Online supplement to

"On Accommodating Spatial Interactions in a Generalized Heterogeneous Data Model (GHDM) of Mixed Types of Dependent Variables"

Chandra R. Bhat (corresp. author), Abdul R. Pinjari, Subodh K. Dubey, and Amin S. Hamdi

SECTION 1

1.1 Description of Matrix M used to Derive the Distribution of Vector *yu*

Create a matrix **M** of size $\left[Q(E+\tilde{G})\times Q(E+\ddot{G})\right]$ whose elemens are all initialized to zero. Then, for every individual *q*, consider a matrix $\mathbf{M}_q$ of size $\left[(E+\tilde{G})\times(E+\ddot{G})\right]$ filled with zeros. Insert an identity matrix of size *E* into the first *E* rows and *E* columns of the matrix $\mathbf{M}_q$. Next, consider the rows from $E+1$ to $E+I_1-1$, and columns from $E+1$ to $E+I_1$. These rows and columns correspond to the first nominal variable. Insert an identity matrix of size (I_1-1) after supplementing with a column of '-1' values in the column corresponding to the chosen alternative. Next, consider the rows $E+I_1$ through $E+I_1+I_2-2$ and columns $E+I_1+1$ through $E+I_1+I_2$ correspond to the second nominal variable. Again position an identity matrix of size (I_2-1) after supplementing with a column of '-1' values in the column corresponding to the chosen alternative for the second nominal variable. Continue this procedure for all *G* nominal variables. Insert this matrix $\mathbf{M}_q$ in the matrix **M** occupying the rows $\left[(q-1)*(E+\tilde{G})+1\right]$ to $\left[(q-1)*(E+\tilde{G})+(E+\tilde{G})\right]$ and columns $\left[(q-1)*(E+\ddot{G})+1\right]$ to $\left[(q-1)*(E+\ddot{G})+(E+\ddot{G})\right]$.

1.2 Rearrangement Matrices R_y and $R_{\tilde{u}}$ used to Partition $\tilde{B}$ and $\tilde{\Omega}$

Consider a rearrangement matrix **R** of size $\left[Q(E+\tilde{G})\times Q(E+\tilde{G})\right]$ filled with zeros. Then, for every individual *q*, consider a identity matrix $\mathbf{R}_q$ of size $(E+\tilde{G})$. Insert first *H* rows of matrix $\mathbf{R}_q$ into matrix **R** occupying rows $\left[(q-1)*H+1\right]$ to $\left[(q-1)*H+H\right]$ and columns $\left[(q-1)*(E+\tilde{G})+1\right]$ to $\left[(q-1)*(E+\tilde{G})+(E+\tilde{G})\right]$. Next insert the remaining $(E+\tilde{G}-H)$ rows of matrix $\mathbf{R}_q$ into matrix **R** occupying rows $\left[QH+(q-1)*(E+\tilde{G}-H)+1\right]$ to $\left[QH+(q-1)*(E+\tilde{G}-H)+(E+\tilde{G}-H)\right]$ and columns $\left[(q-1)*(E+\tilde{G})+1\right]$ to $\left[(q-1)*(E+\tilde{G})+(E+\tilde{G})\right]$. Divide the matrix **R** into two submatrices $\mathbf{R}_{\mathbf{y}}(\mathbf{R}[1{:}QH,1:Q(E+\tilde{G})])$ and $\mathbf{R}_{\tilde{u}}(\mathbf{R}[QH+1:Q(E+\tilde{G}),1:Q(E+\tilde{G})])$. For example: Consider the case with two individuals and one continous, two ordinal, one count and two nominal variables with three alternatives each. Then the matrix **R** may be written as:

$$
\mathbf{R}=\left[\begin{array}{cccccccccccccccc}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
\hline
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1
\end{array}\right]=\left[\begin{array}{c}\mathbf{R}_{\mathrm{y}}\\ \hline \mathbf{R}_{\tilde{\mathbf{u}}}\end{array}\right] \tag{S.1}
$$

Now, consider the following re-arranged vectors and matrices:
$\widetilde{\mathbf{B}}_{\mathrm{y}}=\mathbf{R}_{\mathrm{y}}\widetilde{\mathbf{B}}\ [(QH\times 1)\text{ vector}]$, $\widetilde{\mathbf{B}}_{\tilde{\mathbf{u}}}=\mathbf{R}_{\tilde{\mathbf{u}}}\widetilde{\mathbf{B}}\ \left[(Q\widetilde{E}\times 1)\text{ vector}\right]$, $\widetilde{\mathbf{\Omega}}_{\mathrm{y}}=\mathbf{R}_{\mathrm{y}}\widetilde{\mathbf{\Omega}}\mathbf{R}'_{\mathrm{y}}\ [(QH\times QH)\text{ matrix}]$, $\widetilde{\mathbf{\Omega}}_{\tilde{\mathbf{u}}}=\mathbf{R}_{\tilde{\mathbf{u}}}\widetilde{\mathbf{\Omega}}\mathbf{R}'_{\tilde{\mathbf{u}}}\ \left[(Q\widetilde{E}\times Q\widetilde{E})\text{ matrix}\right]$, and $\widetilde{\mathbf{\Omega}}_{y\tilde{u}}=\mathbf{R}_{y}\widetilde{\mathbf{\Omega}}\mathbf{R}'_{\tilde{u}}\ \left[(QH\times Q\widetilde{E})\text{ matrix}\right]$.

