



Spatial dependence models for HLM2

| | |
|--|---|
| 1. Introduction..... | 1 |
| 2. Description of the data and making the MDM file | 1 |
| 3. Creating the command file..... | 3 |
| 3. Interpreting the output | 4 |

1. Introduction

The spatial dependence option in HLM2 allows researchers to handle nested data collected in spatial settings. In addition to the clustering effects, the spatial HLM2 models accommodate dependence induced by contiguity or proximity in geographic locations. This type of models has applications for clustered data collected from contiguous geographic locations such as school districts, counties, neighborhoods, and countries. Verbitsky-Savitz and Raudenbush (2009), for example, applied these models to exploit the spatial dependence of neighborhood social processes to considerably improve the precision and validity of assessment of neighborhoods.

2. Description of the data and making the MDM file

This example uses data collected by the Project of Human Development in Chicago Neighborhoods (Sampson, Raudenbush, & Earl, 1997) on 7,729 residents living in 342 neighborhoods. It is an unconditional model with a ten-item collective efficacy scale, defined as the fusion of social cohesion and informal social control, as the outcome.

For spatial HLM2 models, the level-1 and level-2 models have the same structure as those described in Section 2.5. These two data files for the example, linked by level-2 neighborhood cluster IDs, are RESIDENT.SAV and NEIGHBOR.SAV. In the level-1 data file, there is one variable, collective efficacy (COLLEFF). In the level-2 data file, a dummy variable is included. The spatial dependence analysis requires another data file with information on spatial contiguity. The information allows the program to create a spatial weight matrix, W , which is a binary contiguity matrix indicating that sites are contiguous to each other. ROOK.SAV, contains such information for our illustrative example. The variables followed by the neighborhood cluster IDs are:

- N1 – N10 (the first to the tenth adjoining neighborhoods, if any)
- COUNT (the total number of contiguous neighborhoods)

The data for the first ten neighborhoods are displayed below. Note that neighborhood 1 (that is, the neighborhood with ID = 1) shares a common boundary with two neighborhoods, specifically, neighborhoods 2 and 3. In contrast, neighborhood 2 shares a boundary with 4 neighborhoods, neighborhoods 8, 6, 3, and 1.

| | id | n1 | n2 | n3 | n4 | n5 | n6 | n7 | n8 | n9 | n10 | count |
|----|----|-----|-----|-----|-----|-----|----|----|----|----|-----|-------|
| 1 | 1 | 2 | 3 | . | . | . | . | . | . | . | . | 2 |
| 2 | 2 | 8 | 6 | 3 | 1 | . | . | . | . | . | . | 4 |
| 3 | 3 | 16 | 5 | 4 | 2 | 1 | . | . | . | . | . | 5 |
| 4 | 4 | 17 | 5 | 3 | . | . | . | . | . | . | . | 3 |
| 5 | 5 | 7 | 6 | 3 | 4 | 17 | . | . | . | . | . | 5 |
| 6 | 6 | 8 | 7 | 5 | 2 | . | . | . | . | . | . | 4 |
| 7 | 7 | 15 | 11 | 5 | 17 | 6 | . | . | . | . | . | 5 |
| 8 | 8 | 11 | 10 | 2 | 6 | . | . | . | . | . | . | 4 |
| 9 | 9 | 334 | 252 | 250 | 241 | 238 | . | . | . | . | . | 5 |
| 10 | 10 | 13 | 12 | 8 | . | . | . | . | . | . | . | 3 |

The file SPATIAL.MDMT stores the commands for creating the two-level multivariate data matrix file, SPATIAL.MDM. The procedure is very similar to those described for a standard HLM2 model. An extra step needed is to instruct the program to include spatial dependence information with the following procedure:

At the **Make MDM – HLM2** dialog box check the box for **Include spatial dependence matrix**. 2. Click **Browse** to select **ROOK.SAV**. Click **Choose Variables** to include the ID and the variables N1-

N10 and COUNT.

