Online supplement to

Autonomous Vehicle Impacts on Travel-Based Activity and Activity-Based Travel

By Katherine A. Dannemiller, Katherine E. Asmussen, Aupal Mondal, and
Chandra R. Bhat (corresponding author)

Mathematical Formulation of the GHDM for the Current Study Involving Ordinal Outcomes and Ranked Outcomes

For ease of presentation, we will suppress the index for decision-makers in our exposition below, and assume that all error terms are independent and identically distributed across decision-makers. Following Bhat's (2015) GHDM formulation, let l be an index for latent variables (l=1,2,…,L). Consider the latent variable z_l^* and write it as a linear function of covariates:

$$z_l^* = \boldsymbol{\alpha}_l' \boldsymbol{w} + \eta_l, \tag{1}$$

where $\boldsymbol{w}$ is a $(\widetilde{D} \times 1)$ vector of observed covariates (excluding a constant), $\boldsymbol{\alpha}_l$ is a corresponding $(\widetilde{D} \times 1)$ vector of coefficients, and η_l is a random error term assumed to be standard normally distributed for identification purpose. Next, define the $(L \times \widetilde{D})$ matrix $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_L)'$, and the $(L \times 1)$ vectors $\boldsymbol{z}^* = (z_1^*, z_2^*, ..., z_L^*)'$ and $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3, \ldots, \eta_L)'$. We allow a multivariate normal (MVN) correlation structure for $\boldsymbol{\eta}$ to accommodate interactions among the unobserved latent variables: $\boldsymbol{\eta} \sim MVN_L[\mathbf{0}_L, \boldsymbol{\Gamma}]$, where $\mathbf{0}_L$ is an $(L \times 1)$ column vector of zeros, and $\boldsymbol{\Gamma}$ is an $(L \times L)$ correlation matrix. In matrix form, we may write Equation (1) as:

$$\boldsymbol{z}^* = \boldsymbol{\alpha w} + \boldsymbol{\eta}. \tag{2}$$

Now consider N ordinal outcomes (indicator variables as well as main outcomes) for the individual, and let n be the index for the ordinal outcomes $(n = 1, 2, ..., N)$. Also, let J_n be the number of categories for the n^{th} ordinal outcome $(J_n \geq 2)$ and let the corresponding index be j_n $(j_n = 1, 2, ..., J_n)$. In our empirical case, N = 14 (corresponding to 12 indicators and the ALT and ADLT dimensions, each with $J_n = 5$). Let $\widetilde{y}_n^*$ be the latent underlying variable whose horizontal partitioning leads to the observed outcome for the n^{th} ordinal variable. Assume that the individual under consideration chooses the a_n^{th} ordinal category. Then, in the usual ordered response formulation, for the individual, we may write:

$$\widetilde{y}_n^* = \widetilde{\boldsymbol{\gamma}}_n' \boldsymbol{x} + \widetilde{\boldsymbol{d}}_n' \boldsymbol{z}^* + \widetilde{\varepsilon}_n, \text{ and } \widetilde{\psi}_{n,a_n-1} < \widetilde{y}_n^* < \widetilde{\psi}_{n,a_n}, \tag{3}$$