SECTION 2

The CML Estimation Approach

To develop the CML function corresponding to Equation (B.2) in Appendix B of the main paper, first we extract from the mean vector $\widetilde{\boldsymbol{B}}$ and covariance matrix $\widetilde{\mathbf{\Omega}}$, relevant components $\widetilde{\boldsymbol{B}}_{qq'}$ and $\widetilde{\mathbf{\Omega}}_{qq'}$ (as well as $\ddot{\boldsymbol{\psi}}_{qq',low}$ from $\ddot{\boldsymbol{\psi}}_{low}$ and $\ddot{\boldsymbol{\psi}}_{qq',up}$ from $\ddot{\boldsymbol{\psi}}_{up}$) for each (and every) pair of individuals q and q'. To do so, define a selection matrix $\mathbf{D}_{qq'}$ of size $[2(E+\widetilde{G})\times Q(E+\widetilde{G})]$and $\mathbf{V}_{qq'}$ of size $(2\widetilde{E}\times Q\widetilde{E})$ (see Section A.1 Step-1 and Step-2 of Supplement Appendix A for construction details). Then the vectors $\widetilde{\boldsymbol{B}}_{qq'}$, $\ddot{\boldsymbol{\psi}}_{qq',low}$, $\ddot{\boldsymbol{\psi}}_{qq',up}$ and matrix $\widetilde{\mathbf{\Omega}}_{qq'}$ can be extracted from $\widetilde{\boldsymbol{B}}$, $\ddot{\boldsymbol{\psi}}_{low}$, $\ddot{\boldsymbol{\psi}}_{up}$ and $\widetilde{\mathbf{\Omega}}$ as follows: $\widetilde{\mathbf{B}}_{qq'}=\mathbf{D}_{qq'}\widetilde{\mathbf{B}}\,[2(E+\widetilde{G})\times 1)\text{ vector}]$, $\ddot{\boldsymbol{\psi}}_{qq',low}=\mathbf{V}_{qq'}\ddot{\boldsymbol{\psi}}_{low}\ (2\widetilde{E}\times 1\text{ vector})$, $\ddot{\boldsymbol{\psi}}_{qq',up}=\mathbf{V}_{qq'}\ddot{\boldsymbol{\psi}}_{up}\ (2\widetilde{E}\times 1\text{ vector})$, and $\widetilde{\mathbf{\Omega}}_{qq'}=\mathbf{D}_{qq'}\widetilde{\mathbf{\Omega}}\mathbf{D}'_{qq'}\ \left[(2(E+\widetilde{G})\times 2(E+\widetilde{G}))\text{matrix}\right]$. Next, for each pair of individuals, we partition vector $\widetilde{\boldsymbol{B}}_{qq'}$ and matrix $\widetilde{\mathbf{\Omega}}_{qq'}$ into components relevant to continuous variable components ($\boldsymbol{y}_{qq'}$) and ordinal, count and nominal variable components ($\widetilde{\boldsymbol{u}}_{qq'}$). To do so, define a selection matrix $\mathbf{R}_{qq'}$ of size $\left[2(E+\widetilde{G})\times Q(E+\widetilde{G})\right]$ (see Supplement

Appendix A.1 Step-3 for construction details). Then vector $\tilde{\boldsymbol{B}}_{qq'}$ and matrix $\tilde{\boldsymbol{\Omega}}_{qq'}$ can be partitioned into the components corresponding to continuous and non-continuous variable components as follows: $\tilde{\boldsymbol{B}}_{qq',y} = \mathbf{R}_{qq',y}\tilde{\boldsymbol{B}}_{qq'}$ $\left[(2H \times 1)\text{ vector}\right], \tilde{\mathbf{B}}_{qq',\tilde{\mathbf{u}}} = \mathbf{R}_{qq',\tilde{\mathbf{u}}}\tilde{\mathbf{B}}_{qq'}$ $\left[(2\tilde{E} \times 1)\text{ vector}\right]$, $\tilde{\boldsymbol{\Omega}}_{qq',\mathbf{y}} = \mathbf{R}_{qq',\mathbf{y}}\tilde{\boldsymbol{\Omega}}_{qq'}\mathbf{R}'_{qq',\mathbf{y}}$ $\left[(2H \times 2H)\text{ matrix}\right], \tilde{\boldsymbol{\Omega}}_{qq',\tilde{\mathbf{u}}} = \mathbf{R}_{qq',\tilde{\mathbf{u}}}\tilde{\boldsymbol{\Omega}}_{qq'}\mathbf{R}'_{qq',\tilde{\mathbf{u}}}$ $\left[(2\tilde{E} \times 2\tilde{E})\text{ matrix}\right]$, and $\tilde{\boldsymbol{\Omega}}_{qq',y\tilde{\boldsymbol{u}}} = \mathbf{R}_{qq',y}\tilde{\boldsymbol{\Omega}}_{qq'}\mathbf{R}'_{qq',\tilde{\boldsymbol{u}}}$ $\left[(2H \times 2\tilde{E})\text{ matrix}\right]$. That is, $\tilde{\boldsymbol{B}}_{qq'} = \begin{bmatrix} \tilde{\boldsymbol{B}}_{qq',y} \\ \tilde{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}} \end{bmatrix}$ and

