
\mathbb{E}\left[\Lambda_n\right] &= \sigma_u^2\mathbb{E}\left[\mathrm{tr}\left\{\mathbf{A}_n\right\}\right]\\
&= \frac{\sigma_u^2}{2n\sqrt{n\left|\mathbf{H}\right|}}\boldsymbol{\alpha}'\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l=1}^{m}\mathbb{E}\left[f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'p_{l,jj}\pi_{ji}\right]=O\left(\sqrt{\left|\mathbf{H}\right|/n}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{i=1}^{n}\mathbb{E}\left[a_{n,ii}^2\right] &= \frac{1}{4n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{l=1}^{m}\sum_{l'=1}^{m}\sum_{j\neq i}\sum_{j'\neq i}\boldsymbol{\alpha}'\mathbb{E}\left[f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\boldsymbol{p}_{l,jj}\boldsymbol{p}_{l',j'j'}\pi_{ji}\pi_{j'i}\right]\boldsymbol{\alpha}\\
&\approx \frac{1}{4n^3\left|\mathbf{H}\right|}\boldsymbol{\alpha}'\sum_{l=1}^{m}\sum_{l'=1}^{m}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{E}\left[p_{l,jj}p_{l',jj}\pi_{ji}^2f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\right]\boldsymbol{\alpha}=O\left(n^{-1}\right),
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\left[\mathrm{tr}\left\{\mathbf{A}_n^2\right\}\right] &= \frac{1}{16n^3\left|\mathbf{H}\right|}\boldsymbol{\alpha}'\sum_{i=1}^{n}\sum_{i'=1}^{n}\sum_{l=1}^{m}\sum_{l'=1}^{m}\mathbb{E}\Bigg[f^{-1}\left(\mathbf{z}_i\right)f^{-1}\left(\mathbf{z}_{i'}\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_{i'}\right)\psi\left(\mathbf{H},z_{i'}\right)'\times\\
&\quad\sum_{j\neq i}\sum_{j'\neq i'}\left(p_{l,jj'}\pi_{j'i}+p_{l,j'j}\pi_{ji}\right)\left(p_{l',jj'}\pi_{j'i'}+p_{l',j'j}\pi_{ji'}\right)\Bigg]\boldsymbol{\alpha}\approx\sum_{i=1}^{n}\mathbb{E}\left[a_{n,ii}^2\right], \qquad \text{(S1.18)}
\end{aligned}
$$

$$
\begin{aligned}
\sum_{i\neq j}\mathbb{E}\left(a_{n,ii}a_{n,jj}\right) &= \frac{1}{4n^3\left|\mathbf{H}\right|}\boldsymbol{\alpha}'\sum_{l=1}^{m}\sum_{l'=1}^{m}\sum_{i\neq j}\sum_{i'\neq i}\sum_{j'\neq j}\mathbb{E}\Big[f^{-1}\left(\mathbf{z}_i\right)f^{-1}\left(\mathbf{z}_j\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_j\right)\psi\left(\mathbf{H},z_j\right)'\times\\
&\quad p_{l,i'i'}p_{l',j'j'}\pi_{i'i}\pi_{j'j}\Big]\boldsymbol{\alpha}\\
&\approx \frac{1}{4n^3\left|\mathbf{H}\right|}\boldsymbol{\alpha}'\sum_{l=1}^{m}\sum_{l'=1}^{m}\sum_{i\neq j}\mathbb{E}\left[f^{-1}\left(\mathbf{z}_i\right)f^{-1}\left(\mathbf{z}_j\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_j\right)\psi\left(\mathbf{H},z_j\right)'p_{l,jj}p_{l',ii}\pi_{ji}^2\right]\boldsymbol{\alpha}\\
&= O\left(n^{-1}\right),
\end{aligned}
$$

so that

$$
2\sigma_u^4\mathbb{E}\left[\mathrm{tr}\left\{\mathbf{A}_n^2\right\}\right]+\left(\mathbb{E}\left[u_1^4\right]-2\sigma_u^4\right)\sum_{i=1}^{n}\mathbb{E}\left[a_{n,ii}^2\right]+\sigma_u^4\sum_{i\neq j}\mathbb{E}\left[a_{n,ii}a_{n,jj}\right]
$$

$$\approx \quad \mathbb{E}\left[u_1^4\right]\sum_{i=1}^{n}\mathbb{E}\left[a_{n,ii}^2\right]+\sigma_u^4\sum_{i\neq j}\mathbb{E}\left[a_{n,ii}a_{n,jj}\right]\approx R_2\left(K\right)\boldsymbol{\alpha}'\boldsymbol{\Omega}_1\boldsymbol{\alpha}/n$$

where

$$\boldsymbol{\Omega}_1=\lim_{n\to\infty}\frac{1}{4n^2}\sum_{l=1}^{m}\sum_{l'=1}^{m}\sum_{i=1}^{n}\sum_{j\neq i}\left\{\mathbb{E}\left[p_{l,jj}\left(\mathbb{E}\left[u_1^4\right]p_{l',jj}+\sigma_u^4 p_{l',ii}\right)f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\right]\right.$$

Moreover, we obtain

$$\begin{aligned}
&\mathbf{b}_n\mathbf{b}_n' \\
=\;&\frac{1}{4n^3\left|\mathbf{H}\right|}\boldsymbol{\alpha}'\sum_{i=1}^{n}\sum_{l=1}^{m}\sum_{l'=1}^{m}f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\bar{\mathbf{C}}(\mathbf{z}_i)'\left[\mathbf{P}_{n,l}\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\right]^s\left[\mathbf{P}_{n,l'}\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\right]^s\times \\
&\bar{\mathbf{C}}(\mathbf{z}_i)\boldsymbol{\alpha}+\frac{1}{n^3\left|\mathbf{H}\right|}\boldsymbol{\alpha}'\sum_{i=1}^{n}\sum_{l=1}^{m}f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\bar{\mathbf{C}}(\mathbf{z}_i)'\left[\mathbf{P}_{n,l}\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\right]^s\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\mathbf{Q}\left(\mathbf{z}_i\right)\Delta_n\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\boldsymbol{\alpha}+ \\
&\frac{1}{n^3\left|\mathbf{H}\right|}\boldsymbol{\alpha}'\sum_{i=1}^{n}f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\Delta_n\left(\mathbf{z}_i\right)'\mathbf{Q}\left(\mathbf{z}_i\right)'\widetilde{\mathbf{K}}_H^2(\mathbf{z}_i)\mathbf{Q}\left(\mathbf{z}_i\right)\Delta_n\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\boldsymbol{\alpha} \\
\equiv\;&\boldsymbol{\alpha}'\left(A_{n1}/4+A_{n2}+A_{n3}\right)\boldsymbol{\alpha},
\end{aligned}$$

where A_{nj} for $j=1,2,3$ are defined in the order of their appearance.