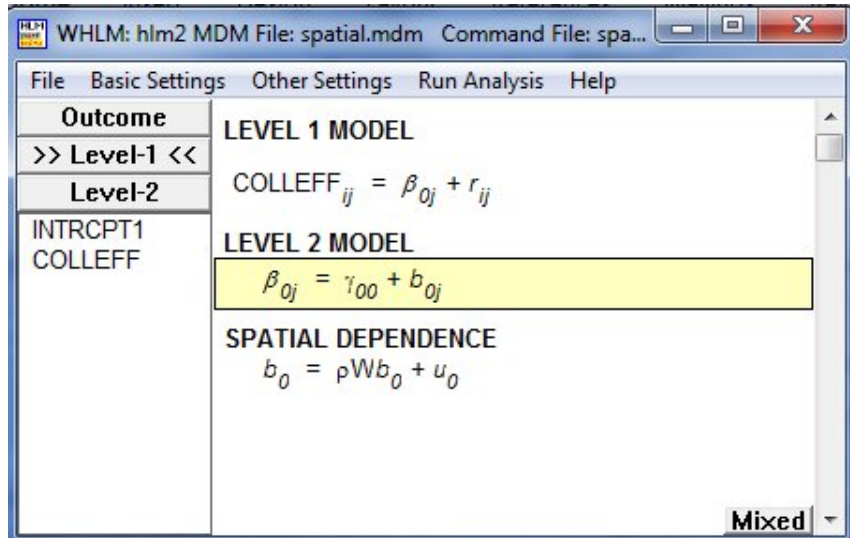
3. Creating the command file

The file SPATIAL.HLM contains the commands for setting up the unconditional model. An additional step is to instruct HLM2 to run the model as a spatial dependence model by the following procedure:

1. Open the **Other Settings** menu and select the **Estimation Settings**.
2. Check the box for **Run as spatial dependence model** (see below).

The screenshot shows the 'Estimation Settings - HLM2' dialog box. The 'Type of Likelihood' section has 'Full maximum likelihood' selected. The 'Adaptive Gaussian Quadrature Iteration Control' section has 'Do adaptive Gaussian iterations' unchecked, with input fields for 'Maximum number of iterations' and 'Number of quadrature points', and 'First derivative' and 'Second derivative' radio buttons. The 'LaPlace Iteration Control' section has 'Do EM Laplace iterations' unchecked and an input field for 'Maximum number of iterations'. Below these are checkboxes for 'Fixed Intercept, Random Coefficient', 'Diagonalize Tau', and 'Run as spatial dependence model' (checked). A row of buttons includes 'Constraint of fixed effects', 'Heterogeneous sigma^2', 'Automatic Imputation', and 'Multiple imputation'. Another row has 'Level-1 Deletion Variables', 'Weighting', 'Latent Variable Regression', and 'Plausible values'. A third row has 'Variable Selection Model', 'ITT effects', and 'IV Effects'. At the bottom, there is a 'Fix sigma^2 to specific value' field with 'computed' entered, and a note: '(Set to "computed" if you want sigma^2 random or if over-dispersion is desired)'. An 'OK' button is in the bottom right corner.

The model window gives the model specifications:



Note that a model for the spatial dependence model is given:

$$b_0 = \rho W b_0 + u$$

where as described in Verbitsky-Savitz and Raudenbush (2009),

- b_0 is a vector of level-2 random spatially autoregressive effects;
- ρ is a spatial correlation parameter with zero indicating no spatial dependence and positive or negative values indicating whether a site is typically surrounded by other sites with similar or different values on the outcome;
- W is the spatial weight matrix used in the analysis. As discussed earlier, it is constructed from ROOK.SAV; and
- u is the level-2 error.

3. Interpreting the output

A spatial dependence analysis using HLM2 provides two sets of results, one for regular HLM and the other HLM with spatial dependence. A comparison test of the fit of these models is performed and the result is given. Below is a partial output of the results of the unconditional model.

Here are the partial results for the regular HLM:

Iterations stopped due to small change in likelihood function

$$\sigma^2 = 0.42136$$

$$\text{Standard error of } \sigma^2 = 0.00693$$

τ
INTRCPT1, β_0 0.08904

Standard error of τ
INTRCPT1, β_0 0.00850

| Random level-1 coefficient | Reliability estimate |
|----------------------------|----------------------|
| INTRCPT1, β_0 | 0.799 |