$$\tilde{\boldsymbol{\Omega}}_{qq'} = \begin{bmatrix} \tilde{\boldsymbol{\Omega}}_{qq',y} & \tilde{\boldsymbol{\Omega}}_{qq',y\tilde{\boldsymbol{u}}} \\ \tilde{\boldsymbol{\Omega}}'_{qq',y\tilde{\boldsymbol{u}}} & \tilde{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}} \end{bmatrix}. \tag{S.2}$$

Thus the conditional distribution of $\tilde{\boldsymbol{u}}_{qq'}$ given $\boldsymbol{y}_{qq'}$, may be written as: mean $\ddot{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}} = \tilde{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}} + \tilde{\boldsymbol{\Omega}}'_{qq',y\tilde{\boldsymbol{u}}}\tilde{\boldsymbol{\Omega}}^{-1}_{qq',y}(\boldsymbol{y}_{qq'} - \tilde{\boldsymbol{B}}_{qq',y})$ and variance $\ddot{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}} = \tilde{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}} - \tilde{\boldsymbol{\Omega}}'_{qq',y\tilde{\boldsymbol{u}}}\tilde{\boldsymbol{\Omega}}^{-1}_{qq',y}\tilde{\boldsymbol{\Omega}}_{qq',y\tilde{\boldsymbol{u}}}$. In the last step, we arrange the elements inside the vector $\ddot{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}}$, $\ddot{\boldsymbol{\psi}}_{qq',low}$, $\ddot{\boldsymbol{\psi}}_{qq',up}$ and matrix $\ddot{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}$ so that the elements corresponding to all ordinal variables of both individuals q and q' (i.e., $2N$ ordinal variables) are stacked together, followed by elements corresponding to all count variables (i.e., $2C$ count variables) and in the end the elements corresponding to all nominal variables (i.e., $2G$ nominal variables). This arrangement makes it easy to enumerate pairs of observed outcomes for forming the CML function. To achieve such ordering, define a matrix $\mathbf{F}_{qq'}$ of size $(2\tilde{E} \times 2\tilde{E})$ (see Supplement Appendix A.1 Step-4 for construction details). Then the elements can be rearranged as follows: $\vec{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}} = \mathbf{F}_{qq'}\ddot{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}}$, $\vec{\boldsymbol{\psi}}_{qq',low} = \mathbf{F}_{qq'}\ddot{\boldsymbol{\psi}}_{qq',low}$, $\vec{\boldsymbol{\psi}}_{qq',up} = \mathbf{F}_{qq'}\ddot{\boldsymbol{\psi}}_{qq',up}$, and $\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}} = \mathbf{F}_{qq'}\ddot{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\mathbf{F}'_{qq'}$. Finally, replace the last $2\tilde{G}$ elements of the lower threshold vector $\vec{\boldsymbol{\psi}}_{qq',low}$ from $-\infty$ to zero. That is, $\vec{\boldsymbol{\psi}}_{qq',low}[2(N+C)+1:2\tilde{E}] = \mathbf{0}_{2\tilde{G}}$.