where $\boldsymbol{x}$ is an $(A \times 1)$ vector of exogenous variables (including a constant) as well as possibly the observed values of other endogenous ordinal variables, and other endogenous ranked-choice variables introduced as dummy variables (thus, in our case, if an individual selected a particular TBA, or a combination of TBAs, within their first three ranked activities, these endogenous variables may be included as dummy variables, though only in a recursive fashion and not in a cyclic manner) , $\tilde{\boldsymbol{\gamma}}_{\boldsymbol{n}}$ is a corresponding vector of coefficients to be estimated, $\tilde{\boldsymbol{d}}_{\boldsymbol{n}}$ is an $(L \times 1)$ vector of latent variable loadings on the n^{th} ordinal outcome, the $\tilde{\psi}$ terms represent thresholds, and $\tilde{\varepsilon}_n$ is the standard normal random error for the n^{th} ordinal outcome (note, however, that for the indicators (but not the main outcomes), typically the $\boldsymbol{x}$ vector will not appear on the right side of Equation (3); also, there are specific identification conditions for the number of non-zero elements of $\tilde{\boldsymbol{d}}_{\boldsymbol{n}}$ that can be present in each indicator equation and across all indicator equations; please see Bhat, 2015 for additional details). For each ordinal outcome, $\tilde{\psi}_{n,0} < \tilde{\psi}_{n,1} < \tilde{\psi}_{n,2} ... < \tilde{\psi}_{n,J_n-1} < \tilde{\psi}_{n,J_n}$; $\tilde{\psi}_{n,0} = -\infty$, $\tilde{\psi}_{n,1} = 0$, and $\tilde{\psi}_{n,J_n} = +\infty$. For later use, let $\tilde{\boldsymbol{\psi}}_{\boldsymbol{n}} = (\tilde{\psi}_{n,2}, \tilde{\psi}_{n,3} ..., \tilde{\psi}_{n,J_n-1})'$ and $\tilde{\boldsymbol{\psi}} = (\tilde{\boldsymbol{\psi}}_{\boldsymbol{1}}', \tilde{\boldsymbol{\psi}}_{\boldsymbol{2}}', ..., \tilde{\boldsymbol{\psi}}_N)'$. Stack the N underlying continuous variables $\tilde{y}_n^*$ into an $(N \times 1)$ vector $\tilde{\boldsymbol{y}}^*$, and the N error terms $\tilde{\varepsilon}_n$ into another $(N \times 1)$ vector $\tilde{\boldsymbol{\varepsilon}}$. Define $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \tilde{\gamma}_2, ..., \tilde{\gamma}_H)'$ $[(N \times A)$ matrix] and $\tilde{\boldsymbol{d}} = \left(\tilde{\boldsymbol{d}}_{\boldsymbol{1}}, \tilde{\boldsymbol{d}}_{\boldsymbol{2}}, ..., \tilde{\boldsymbol{d}}_N\right)$ $[(N \times L)$ matrix], and let $\mathbf{IDEN}_N$ be the identity matrix of dimension N representing the correlation matrix of $\tilde{\boldsymbol{\varepsilon}}$ (the unit diagonals are needed for identification; for convergence stability and parsimony, we assume that the elements of the $\tilde{\boldsymbol{\varepsilon}}$ vector are uncorrelated with each other, though specific elements of the $\tilde{\boldsymbol{y}}^*$ vector can still be correlated through the stochatic latent constructs). Finally, stack the lower thresholds for the decision-maker $\tilde{\psi}_{n,a_n-1}(n = 1, 2, ..., N)$ into an $(N \times 1)$ vector $\tilde{\boldsymbol{\psi}}_{\boldsymbol{low}}$ and the upper thresholds $\tilde{\psi}_{n,a_n}(n = 1, 2, ..., N)$ into another vector $\tilde{\boldsymbol{\psi}}_{up}$. Then, in matrix form, the measurement equation for the ordinal outcomes (indicators) for the decision-maker may be written as:

$$\tilde{\boldsymbol{y}}^* = \tilde{\boldsymbol{\gamma}}\boldsymbol{x} + \tilde{\boldsymbol{d}}\boldsymbol{z}^* + \tilde{\boldsymbol{\varepsilon}}, \quad \tilde{\boldsymbol{\psi}}_{\boldsymbol{low}} < \tilde{\boldsymbol{y}}^* < \tilde{\boldsymbol{\psi}}_{\boldsymbol{up}}. \tag{4}$$

Now let there be G ranked outcome variables for an individual, and let g be the index for the ranked variables $(g = 1, 2, 3, ..., G)$. Also, let I_g be the number of alternatives corresponding to the g^{th} ranked variable ($I_g \geq 3$) and let i_g be the corresponding index $(i_g = 1, 2, 3, ..., I_g)$. In our case, G=1 and I_1 =7; however we present the framework for any number of ranked otcomes. Consider the g^{th} ranked variable and assume the usual random utility structure for each alternative i_g.