Applying standard calculations, we obtain

$$\begin{aligned}
\mathbb{E}\left[A_{n1}\right] &= \frac{1}{4n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{l=1}^{m}\sum_{l'=1}^{m}\mathbb{E}\Bigg[f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\times \\
&\quad\sum_{j\neq i}\sum_{j'\neq i}\sum_{j''\neq i}\left(p_{l,jj'}\pi_{j'i}+p_{l,j'j}\pi_{ji}\right)\left(p_{l',jj''}\pi_{j''i}+p_{l',j''j}\pi_{ji}\right)\bar{\mathbf{m}}_{j'}'\boldsymbol{\Pi}(\mathbf{z}_{j'}^*,\mathbf{z}_i)\bar{\mathbf{m}}_{j''}'\boldsymbol{\Pi}(\mathbf{z}_{j''}^*,\mathbf{z}_i)\Bigg] \\
&\approx \frac{1}{n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l=1}^{m}\sum_{l'=1}^{m}\mathbb{E}\left[f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\times\right. \\
&\quad\left.p_{l,jj}p_{l',jj}\pi_{ji}^2\left(\bar{\mathbf{m}}_j'\boldsymbol{\Pi}(\mathbf{z}_j^*,\mathbf{z}_i)\right)^2\right]=O\left(\left\|\mathbf{H}\right\|^4/n\right),
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left[A_{n2}\right] &= \frac{1}{n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{l=1}^{m}\sum_{l'=1}^{m}\mathbb{E}\left[f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\bar{\mathbf{C}}(\mathbf{z}_i)'\left[\mathbf{P}_{n,l}\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\right]^s\times\right. \\
&\quad\left.\widetilde{\mathbf{K}}_H(\mathbf{z}_i)\mathbf{Q}\left(\mathbf{z}_i\right)\Delta_n\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\right] \\
&= \frac{1}{n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{l=1}^{m}\sum_{l'=1}^{m}\sum_{j\neq i}\sum_{j'\neq i}\mathbb{E}\left\{f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\pi_{ji}\left(p_{l,jj'}\pi_{j'i}+p_{l,j'j}\pi_{ji}\right)\bar{\mathbf{m}}_{j'}'\boldsymbol{\Pi}(\mathbf{z}_{j'}^*,\mathbf{z}_i)\times\right. \\
&\quad\left.\left[\begin{array}{cc}\mathbf{Q}_{n,j}'\mathbf{E}_1\left(\mathbf{z}_i\right) & \mu_{12}\left(k\right)\left[\mathbf{H}^{-1}\left(\mathbf{z}_j-\mathbf{z}_i\right)\right]'\otimes\left[\mathbf{Q}_{n,j}'\mathbf{E}_{1m}\left(\mathbf{z}_i\right)\right]\end{array}\right]\psi\left(\mathbf{H},z_i\right)'\right\} \\
&\approx \frac{2}{n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l=1}^{m}\sum_{l'=1}^{m}\mathbb{E}\left\{f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'p_{l,jj}\pi_{ji}^2\bar{\mathbf{m}}_j'\boldsymbol{\Pi}(\mathbf{z}_j^*,\mathbf{z}_i)\times\right. \\
&\quad\left.\left[\begin{array}{cc}\mathbf{Q}_{n,j}'\mathbf{E}_1\left(\mathbf{z}_i\right) & \mu_{12}\left(k\right)\left[\mathbf{H}^{-1}\left(\mathbf{z}_j-\mathbf{z}_i\right)\right]'\otimes\left[\mathbf{Q}_{n,j}'\mathbf{E}_{1m}\left(\mathbf{z}_i\right)\right]\end{array}\right]\psi\left(\mathbf{H},z_i\right)'\right\}=O\left(\left\|\mathbf{H}\right\|^2/n\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[A_{n3}\right] &= \frac{1}{n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\mathbb{E}\left[f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\Delta_n\left(\mathbf{z}_i\right)'\mathbf{Q}\left(\mathbf{z}_i\right)'\widetilde{\mathbf{K}}_H^2(\mathbf{z}_i)\mathbf{Q}\left(\mathbf{z}_i\right)\Delta_n\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\right] \\
&= \frac{1}{n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{E}\left[\pi_{ij}^2 f^{-2}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\Delta_n\left(\mathbf{z}_i\right)'\left[\mathcal{Z}_j\left(\mathbf{z}_i\right)\mathcal{Z}_j\left(\mathbf{z}_i\right)'\right]\otimes\left(\mathbf{Q}_{n,j}\mathbf{Q}_{n,j}'\right)\Delta_n\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\right] \\
&\approx R_2\left(K\right)\mathbf{S}_3/n = O\left(n^{-1}\right),
\end{aligned}$$

where

$$\mathbf{S}_3 = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\boldsymbol{\Omega}_Q\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)'\right].$$

Combing all the above results, we obtain

$$\mathbb{E}\left[\mathbf{b}_n\mathbf{b}_n'\right] \approx R_2\left(K\right)\boldsymbol{\alpha}'\mathbf{S}_3\boldsymbol{\alpha}/n. \tag{S1.19}$$

Moreover, we have

$$\begin{aligned}
&\sum_{i=1}^{n}\mathbb{E}\left[a_{n,ii}b_{n,i}\right] \\
=\ & \frac{1}{8n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{j'\neq i}\sum_{l=1}^{m}\sum_{l'=1}^{m}\boldsymbol{\alpha}'\mathbb{E}\left[f^{-1}\left(\mathbf{z}_j\right)f^{-1}\left(\mathbf{z}_{j'}\right)\psi\left(\mathbf{H},z_j\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_j\right)'\widetilde{\mathbf{F}}_{2,l'}\left(\mathbf{H},\mathbf{z}_{j'}\right)\psi\left(\mathbf{H},z_{j'}\right)'\times\right. \\
& \left. p_{l,ii}p_{l',ij'}\pi_{ij}\pi_{j'i}\bar{\mathbf{m}}_{j'}'\boldsymbol{\Pi}(\mathbf{z}_{j'}^*,\mathbf{z}_i)\right]\boldsymbol{\alpha} + \frac{1}{2n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{j'\neq i}\sum_{l=1}^{m}\boldsymbol{\alpha}'\mathbb{E}\left[f^{-1}\left(\mathbf{z}_j\right)f^{-1}\left(\mathbf{z}_{j'}\right)\psi\left(\mathbf{H},z_j\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_j\right)'\times\right. \\
& \left. p_{l,ii}\pi_{ij}\pi_{j'i}\left[\begin{array}{cc} Q_{n,j'}'\mathbf{E}_1\left(\mathbf{z}_i\right) & \mu_{12}\left(k\right)\left[\mathbf{H}^{-1}\left(\mathbf{z}_{j'}-\mathbf{z}_i\right)\right]'\otimes\left[Q_{n,j'}'\mathbf{E}_{1m}\left(\mathbf{z}_i\right)\right]\end{array}\right]\psi\left(\mathbf{H},z_{j'}\right)'\right]\boldsymbol{\alpha} \\
\approx\ & \frac{1}{2n^3\left|\mathbf{H}\right|}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{l=1}^{m}\boldsymbol{\alpha}'\mathbb{E}\left[f^{-2}\left(\mathbf{z}_j\right)\psi\left(\mathbf{H},z_j\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_j\right)'\times\right. \\
& \left. p_{l,ii}\pi_{ij}^2\left[\begin{array}{cc} \mathbf{Q}_{n,j}'\mathbf{E}_1\left(\mathbf{z}_i\right) & \mu_{12}\left(k\right)\left[\mathbf{H}^{-1}\left(\mathbf{z}_j-\mathbf{z}_i\right)\right]'\otimes\left[\mathbf{Q}_{n,j}'\mathbf{E}_{1m}\left(\mathbf{z}_i\right)\right]\end{array}\right]\psi\left(\mathbf{H},z_j\right)'\right]\boldsymbol{\alpha} \\
\approx\ & R_2\left(K\right)\boldsymbol{\alpha}'\mathbf{S}_4\boldsymbol{\alpha}/\left(2n\right) = O\left(n^{-1}\right),
\end{aligned}$$

where, denoting $s\left(z\right)=\mathbb{E}\left[\mathbf{Q}_{n,j}'|\mathbf{z}_j=z\right]$, we have

$$\mathbf{S}_4 = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{l=1}^{m}\mathbb{E}\left[p_{l,ii}f^{-1}\left(\mathbf{z}_i\right)\psi\left(\mathbf{H},z_i\right)\widetilde{\mathbf{F}}_{2,l}\left(\mathbf{H},\mathbf{z}_i\right)'\left[\begin{array}{cc} s\left(\mathbf{z}_i\right)\mathbf{E}_1\left(\mathbf{z}_i\right) & \mathbf{0}_{d\times[q(p_2+1)]}\end{array}\right]\psi\left(\mathbf{H},z_i\right)'\right].$$