Now, to explicitly write the CML function in terms of standard (multivariate) normal density and cumulative distribution functions, define $\boldsymbol{\omega}_{\Delta}$ as the diagonal matrix of standard deviations of the matrix $\boldsymbol{\Delta}$, using $\phi_R(.;\boldsymbol{\Delta}^*)$ for the multivariate standard normal density function of dimension R and correlation matrix $\boldsymbol{\Delta}^*$ ($\boldsymbol{\Delta}^* = \boldsymbol{\omega}_{\Delta}^{-1}\boldsymbol{\Delta}\boldsymbol{\omega}_{\Delta}^{-1}$), and $\Phi_E(.;\boldsymbol{\Delta}^*)$ for the multivariate standard normal cumulative distribution function of dimension E and correlation matrix $\boldsymbol{\Delta}^*$. Also, define a set of two selection matrices: $\mathbf{H}_{vg}$ of size $[I_g \times (2\tilde{E})]$ and $\mathbf{Y}_{gg'}$ of size $[(I_g + I_{g'} - 2) \times (2\tilde{E})]$, constructed as described in Supplement Appendix A.2. Using these selection matrices, let $\hat{\boldsymbol{\Omega}}_{qq',vg} = \mathbf{H}_{vg}\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\mathbf{H}'_{vg}$, $\vec{\boldsymbol{\Omega}}_{qq',gg'} = \mathbf{Y}_{gg'}\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\mathbf{Y}'_{gg'}$, $\mu_{v,up} = \dfrac{\left[\vec{\boldsymbol{\psi}}_{qq',\boldsymbol{up}}\right]_v - \left[\vec{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}}\right]_v}{\sqrt{\left[\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\right]_{vv}}}$, $\mu_{v,low} = \dfrac{\left[\vec{\boldsymbol{\psi}}_{qq',\boldsymbol{low}}\right]_v - \left[\vec{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}}\right]_v}{\sqrt{\left[\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\right]_{vv}}}$, and $\rho_{vv'} = \dfrac{\left[\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\right]_{vv'}}{\sqrt{\left[\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\right]_{vv}\left[\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\right]_{v'v'}}}$, where $\left[\vec{\boldsymbol{\psi}}_{qq',\boldsymbol{up}}\right]_v$ represents the v^{th} element of $\vec{\boldsymbol{\psi}}_{qq',up}$ (and similarly for other vectors), and $\left[\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}\right]_{vv'}$ represents the vv'^{th} element of the matrix $\vec{\boldsymbol{\Omega}}_{qq',\tilde{\boldsymbol{u}}}$. Then, one may write the CML function as follows:

$$L_{CML}(\boldsymbol{\lambda}) = \prod_{q=1}^{Q-1} \prod_{q'=q+1}^{Q} L_{CML,qq'}(\boldsymbol{\lambda}),$$

$$\text{where } L_{CML,qq'}(\boldsymbol{\lambda}) = \left(\prod_{h=1}^{2H} \boldsymbol{\omega}_{\tilde{\boldsymbol{\Omega}}_{qq',y}} \right)^{-1} \varphi_{2H}\left(\left[\boldsymbol{\omega}_{\tilde{\boldsymbol{\Omega}}_{qq',y}} \right]^{-1} \left[\boldsymbol{y}_{qq'} - \tilde{\boldsymbol{B}}_{qq',\boldsymbol{y}} \right]; \tilde{\boldsymbol{\Omega}}^{*}_{qq',\boldsymbol{y}} \right) \times$$

$$\left(\prod_{v=1}^{2N+2C-1} \prod_{v'=v+1}^{2N+2C} \left[\begin{array}{l} \Phi_2(\mu_{v,up}, \mu_{v',up}, \rho_{vv'}) - \Phi_2(\mu_{v,up}, \mu_{v',low}, \rho_{vv'}) \\ - \Phi_2(\mu_{v,low}, \mu_{v',up}, \rho_{vv'}) + \Phi_2(\mu_{v,low}, \mu_{v',low}, \rho_{vv'}) \end{array} \right] \right) \times$$

$$\left(\prod_{v=1}^{2N+2C} \prod_{g=1}^{2G} \Phi_{I_g}\left[\boldsymbol{\omega}^{-1}_{\hat{\boldsymbol{\Omega}}_{qq',vg}} \mathbf{H}_{vg} \left\{ \vec{\boldsymbol{\psi}}_{qq',up} - \vec{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}} \right\}; \hat{\boldsymbol{\Omega}}^{*}_{qq',vg} \right] - \Phi_{I_g}\left[\boldsymbol{\omega}^{-1}_{\hat{\boldsymbol{\Omega}}_{qq',vg}} \mathbf{H}_{qq',vg} \left\{ \vec{\boldsymbol{\psi}}_{qq',low} - \vec{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}} \right\}; \hat{\boldsymbol{\Omega}}^{*}_{qq',vg} \right] \right) \times$$

$$\left(\prod_{g=1}^{2G-1} \prod_{g'=g+1}^{2G} \Phi_{I_g + I_{g'} - 2}\left[\boldsymbol{\omega}^{-1}_{\hat{\boldsymbol{\Omega}}_{qq',gg'}} \mathbf{Y}_{\boldsymbol{gg'}} \left\{ -\vec{\boldsymbol{B}}_{qq',\tilde{\boldsymbol{u}}} \right\}; \hat{\boldsymbol{\Omega}}^{*}_{qq',gg'} \right] \right). \tag{S.3}$$

In Equation (S.3), for each pair of individuals q and q', the first component corresponds to the marginal likelihood of the continuous outcomes for the two individuals, the second component corresponds to the likelihood of pairs of outcomes across all ordinal and count outcomes, the third component corresponds to the pairwise likelihood of ordinal/count outcomes and nominal outcomes, and the last component corresponds to the pairwise likelihood for the nominal outcomes.

