

# **Non-decreasing threshold variances in mixed generalized ordered response models: A negative correlations approach to variance reduction**

**Suryaprasanna Kumar Balusu**

Doctoral Student

Department of Civil and Environmental Engineering

University of South Florida, 4202 E. Fowler Avenue, ENB 118, Tampa, FL 33620

Email: balusu@mail.usf.edu

**Abdul Rawoof Pinjari** (corresponding author)

Associate Professor

Department of Civil Engineering

Centre for infrastructure, Sustainable Transportation, and Urban Planning (CiSTUP)

Indian Institute of Science (IISc), Bangalore 560012, India

Email: abdul@iisc.ac.in

**Fred L. Mannering**

Associate Dean for Research, College of Engineering

Professor of Civil and Environmental Engineering

University of South Florida, 4202 E. Fowler Avenue, ENC 3506, Tampa, FL 33620

Email: flm@usf.edu

**Naveen Eluru**

Associate Professor

Department of Civil, Environmental and Construction Engineering

University of Central Florida, 12800 Pegasus Drive, Room 301D, Orlando, FL 32816

Email: naveen.eluru@ucf.edu

September 2018

## **Abstract**

Mixed Generalized Ordered Response (MGOR) models, that allow random heterogeneity in thresholds, are widely used to model ordered outcomes such as accident injury severity. This study highlights a potential limitation of these models, as applied in most empirical research, that the variances of the random thresholds are implicitly assumed to be in a non-decreasing order. This restriction is unnecessary and can lead to difficulty in estimation of random parameters in higher order thresholds. In this study, we investigate the use of negative correlations between random parameters as a variance reduction technique to relax the property of non-decreasing variances of thresholds in MGOR models. To this end, a simulation-based approach was used (where multiple datasets were simulated assuming a known negative correlation structure between the true parameters), and two models were estimated on each dataset – one allowing correlations between random parameters and the other not allowing such correlations. Allowing negative correlations helped relax the non-decreasing variance property of MGOR models. However, maximum simulated likelihood estimation of parameters on data with correlations occasionally encountered model convergence and parameter identification issues. Comparison of the models that did converge suggests that ignoring correlations leads to an estimation of fewer random parameters in the higher order thresholds and results in bias and/or loss of precision for a few parameter estimates. However, ignoring correlations leads to an adjustment of other parameter estimates such that overall likelihood values, predicted percentage shares, and the marginal effects are similar to those from the models with correlations.

**Keywords:** generalized ordered response models, random thresholds, variance reduction, negative correlation, parameter identification

## 1. Introduction

Ordered outcomes, such as those encountered in accident-injury severity (no injury, injury, fatality), measurements of satisfaction (highly dissatisfied, dissatisfied, neutral, satisfied, highly satisfied), measurements of levels of agreement or disagreement (strongly disagree, disagree, neutral, agree, strongly agree), and so on, are often modeled using ordered response models. These models have a potential advantage over unordered response models, such as the multinomial logit model and its variants, because ordered models recognize the inherent ordinal pattern of outcome responses. Standard ordered response models are based on an underlying continuous latent propensity function that is assumed to be a function of observed explanatory variables and an unobserved random component (Aitchison and Silvey, 1957; McKelvey and Zavoina, 1975; Washington et al., 2011). The latent propensity function is mapped to observed outcomes using a set of thresholds that are increasing in order. The major drawback associated with this standard ordered response (SOR) model is that it assumes the thresholds to be same for all individuals, which might not be appropriate in all applications.

To overcome this threshold restriction in the standard ordered response models, Maddala (1986) and Ierza (1985) proposed a generalized-thresholds version of the ordered response model where the thresholds were expressed as a linear function of explanatory variables. As an extension to this model structure, Srinivasan (2002) expressed the thresholds as correlated random variables with their mean as a linear function of observed explanatory variables. However, this linear specification of thresholds does not ensure the increasing order of thresholds and might result in negative probabilities (Greene and Hensher 2010a). To address this issue, Eluru et al. (2008) and Greene and Hensher (2010b) used a nonlinear specification for thresholds where each threshold was obtained by adding a non-negative term to the preceding threshold, so that the ordering of