Therefore, $\mathbb{V}ar\left[\Lambda_n\right]=O\left(n^{-1}\right)$ and (S1.17) holds true with $\boldsymbol{\Omega}_0=\boldsymbol{\Omega}_1+\sigma_u^2\mathbf{S}_3+\mathbb{E}\left[u_1^3\right]\mathbf{S}_4$. This completes the proof of this lemma. ■

Proof of Theorem 6. The proof of this theorem closely follows that of Theorem 5. Now, $\boldsymbol{\theta}_n=\xi_n\left[\begin{array}{cc}\widehat{\boldsymbol{\gamma}}\left(z\right)-\boldsymbol{\gamma}\left(z\right) & \left[\triangledown\widehat{\boldsymbol{\gamma}}\left(z\right)-\triangledown\boldsymbol{\gamma}\left(z\right)\right]\mathbf{H}\end{array}\right]$ is a $p_2\times(q+1)$ matrix, $\boldsymbol{\vartheta}_n\equiv\xi_n\left[\widehat{\rho}\left(z\right)-\rho_0,\widehat{\boldsymbol{\beta}}_1\left(z\right)-\boldsymbol{\beta}_1\right]$, $\mathbf{u}\left(\boldsymbol{\theta},\boldsymbol{\vartheta}\right)$ is an $n\times 1$ vector with a typical element being equal to $u_i\left(\boldsymbol{\theta},\boldsymbol{\vartheta}\right)=y_i^*-\xi_n^{-1}\mathbf{m}_{1i}'\boldsymbol{\vartheta}-\xi_n^{-1}\mathbf{m}_i'\boldsymbol{\theta}\mathcal{Z}_i(z)$, and $\{\xi_n\}$ again is a sequence of positive constants such that $0<M_1<\|\boldsymbol{\theta}_n\|$, $\|\boldsymbol{\vartheta}_n\|$

$< M_2 < \infty$. Denoting $\widetilde{\mathbf{M}}(z) \equiv [\mathbf{Wy}, \mathbf{X}_1, \mathbf{M}(z)]$ and $\boldsymbol{\alpha}_n \equiv \left[\operatorname{vec}(\boldsymbol{\vartheta}_n)', \operatorname{vec}(\boldsymbol{\theta}_n)'\right]'$, we again have $\boldsymbol{\alpha}_n = B_n^{-1}(z) A_n(z)'$. In addition, we have $y_i^* = u_i + \mathbf{x}_{2i}'\boldsymbol{\Pi}(\mathbf{z}_i^*)/2$, and $\boldsymbol{\Pi}(z_i^*)$ becomes a $p_2 \times 1$ vector with its lth element being equal to $\Pi_l(\mathbf{z}_i^*) = (\mathbf{z}_i - z)' \bigtriangledown^2 \boldsymbol{\beta}_{2,l}(\mathbf{z}_i^*)(\mathbf{z}_i - z)$, $l = 1, ..., p_2$.

In Lemmas 12–14, we show that (A.2)–(A.3) continue to hold with the newly defined $\varkappa_A(\mathbf{H}, z)$, $\varkappa_B(\mathbf{H}, z)$ and $\boldsymbol{\Omega}(z)$ and that ξ_n is of order $\sqrt{n|\mathbf{H}|}$.

Next, letting $\mathbf{S}_{p_1+1,p_1+(q+1)p_2+1}$ be the first $p_1 + 1$ rows of the identity matrix $\mathbf{I}_{p_1+1+(q+1)p_2}$, under Assumption 5, we have

$$\begin{aligned}
\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0 &= \frac{1}{n\sqrt{n|\mathbf{H}|}} \sum_{i=1}^{n} \mathbf{S}_{p_1+1,p_1+(q+1)p_2+1} B_n^{-1}(\mathbf{z}_i) A_n(\mathbf{z}_i) \\
&\approx \frac{1}{n} \sum_{i=1}^{n} \mathbf{S}_{p_1+1,p_1+(q+1)p_2+1} \varkappa_B(\mathbf{H}, z_i)^{-1} \varkappa_A(\mathbf{H}, \mathbf{z}_i)' + \\
&\quad \frac{1}{n(n|\mathbf{H}|)^{3/2}} \sum_{i=1}^{n} f^{-2}(\mathbf{z}_i) \mathbf{S}_{p_1+1,p_1+(q+1)p_2+1} \varkappa_B(\mathbf{H}, \mathbf{z}_i)^{-1} A_{n2}(\mathbf{z}_i)' \\
&= \frac{1}{n(n|\mathbf{H}|)^{3/2}} \sum_{i=1}^{n} f^{-2}(\mathbf{z}_i) \mathbf{S}_{p_1+1,p_1+(q+1)p_2+1} \varkappa_B(\mathbf{H}, \mathbf{z}_i)^{-1} A_{n2}(\mathbf{z}_i)' + O_p\left(\|\mathbf{H}\|^2\right).
\end{aligned}$$

By Lemma 15, we then obtain that

$$\sqrt{n}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) \overset{d}{\to} \mathbb{N}(\mathbf{0}_{p_1}, \ \boldsymbol{\Omega}_1)$$

if $n\|\mathbf{H}\|^4 \to 0$. This completes the proof of this theorem. ■

Lemma 12 *Under Assumptions **1**–**3** and **9**, we obtain that* $(n|\mathbf{H}|)^{-2} \xi_n A_{n1}(z) = f^2(z) \varkappa_A(\mathbf{H}, z)' + o_p\left(\|\mathbf{H}\|^2\right)$.

(i) If $\boldsymbol{\beta}_2(z) \neq \mathbf{0}_p$ *holds over at least one non-empty subset,*

$$\varkappa_{A,Q}(\mathbf{H}, z) = \frac{\mu_{1,2}(k)}{2} \left[\begin{array}{cc} \mathbf{E}_2(\mathbf{H}, z) \mathbf{E}_1(z) & \mathbf{0}_{qp_2}' \end{array} \right] \tag{S1.20}$$

$$\varkappa_{A,P}(\mathbf{H}, z) = \sum_{l=1}^{m} \widetilde{\mathbf{F}}_{3,l}(\mathbf{H}, z) \left[\widetilde{\mathbf{F}}_{1,l}(\mathbf{H}, z) + \frac{1}{4}\widetilde{\mathbf{F}}_{2,l}(z)\right]; \tag{S1.21}$$

(ii) If $\boldsymbol{\beta}_1 = \mathbf{0}_{p_1}$ *and* $\boldsymbol{\beta}_2(z) = \mathbf{0}_{p_2}$ *hold over their domains, we have* $\varkappa_{A,Q}(\mathbf{H}, z) = \mathbf{0}_{p_1+1+(q+1)p_2}$ *and* $\varkappa_{A,P}(\mathbf{H}, z) = O_p\left(\left(n\sqrt{|\mathbf{H}|}\right)^{-1}\right)$.