Among all the multivariate normal cumulative distribution (MVNCD) integrals in Equation (S.3), the maximum dimension of integration is the sum of alternatives of nominal variables with the two highest numbers of alternatives minus 2. To solve such MVNCD integrals within the CML estimation routine, Bhat (2011) proposed the use of an analytic approximation method. This approach, labelled the MACML approach, as demonstrated by Bhat and Sidharthan (2011), is at least as accurate as simulation based approaches in retrieving model parameters, albeit computationally much faster and robust in that the approximate CML surface is smoother and easier to maximize than the traditional simulation-based likelihood surfaces.

We write the resulting equivalent of Equation (S.3) computed using the analytic approximation for MVNCD as $L_{MACML}(\boldsymbol{\lambda}) = \prod_{q=1}^{Q-1} \prod_{q'=q+1}^{Q} L_{MACML,qq'}(\boldsymbol{\lambda})$. The MACML estimator is then obtained by maximizing the logarithm of $L_{MACML}(\boldsymbol{\lambda})$, which involves computation of pairwise log-likelihoods (i.e., $\log L_{MACML,qq'}(\boldsymbol{\lambda})$) for $Q(Q-1)/2$ pairs. The asymptotic covariance matrix of the parameters $\mathbf{V}_{\text{MACML}}(\boldsymbol{\lambda})$ may be estimated by the inverse of Godambe's (1960) sandwich information matrix.

$$\mathbf{V}_{MACML}(\boldsymbol{\lambda}) = \frac{\left[\hat{\boldsymbol{G}}(\boldsymbol{\lambda})\right]^{-1}}{Q} = \frac{\left[\hat{\boldsymbol{H}}^{-1}\right]\left[\hat{\boldsymbol{J}}\right]\left[\hat{\boldsymbol{H}}^{-1}\right]'}{Q} \tag{S.4}$$

The reader is referred to Zhao and Joe (2005), Bhat (2014), and Sidharthan and Bhat (2012) for more details on the calculations of the Hessian matrix ($\hat{\boldsymbol{H}}$) and the Jacobian matrix ($\hat{\boldsymbol{J}}$) in the above expression (Sidharthan and Bhat, 2012 provide these details for spatial models).

As spatial dependency decreases with an increase in the distance between any two individuals, one can reduce the number of pairings (between individual observations) in the MACML function by neglecting all the pairings beyond a certain threshold distance. To determine this threshold distance, analyst can estimate the model with different threshold distances and choose the one that minimizes the total variance across all parameters as given by the trace of the asymptotic covariance matrix $\mathbf{V}_{MACML}(\boldsymbol{\lambda})$.

One final consideration relevant to model estimation is that the matrix $\vec{\boldsymbol{\Sigma}}$ for each observation has to be positive definite. The simplest way to guarantee this is to ensure that the (L×L) correlation matrix $\boldsymbol{\Gamma}$ is positive definite, and each matrix $\breve{\boldsymbol{\Lambda}}_{\mathbf{g}}$ (g=1,2,…,G) is also positive definite. To do so, we parameterize the CML function in terms of the Cholesky parameters for these matrices. Further, because the matrix $\boldsymbol{\Gamma}$ is a correlation matrix, we write each diagonal element (say the aa^{th} element) of the lower triangular Cholesky matrix of $\boldsymbol{\Gamma}$ as $\sqrt{1-\sum_{j=1}^{a-1} p_{aj}^2}$, where the p_{aj} elements are the Cholesky factors to be estimated. In addition, note that the top diagonal element of each $\breve{\boldsymbol{\Lambda}}_{\mathbf{g}}$ matrix has to be normalized to one (as discussed in Appendix A of the main paper), which implies that the first element of the Cholesky matrix of each $\breve{\boldsymbol{\Lambda}}_{\mathbf{g}}$ is fixed to the value of one. Also, the spatial autoregressive parameter $\delta_l\,(l=1,2,...,L)$ should be constrained between 0 and 1 $(-1<\delta_l<1\,\forall l)$, for which we parameterize the spatial autoregressive parameter as $\delta_l=\pm\left(1/[1+\exp(\widehat{\delta}_l)]\right)\forall l$. However, in our case, we expect the parameter to be positive, because the spatial dependence is being introduced in the latent pschological constructs, and this impose the '+' sign.

References

Bhat, C.R. (2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B*, 45(7), 923-939.

Bhat, C.R. (2014). The composite marginal likelihood (CML) inference approach with applications to discrete and mixed dependent variable models. *Foundations and Trends in Econometrics*, 7(1), 1-117.

Bhat, C.R., Sidharthan, R. (2011). A simulation evaluation of the maximum approximate composite marginal likelihood (MACML) estimator for mixed multinomial probit models. *Transportation Research Part B*, 45(7), 940-953.