$$U_{gi_g} = \boldsymbol{b}_{gi_g}'\boldsymbol{x} + \boldsymbol{\vartheta}_{gi_g}'(\boldsymbol{\beta}_{gi_g}\boldsymbol{z}^*) + \varsigma_{gi_g}, \tag{5}$$

where $\boldsymbol{x}$ is an $(A \times 1)$ vector of exogenous variables (including a constant) as well as possibly the observed values of other endogenous ordinal variables (introduced in a recursive fashion), as defined earlier, $\boldsymbol{b}_{gi_g}$ is an $(A \times 1)$ column vector of corresponding coefficients, and ς_{gi_g} is normal error term. $\boldsymbol{\beta}_{gi_g}$ is an $(N_{gi_g} \times L)$-matrix of variables interacting with latent variables to influence the utility of alternative i_g, and $\boldsymbol{\vartheta}_{gi_g}$ is an $(N_{gi_g} \times 1)$-column vector of coefficients capturing the effects of latent variables and their interaction effects with other exogenous variables. If each of the latent variables impacts the utility of the alternatives for each ranked variable purely through a constant shift in the utility function, $\boldsymbol{\beta}_{gi_g}$ will be an identity matrix of size *L*, and each element of $\boldsymbol{\vartheta}_{gi_g}$ will capture the effect of a latent variable on the constant specific to alternative i_g of nominal variable g. Let $\boldsymbol{\varsigma_g} = (\varsigma_{g1}, \varsigma_{g2}, ...\varsigma_{gI_g})'$ ($I_g \times 1$ vector), and $\boldsymbol{\varsigma_g} \sim MVN_{I_g}(\mathbf{0}, \boldsymbol{\Lambda_g})$. Taking the difference with respect to the first alternative, the only estimable elements are found in the covariance matrix $\breve{\boldsymbol{\Lambda}}_g$ of the error differences, $\breve{\boldsymbol{\varsigma}}_{\boldsymbol{g}} = (\breve{\varsigma}_{g2}, \breve{\varsigma}_{g3}, ..., \breve{\varsigma}_{gI_g})$ (where $\breve{\varsigma}_{gi} = \varsigma_{gi} - \varsigma_{g1}, i \neq 1$). Further, the variance term at the top left diagonal of $\breve{\boldsymbol{\Lambda}}_g$ $(g = 1, 2, ..., G)$ is set to 1 to account for scale invariance. $\boldsymbol{\Lambda_g}$ is constructed from $\breve{\boldsymbol{\Lambda}}_{\mathbf{g}}$ by adding a row on top and a column to the left. All elements of this additional row and column are filled with values of zero. In addition, the usual identification restriction is imposed such that one of the alternatives serves as the base when introducing alternative-specific constants and variables that do not vary across alternatives (that is, whenever an element of $\boldsymbol{x}$ is individual-specific and not alternative-specific, the corresponding element in $\boldsymbol{b}_{gi_g}$ is set to zero for at least one alternative i_g). To proceed, define $\boldsymbol{U}_g = (U_{g1}, U_{g2}, ..., U_{gI_g})'$ ($I_g \times 1$ vector), $\boldsymbol{b}_g = (\boldsymbol{b}_{g1}, \boldsymbol{b}_{g2}, \boldsymbol{b}_{g3}, ..., \boldsymbol{b}_{gI_g})'$ ($I_g \times A$ matrix), and $\boldsymbol{\beta}_g = (\boldsymbol{\beta}'_{g1}, \boldsymbol{\beta}'_{g2}, ..., \boldsymbol{\beta}'_{gI_g})'$ $\left(\sum_{i_g=1}^{I_g} N_{gi_g} \times L\right)$ matrix. Also, define the $\left(I_g \times \sum_{i_g=1}^{I_g} N_{gi_g}\right)$ matrix $\boldsymbol{\vartheta_g}$, which is initially filled with all zero values. Then, position the $(1 \times N_{g1})$ row vector $\boldsymbol{\vartheta}'_{g1}$ in the first row to occupy columns 1 to N_{g1}, position the $(1 \times N_{g2})$ row vector $\boldsymbol{\vartheta}'_{g2}$ in the second row to occupy columns $N_{g1}+1$ to $N_{g1} + N_{g2}$, and so on until the $(1 \times N_{gI_g})$ row vector $\boldsymbol{\vartheta}'_{gI_g}$ is appropriately positioned. Further, define $\boldsymbol{\varpi}_g = (\boldsymbol{\vartheta}_g \boldsymbol{\beta}_g)$ ($I_g \times L$ matrix), $\vec{G} = \sum_{g=1}^{G} I_g$, $\tilde{G} = \sum_{g=1}^{G} (I_g - 1)$, $\boldsymbol{U} = \left(\boldsymbol{U}'_1, \boldsymbol{U}'_2, \dots, \boldsymbol{U}'_G\right)'$ ($\vec{G} \times 1$ vector), $\boldsymbol{\varsigma} = (\boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_2, ...\boldsymbol{\varsigma}_G)'$ ($\vec{G} \times 1$ vector), $\boldsymbol{b} = (\boldsymbol{b}'_1, \boldsymbol{b}'_2, ..., \boldsymbol{b}'_{\boldsymbol{G}})'$ ($\vec{G} \times A$ matrix), $\boldsymbol{\varpi} = (\boldsymbol{\varpi}'_1, \boldsymbol{\varpi}'_2, ..., \boldsymbol{\varpi}'_G)'$ ($\vec{G} \times L$ matrix), and $\boldsymbol{\vartheta} = \text{Vech}(\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2, ..., \boldsymbol{\vartheta}_G)$ (that is, $\boldsymbol{\vartheta}$ is a column vector that includes all elements of the matrices $\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2, ..., \boldsymbol{\vartheta}_G$). Then, in matrix form, we may write Equation (5) as:

