Advances in methods to support store location and design decisions

Hunneman, Auke

Publication date:
2011

Appendix: Model Specification and Estimation

A.1. The Spatial Error Random Effects Hierarchical Model

Assume that observations are sorted by store and, for each store, by time and then by zip code. Let Y_i (which can be $ln\left(NV_i\right)$ or $ln\left(EXP_i\right)$), X_i, and Z_i denote the observations, and let ε_i indicate the disturbance terms that are stacked for a particular store. The length of these vectors or matrices is store specific, because the number of zip codes within each store's trade area differs, and that number might change over time. Consequently, each vector or matrix consists of $\sum_{t=1}^{T_i} J_{it}$ observations. The full model can be written as:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_I \end{bmatrix} = \begin{bmatrix} \iota_1 \\ \iota_2 \\ \vdots \\ \iota_I \end{bmatrix} \gamma^{NV} + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_I \end{bmatrix} \alpha + \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_I \end{bmatrix} \beta + \begin{bmatrix} \iota_1 & 0 & . & 0 \\ 0 & \iota_2 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & \iota_I \end{bmatrix} \begin{bmatrix} \nu_1^{NV} \\ \nu_2^{NV} \\ \vdots \\ \nu_I^{NV} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_I \end{bmatrix}. \qquad \text{(A.1)}$$

Because $Var\left(\nu_i^{NV}\right) = \phi_{NV}^2$ and the covariance matrix Φ of the composite disturbance term $diag\left(\iota_1, \ldots, \iota_I\right) \times \nu_i^{NV} + \varepsilon$ is block diagonal, the i^{th} block diagonal is given by:

$$\Phi_i = \phi_{NV}^2 \iota_i \iota_i^{'} + \sigma_{NV}^2 \begin{bmatrix} \Omega_{iJ_{i1}} & 0 & . & 0 \\ 0 & \Omega_{iJ_{i2}} & . & 0 \\ . & . & . & . \\ 0 & 0 & . & \Omega_{iJ_{iT_i}} \end{bmatrix} = \phi_{NV}^2 \iota_i \iota_i^{'} + \sigma_{NV}^2 \Delta_i, \tag{A.2}$$

where $\Omega_{iJ_{it}}$ is the $\left(J_{it} \times J_{it}\right)$ covariance matrix for each cross-section of zip codes that belong to the trade area of store i at time t. This matrix takes the form $\Omega_{it} = \left[\left(I_{J_{it}} - \lambda^{NV} W_{it}\right)^{'}\left(I_{J_{it}} - \lambda^{NV} W_{it}\right)\right]^{-1}$.

The inverse of Φ_i is $\Phi_i^{-1} = \frac{\Delta_i^{-1}}{\sigma_{NV}^2} - \frac{\Delta_i^{-1}\iota_i}{\sigma_{NV}^2}\left[\frac{1}{\phi_{NV}^2} + \frac{\iota_i^{'}\Delta_i^{-1}\iota_i}{\sigma_{NV}^2}\right]^{-1}\frac{\iota_i^{'}\Delta_i^{-1}}{\sigma_{NV}^2}$ (Frees 2004).

If the total number of observations is denoted $nobs = \sum_{i=1}^{I}\sum_{t=1}^{T_i} J_{it}$, and $\varphi^2 = \left(\phi_{NV}^2 / \sigma_{NV}^2\right)$, the log-likelihood function can be written as:

$$\begin{aligned} LogL = \quad & -\frac{nobs}{2}\log\left(2\pi\sigma_{NV}^2\right) - \frac{1}{2}\sum_{i=1}^{I}\log\left|\varphi^2 \iota_i \iota_i^{'} + \Delta_i\right| \\ & -\frac{1}{2\sigma_{NV}^2}\sum_{i=1}^{I}\tilde{e}_i^{'}\left(\varphi^2 \iota_i \iota_i^{'} + \Delta_i\right)^{-1}\tilde{e}_i, \end{aligned} \tag{A.3}$$

where $\tilde{e}_i = Y_i - \left[\iota_i \; X_i Z_i\right]\left[\gamma_{NV} \; \alpha^{'} \beta^{'}\right]^{'}$, γ_{NV} is a scalar, α is a $\left(K \times 1\right)$ vector of α_k^{NV} s, and β is a $\left(N \times 1\right)$ vector of β_n^{NV} s.

