**Code Documentation for the Estimation of a Spatial Multiple Discrete-Continuous Probit Model**

This addendum serves to document the estimation code for a spatial multiple discrete continuous (MDC) probit model, which allows the dependent variable to exist in multiple discrete states with an intensity associated with each discrete state and at the same time explicitly accommodating spatial dependencies. An explanation of the parameters in the model and the different settings available within the code are documented below. Please refer to Bhat, Dubey, Jobair Bin Alam and Khushefati (2013) for the notations and the model structure.

Bhat, C.R., S.K. Dubey, M. Jobair Bin Alam and W.H. Khushefati (2013) "**A New Spatial Multiple Discrete-Continuous Modeling Approach to Land Use Change Analysis**," Technical paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, 2013.

1. **Data Generation Settings**

* The current version of the code allows a user to generate the data within the code as well as specify the external dataset.
* If the user wants to simulate the data inside the code, choose the option “Simulation\_data” (line 21). On the other hand, if the user wants to specify his or her own dataset, choose the option “Own\_data” (line 19) and specify the file name in the section “How to specify your own dataset” in the code. The user can also change the seed for data generation process under the simulation option by changing the value of variable “Run\_No” (line 50).

1. **Dataset Specifications**

The dataset should be a csv file with no header. The dataset should include the columns in the following order.

1. Columns of exogenous variables for each of the inside alternatives.
2. Columns of exogenous variables for the outside alternative.
3. A column filled with the value 1.
4. A column filled with the value 0.
5. Columns of consumption value for all the inside alternatives.
6. Column of consumption value of the outside alternative.
7. **Simulation Data**

The data in the code is generated using the simulation configuration discussed in the paper titled “**A New Spatial Multiple Discrete-Continuous Modeling Approach to Land Use Change Analysis”.**  Please refer to the paper for details of simulation configuration.

1. **Code Settings**

The user must specify the value of following variables (lines 30 – 40).

**Table 1: Description of Variables**

|  |  |
| --- | --- |
| **Variables** | **Description** |
| nind | Number of observations |
| nvar | Number of variables in the utility equation per alternative |
| nc | Number of alternatives in the model |
| nran | Number of random parameters in the model |
| cutoff\_distance | Distance band for pairing of observations in the likelihood expression |

1. **Sample Data**

The data consist of 2000 individual record (thus, 2000 grids). The number of alternatives in the dataset is equal to four with three explanatory variables each. The following table presents the content and the structure of the sample dataset.

**Table 2: Sample Data Description**

|  |  |
| --- | --- |
| **Column No** | **Explanation** |
| 1 to 12 | Explanatory variables for each of the four alternatives in the  order specified in the section dataset specification |
| 13 | A column of 1’s |
| 14 | A column of 0’s |
| 15 to 18 | Columns of consumption value for each of the four alternative in  the order specified in the section dataset specification |

We do not provide labels for the explanatory variables because the labels are generated internally by the program. Please refer to line number 122 for how to define labels for variables. The sample data follows the same structure as discussed in the simulation exercise of the paper.

1. **Estimation Results**

The code has two likelihood expressions as some sort of parameterization is necessary to ensure that all the spatial parameters have a value between zero and one at the end of the estimation. Thus, the first likelihood expression estimates the parameterized coefficients and then passes the un-parameterized values to the second likelihood expression.

Since the standard error calculations for spatial models are slightly different from non-spatial models, please ignore all the standard error values printed by Gauss. The correct standard error and the corresponding t-stat value for the un-parameterized coefficients are printed at the end under the section titled “Final Result,” as shown below.

Here we assume generic coefficients for three exogenous variables. We also consider the first variable to be fixed, but allow the remaining two variables to be normally distributed. Thus, we estimate three generic coefficients (mean effect represented by X01, X02 and X03), three Cholesky elements (lower triangular matrix) of the random parameter covariance matrix (represented by OmTild01, OmTild02 and OmTild03), three translation parameters for each of the three inside alternatives (Gamma01, Gamma02 and Gamma03), five Cholesky elements (lower triangular matrix) of the error differenced covariance matrix (Psi02, Psi03, Psi04, Psi05 and Psi06) and three spatial lag coefficients (Del01, Del02 and Del03) for each of three inside alternatives. Please note that for all the Cholesky decomposition elements, the corresponding position of the elements in the matrix is added at the end of their name. For example: OmTild01 indicates that this element is the 1st element of the Cholesky decomposition (lower triangular matrix) of the random parameter covariance matrix.

-----------------Final Result---------------------------------------------------------------------------------

Log-likelihood value : -108796.08

Parameter Estimate ST.Error T-Stat

X01 0.443 0.0088 50.20

X02 -1.069 0.0085 -125.41

X03 1.000 0.0062 161.20

OmTild01 0.892 0.0092 97.23

OmTild02 0.603 0.0116 51.96

OmTild03 0.729 0.0121 60.35

Gamma01 0.921 0.0338 27.21

Gamma02 0.340 0.0127 26.68

Gamma03 0.047 0.0015 30.70

Psi02 0.750 0.0086 87.25

Psi03 0.939 0.0060 155.70

Psi04 0.940 0.0099 94.82

Psi05 0.195 0.0081 23.95

Psi06 0.963 0.0063 153.12

Del01 0.777 0.0122 63.95

Del02 0.618 0.0281 21.97

Del03 0.661 0.0192 34.38