$$\boldsymbol{U} = \boldsymbol{bx} + \boldsymbol{\varpi}\, \boldsymbol{z}^* + \boldsymbol{\varsigma}, \tag{6}$$

where $\boldsymbol{\varsigma} \sim MVN_{\vec{G}}(\mathbf{0}_{\vec{G}}, \boldsymbol{\Lambda})$. As earlier, to ensure identification, we specify $\boldsymbol{\Lambda}$ as follows:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_2 & \mathbf{0} & \mathbf{0} \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Lambda}_3 & \mathbf{0} \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \ldots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \ldots & \boldsymbol{\Lambda}_G \end{bmatrix} (\vec{G} \times \vec{G} \text{ matrix}). \tag{7}$$

In the general case, this allows the estimation of $\sum_{g=1}^{G}\left(\frac{I_g*(I_g-1)}{2}-1\right)$ terms across all the G nominal variables, as originating from $\left(\frac{I_g*(I_g-1)}{2}-1\right)$ terms embedded in each $\breve{\mathbf{\Lambda}}_g$ matrix; (g=1,2,…,G) .

Let $\boldsymbol{\delta}$ be the collection of parameters to be estimated: $\boldsymbol{\delta}=[\text{Vech}(\boldsymbol{\alpha}), \text{Vechup}(\boldsymbol{\Gamma}), \text{Vech}(\tilde{\boldsymbol{\gamma}}), \text{Vech}(\tilde{\boldsymbol{d}}), \tilde{\boldsymbol{\psi}}, \text{Vech}(\boldsymbol{b}), \boldsymbol{\vartheta}, \text{Vech}(\boldsymbol{\Lambda})]$, where the operator "Vech(.)" vectorizes all the non-zero elements of the matrix/vector on which it operates and "Vechup(.)" indicates strictly upper diagonal elemenets.