thresholds was ensured. The non-negative term was specified as an exponential function of a linear function of explanatory variables. Researchers have termed this generalized-thresholds version as the generalized ordered response model. To avoid confusion with the model names used in the literature, we term the linear-thresholds specification models as the ordered mixed response (OMR) model and the nonlinear-thresholds specification generalized ordered response (GOR) model. With regard to the GOR model, to account for heterogeneity in the parameter estimates due to unobserved factors, researchers have considered random parameters in both the propensity function and the thresholds. This model structure is referred to as the mixed generalized ordered response (MGOR) model by Eluru et al. (2008) and hierarchical ordered probit (HOPIT) model by Greene and Hensher (2010a). We use the term “MGOR” hereafter to avoid confusion with the model names. It is worth noting here that the random parameters in the thresholds are typically assumed to follow distributions with an unbounded support, such as the normal distribution.

Due to the flexibility offered by generalized ordered response (GOR) and mixed generalized ordered response (MGOR) models relative to the standard ordered response (SOR) model, many researchers (Yasmin et al. 2015a, 2015b; Forbes and Habib 2015; Fountas and Anastasopoulos 2017) have used these models in various contexts. Chiou et al. (2013) proposed a bivariate generalized ordered probit model and used it to model accident-injury severities in two-vehicle crashes. Castro et al. (2013) developed spatial random parameters generalized ordered probit model to accommodate the spatial dependencies in the accident-injury severity levels. Yasmin et al. (2014) proposed a latent segmentation based generalized ordered logit model assuming the presence of different latent groups of observations. Table 1 summarizes various

studies that have used the GOR family of models in the context of modeling traffic accident injury severity outcomes<sup>1</sup>.

Despite the above-discussed evolution of the MGOR family of models, to the best of our knowledge, all implementations of the MGOR models to date impose an implicit restriction on the order of variances of thresholds. Specifically, as discussed earlier, the thresholds in ordered response models must be in an increasing order, which is ensured in MGOR models by specifying a higher order threshold as a sum of its preceding threshold and a non-negative random term that is typically in the form of an exponential function. Such a hierarchical specification of thresholds with random parameters leads to the restriction that the variances of thresholds are also in a non-decreasing order. However, this restriction is not necessary and can potentially lead to difficulty in the estimation of random parameters in higher order thresholds (more later).

To be sure, the MGOR model structure, in its very general form, does allow the analyst to relax the non-decreasing order of threshold variances. This can be done in at least two ways. The first approach is to allow negative correlations between the random parameters of different thresholds. Since a higher order threshold is specified as a sum of two terms (its preceding threshold with random parameters and an exponential term with random parameters), negative correlation between the two terms allows for the overall variance of the higher order threshold to be lower than the variance of its preceding threshold. The second approach is to use truncated distributions for thresholds, where the distribution of a higher order threshold is left-truncated by the distribution of its preceding threshold. Between these two approaches, the former is easier to

---

<sup>1</sup> Mannering et al. (2016), provide a general discussion of unobserved heterogeneity in accident injury-severity modeling. Apart from the accident injury-severity modeling, there are other research areas (sociology, psychology and economics) that have used OMR and MGOR model structures (Pudney and Shields, 2000; Boes and Winkelmann, 2006; Greene et al., 2008; Baba, 2009; Mentzakis and Moro, 2009; Boes and Winkelmann, 2010; Stanley et al., 2011; Greene et al., 2014; and Shabanpour et al., 2017).

implement. The latter approach is a non-trivial<sup>2</sup> modification of the MGOR structure, albeit it is a fruitful avenue for future research. Even in the context of the former approach, we are not aware of studies in the literature that explored correlated random parameters in MGOR models.

The intent of the current paper is to highlight the above-discussed implicit assumption made by most implementations of MGOR models that the thresholds follow a non-decreasing order of variances. In addition, the paper explores the use of negative correlations as a variance reduction technique for relaxing the non-decreasing variances restriction in the MGOR family of models. We explore these variance and correlation issues through a simulation experiment. Specifically, we simulate ordinal outcome datasets with known negative correlation structures among an underlying true set of random parameters in threshold functions. For each of the simulated datasets, two models were estimated; one allowing correlations between random parameters and the other not allowing such correlations. The impact of ignoring correlations is then evaluated by comparing the two models using various evaluation criteria to assess the efficacy of introducing negative correlations as a variance reduction technique in the thresholds of MGOR models.