Proof. Since the proof closely follows that of Lemma 1, below we only provide new results for the terms with limit results that differ from those given in the proof of Lemma 1.

We begin with case *(i)*. Applying straightforward calculation gives

$$\begin{aligned}
\frac{1}{n|\mathbf{H}|}\mathbf{Q}(z)'\mathbf{K}_H(z)\widetilde{\mathbf{M}}(z) &= \frac{1}{n|\mathbf{H}|} \sum_{i=1}^{n} \pi_i \left[\mathcal{Z}_i(z) \otimes \mathbf{Q}_{n,i}\right] \left[\mathbf{m}_{1i}', \mathcal{Z}_i(z)' \otimes \mathbf{m}_i'\right] \\
&= f(z) \left[\begin{array}{ccc} \mathbf{E}_{1x}(z) & \mathbf{E}_{1m}(z) & \mathbf{0}_{d\times(qp_2)}' \\ \mathbf{0}_{(dq)\times p_1} & \mathbf{0}_{d(q\times p_2)} & \mu_{1,2}(k) \mathbf{I}_q \otimes \mathbf{E}_{1m}(z) \end{array} \right] + \\
&\quad O_p\left(\|\mathbf{H}\|^2 + (n|\mathbf{H}|)^{-1/2}\right),
\end{aligned} \tag{S1.22}$$

where $\mathbf{E}_{1x}(z) \equiv n^{-1}\sum_{i=1}^{n}\mathbb{E}[\mathbf{Q}_{n,i}\mathbf{m}_{1i}'|\mathbf{z}_i = z]$, $\mathbf{E}_{1m}(z) \equiv n^{-1}\sum_{i=1}^{n}\mathbb{E}[\mathbf{Q}_{n,i}\mathbf{m}_i'|\mathbf{z}_i = z]$ and $\mathbf{E}_1(z) \equiv \begin{bmatrix} \mathbf{E}_{1x}(z) & \mathbf{E}_{1m}(z) \end{bmatrix}$. Also, we have

$$\mathbf{E}_2(\mathbf{H}, z) \equiv n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[\operatorname{tr}\left\{\mathbf{H}\frac{\partial^2[\beta_2(z)'\mathbf{x}_{2,i}]}{\partial z\partial z'}\mathbf{H}\right\}\mathbf{Q}_{n,i}'\middle|\mathbf{z}_i = z\right]. \tag{S1.23}$$

Next, denoting $\bar{\mathbf{u}} \equiv \mathbf{G}_n\mathbf{u}$ and $\bar{\mathbf{y}} \equiv \mathbf{G}_n[\mathbf{X}_1\boldsymbol{\beta}_1 + \operatorname{mtx}\{\mathbf{X}_2, \boldsymbol{\beta}_2(\mathbf{Z})\}]$, we have $\mathbf{W}\mathbf{y} = \bar{\mathbf{y}} + \bar{\mathbf{u}}$, where the ith elements of the $n\times 1$ vectors $\bar{\mathbf{u}}$ and $\bar{\mathbf{y}}$ are respectively denoted by $\bar{u}_i$ and $\bar{y}_i$. Then, for any given l, we have

$$\begin{aligned}\frac{\widetilde{\Psi}_{1,l}}{n|\mathbf{H}|} &= \frac{1}{n|\mathbf{H}|}\sum_{i=1}^{n}\sum_{j=1}^{n}(p_{l,ij}\pi_j + p_{l,ji}\pi_i)u_i\left[\bar{u}_i, \mathbf{x}_{1i}', \left(\mathcal{Z}_j'(z)\otimes\mathbf{m}_j'\right)\right] \\ &= 2f(z)\widetilde{\mathbf{F}}_{1,l}(\mathbf{H}, z) + o_p(1)\end{aligned} \tag{S1.24}$$

and

$$\begin{aligned}\frac{\widetilde{\Psi}_{2,l}}{n|\mathbf{H}|} &= \frac{1}{2n|\mathbf{H}|}\sum_{i=1}^{n}\sum_{j=1}^{n}(p_{l,ij}\pi_j + p_{l,ji}\pi_i)\mathbf{m}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\left[\mathbf{m}_{1i}', \left(\mathcal{Z}_j'(z)\otimes\mathbf{m}_j'\right)\right] \\ &= \frac{1}{2n|\mathbf{H}|}\sum_{i=1}^{n}\sum_{j=1}^{n}p_{l,ij}\pi_j\mathbf{m}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\left[\bar{u}_i, \mathbf{x}_{1i}', \left(\mathcal{Z}_j'(z)\otimes\mathbf{m}_j'\right)\right] + o_p(1) \\ &= \frac{1}{2}f(z)\widetilde{\mathbf{F}}_{2,l}(\mathbf{H}, z) + o_p(1),\end{aligned} \tag{S1.25}$$

where $\mathbf{F}_{21,l}(z) = n^{-1}\sum_{i\neq j}\mathbb{E}\left[p_{l,ij}\pi_j\mathbf{m}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\mathbf{x}_{1j}'|\mathbf{z}_j = z\right]$, $\mathbf{F}_{22,l}(z) = n^{-1}\sum_{i\neq j}\mathbb{E}\left[p_{l,ij}\mathbf{m}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\mathbf{m}_j'|\mathbf{z}_j = z\right]$,

$$\begin{aligned}\widetilde{\mathbf{F}}_{1,l}(\mathbf{H}, z) &= \left[\sigma_u^2(n|\mathbf{H}|f(z))^{-1}\sum_{i=1}^{n}p_{l,ii}g_{ii}\pi_i, \mathbf{0}_{p_1}', [1, \mathbf{0}_q']\otimes\mathbf{F}_{1,l}(\mathbf{H},z)\right] \\ \widetilde{\mathbf{F}}_{2,l}(\mathbf{H}, z) &= \left[0, \mathbf{F}_{21,l}(z), [1, \mathbf{0}_q']\otimes\mathbf{F}_{22,l}(z)\right]\end{aligned} \tag{S1.26}$$

and

$$\widetilde{\mathbf{F}}_{3,l}(\mathbf{H}, z) = \frac{1}{n}\sum_{i\neq j}\operatorname{tr}\left\{\mathbf{H}\mathbb{E}\left[p_{l,ij}\mathbf{m}_i'\boldsymbol{\Pi}(\mathbf{z}_i^*)\frac{\partial^2\left(\mathbf{x}_{2j}'\boldsymbol{\beta}_2(z)\right)}{\partial z\partial z'}\middle|\mathbf{z}_j = z\right]\mathbf{H}\right\}. \tag{S1.27}$$

In addition, we have $\widetilde{\Gamma}_{1,l}/(n|\mathbf{H}|) = O\left(\left(n\sqrt{\mathbf{H}}\right)^{-1}\right)$ since $\mathbf{F}_{4,l}(\mathbf{H}, z) = 0$.