According to Elhorst and Zeilstra (2007), the maximum likelihood estimators of the response parameters γ, α, and β (provided that X and Z do not include a lagged dependent or any endogenous explanatory variables) are equal to the generalized least squares (GLS) estimator:

$$\begin{bmatrix}\gamma\\ \alpha\\ \beta\end{bmatrix}_{GLS}=\left[\sum_{i=1}^{I}\begin{bmatrix}\iota_i & X_i & Z_i\end{bmatrix}^{'}\Phi_i^{-1}\begin{bmatrix}\iota_i & X_i & Z_i\end{bmatrix}\right]^{-1}$$
$$\times\left[\sum_{i=1}^{I}\begin{bmatrix}\iota_i & X_i & Z_i\end{bmatrix}^{'}\Phi_i^{-1}Y_i\right]$$
$$=\left[\sum_{i=1}^{I}\sum_{t=1}^{T_i}S_{it}^{*'}S_{it}^{*}-\sum_{i=1}^{I}\sum_{t=1}^{T_i}S_{it}^{*'}\iota_{it}^{*}\left[\frac{1}{\varphi^2}+\sum_{i=1}^{I}\sum_{t=1}^{T_i}\iota_{i't'}^{*'}\iota_{i't'}^{*}\right]^{-1}\iota_{it}^{*'}S_{it}^{*}\right]^{-1} \tag{A.4}$$
$$\times\left[\sum_{i=1}^{I}\sum_{t=1}^{T_i}S_{it}^{*'}Y_{it}^{*}-\sum_{i=1}^{I}\sum_{t=1}^{T_i}S_{it}^{*'}\iota_{it}^{*}\left[\frac{1}{\varphi^2}+\sum_{i=1}^{I}\sum_{t=1}^{T_i}\iota_{i't'}^{*'}\iota_{i't'}^{*}\right]^{-1}\iota_{it}^{*'}Y_{it}^{*}\right],$$

where $S_{it}^{*}=\left[\iota_{it}^{*}\ X_{it}^{*}Z_{it}^{*}\right]$, and the superscript * denotes the transformation $S_{it}^{*}=\left(I_{J_{it}}-\lambda^{NV}W_{it}\right)S_{it}$, applied to the variables ι_{it}, X_{it}, and Z_{it}. In addition, $\sigma_{NV}^{2}={}^{1}\!/_{nobs}\sum_{i=1}^{I}\tilde{e}_i^{'}\left(\varphi^2\iota_i\iota_i^{'}+\Delta_i\right)^{-1}\tilde{e}_i$.

In contrast, there is no closed-form solution for λ^{NV} and φ^2. Therefore, we develop an iterative, two-step estimation procedure, in which the two sets of parameters are estimated alternately until convergence; γ^{NV}, α, β, and σ_{NV}^2, given λ^{NV} and φ^2, can be estimated using the GLS estimator in Equation A.4, whereas λ^{NV} and φ^2, given γ^{NV}, α, β, and σ_{NV}^2, can be estimated by maximizing the log-likelihood in Equation A.3. A Matlab routine of this estimation procedure can be downloaded from the Web site http://www.regroningen.nl/elhorst/.

A.2. The Spatial Error Random Effects Model

Assume that the observations are sorted by time and then, for each time period, by zip code. Let Y_t ($=logit\left(PR_t\right)$) and $U_t=\left[\iota_t X_t\ Z_t\right]$ denote observations stacked within a particular time period (Y_t and U_t consist of J observations). Further assume that $\theta = \left[\gamma^{PR}\ \alpha^{'} \beta^{'}\right]^{'}$, where α is a $(K \times 1)$ vector of α_k^{PR}s, and β is a $(N \times 1)$ vector of β_n^{PR}s. The model then can be written as:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_T \end{bmatrix} \theta + \delta, \ \ \delta = \left(\iota_T \otimes I_J\right)\nu + \left(I_T \otimes B^{-1}\right)\varepsilon, \qquad \text{(A.5)}$$

where ι_T is a $(T \times 1)$ vector of unit elements, $\nu = \left(\nu_1^{PR}, \ldots, \nu_J^{PR}\right)$, ε is a $(TJ \times 1)$ vector of disturbance terms ε_{jt}^{PR}, and $B = \left(I_J - \lambda^{PR} W\right)$.