The remainder of our paper is structured as follows. Section 2 presents the model structure of SOR and MGOR models. Section 3 provides a simple, mathematical proof for the non-decreasing order of variance of thresholds in the absence of correlation between random parameters in thresholds. In addition, this section presents a hypothetical scenario to explain how such a restriction (due to ignoring correlations) can potentially lead to difficulties in estimating random parameters in higher order thresholds. Section 4 describes the methodology (simulation

---

<sup>2</sup> The left truncation point of the distribution for a higher order threshold is another random variable (given by the distribution of the preceding threshold), as opposed to a deterministic value. Therefore, deriving an MGOR model structure with randomly truncated threshold distributions is a non-trivial extension and beyond the scope of this paper.

experiments) adopted to evaluate correlations as a variance reduction technique in the thresholds of MGORL models. Section 5 presents the results and findings. Section 6 concludes the study.

## 2. Model structure

In this section, we present the basic model structures for the standard ordered response model (SOR) and the mixed generalized ordered response model (MGOR).

### 2.1 Standard ordered response model (SOR)

Let  $n$  (1, 2, 3... $N$ ) denote each observation and  $k$  (1, 2, 3... $K$ ) denote ordered outcomes. The latent propensity function  $y_n^*$  for observation  $n$  is expressed as

$$y_n^* = \boldsymbol{\beta} \mathbf{X}_n + \varepsilon_n \quad (1)$$

where  $\mathbf{X}_n$  is a vector of explanatory variables that influence  $y_n^*$ ,  $\boldsymbol{\beta}$  is a corresponding vector of estimable parameters, and  $\varepsilon_n$  is an unobserved random term which is assumed to follow a known probability distribution. Observed ordinal outcome  $y_n$  is then defined by the latent propensity function  $y_n^*$  using a set of threshold parameters as follows:

$$y_n = k, \text{ if } \Psi_{k-1} < y_n^* < \Psi_k \quad (2)$$

where  $\Psi_{k-1}$  and  $\Psi_k$  are a pair of estimable thresholds associated with  $k^{th}$  ordered outcome. All the thresholds are restricted to be in an increasing order, and the lower most and upper most thresholds are assumed to be negative infinity and positive infinity ( $-\infty < \Psi_1 < \Psi_2 < \dots < \Psi_{K-1} < \infty$ ), respectively (this is assuming that the latent propensity function  $y_n^*$  follows an unbounded distribution). For identification reasons, either the constant in the propensity function or any one of the thresholds must be fixed to zero. In this exposition, the constant in the propensity function

is fixed to zero, and all the  $K-1$  thresholds are parameters to be estimated. The log-likelihood ( $LL_{nk}$ ) for observation  $n$  facing  $k^{th}$  ordered outcome is:

$$LL_{nk} = Pr(y_n = k | \mathbf{X}_n) = \Gamma[\Psi_k - \boldsymbol{\beta}\mathbf{X}_n] - \Gamma[\Psi_{k-1} - \boldsymbol{\beta}\mathbf{X}_n] \quad (3)$$

where  $\Gamma[\cdot]$  is the cumulative distribution function of the random error term  $\varepsilon_n$ .

## 2.2 Mixed generalized ordered response (MGOR) model

The MGOR model structure is an extension of the SOR model structure with the thresholds parameterized as a function of explanatory variables, and the inherent ordering of the thresholds is ensured using a nonlinear specification for thresholds where each threshold is specified by adding a non-negative term to the preceding threshold. Moreover, to account for unobserved heterogeneity in the parameter estimates across observations, random parameters are included in the propensity function and the threshold functions as shown in Eq. (4), (5), and (6).