Godambe, V.P. (1960). An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics* 31(4), 1208-1211.

Sidharthan, R., Bhat, C.R. (2012). Incorporating spatial dynamics and temporal dependency in land use change models. *Geographical Analysis*, 44(4), 321-349.

Zhao, Y., Joe, H. (2005). Composite likelihood estimation in multivariate data analysis. *The Canadian Journal of Statistics* 33(3), 335-356.

Supplement Appendix A: Construction of Selection Matrices

A.1. To Extract Relevant Components from $\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{\Omega}}$ for a pair of individuals q and q', and rearrange elements in the order of ordinal, count and nominal variables

To build the CML function, we use a set of selection matrices such that one can extract the relevant components from vectors $\tilde{\boldsymbol{B}}$ and matrix $\tilde{\boldsymbol{\Omega}}$ for a pair of individuals q and q'. The selection matrices are described below:

Step-1: A matrix $\mathbf{D}_{qq'}$ of size $\left[2(E+\tilde{G})\times Q(E+\tilde{G})\right]$ filled with zeros. Insert an identity matrix of size $(E+\tilde{G})$ in first $(E+\tilde{G})$ rows and columns $\left[(q-1)*(E+\tilde{G})+1\right]$ to $\left[(q-1)*(E+\tilde{G})+(E+\tilde{G})\right]$. Next, insert another identity matrix of size $(E+\tilde{G})$ in rows $(E+\tilde{G}+1)$ to $[2(E+\tilde{G})]$ and columns $\left[(q'-1)*(E+\tilde{G})+1\right]$ to $\left[(q'-1)*(E+\tilde{G})+(E+\tilde{G})\right]$.

Step-2: A matrix $\mathbf{V}_{qq'}$ of size $(2\tilde{E}\times Q\tilde{E})$ filled with zeros. Insert an identity matrix of size $(\tilde{E})$ in first $(\tilde{E})$ rows and columns $\left[(q-1)*(\tilde{E})+1\right]$ to $\left[(q-1)*(\tilde{E})+\tilde{E}\right]$. Next, insert another identity matrix of size $(\tilde{E})$ in rows $(\tilde{E}+1)$ to $(2\tilde{E})$ and columns $\left[(q'-1)*(\tilde{E})+1\right]$ to $\left[(q'-1)*(\tilde{E})+\tilde{E}\right]$.

Step-3: A matrix $\mathbf{R}_{qq'}$ of size $\left[2(E+\tilde{G})\times Q(E+\tilde{G})\right]$ filled with zeros. Then, for individual q, consider an identity matrix $\mathbf{R}_q$ of size $(E+\tilde{G})$. Insert first H rows of the matrix $\mathbf{R}_q$ into matrix $\mathbf{R}_{qq'}$ occupying first H rows and columns $\left[(q-1)*(E+\tilde{G})+1\right]$ to $\left[(q-1)*(E+\tilde{G})+(E+\tilde{G})\right]$. Next insert the remaining $(E+\tilde{G}-H)$ rows of matrix $\mathbf{R}_q$ into matrix $\mathbf{R}_{qq'}$ occupying rows $[2H+1]$ to $\left[2H+(E+\tilde{G}-H)\right]$ and columns $\left[(q-1)*(E+\tilde{G})+1\right]$ to $\left[(q-1)*(E+\tilde{G})+(E+\tilde{G})\right]$. For individual q', Insert first H rows of matrix $\mathbf{R}_{q'}$ into matrix $\mathbf{R}_{qq'}$ occupying rows $(H+1)$ to $(2H)$ and remaining $(E+\tilde{G}-H)$ rows of matrix $\mathbf{R}_{q''}$ into matrix $\mathbf{R}_{qq'}$ occupying rows $\left[2H+\tilde{E}+1\right]$ to $\left[2H+2\tilde{E}\right]$ and columns $\left[(q'-1)*(E+\tilde{G})+1\right]$ to $\left[(q'-1)*(E+\tilde{G})+(E+\tilde{G})\right]$. Divide the matrix $\mathbf{R}_{qq'}$ into two submatrices $\mathbf{R}_{qq',y}(\mathbf{R}_{qq'}[1{:}2H,1:Q(E+\tilde{G})])$ and $\mathbf{R}_{qq',\tilde{u}}(\mathbf{R}_{qq'}[2H+1:2(E+\tilde{G}),1:Q(E+\tilde{G})])$.

Step-4: Now, define a matrix $\mathbf{F}_{qq'}$ of size $(2\tilde{E}\times 2\tilde{E})$ filled with zeros. Next, perform the steps described below to arrange the elements inside the vector $\ddot{\boldsymbol{B}}_{qq',\tilde{u}}$ and matrix $\ddot{\boldsymbol{\Omega}}_{qq',\tilde{u}}$ in the order of ordinal, count and nominal variables.