With the matrix definitions above, the continuous components of the model system may be written compactly as:

$$\boldsymbol{z}^* = \boldsymbol{\alpha w} + \boldsymbol{\eta}, \tag{8}$$

$$\tilde{\boldsymbol{y}}^* = \tilde{\boldsymbol{\gamma}}\boldsymbol{x} + \tilde{\boldsymbol{d}}\boldsymbol{z}^* + \tilde{\boldsymbol{\varepsilon}}, \text{ with Var } (\tilde{\boldsymbol{\varepsilon}}) = \mathbf{IDEN}_N \;\; (N \times N \text{ matrix}) , \tag{9}$$

$$\boldsymbol{U} = \boldsymbol{bx} + \boldsymbol{\varpi}\boldsymbol{z}^* + \boldsymbol{\varsigma}. \tag{10}$$

To develop the reduced form equations, replace the right side of Equation (8) for $\boldsymbol{z}^*$ in Equations (9) and (10) to obtain the following system:

$$\tilde{\boldsymbol{y}}^* = \tilde{\boldsymbol{\gamma}}\boldsymbol{x} + \tilde{\boldsymbol{d}}\boldsymbol{z}^* + \tilde{\boldsymbol{\varepsilon}} = \tilde{\boldsymbol{\gamma}}\boldsymbol{x} + \tilde{\boldsymbol{d}}(\boldsymbol{\alpha w} + \boldsymbol{\eta}) + \tilde{\boldsymbol{\varepsilon}} = \tilde{\boldsymbol{\gamma}}\boldsymbol{x} + \tilde{\boldsymbol{d}}\boldsymbol{\alpha w} + \tilde{\boldsymbol{d}}\boldsymbol{\eta} + \tilde{\boldsymbol{\varepsilon}}, \tag{11}$$

$$\boldsymbol{U} = \boldsymbol{bx} + \boldsymbol{\varpi}\boldsymbol{z}^* + \boldsymbol{\varsigma} = \boldsymbol{bx} + \boldsymbol{\varpi}(\boldsymbol{\alpha w} + \boldsymbol{\eta}) + \boldsymbol{\varsigma} = \boldsymbol{bx} + \boldsymbol{\varpi}\boldsymbol{\alpha w} + \boldsymbol{\varpi}\boldsymbol{\eta} + \boldsymbol{\varsigma}.$$

Now, consider the $[(N+\ddot{G})\times 1)]$ vector $\boldsymbol{yU} = \left[[\tilde{\boldsymbol{y}}^*]', \boldsymbol{U}'\right]'$. Define

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{\gamma}}\boldsymbol{x} + \tilde{\boldsymbol{d}}\boldsymbol{\alpha w} \\ \boldsymbol{bx} + \boldsymbol{\varpi}\boldsymbol{\alpha w} \end{bmatrix} \text{and}\, \boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_1 & \boldsymbol{\Omega}_{12}' \\ \boldsymbol{\Omega}_{12} & \boldsymbol{\Omega}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{d}}\boldsymbol{\Gamma}\tilde{\boldsymbol{d}}' + \mathbf{IDEN_N} & \tilde{\boldsymbol{d}}\boldsymbol{\Gamma}\boldsymbol{\varpi}' \\ \boldsymbol{\varpi}\boldsymbol{\Gamma}\tilde{\boldsymbol{d}}' & \boldsymbol{\varpi}\boldsymbol{\Gamma}\boldsymbol{\varpi}' + \boldsymbol{\Lambda} \end{bmatrix}. \tag{12}$$

Then $\boldsymbol{yU} \sim \boldsymbol{MVN}_{N+\ddot{G}}(\boldsymbol{B},\boldsymbol{\Omega})$.

We now focus on the estimation of the model. For the case of ranked nominal variables, the utility differentials are arrived at based on the order of the ranking (for ease in presentation, we first present the estimation formulation for the case of a unique ranking scenario, i.e., when there are no tied-rankings, and subsequently provide the changes needed to accommodate the case of tied ranks, which is what is encountered in our actual empirical study). In particular, let $\boldsymbol{r_g}$ be a specific rank ordering of the alternatives corresponding to the g^{th} nominal variable. That is, $r_g^{\,1}$ is the first-ranked alternative, $r_g^{\,2}$ is the second-ranked alternative and so on. $R_{\boldsymbol{r},g}$ denotes the event that the alternatives are ranked in the order $\boldsymbol{r_g}$ for the ranked variable g by the individual. According to the random utility maximization framework, the following relationship must hold for $R_{\boldsymbol{r},g}$,

$$R_{\boldsymbol{r},g} : U_{i_g r^2} - U_{i_g r^1} < 0, U_{i_g r^3} - U_{i_g r^2} < 0, \ldots, U_{i_g r^{I_g}} - U_{i_g r^{I_g - 1}} < 0$$