$$y_n^* = \boldsymbol{\beta}\mathbf{X}_n + \boldsymbol{\gamma}_n\mathbf{Y}_n + \varepsilon_n \quad (4)$$

$$\Psi_{nk} = \boldsymbol{\alpha}_k\mathbf{U}_{nk} + \boldsymbol{\theta}_{nk}\mathbf{V}_{nk}, \text{ if } k = 1 \quad (5)$$

$$\Psi_{nk} = \Psi_{n,k-1} + \exp(\boldsymbol{\alpha}_k\mathbf{U}_{nk} + \boldsymbol{\theta}_{nk}\mathbf{V}_{nk}), \forall k > 1 \quad (6)$$

where  $\mathbf{X}_n$  and  $\mathbf{Y}_n$  are vectors of exogenous variables in the propensity function,  $\boldsymbol{\beta}$  is a vector of fixed parameters and  $\boldsymbol{\gamma}_n$  is a vector of random parameters in the propensity function. Similarly,  $\mathbf{U}_{nk}$  and  $\mathbf{V}_{nk}$  are vectors of exogenous variables,  $\boldsymbol{\alpha}_k$  and  $\boldsymbol{\theta}_{nk}$  are vectors of fixed and random parameters, respectively, in the  $k^{th}$  threshold. For identification reasons, and without loss of generality, all the parameters in the first threshold except a constant are set to zero ( $\Psi_{n1} = \alpha_1$ ).

As discussed earlier, the hierarchical specification of thresholds in Eq. (6), where a higher order threshold ( $\Psi_{nk}$ ) is specified as a sum of its preceding threshold ( $\Psi_{n,k-1}$ ) plus a non-negative random term,  $\exp(\boldsymbol{\alpha}_k\mathbf{U}_{nk} + \boldsymbol{\theta}_{nk}\mathbf{V}_{nk})$ , ensures that the thresholds are in an increasing order.



The random parameters vectors  $\boldsymbol{\gamma}_n$  and  $\boldsymbol{\theta}_n$ , where  $\boldsymbol{\theta}_n$  is obtained by stacking the  $\boldsymbol{\theta}_{nk}$  vectors across all  $k$ , are realizations from multivariate distributions  $f(\boldsymbol{\gamma})$  and  $f(\boldsymbol{\theta})$ . The log-likelihood ( $LL_{nk}$ ) for observation  $n$  facing  $k^{th}$  ordered outcome is written as,

$$LL_{nk} = \int_{\boldsymbol{\gamma}} \int_{\boldsymbol{\theta}} \Gamma[(\Psi_{nk}|\boldsymbol{\theta}) - (\boldsymbol{\beta}X_n|\boldsymbol{\gamma})] - \Gamma[(\Psi_{n,k-1}|\boldsymbol{\theta}) - (\boldsymbol{\beta}X_n|\boldsymbol{\gamma})] f(\boldsymbol{\theta}) f(\boldsymbol{\gamma}) d(\boldsymbol{\theta}) d(\boldsymbol{\gamma}) \quad (7)$$

Note that  $f(\boldsymbol{\gamma})$  and  $f(\boldsymbol{\theta})$  are multivariate distributions. Therefore, the structure of the MGOR model allows for correlations among the random parameters, in the latent propensity function as well as in the threshold functions. However, it is a common practice in empirical research to ignore such correlations; as indicated earlier, we are not aware of empirical studies that explored correlations between random parameters in the threshold functions of ordered response models.

### **3. Non-decreasing order of variances of thresholds in MGOR models with uncorrelated random parameters**

In this section, we prove the non-decreasing order of variances of thresholds in MGOR models with uncorrelated random parameters and demonstrate, through a hypothetical example, how such restriction might lead to difficulties in estimating random parameters in higher order thresholds.

#### *3.1 The order of variance of thresholds*

Let  $VAR(\cdot)$  and  $E(\cdot)$  represent the variance and the expected value of a random variable, respectively, and let  $COV(\cdot)$  represent the covariance between any two random variables. Thresholds and non-negative terms are random in the presence of random parameters and the variance of a threshold  $\Psi_{nk}$  is expressed as  $VAR(\Psi_{nk}) = VAR(\Psi_{n,k-1}) + VAR(\Delta_{nk}) + 2COV(\Psi_{n,k-1}, \Delta_{nk})$ , where  $\Delta_{nk}$  is a non-negative term that is added to  $\Psi_{n,k-1}$  to obtain  $\Psi_{nk}$ . As