Next, we consider case *(ii)*. In this instance, $\mathbf{m}_{1i} = [\bar{u}_i, \mathbf{x}_{1i}']'$ and $\mathbf{E}_1(z) = \left[0, n^{-1}\sum_{i=1}^{n}\mathbb{E}[\mathbf{Q}_{n,i}\mathbf{x}_{1i}'|\mathbf{z}_i = z], n^{-1}\sum_{i=1}^{n}\mathbb{E}[\mathbf{Q}_{n,i}\mathbf{x}_{2i}'|\mathbf{z}_i = z]\right]$. Also, if $\boldsymbol{\beta}_2(z) = \mathbf{0}_p$ holds over its domain, we have $\mathbf{C}(z) = \mathbf{0}_n$, which implies $\Gamma_{1,l} = 0$, $\Gamma_{2,l} = 0$ and $\widetilde{\Psi}_{2,l} = \mathbf{0}_{[p_1+1+p_2(q+1)]}'$. Hence, $(n|\mathbf{H}|)^{-2}\xi_n A_{n1}(z) = (2n|\mathbf{H}|)^{-2}\sum_{l=1}^{m}\Gamma_{3,l}\widetilde{\Psi}_{1,l}' = O_p\left(\left(n\sqrt{|\mathbf{H}|}\right)^{-1}\right)$. This completes the proof of this lemma. ■

Lemma 13 *Under Assumptions **1**–**3** and **9**, we obtain*

$$\frac{\xi_n^2}{(n|\mathbf{H}|)^2}B_n(z) = f^2(z)\varkappa_B(\mathbf{H}, z) + O_p\left((\xi_n\|\mathbf{H}\|)^{-1} + (\xi_n\|\mathbf{H}\|)^{-2}\right) + o_p(1), \tag{S1.28}$$

where

$$\begin{aligned}
\varkappa_{B,Q}(\mathbf{H},z) &= \left[\begin{array}{ccc}
\mathbf{E}_{1x}(z)'\mathbf{E}_{1x}(z) & \mathbf{0}_{p_1\times(p_2+1)} & \mathbf{0}_{p_1\times(q(p_2+1))} \\
\mathbf{0}_{d\times p_1} & \mathbf{E}_{1m}(z)\mathbf{E}_{1m}(z)' & \mathbf{0}_{d\times(q(p_2+1))} \\
\mathbf{0}_{(dq)\times p_1} & \mathbf{0}_{(dq)\times(p_2+1)} & \mu_{1,2}^2(k)\mathbf{I}_q\otimes\left[\mathbf{E}_{1m}(z)\mathbf{E}_{1m}(z)'\right]
\end{array}\right] \\
\varkappa_{B,P}(\mathbf{H},z) &= \sum_{l=1}^{m}\left[2\widetilde{\mathbf{F}}_{1,l}(\mathbf{H},z)+\frac{1}{2}\widetilde{\mathbf{F}}_{2,l}(z)\right]'\left[2\widetilde{\mathbf{F}}_{1,l}(\mathbf{H},z)+\frac{1}{2}\widetilde{\mathbf{F}}_{2,l}(z)\right].
\end{aligned}$$

Proof. The proof is omitted since it closely follows that of Lemma 5. ■

Lemma 14 *Under Assumptions **1**–**3** and **9**, we obtain*

$$\frac{\xi_n A_{n,2}(z)}{(n|\mathbf{H}|)^{3/2}} \xrightarrow{d} \mathbb{N}\Big(\mathbf{0},\ f^3(z)R_2(K)\mathbf{\Omega}(z)\Big), \tag{S1.29}$$

where $\mathbf{\Omega}(z)$ *is given in the proof.*

Proof. Since the proof closely follows that of Lemma 6, we omit the details to conserve space. In this lemma, we have

$$\Delta_n(z) = \left[\begin{array}{ccc}
\mathbf{E}_{1x}(z) & \mathbf{E}_{1m}(z) & \mathbf{0}'_{d\times(qp_2)} \\
\mathbf{0}_{(dq)\times p_1} & \mathbf{0}_{d(q\times p_2)} & \mu_{1,2}(k)\mathbf{I}_q\otimes\mathbf{E}_{1m}(z)
\end{array}\right],$$

so that

$$\mathbf{\Omega}_Q(z) = \left[\begin{array}{cc}
\mathbf{E}_1(z)'\mathbf{E}_3(z)\mathbf{E}_1(z) & \mathbf{0}_{p_1\times[q(p_2+1)]} \\
\mathbf{0}_{[q(p_2+1)]\times p_1} & R_2^{-1}(K)\mu_{1,2}^2(k)\mu_{2,2}(k)\mathbf{I}_q\otimes\left[\mathbf{E}_{1m}(z)'\mathbf{E}_3(z)\mathbf{E}_{1m}(z)\right]
\end{array}\right]. \tag{S1.30}$$

In addition, we obtain

$$\begin{aligned}
\mathbf{\Omega}_P(z) &= \sum_{l=1}^{m}\sum_{l'=1}^{m}\widetilde{\mathbf{F}}_{2,l}(z)'\widetilde{\mathbf{F}}_{2,l'}(z)\Bigg[\mathbb{V}ar\left(u_1^2\right)\lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[p_{l,ii}p_{l',ii}|\mathbf{z}_i=z\right]+ \\
&\quad \frac{1}{4}\lim_{n\to\infty}\frac{1}{n}\sum_{i'\neq i\neq j}\mathbb{E}\left[p_{l,ij}p_{l',i'j}\bar{\mathbf{m}}_i'\mathbf{\Pi}(\mathbf{z}_i^*)\bar{\mathbf{m}}_{i'}'\mathbf{\Pi}(\mathbf{z}_{i'}^*)|\mathbf{z}_j=z\right]\Bigg]
\end{aligned} \tag{S1.31}$$

and

$$\begin{aligned}
\mathbf{\Omega}_{Q,P}(z) &= 2^{-1}\sum_{l=1}^{m}\widetilde{\mathbf{F}}_{2,l}(z)'\Bigg[\Bigg(\sum_{l'=1}^{m}\lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}\mathbb{E}\left(p_{l',ii}\mathbf{Q}_{n,i}'|\mathbf{z}_i=z\right)+ \\
&\quad \lim_{n\to\infty}\frac{1}{n}\sum_{i\neq j}\mathbb{E}\left[p_{l,ij}\bar{\mathbf{m}}_i'\mathbf{\Pi}(\mathbf{z}_i^*)\mathbf{Q}_{n,j}'|\mathbf{z}_j=z\right]\Bigg)\lim_{n\to\infty}\mathbf{E}_1(z),\mathbf{0}_{qp_2}'\Bigg].
\end{aligned} \tag{S1.32}$$

This completes the proof of this lemma. ■

Lemma 15 *Under Assumptions **1–3**, **5** and **9**, we obtain*

$$\frac{1}{n^2 \left|\mathbf{H}\right|^{3/2}} \sum_{i=1}^{n} f^{-2}\left(\mathbf{z}_i\right) \mathbf{S}_{p_1,p_1+q(p_2+1)} \varkappa_B \left(\mathbf{H}, z_i\right)^{-1} A_{n2}\left(\mathbf{z}_i\right) \overset{d}{\to} \mathbb{N}\Big(\mathbf{0}_{p_1},\ \mathbf{\Omega}_1\Big), \tag{S1.33}$$

where $\mathbf{\Omega}_1$ has the same formula as does $\mathbf{\Omega}_0$ given in Lemma 11 with the newly defined $\Delta_n(z)$, $\widetilde{\mathbf{F}}_{2,l}(\mathbf{H}, z)$ and $\chi_B(\mathbf{H}, z)$.