The above latent utility differentials for the rank-ordered outcome g are stacked as $\boldsymbol{u}_g = \left[\left(u_{gr^2r^1}, u_{gr^3r^2}, ..., u_{gr^{I_g}r^{I_{g-1}}}\right)'\right]$. Now, define $\boldsymbol{u} = \left(\left[\boldsymbol{u}_1\right]', \left[\boldsymbol{u}_2\right]', ..., \left[\boldsymbol{u}_G\right]'\right)'$, We now need to develop the distribution of the vector $\boldsymbol{yu} = \left([\tilde{\boldsymbol{y}}^*]', \boldsymbol{u}'\right)'$ from that of $\boldsymbol{yU} = \left([\tilde{\boldsymbol{y}}^*]', \boldsymbol{U}'\right)'$. To do so, define a matrix $\mathbf{M}$ of size $\left[N+\tilde{G}\right] \times \left[N+\ddot{G}\right]$. Fill this matrix with values of zero. Then, insert an identity matrix of size N into the first N rows and N columns of the matrix $\mathbf{M}$. Next, consider the rows from $N+1$ to $N+I_1-1$, and columns from $N+1$ to $N+I_1$. These rows and columns correspond to the first ranked variable. We do the following in this sub-matrix: place a value of '–1' at the column corresponding to the first ranked alternative and '1' at the column corresponding to the second ranked alternative. Similarly, in the second row, place a value of '–1' at the column corresponding to the second ranked alternative and '1' at the column corresponding to the third ranked alternative. Continue this procedure for $(\boldsymbol{I}_1-1)$ rows. . Next, rows $N+I_1$ through $N+I_1+I_2-2$ and columns $N+I_1+1$ through $N+I_1+I_2$ correspond to the second ranked variable. Repeat the above process. Continue this procedure for all G ranked variables. With the matrix $\mathbf{M}$ as defined, we can write $\boldsymbol{yu} \sim \boldsymbol{MVN}_{N+\tilde{G}}(\tilde{\boldsymbol{B}}, \tilde{\boldsymbol{\Omega}})$, where $\tilde{\boldsymbol{B}} = \mathbf{M}\boldsymbol{B}$ and $\tilde{\boldsymbol{\Omega}} = \mathbf{M\Omega M}'$.

However, to deal with the cases where different alternatives have identical rankings, the likelihood is calculated as the probability of all utility values that can result in the rank ordering depicted by the respondent. For example, if an individual q assigns the first rank to alternative 3, second rank to two alternatives (say, 2 and 4), and third rank to alternative 1, pertaining to the ranked variable g, the sub-matrix pertaining to this ranked outcome within the contrast matrix $\mathbf{M}$ is structured to represent the following four conditions (suppressing q in the equation):

$$U_{i_g r^2} - U_{i_g r^3} < 0, U_{i_g r^4} - U_{i_g r^3} < 0, U_{i_g r^1} - U_{i_g r^2} < 0, U_{i_g r^1} - U_{i_g r^4} < 0$$

This is equivalent to the following:

$$\mathbf{M}_g = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$$

Note that the number of rows in $\mathbf{M}_g$ varies depending on the number of ties at different rank levels. Therefore, let $\breve{G}_g$ be the number of rows for the contrast matrix produced by ranked variable g (this will depend on the ranking preferences provided by each individual and will not be constant across all individuals unlike the case of non-tied rankings). Therefore, the total number of rows for the contrast matrix pertaining to the ranked variables will be:

$$\widehat{G} = \sum_{g=1}^{G} \breve{G}_g$$

Therefore, with the new matrix $\mathbf{M}$ of size $\left[N+\widehat{G}\right]\times\left[N+\ddot{G}\right]$ as defined, we can write $\boldsymbol{yu} \sim MVN_{N+\hat{G}}(\tilde{\boldsymbol{B}},\tilde{\boldsymbol{\Omega}})$, where $\widetilde{\boldsymbol{B}} = \mathbf{M}\boldsymbol{B}$ and $\widetilde{\boldsymbol{\Omega}} = \mathbf{M\Omega M}'$, for the case of tied ranking