can be observed in Eq. (6), the non-negative term  $\Delta_{nk}$  is  $\exp(\boldsymbol{\alpha}_k \mathbf{U}_{nk} + \boldsymbol{\theta}_{nk} \mathbf{V}_{nk})$ . If the correlations between the random parameters in  $\Psi_{n,k-1}$  and  $\Delta_{nk}$  are ignored or restricted to zero, the covariance term,  $COV(\Psi_{n,k-1}, \Delta_{nk})$ , becomes zero and forces the variance of each threshold to be either greater than or equal<sup>3</sup> to the variance of the preceding threshold. On the other hand, a negative correlation between the random parameters in  $\Psi_{n,k-1}$  and  $\Delta_{nk}$  allows for a possibility that  $VAR(\Psi_{nk}) < VAR(\Psi_{n,k-1})$ , depending on the level of correlation between the random parameters and the magnitude of the deterministic components.

Considering normally distributed random parameters, which are generally employed in empirical research involving MGOR models, the thresholds can be viewed as a sum of multiple log-normally distributed random variables.<sup>4</sup> Following the notation used in section 2.2, the expressions for the variance of first three thresholds with normally distributed random parameters can be written as shown below (see Appendix A for a detailed derivation).

$$\begin{aligned} VAR(\Psi_{n1}) &= 0, \\ VAR(\Psi_{n2}) &= VAR(\exp(\boldsymbol{\alpha}_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2})), \end{aligned} \tag{8}$$

$$VAR(\Psi_{n3}) = VAR(\exp(\boldsymbol{\alpha}_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2})) + VAR(\exp(\boldsymbol{\alpha}_3 \mathbf{U}_{n3} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})) + C_{n23},$$

where  $C_{n23} = \exp[\boldsymbol{\alpha}_2 \mathbf{U}_{n2} + \boldsymbol{\alpha}_3 \mathbf{U}_{n3} + E(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3}) + \frac{VAR(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2}) + VAR(\boldsymbol{\theta}_{n3} \mathbf{V}_{n3})}{2}] \times [\exp(COV(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2}, \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})) - 1]$ .

If the random parameters are uncorrelated, the covariance term  $COV(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2}, \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})$  is equal to zero, which implies that  $C_{n23} = 0$  and  $VAR(\Psi_{n3}) = VAR(\Psi_{n2}) + VAR(\exp(\boldsymbol{\alpha}_3 \mathbf{U}_{n3} +$

---

<sup>3</sup> The variances would be equal when the non-negative term  $\Delta_{nk}$  does not have random parameters. For example, when an empirical specification does not yield statistically significant random parameters in  $\Delta_{nk}$ .

<sup>4</sup> This is assuming that the first threshold, which is not an exponential function, is not a random parameter (for identification reasons). However, if the first threshold is a normally distributed random parameter (with other normalization restrictions for identification), then the subsequent thresholds will become a sum of normal distribution and lognormal distributions.

$\theta_{n3}V_{n3}$ )). Therefore, in the presence of uncorrelated random parameters in the thresholds, the variance of thresholds are restricted to be in a non-decreasing order for a given observation. But there is no need for imposing such a restriction on the order of variance of thresholds. For example, in the ordered response model (with the linear specification for thresholds and with correlation between the random parameters in thresholds) estimated by Srinivasan (2002), the variances of the estimated thresholds do not follow any order.

This inherent restriction on the variance of thresholds can be relaxed by allowing correlations between random parameters in thresholds, which indeed makes the covariance term  $COV(\theta_{n2}V_{n2}, \theta_{n3}V_{n3}) \neq 0$ . Depending on the sign and level of correlation between the random parameters, and the magnitude of the deterministic component in the thresholds, the variance of a threshold can be less than the variance of the preceding threshold. Specifically, negative correlations allow for the possibility of non-decreasing order in the variances of the thresholds. Although we derived these expressions for normally-distributed random parameters, the above-discussed restriction on the order of variance of the thresholds will occur with the other parametric distributions as well when correlations are not allowed.

### *3.2 Potential issues with estimation of random parameters in thresholds*

Due to the hierarchical specification of thresholds in the MGOR model, observed variables and their coefficients (along with the random components) entering the threshold function of a lower order threshold (say, the  $k^{\text{th}}$  threshold) also enter the threshold functions of all higher order thresholds (greater than  $k^{\text{th}}$  thresholds). Therefore, an independent variable entering  $k^{\text{th}}$  threshold function influences the probabilities of the corresponding  $k^{\text{th}}$  order outcome as well as potentially all higher order outcome responses.