Proof. The proof closely follows that of Lemma 11 and is therefore omitted. ■

S2 Empirical Application: Hedonic House Pricing

Hedonic valuation has become a standard approach to analyzing the effects of various qualitative attributes on house prices. Such qualitative factors primarily include structural, accessibility and neighborhood characteristics of the property. Following the first wave of studies that relied on linear specifications, in order to obtain more realistic marginal valuations of housing attributes, many researches have later attempted to allow for potential nonlinearities in hedonic price functions by resorting to semi- and/or nonparametric kernel methods [e.g., see Anglin & Gençay (1996), Parmeter et al. (2007) and the references therein]. At the same time, following Straszheim's (1975) argument that the location-wise variation in both housing attributes *and prices* is a crucial characteristic of the housing market, applied researchers have also become more aware of the importance of accounting for spatial effects in the housing data (Pace, Barry & Sirmans, 1998). This has led to a wide adoption of (linear) spatial econometric methods in the hedonic house pricing literature (e.g., Anselin & Lozano-Gracia, 2009; Baltagi & Bresson, 2011).

Given that our estimator constitutes a fusion of the two above approaches to hedonic house pricing, the latter appears to be a natural application to showcase the empirical relevance of our model. Specifically, we estimate a semiparametric spatial hedonic house price function using the well-known Harrison & Rubenfeld's (1978) house price data for 506 census tracts in the Boston Standard Metropolitan Statistical Area. In this paper, we use the corrected version of this dataset from Gilley & Pace (1996). Out of 14 non-constant explanatory variables, we limit our analysis to only those 10 that Gilley & Pace (1996) find to be statistically significant.

We model the log of census-tract-wise median house price (our outcome "y" variable) as a function of the following variables grouped into three categories: (i) structural characteristics – the number of rooms (squared); (ii) accessibility characteristics – distance to employment centers (logged) and the index of accessibility to radial highways (logged); and (iii) neighborhood characteristics – property tax rates, pupil-to-teacher ratio, black share of population, "lower status" share of population, crime rate, the Charles river dummy and the average nitrogen oxide ($\mathrm{NO_x}$) concentration in the air.[9] The first 9 explanatory variables are modeled linearly as "x" variables, whereas the $\mathrm{NO_x}$ concentration is permitted to enter the hedonic function in an unspecified nonlinear fashion as a contextual "z" variable that may also affect marginal valuations of other covariates. Our choice of the $\mathrm{NO_x}$ concentration as a smoothing variable, which we model in a flexible nonparametric way, is motivated by Harrison & Rubenfeld's (1978) findings in favor of nonlinearity in the relationship

[9]The black share of population is defined as $1000 \times (B - 0.63)^2$ where B is the proportion of blacks. The lower status share of population is defined as one half times the sum of the proportion of adults without some high school education and the proportion of male workers classified as laborers. See Harrison & Rubenfeld (1978) and Gilley & Pace (1996) for details about the data.

Table 5. Coefficient Function Point Estimates

Coefficient	*Semiparametric* 1st Qu.	Median	3rd Qu.	*Parametric*
Spatial Lag	0.0477	0.1594	0.3206	0.2502
	(0.1687)	(0.0962)	(0.1111)	(0.1001)
Constant	1.7244	2.2697	4.0294	2.6052
	(0.4530)	(0.4027)	(0.6585)	(0.4596)
# Rooms (Square)	–0.0029	0.0123	0.0239	0.0098
	(0.0050)	(0.0021)	(0.0020)	(0.0017)
Distance to Employment (Log)	–0.0256	–0.0041	0.0250	–0.1075
	(0.3906)	(0.2382)	(0.2430)	(0.0274)
Access to Highways (Log)	–0.0280	0.0115	0.0444	0.0579
	(0.0636)	(0.0513)	(0.0427)	(0.0218)
Property Tax Rate	–0.0008	–0.0007	0.0001	–0.0003
	(0.0005)	(0.0003)	(0.0006)	(0.0001)
Pupil-to-Teacher Ratio	–0.0892	–0.0235	–0.0009	–0.0185
	(0.0273)	(0.0144)	(0.0161)	(0.0079)
Black Share of Population	0.0004	0.0012	0.0017	0.0003
	(0.0002)	(0.0004)	(0.0007)	(0.0001)
Lower Status Share of Population	–0.0226	–0.0175	–0.0119	–0.0244
	(0.0055)	(0.0037)	(0.0047)	(0.0035)
Crime Rate	–0.0029	0.0056	0.0070	–0.0076
	(0.0056)	(0.0328)	(0.0138)	(0.0026)
Charles River	–0.0633	0.1834	0.4067	0.0484
	(0.0514)	(0.0829)	(0.1590)	(0.0457)

Notes: "Semiparametric" refers to a semiparametric smooth coefficient spatial autoregressive model in (2.1), whereas "parametric" refers to a standard linear spatial autoregressive model which treats coefficients as fixed parameters. Wild bootstrap standard errors are in parentheses.

between the air quality and house prices.[10]

We also explicitly recognize the spatial correlation between house prices and include the spatially lagged house price as an additional explanatory variable in the hedonic price function. By doing so, we are able to (at least partly) control for our lack of information on many relevant structural indicators as well as information on local amenities which all have effects on the house pricing. Since neighborhoods tend to be developed at the same time, houses are likely to have similar characteristics which would be incorporated in their prices. Furthermore, houses are usually valued on the basis of real estate agents' knowledge of prices in the neighborhood. Inclusion of the spatial lag term in the equation can thus accommodate these intrinsic characteristics of the housing market.

Before we can estimate our model, we first have to specify the spatial weighting matrix to be used in order to capture the spatial relationship between house prices. Given that our data are at the census track (as opposed to the housing unit) level, it appears that contiguity-based weights are a natural choice. Specifically, we use the first-order queen definition in order to construct the

[10]In their paper, Harrison & Rubenfeld (1978) limit their investigation of potential nonlinearities to the polynomial form, which leads them to settle on a quadratic specification. Our approach is far more flexible.

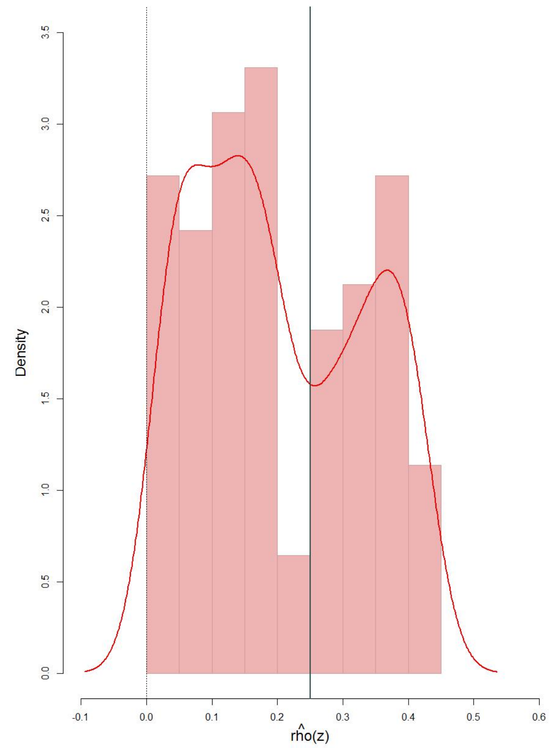

Figure 1. Kernel Density of the Spatial Lag Coefficient Function. [Note: The solid vertical line is the parametric point estimate of the coefficient.]

weights.[11] We estimate our model via our estimator in (3.8)[12] using the optimal bandwidth for the $\mathrm{NO_x}$ concentration, which we select using the data-driven leave-one-out cross-validation. Further, given that the number of instruments is greater than the number of unknown parameter functions, for potential efficiency gains, we use $\mathcal{P}_Q(z) = \mathbf{Q}(z)\left[\mathbf{Q}'(z)\mathbf{Q}(z)\right]^{-1}\mathbf{Q}(z)'$ as the feasible instrument matrix,[13] which implies that $\mathbf{\Xi}_H(z) \equiv \mathbf{K}_H(z)'\mathcal{P}_Q(z)\mathbf{K}_H(z)$ in (3.8).