Step-4.1 Insert an identity matrix of size *N* in first *N* rows and *N* columns. Insert another identity matrix of size *N* in rows (*N+1*) to (2*N*) and columns $(\tilde{E}+1)$ to $(\tilde{E}+N)$.

Step-4.2 Again, insert an identity matrix of size C in rows $(2N+1)$ to $(2N+C)$ and columns $(N+1)$ to $(N+C)$. Insert another identity matrix of size C in rows $(2N+C+1)$ to $(2N+2C)$ and columns $(\widetilde{E}+N+1)$ to $(\widetilde{E}+N+C)$.

Step-4.3 Finally, insert an identity matrix of size $\widetilde{G}$ in rows $(2N+2C+1)$ to $(2N+2C+\widetilde{G})$ and columns $(N+C+1)$ to $(\widetilde{E})$. Insert another identity matrix of size $\widetilde{G}$ in rows $(2N+2C+\widetilde{G}+1)$ to $(2\widetilde{E})$ and columns $(\widetilde{E}+N+C+1)$ to $(2\widetilde{E})$.

A.2. To write the CML function in terms of standard normal density and cumulative distribution functions

Define a set of two selection matrices as follows: (1) $\mathbf{H}_{vg}$ is a $I_g \times (2\widetilde{E})$ selection matrix with an entry of '1' in the first row and the v^{th} column and an identity matrix of size $I_g - 1$ occupying the last $I_g - 1$ rows and the $2N+2C+\left[\sum_{j=1}^{g-1}(I_j-1)+1\right]^{th}$ through $2N+2C+\left[\sum_{j=1}^{g}(I_j-1)\right]^{th}$ columns (with the convention that $\sum_{j=1}^{0}(I_j-1)=0$), and entries of '0' everywhere else, (2) $\mathbf{Y}_{gg'}$ is a $(I_g+I_{g'}-2)\times(2\widetilde{E})$ selection matrix with an identity matrix of size (I_g-1) occupying the first (I_g-1) rows and the $2N+2C+\left[\sum_{j=1}^{g-1}(I_j-1)+1\right]^{th}$ through $2N+2C+\left[\sum_{j=1}^{g}(I_j-1)\right]^{th}$ columns (with the convention that $\sum_{j=1}^{0}(I_j-1)=0$), and another identity matrix of size $(I_{g'}-1)$ occupying the last $(I_{g'}-1)$ rows and the $2N+2C+\left[\sum_{j=1}^{g'-1}(I_j-1)+1\right]^{th}$ through $2N+2C+\left[\sum_{j=1}^{g'}(I_j-1)\right]^{th}$ columns; all other elements of $\boldsymbol{Y}_{gg'}$ take a value of zero.

SECTION 3

Table 1. Descriptive statistics of independent variables

Variable	Categories	Percentage
Age of school going children	5-10 years old	47.37
	11-15 years old	40.83
	16-18 years old	11.80
Gender of school going children	Boys	51.80
	Girls	48.20
Household monthly income	Less than 25K	7.70
	25K – 49,999	16.18
	50K – 74,999	16.68
	75K – 99,999	16.77
	100K or more	42.67
Race	Caucasian	69.78
	African-American	7.39
	Asian	9.72
	Hispanic	12.07
	Others	1.04
Households with fraction of adults (25 or more) with	High school degree or less	25.45
	Some college degree	25.73
	Bachelor's degree	27.69
	Graduate degree	21.13
Households with fraction of adults in age group	19-30 years	7.87
	31-45 years	51.70
	46-60 years	36.43
	61 or more	4.00
Households with number of full-time workers	0	5.21
	1	58.66
	2	33.36
	3 or more	2.77
Households with number of part-time workers	0	67.93
	1	28.80
	2 or more	3.27
Households with number of workers with the option to work from home	0	75.62
	1	21.43
	2 or more	2.95
Households with number of workers with flexible work time	0	42.26
	1	44.52
	2 or more	13.22
Family type	Nuclear	93.41
	Single-parent	6.59
Housing type	Detached	81.75
	Duplex	6.54
	Apartment or townhouse	11.71
Tenure	Own	80.05
	Rent	19.95
Distance to school	Less than ¼ mile	11.29
	Between ¼ to ½ mile	10.37
	Greater than ½ mile and less than or equal to 1 mile	14.47
	Greater than 1 mile and less than or equal to 2 miles	21.43
	More than 2 miles	42.44

SECTION 4

Table 2. Parameter estimates of exogenous variable effects on non-nominal variables