For example, let us consider the injury severity of motorcycle crashes on freeways. Let the outcome injury severity levels be no injury, non-incapacitating injury, incapacitating injury, and fatal injury and the corresponding thresholds are  $\Psi_{1n}$ ,  $\Psi_{2n}$ , and  $\Psi_{3n}$ . Define a variable  $PC_n$  for protective clothing, which is equal to one if the motorcyclist wore protective clothing during the crash, and zero otherwise (protective clothing may include jacket, gloves, heavy pants, boots, knee pads and elbow guards). Let the threshold specifications be as  $\Psi_1 = \alpha_1$ ,  $\Psi_{2n} = \alpha_1 + \exp(\alpha_2 + \theta_2 PC_n)$  and  $\Psi_{3n} = \alpha_1 + \exp(\alpha_2 + \theta_2(PC_n)) + \exp(\alpha_3)$ , where,  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  are fixed constants and  $\theta_2$  is a positive fixed parameter on protective clothing indicator variable entering directly only in the second threshold<sup>5</sup>. Figure 1a shows the position of thresholds for two groups of individuals: (a) who wore a protective clothing during the crash and (b) who did not wear it during the crash. In the case when  $PC_n$  enters only the second threshold ( $\Psi_{2n}$ ) and takes the value 1, the second and third thresholds move to the right by  $\Delta_x (= \exp(\alpha_2 + \theta_2) - \exp(\alpha_2))$  on the propensity scale, resulting in an increase in the probability of non-incapacitating injury and decrease in the probability of fatal injury (while all the independent variables and parameters in the propensity function remains same).

In reality, however, wearing protective clothing during a high impact crash, such as those likely to occur on freeways, will likely reduce the injury severity level from incapacitating injury to non-incapacitating injury, but may have little influence on reducing higher injury-severity levels (fatal injury level in our case). See, for example, Erdogan et al. (2013) for a supporting finding that wearing protective clothing protects from soft tissue injuries but not from severe fractures. To incorporate such differential effect of protective clothing on incapacitating and fatal injury levels, the  $PC_n$  variable should enter directly into the  $\Delta_{nk}$  term of the third threshold as well (along with

---

<sup>5</sup> Note that the protective clothing variable enters the third threshold through the second threshold but not directly.

its entry through the second threshold), but with a negative coefficient, that is, when the third threshold is specified as  $\Psi_{3n} = \exp(\alpha_2 + \theta_2(PC_n)) + \exp(\alpha_3 + \theta_3(PC_n))$ , where  $\theta_3 < 0$ . With such specification, when  $PC_n$  takes the value 1, the second term of  $\Psi_{3n}$  shrinks by  $-\Delta_y$  ( $= \exp(\alpha_3 + \theta_3) - \exp(\alpha_3)$ ) making the overall shift in the third threshold equal to  $\Delta_x - \Delta_y$  (as shown in figure 1b). That is, a rightward shift of  $\Psi_{3n}$  through a positive coefficient ( $\theta_2$ ) will be counteracted by a leftward shift through a negative coefficient ( $\theta_3$ ). Naturally, the higher is the value of  $\theta_2$ , the lower should be the value of  $\theta_3$  (a negative number of larger magnitude) for counteracting the influence of  $\theta_2$  on  $\Psi_{3n}$ .

Now, let us extend this discussion when we have random parameters on  $PC_n$ . Let the parameter estimates of  $PC_n$  in second and third thresholds be represented by two random parameters  $\theta_{2n}$  and  $\theta_{3n}$ , respectively. Analogous to the discussion above, a negative dependency can be allowed between  $\theta_{2n}$  and  $\theta_{3n}$  through a negative correlation between the two random parameters. Ignoring such dependency (or negative correlation), as discussed earlier, imposes that the variance of  $\Psi_{3n}$  is greater than that of  $\Psi_{2n}$ . Since variability in the influence of unobserved influences on the third threshold need not always be greater than that on the second threshold, ignoring negative correlation between  $\theta_{2n}$  and  $\theta_{3n}$  might make it difficult to estimate a statistically significant variance parameter for  $\theta_{3n}$ . In other words, the above-discussed restriction might make it difficult to estimate statistically significant random parameters in higher order thresholds, simply because random parameters from lower order thresholds would simply carry forward to higher order thresholds.