Next, define threshold vectors as follows:

$\ddot{\boldsymbol{\psi}}_{low} = \left[\tilde{\boldsymbol{\psi}}'_{low}, \left(-\boldsymbol{\infty}_{\widehat{G}}\right)'\right]'$ ($[(N+\widehat{G})\times 1]$ vector) and $\ddot{\boldsymbol{\psi}}_{up} = \left[\tilde{\boldsymbol{\psi}}'_{up}, \left(\mathbf{0}_{\widehat{G}}\right)'\right]'$ ($[(N+\widehat{G})\times 1]$ vector),

where $-\boldsymbol{\infty}_{\widehat{G}}$ is a $\widehat{G}\times 1$-column vector of negative infinities, and $\mathbf{0}_{\widehat{G}}$ is another $\widehat{G}\times 1$-column vector of zeros. Then the likelihood function may be written as:

$$L(\boldsymbol{\delta}) = \Pr\left[\ddot{\boldsymbol{\psi}}_{low} \le \boldsymbol{yu} \le \ddot{\boldsymbol{\psi}}_{up}\right], \qquad (13)$$

$$= \int_{D_r} f_{N+\widehat{G}}(\boldsymbol{r} \mid \tilde{\boldsymbol{B}},\tilde{\boldsymbol{\Omega}})dr,$$

where the integration domain $D_r = \{\boldsymbol{r} : \ddot{\boldsymbol{\psi}}_{low} \le \boldsymbol{r} \le \ddot{\boldsymbol{\psi}}_{up}\}$ is simply the multivariate region of the elements of the $\boldsymbol{yu}$ vector determined by the observed ordinal outcomes, and the range $(-\boldsymbol{\infty}_{\tilde{G}}, \mathbf{0}_{\tilde{G}})$ for the utility differences taken with respect to the utility of the ranked preference for the ranked outcome. The likelihood function for a sample of Q decision-makers is obtained as the product of the individual-level likelihood functions.

Since a closed form expression does not exist for this integral and evaluation using simulation techniques can be time consuming, we used the One-variate Univariate Screening technique proposed by Bhat (2018) for approximating this integral. The estimation of parameters was carried out using the *maxlik* library in the GAUSS matrix programming language.

References

Bhat, C.R., 2015. A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. *Transportation Research Part B* 79, 50-77.

Bhat, C.R., 2018. New matrix-based methods for the analytic evaluation of the multivariate cumulative normal distribution function. *Transportation Research Part B* 109, 238-256.

Table 1. Loading of Indicators on Latent Constructs

Attitudinal Indicators	Loading of Indicators on Latent Constructs							
	Tech Savviness		Safety Concern		Being Chill		IPTT	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
I like to be the first to have the latest technology	0.516	10.71						
Learning how to use new tech is often frustrating to me	-0.413	-7.45						
Having internet connectivity everywhere I go is important to me	0.310	7.72						
I would feel comfortable having an AV pick up/drop off children without adult supervision			-0.795	-19.08				
I am concerned about the potential failure of AV sensors, equipment, technology, or programs			0.457	11.84				
I would feel comfortable sleeping while traveling in an AV			-0.792	-21.24				
AVs would make me feel safer on the street as a pedestrian or cyclist			-0.662	-18.34				
Having to wait can be a good pause in a day					0.600	15.94		
I prefer to do one thing at a time					0.163	6.16		
The time spent traveling to places provides a useful transition between activities					0.653	15.90		
I try to make good use of my time traveling							0.810	14.36
The level of congestion on my daily travel bothers me							0.209	7.66