Anselin (1988) and Baltagi (2005) show that the log-likelihood function of this model is:

$$\begin{aligned} LogL = \ & -\frac{JT}{2}\log\left(2\pi\sigma_{PR}^2\right) - \frac{1}{2}\log\left|T\kappa^2 I_J + \left(B^{'}B\right)^{-1}\right| \\ & +(T-1)\log\left|B\right| - \frac{1}{2\kappa^2} e^{'}\left(\frac{1}{T}\iota_T\iota_T^{'} \otimes \left(T\kappa^2 I_J + \left(B^{'}B\right)^{-1}\right)\right)^{-1} e \qquad \text{(A.6)} \\ & -\frac{1}{2\sigma_{PR}^2} e^{'}\left(I_T - \frac{1}{T}\iota_T\iota_T^{'}\right) \otimes \left(B^{'}B\right) e, \end{aligned}$$

where $e = \left(e_1, \ldots, e_T\right)^{'}$, $e_t = Y_t - U_t\theta$, and $\kappa^2 = \left(\sigma_\nu^2 / \sigma_{PR}^2\right)$.

Elhorst (2003) also shows that if the determinants of the matrices $T\kappa^2 I_J + (B^{'}B)^{-1}$ and B are expressed as a function of the characteristic roots of W, denoted by ω_j ($j = 1, \ldots, J$), the log-likelihood function can be rewritten as:

$$\begin{aligned} LogL = & -\frac{JT}{2}\log\left(2\pi\sigma_{PR}^2\right) - \frac{1}{2}\sum_{j=1}^{J}\log\left(1 + T\kappa^2\left(1 - \lambda^{PR}\omega_j\right)^2\right) \\ & + T\sum_{j=1}^{J}\log\left(1 - \lambda^{PR}\omega_j\right) - \frac{1}{2\sigma_{PR}^2}\sum_{t=1}^{T}\tilde{e}_t^{'}\tilde{e}_t, \end{aligned} \qquad \text{(A.7)}$$

where $\tilde{e}_t = Y_t^* - U_t^*\theta$,

$$Y_t^* = P\overline{Y} + B\left(Y_t - \overline{Y}\right) = BY_t + \left(P - B\right)\overline{Y} = \left(I_J - \lambda^{PR}W\right)Y_t$$
$$+\left(P - \left(I_J - \lambda^{PR}W\right)\right)\overline{Y},$$

$U_t^* = \left(I_J - \lambda^{PR}W\right)U_t + \left(P - \left(I_J - \lambda^{PR}W\right)\right)\overline{U}$, and

P is such that $P'P = \left(T\kappa^2 I_J + \left(B'B\right)^{-1}\right)^{-1}$.

Consequently, the GLS estimator of the response parameters θ and σ_{PR}^2 can be computed as:

$$\theta = \left(u^{*'}u^*\right)^{-1}\left(u^{*'}y^*\right) \text{ and } \sigma_{PR}^2 = \frac{\sum_{t=1}^{T}\tilde{e}_t'\tilde{e}_t}{JT}, \tag{A.8}$$

where $u^* = \begin{bmatrix} U_1^* \\ \vdots \\ U_T^* \end{bmatrix}$, and $y^* = \begin{bmatrix} Y_1^* \\ \vdots \\ Y_T^* \end{bmatrix}$.

Just as in the previous case, there is no closed-form solution for λ^{PR} and κ^2, so we need an iterative two-step estimation procedure in which the two sets of parameters get estimated alternately until convergence occurs. A Matlab routine of this estimation procedure can be downloaded from the Web site http://www.regroningen.nl/elhorst/.