The value of the log-likelihood function at iteration 6 = -7.911855E+003

Final estimation of fixed effects:

| Fixed Effect | Coefficient | Standard error | t-ratio | Approx. d.f. | p-value |
|---|-------------|----------------|---------|--------------|---------|
| For INTRCPT1, β_0 INTRCPT2, γ_{00} | 3.433243 | 0.018056 | 190.142 | 341 | <0.001 |

**Final estimation of fixed effects
(with robust standard errors)**

| Fixed Effect | Coefficient | Standard error | t-ratio | Approx. d.f. | p-value |
|---|-------------|----------------|---------|--------------|---------|
| For INTRCPT1, β_0 INTRCPT2, γ_{00} | 3.433243 | 0.018056 | 190.144 | 341 | <0.001 |

Final estimation of variance components

| Random Effect | Standard Deviation | Variance Component | d.f. | χ^2 | p-value |
|-----------------|--------------------|--------------------|------|------------|---------|
| INTRCPT1, u_0 | 0.29839 | 0.08904 | 341 | 1870.37148 | <0.001 |
| level-1, r | 0.64913 | 0.42136 | | | |

Statistics for the current model

Deviance = 15823.710765
Number of estimated parameters = 3

Results for the spatial dependence model follows next.

HLM with Spatial Dependence Model Results - Iteration 135

The value of the log-likelihood function at iteration 135 = -7.835990E+003
Iterations stopped due to small change in likelihood function

$$\sigma^2 = 0.42149$$

τ
INTRCPT1, β 0.03477

ρ
INTRCPT1, β 0.81701

Final estimation of fixed effects:

| Fixed Effect | Coefficient | Standard error | t-ratio | Approx. d.f. | p-value |
|---|-------------|----------------|---------|--------------|---------|
| For INTRCPT1, β_0 INTRCPT2, γ_{00} | 3.404181 | 0.056443 | 60.312 | 341 | <0.001 |

Statistics for the current model

Deviance = 15671.980461
Number of estimated parameters = 4

Regular HLM vs. HLM with spatial dependence model comparison test

χ^2 statistic = 151.73030
Degrees of freedom = 1
p-value = <0.001

Average Level-2 Variance = 0.088502

Average Level-2 Covariance = 0.005961

The average level-2 variance is the average of the neighborhood-specific variance. These depend on τ , but also on the magnitude of the spatial dependence correlation, ρ , and the configuration of neighborhoods near that neighborhood. The average level-2 covariance is the average covariance between pairs of contiguous neighborhoods.

Two features of the results are noteworthy:

- The result of the comparison test provides evidence that the HLM with spatial dependence provides a better fit, as indicated by the χ^2 statistic of 151.73, $df=1$, $p < .001$.
- A comparison of the standard errors for $\hat{\gamma}_{00}$ the regular HLM and HLM with spatial dependence (.018 vs .056) suggests that, given $\hat{\rho}$ is equal to .8, that there is an underestimation of the standard errors when spatial dependence is ignored.

Users can also obtain spatial empirical Bayes estimates of the neighborhood collective efficacy measures by following the procedure as specified in Section 2.5.4.2. The figure below shows the ten records of the residual file for the unconditional model.

| | l2id | nj | u_intrcp | b_intrcp | fwintrcp |
|----|------|----|----------|----------|----------|
| 1 | 1 | 39 | -.213 | -.420 | 3.404 |
| 2 | 2 | 43 | -.168 | -.292 | 3.404 |
| 3 | 3 | 15 | .021 | -.214 | 3.404 |
| 4 | 4 | 41 | -.128 | -.203 | 3.404 |
| 5 | 5 | 15 | .101 | -.021 | 3.404 |
| 6 | 6 | 13 | -.099 | -.155 | 3.404 |
| 7 | 7 | 39 | -.113 | -.138 | 3.404 |
| 8 | 8 | 14 | .192 | .178 | 3.404 |
| 9 | 9 | 45 | .287 | .328 | 3.404 |
| 10 | 10 | 17 | .281 | .376 | 3.404 |

Data View Variable View

U_INTRCP and B_INTRCP are the two Empirical Bayes for the regular HLM and the HLM with spatial dependence. For a discussion of the properties of the empirical Bayes estimator that exploits spatial dependence, see Verbitsky-Savitz and Raudenbush (2009). Spatial dependence models handle continuously distributed as well as discrete outcomes, including binary outcomes, counted data, ordered categories, and multinomial outcomes.