Dependent variables	Constants		Thresholds for ordinal variables between… A little bit of an issue & somewhat of an issue		Somewhat of an issue & very much an issue		Very much an issue & a serious issue		Dispersion parameter*		Variance	
	Coeff	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**
Household average commute distance (miles)	2.428	3.80	----	----	----	----	----	----	----	----	0.770	4.06
Walk/bike issue: violence/crime along the route	0.072	1.67	0.419	10.22	0.779	13.91	1.035	17.54	----	----	----	----
Walk/bike issue: speed of traffic along the route	3.543	2.37	1.154	6.71	2.757	11.16	4.117	15.65	----	----	----	----
Walk/bike issue: amount of traffic along the route	3.667	2.52	1.353	10.33	2.848	15.56	4.268	18.16	----	----	----	----
Number of biking episodes in past week	-2.355	-3.01	----	----	----	----	----	----	0.062	86.63	----	----
Number of walking episodes in past week	----	----	----	----	----	----	----	----	----	----	----	----
Number of times public transit used in past week	0.424	2.65	----	----	----	----	----	----	0.098	86.00	----	----
Vehicle ownership	1.384	4.79	----	----	----	----	----	----	----	----	----	----
At least one working adult uses public transit/walk/bike as mode to work	-0.097	-2.31	----	----	----	----	----	----	----	----	----	----

* The t-stat for dispersion parameter is calculated with respect to numerical value of 5 as oppose to zero. This is due to the fact that negative binomial count model collapses to Poisson count model for a dispersion parameter value of 5 or more.

Table 2 (cont.) Parameter estimates of exogenous variable effects on non-nominal variables

Dependent variables	Number of children (18 years or less) in the household		Household income (base: 75K or more)						Number of full-time workers		Number of part-time workers	
			Less than 25K		25K-49,999		50K-74,999					
	Coeff	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**
Household average commute distance (miles)	0.052	1.73	-0.720	-6.37	-0.372	-4.96	-0.118	-2.66	----	----	-0.162	-3.38
Walk/bike issue: violence/crime along the route	----	----	----	----	----	----	----	----	----	----	----	----
Walk/bike issue: speed of traffic along the route	----	----	----	----	----	----	----	----	----	----	----	----
Walk/bike issue: amount of traffic along the route	----	----	----	----	----	----	----	----	----	----	----	----
Number of biking episodes in past week	----	----	----	----	----	----	----	----	----	----	----	----
Number of walking episodes in past week	----	----	----	----	----	----	----	----	----	----	----	----
Number of times public transit used in past week	----	----	----	----	----	----	----	----	----	----	----	----
Vehicle ownership	0.017	2.27	-0.308	-1.91	-0.110	-2.18	----	----	0.229	2.63	0.234	4.78
At least one working adult uses public transit/walk/bike as mode to work	----	----	----	----	----	----	----	----	----	----	----	----

Table 2 (Cont.) Parameter estimates of exogenous variable effects on non-nominal variables

Dependent variables	Number of workers with the option to work from home		Housing type (base: detached)				Tenure (base: owned)	
			Duplex		Apartment or Townhouse		Rented	
	Coeff	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**
Household average commute distance (miles)	----	----	----	----	----	----	----	----
Walk/bike issue: violence/crime along the route	----	----	----	----	----	----	----	----
Walk/bike issue: speed of traffic along the route	----	----	----	----	----	----	----	----
Walk/bike issue: amount of traffic along the route	----	----	----	----	----	----	----	----
Number of biking episodes in past week	----	----	----	----	----	----	----	----
Number of walking episodes in past week	----	----	----	----	----	----	----	----
Number of times public transit used in past week	----	----	----	----	----	----	----	----
Vehicle ownership	-0.039	-3.08	-0.257	-2.42	-0.610	-5.65	-0.176	-2.34
At least one working adult uses public transit/walk/bike as mode to work	----	----	----	----	----	----	----	----

Table 3. Parameter estimates of exogenous variable effects on residential location choice

Variables	Residential location (base: less than 1000 housing units per sq mile)					
	1000-1999		2000-3999		4000 or more	
	Coeff	**T-stat**	**Coeff**	**T-stat**	**Coeff**	**T-stat**
Constants	0.584	6.42	0.563	11.04	0.238	1.75
Family type (base: nuclear family)						
Single-parent	----	----	0.018	1.44	0.018	1.44
Household income (base: less than 100K)						
100K or more	----	----	-0.037	-2.47	----	----
Housing type (base: detached)						
Duplex	----	----	----	----	0.099	6.60
Apartment or Townhouse	----	----	0.072	2.44	0.302	4.65
Number of workers with the option to work from home	----	----	-0.028	-2.15	-0.028	-2.15

In addition to the above observed effects of exogenous variables, we estimated the following covariance matrix (t-statistics in parenthesis) between the differences of error terms with respect to the error term of the lowest density residential location alternative:

$$\begin{bmatrix} 1.00\,(\text{fixed}) & & \\ 0.97\,(1.98) & 1.13\,(2.61) & \\ 0.59\,(3.37) & 0.76\,(2.74) & 0.59\,(1.70) \end{bmatrix}$$

Note that the terms in the parentheses represent the t-statistic values. Since this is a differenced error-covariance matrix (**Λ**), a clear interpretation of the elements of this matrix is not straightforward.