In addition to our semiparametric spatial hedonic model, we also estimate its parametric (linear) counterpart, which we use as a baseline for comparison. That is, we estimate a standard spatial hedonic price function in the form of a spatial autoregressive model under the assumption of constancy of all coefficients in the equation. In the latter case, the house price is modeled as a function of its spatial lag and the 9 "x" variables, with the contextual $\mathrm{NO_x}$ variable suppressed from the equation. Table 5 reports the interquartile range and median of the coefficient function point estimates from our semiparametric model as well as point estimates of the coefficients from a baseline parametric model. The estimates are augmented with their respective standard errors which we compute via wild bootstrap with 499 replications.

We first focus on the spatial lag appearing in the hedonic price function. While the baseline

[11] We choose the order of the queen spatial weighting matrix that yields the highest R^2 measure, defined as the square of the correlation coefficient between the outcome variable and its fitted values. The *first*-order queen matrix produces the highest R^2 of 0.81, with the second highest R^2 of 0.57 produced by the third-order queen matrix.

[12] We choose (3.8) over (3.6) primarily for the sake of computational ease since the former has a closed-form solution. For simplicity, we also limit ourselves to the first-stage estimation only.

[13] In the spirit of the two-stage least squares estimator.

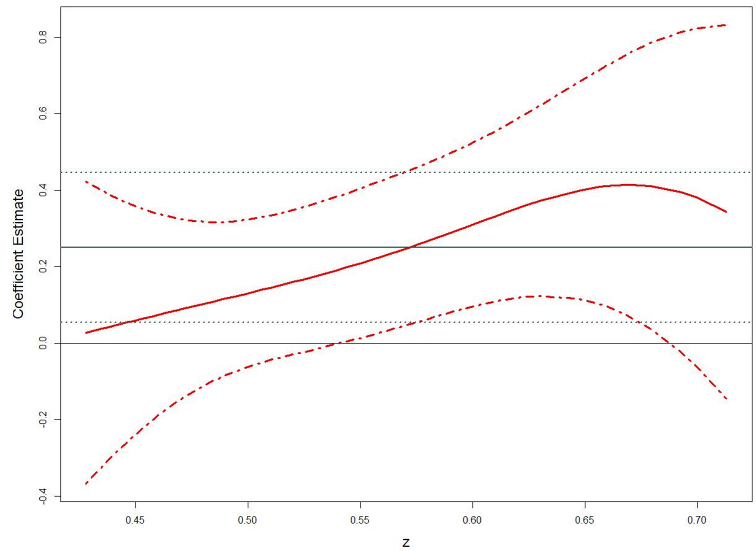

Figure 2. Spatial Lag Coefficient Function Estimates with the 95% Bootstrap Confidence Intervals. [Note: The solid horizontal line is the parametric point estimate of the coefficient with its corresponding confidence interval (dotted).]

parametric estimator indicates a statistically significant (at the 5% significance level) positive spatial dependence between house prices at the magnitude of 0.25, our preferred semiparametric model suggests the presence of stark heterogeneity in the degree of this dependence across neighborhoods. The semiparametric point estimates of the $\rho(\cdot)$ coefficient vary greatly across census tracks, with the corresponding interquartile range lying between 0.05 (suggesting no dependence) and 0.32 (suggesting relatively strong spatial autocorrelation). Figure 1, which plots the kernel density of these estimates, also points to bimodality of the distribution of $\rho(\cdot)$ that a more restrictive parametric model is not capable of detecting. Nonetheless, our semiparametric model also lends support to the premise of a positive spatial dependence in the housing market. The latter is an expected qualitative result and is in line with theoretical predictions. However, the spatial dependence in house prices is not statistically significant for all neighborhoods. To clearly see this, we plot the estimated spatial lag coefficient against the NO_x concentration in the air in Figure 2. The figure shows that spatial dependence between house prices is statistically significant only at higher values of the NO_x concentration in the air (of magnitudes between 0.55 and 0.68 pphm) and that the degree of this spatial dependence, on average, increases as the air quality declines. This finding suggests that locational similarity may matter little for house valuations in pollution-free localities.

We next proceed to the discussion of the remaining coefficients in the hedonic price function. Both the semiparametric and parametric point estimates are summarized in Table 5 and Figure 3. The baseline parametric specification produces empirical results that are qualitatively consistent with those reported in Gilley & Pace (1996) except for the coefficient on the Charles river dummy which the model finds to be insignificant. While our semiparametric model yet again uncovers remarkable heterogeneity in the coefficient function estimates across neighborhoods, our findings are quite qualitatively similar to those from a parametric model *at the median values*, except for distance to employment centers, crime rate and the Charles river variables. Our semiparametric model suggests the coefficient functions on the distance to employment and local crime rate are

