Economics Bulletin

Appendix and Supplemental material not intended for publication-Round 3

Submission Number:EB-18-00997

Appendix

Email: nmrivera@alaska.edu. Part of this work was done during Rivera's doctoral studies at Michigan State University.
Submitted: Dec 14 2018. **Revised:** May 02, 2019.

A. Appendix

Let y, q, q^*, $\boldsymbol{\epsilon}$, and $\boldsymbol{\nu}'$ be $N \times 1$ vectors, and x be a $N \times 2$ vector (includes the unity). Let $\boldsymbol{\gamma}_1$ be a 2×1 vector containing γ_0 and γ_1, and M_x be the $N \times N$ complementary projection matrix of $x(x^T x)^{-1} x^T$. Using the FWL theorem, the OLS estimator of γ_2 can be expressed as:

$$\hat{\gamma_2} = (q^{*T} M_x q^*)^{-1} (q^{*T} M_x y) \tag{1}$$

$$\hat{\gamma_2} = (q^{*T} M_x q^*)^{-1} (q^{*T} M_x (x\boldsymbol{\gamma}_1 + q\gamma_2 + \boldsymbol{\epsilon})). \tag{2}$$

Using $q = x\boldsymbol{\beta}_1' + q^* \beta_2' + \boldsymbol{\nu}'$ and considering that $M_x x = 0$:

$$\hat{\gamma_2} = (q^{*T} M_x q^*)^{-1} (q^{*T} M_x ((q^* \beta_2' + \boldsymbol{\nu}')\gamma_2 + \boldsymbol{\epsilon})) \tag{3}$$

$$\hat{\gamma_2} = \gamma_2 \beta_2' + (q^{*T} M_x q^*)^{-1} (q^{*T} M_x (\boldsymbol{\nu}' \gamma_2 + \boldsymbol{\epsilon})). \tag{4}$$

In summation notation:

$$\hat{\gamma_2} = \gamma_2 \beta_2' + \left(n^{-1} \sum_j \sum_i q_j^* M_{x_{ij}} q_j^* \right)^{-1} n^{-1} \sum_j \sum_i q_j^* M_{x_ij} (\gamma_2 \nu_{ij}' + \epsilon_{ij}). \tag{5}$$

Taking the probability limit of $\hat{\gamma_2}$ as $n \to \infty$:

$$plim(\hat{\gamma_2}) = \gamma_2 \beta_2' + plim\left(\left(n^{-1} \sum_j \sum_i q_j^* M_{x_{ij}} q_j^* \right)^{-1} \right) plim\left(n^{-1} \sum_j \sum_i q_j^* M_{x_ij} (\gamma_2 \nu_{ij}' + \epsilon_{ij}) \right) \tag{6}$$

$$\begin{aligned} plim(\hat{\gamma_2}) = \gamma_2 \beta_2' + \gamma_2 & \underbrace{plim\left(n^{-1} \sum_j \sum_i q_j^* M_{x_{ij}} q_j^* \right)^{-1} plim\left(n^{-1} \sum_j \sum_i q_j^* M_{x_{ij}} \nu_{ij}' \right)}_{(*)} \\ & + \underbrace{plim\left(n^{-1} \sum_j \sum_i q_j^* M_{x_{ij}} q_j^* \right)^{-1} plim\left(n^{-1} \sum_j \sum_i q_j^* M_{x_{ij}} \epsilon_{ij} \right)}_{(*)} \end{aligned} \tag{7}$$

$$plim(\hat{\gamma_2}) = \gamma_2 \beta_2', \tag{8}$$

where $(*)$ uses $cov(q_j^*, \nu_{ij}') = cov(q_j^*, \epsilon_{ij}) = 0$, and $cov(x_{ij}, \nu_{ij}') = cov(x_{ij}, \epsilon_{ij}) = 0$.

Similarly, let M_{q^*} be the complementary projection matrix of $q^*(q^{*T} q^*)^{-1} q^{*T}$. Using the FWL theorem, the OLS estimator of vector $\boldsymbol{\gamma}_1$ is:

$$\hat{\boldsymbol{\gamma}_1} = (x^T M_{q^*} x)^{-1} (x^T M_{q^*} y) \tag{9}$$

$$\hat{\boldsymbol{\gamma}_1} = (x^T M_{q^*} x)^{-1} (x^T M_{q^*} (x\boldsymbol{\gamma}_1 + q\gamma_2 + \boldsymbol{\epsilon})). \tag{10}$$

Using $q = x\boldsymbol{\beta}_1' + q^* \beta_2' + \boldsymbol{\nu}'$, and considering that $M_{q^*} q^* = 0$:

$$\hat{\boldsymbol{\gamma}_1} = \boldsymbol{\gamma}_1 + \beta_1' \gamma_2 + (x^T M_{q^*} x)^{-1} (x^T M_{q^*} (\boldsymbol{\nu}' \boldsymbol{\gamma}_2 + \boldsymbol{\epsilon})). \tag{11}$$

In summation notation:

$$\hat{\gamma_1} = \gamma_1 + \beta_1^{'}\gamma_2 + \left(n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}x_{ij}\right)^{-1} n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}(\gamma_2\nu_{ij}^{'} + \epsilon_{ij}). \tag{12}$$

Taking the probability limit of $\hat{\gamma_1}$ as $n \to \infty$:

$$plim(\hat{\gamma_1}) = \gamma_1 + \beta_1^{'}\gamma_2 + plim\left(\left(n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}x_{ij}\right)^{-1}\right) plim\left(n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}(\gamma_2\nu_{ij}^{'} + \epsilon_{ij})\right) \tag{13}$$

$$plim(\hat{\gamma_1}) = \gamma_1 + \beta_1^{'}\gamma_2 + \gamma_2 \underbrace{plim\left(n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}x_{ij}\right)^{-1} plim\left(n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}\nu_{ij}^{'}\right)}_{(*)} + \underbrace{plim\left(n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}x_{ij}\right)^{-1} plim\left(n^{-1}\sum_j\sum_i x_{ij}M_{q_j^*}\epsilon_{ij}\right)}_{(*)} \tag{14}$$

$$plim(\hat{\gamma_1}) = \gamma_1 + \beta_1^{'}\gamma_2, \tag{15}$$

where $(*)$ uses $cov(x_{ij}, \nu_{ij}^{'}) = cov(x_{ij}, \epsilon_{ij}) = 0$, and $cov(q_{ij}^*, \nu_{ij}^{'}) = cov(q_{ij}^*, \epsilon_{ij}) = 0$.

A.1 Spatial Aggregation Bias in Presence of Omitted Variables

Previous results hold under the assumption that $cov(q_{ij}^*, \epsilon_{ij}) = 0$. When this assumption fails, and relevant variables affecting the outcome variable y_{ij} and related at the geographical level are omitted from the analysis, there is an additional source of bias. Let the regression of ϵ_{ij} on x_{ij} and q_j^* be as follows:

$$\epsilon_{ij} = x_{ij}\boldsymbol{\alpha}_1 + \alpha_2 q_j^* + \mu_{ij}. \tag{16}$$

When the assumption $cov(q_{ij}^*, \epsilon_{ij}) = 0$ fails, then $\alpha_2 \neq 0$ and so $plim(\hat{\gamma_2}) = \gamma_2\beta_2^{'} + \alpha_2$. Whether the omitted variable bias raises the magnitude of the estimated coefficient depends on the sign of α_2.