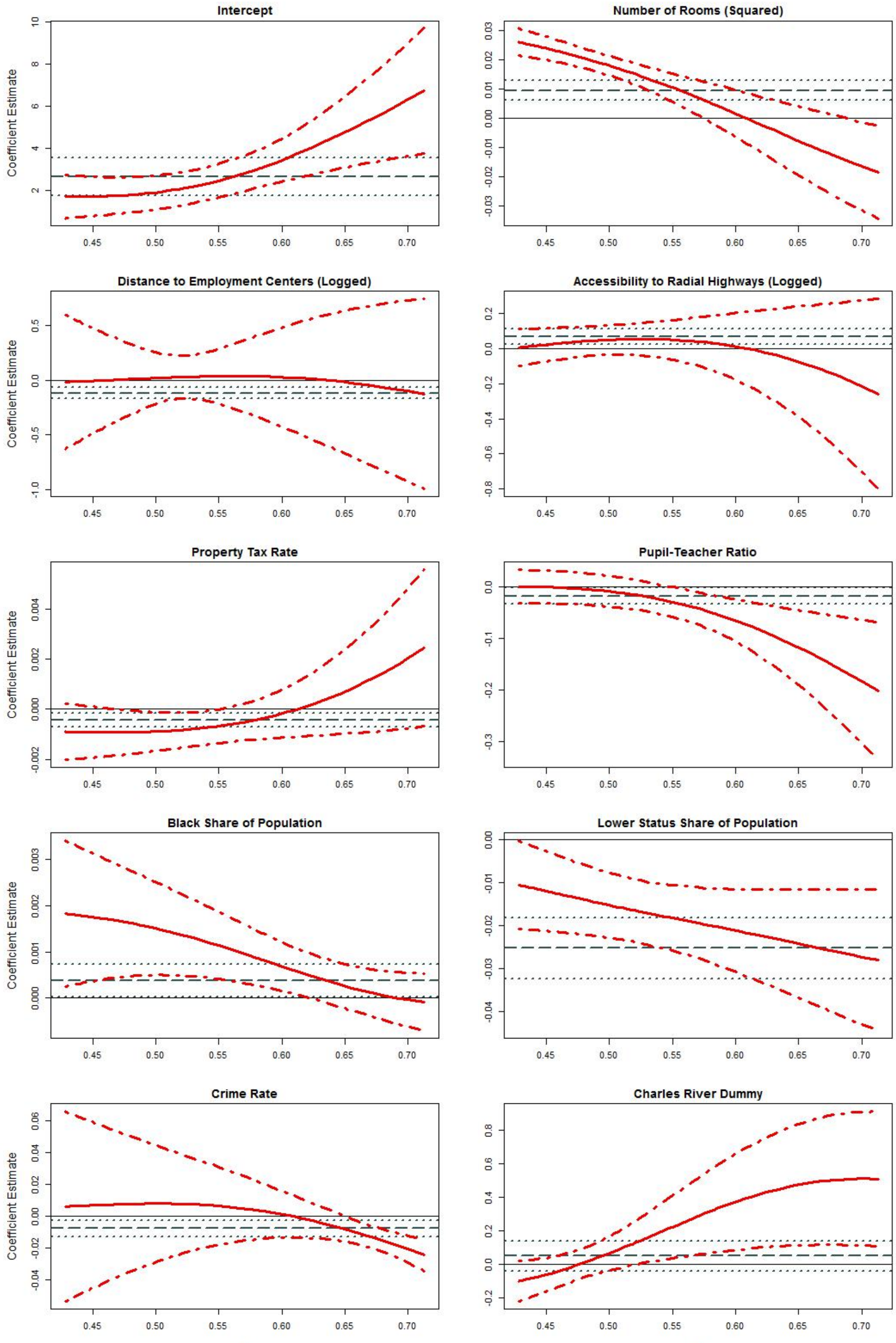

Figure 3. Coefficient Function Estimates with the 95% Bootstrap Confidence Intervals. [Note: The solid horizontal lines are parametric point estimates of the coefficients with their corresponding confidence intervals (dotted).]

largely statistically insignificant, whereas the opposite is true for the Charles river dummy. These findings are more evident in Figure 3. Also, from the figure we see that the coefficients on the number of rooms, pupil-to-teacher ratio, black and lower status shares of population tend to decline as the NO_x concentration in the air rises, indicating that these property characteristics seem to matter little (or even have an adverse impact) for house prices in neighborhoods with unclean air. In contrast, the (positive) coefficient on the Charles river dummy tends to increase with the NO_x pollution, which suggests that the river front may be regarded as a compensation for unclean air, i.e., dwellers might be willing to tolerate higher air pollution levels if living close to such a desirable "amenity" as the river.

However, the coefficient function estimates, which we have focused on until now, cannot be interpreted as representing marginal valuations (effects) of hedonic characteristics on price houses. To obtain the latter, we rewrite our model in a reduced form (2.2) as follows:

$$\mathbf{y} = \mathbf{S}_n(\mathbf{Z}) \times \text{mtx}\left\{\mathbf{X}, \boldsymbol{\beta}(\mathbf{Z})\right\} + \mathbf{S}_n(\mathbf{Z})\mathbf{u} = \sum_{j=1}^{p} \mathbf{S}_n(\mathbf{Z}) \left[\boldsymbol{\beta}_j(\mathbf{Z}) \odot \mathbf{x}_j\right] + \mathbf{S}_n(\mathbf{Z})\mathbf{u}, \tag{S2.1}$$

where "$\odot$" denotes the Hadamard product. In the case of our application, $\mathbf{y}$ is the 506×1 vector of median house prices; $\mathbf{X}$ is the 506×10 matrix of a unit vector (for the intercept) and 9 non-constant explanatory variables listed in the beginning of this section; $\mathbf{Z}$ is the 506×1 vector of the NO_x concentration. From (S2.1), the matrix of marginal valuations (effects) is given an $n \times n$ matrix $\partial \mathbb{E}\left[\mathbf{y}|\mathbf{X}, \mathbf{Z}\right] / \partial \mathbf{x}_j = \mathbf{S}_n(\mathbf{Z}) \times \text{diag}\left\{\boldsymbol{\beta}_j(\mathbf{Z})\right\}$, where, in the spirit of LeSage & Pace (2009), we label the diagonal elements of $\partial \mathbb{E}\left[\mathbf{y}|\mathbf{X}, \mathbf{Z}\right] / \partial \mathbf{x}_j$ as neighborhood-specific "direct marginal effects" (DMEs) and off-diagonal elements as neighborhood-pair-specific "indirect marginal effects". Given that each locality has $(n-1)$ indirect marginal effects, for the ease of exposition, we compute the neighborhood-specific averages of these marginal effects, which we label as "average indirect marginal effects" (AIMEs).

First, as one would have anticipated, we find that AIMEs tend to be incomparably smaller in magnitudes than DMEs across all variables and models. Second, DMEs and the coefficient (function) estimates are largely of the same magnitudes. Third, a parametric model predicts little (if any) variation in DMEs and AIMEs across neighborhoods. Lastly, based on our preferred semiparametric model, the two factors that seem to have the largest (magnitude-wise) statistically significant marginal effects on house prices are the pupil-to-teacher ratio and the Charles river dummy. Their respective mean DMEs are –0.07 and 0.18, suggesting that an increase in the pupil-to-teacher ratio by a unit leads to a decline in the median house prices by 7% whereas the river front constitutes a house price premium of 18% on average.

One may also be interested in the marginal effect of the NO_x concentration on the price of a house. However, obtaining a closed-form solution for $\partial \mathbb{E}\left[\mathbf{y}|\mathbf{X}, \mathbf{Z}\right] / \partial \mathbf{z}_r \ \forall\ r = 1 \ldots, q$ from (S2.1) is a quite tedious task since $\mathbf{z}_r$ appears inside the inverted matrix $\mathbf{S}_n(\cdot)$. Instead, we compute $\partial \mathbb{E}\left[\mathbf{y}|\mathbf{X}, \mathbf{Z}\right] / \partial \mathbf{z}$ using the numerical Jacobian. The estimated median and mean DME (AIME) of a unit increase in the NO_x concentration on house prices are –11% and –57% (0.9% and 0.5%), suggesting a strong negative net effect of poor air quality on house prices.

We complete the comparison of the two models with a formal specification test J_n. We bootstrap the test statistic 399 times to obtain the bootstrap p-value from its null distribution. The p-value is less than 10^{-10}, providing a strong evidence against a standard linear spatial autoregressive model of hedonic house pricing.

References

Anglin, P. M. & Gençay, R. (1996). Semiparametric estimation of a hedonic price function. *Journal of Applied Econometrics*, *11*, 633–648.

Anselin, L. & Lozano-Gracia, N. (2009). Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, *34*, 5–34.

Baltagi, B. & Bresson, G. (2011). Maximum likelihood estimation and Lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: An application to hedonic housing prices in Paris. *Journal of Urban Economics*, *69*, 24–42.

Gilley, O. W. & Pace, R. K. (1996). On the Harrison and Rubenfeld data. *Journal of Environmental Economics and Management*, *31*, 403–405.

Harrison, D. & Rubenfeld, D. L. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management*, *5*, 81–102.

Kelejian, H. H. & Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics*, *104*, 219–257.

LeSage, J. & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Boca Raton: Taylor & Francis CRC Press.

Pace, R. K., Barry, R., & Sirmans, C. F. (1998). Spatial statistics and real estate. *Journal of Real Estate Finance and Economics*, *17*, 5–13.

Parmeter, C. F., Henderson, D. J., & Kumbhakar, S. C. (2007). Nonparametric estimation of a hedonic price function. *Journal of Applied Econometrics*, *22*, 695–699.

Straszheim, M. R. (1975). *An Econometric Analysis of the Urban Housing Market*. New York: National Bureau of Economic Research.