As evident from the literature reviewed in Table 1, Eluru et al. (2008), Yasmin et al. (2015a) and Xin et al. (2017) tried to estimate random parameters in the thresholds without allowing correlations between them. However, perhaps due in part to the above-mentioned

reasons, they were unable to find statistically significant random parameters in thresholds. Therefore, ideally, the model structure should not restrict an order on the variances of thresholds while specifying and testing the model.

#### 4. Experimental design

To evaluate the efficacy of introducing negative correlations as a variance reduction technique in the thresholds of MGOR models, we simulated motorcycle crash datasets with known negative correlation structures among an underlying true set of random parameters in threshold functions. While one may use a real data to evaluate the technique, it is difficult to control for the unobserved internal relationships which might exist between the independent variables and injury outcomes. In order to avoid such issues, a simulation-based approach is used in this study. For each of the simulated datasets, we estimated two models – one allowing correlations between random parameters and the other not allowing correlations. The impact of ignoring correlations was then evaluated by comparing the two models using various evaluation criteria.

The outcome injury severity levels in the simulated datasets were no injury, non-incapacitating injury, incapacitating injury, and fatal injury. We assumed that three explanatory variables – age of motorcyclist, male indicator (1 if a motorcyclist is male, zero otherwise), and intersection indicator (1 if a crash occurred at an intersection, zero otherwise) – influence the latent injury risk propensity of a motorcyclist. Specifically, the latent propensity function is specified as:

$$y_n^* = (\beta_1 \times age_n) + (\beta_2 \times male_n) + (\beta_3 \times intersection_n) + \varepsilon_n \quad (9)$$

where  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  are the parameters and  $\varepsilon_n$  is the random component of the propensity function.

To simulate data with the above propensity function, values for the age variable were drawn from a truncated normal distribution with mean 40 years, standard deviation 15 years, and 16 years

and 75 years as the left and right truncation limits, respectively. Values for the indicator variables were drawn from Bernoulli distributions with mean 0.5. Values for the error term  $\varepsilon_n$  in the propensity function were drawn from a standard logistic distribution.

#### 4.1 Threshold scenarios

We simulated five different scenarios for the thresholds, as summarized in Table 2. The second column of the table specifies the propensity function and the threshold functions used, including the parameter values assumed, in each scenario. The parameter of the age variable in the propensity function is assumed to be positive considering that the older people tend to be susceptible to a higher injury severity levels relative to younger people (Savolainen and Mannering 2007). Literature suggests that males tend to sustain a lower injury severities relative to females and therefore a negative parameter is selected for the male indicator variable (Quddus et al. 2002; Rifaat et al. 2012). Similarly, a negative parameter is considered for the intersection indicator variable assuming that the crashes occurring at intersections tend to be less severe due to driver caution and other factors (Savolainen and Mannering 2007).

As can be observed from the second column of Table 2, scenarios S1, S2, S3, and S5 simulated threshold functions based on the example discussed in Section 3.2. Specifically, thresholds were assumed to depend on whether a motorcyclist was wearing a protective clothing or not, using an indicator variable ( $PC_n$ ) that was equal to 1 if the motorcyclist was wearing a protective clothing when the crash happened (zero otherwise). This indicator variable was assumed to be Bernoulli distributed with mean 0.5. Random parameters were allowed on the protective-clothing indicator variable in the second and third thresholds while keeping the first threshold fixed (to  $\alpha_1$ ), as in Equation (10) below:

$$\begin{aligned}
\Psi_1 &= \alpha_1, \\
\Psi_{2n} &= \alpha_1 + \exp(\alpha_2 + \theta_{2n} \times PC_n), \\
\Psi_{3n} &= \alpha_1 + \exp(\alpha_2 + \theta_{2n} \times PC_n) + \exp(\alpha_3 + \theta_{3n} \times PC_n).
\end{aligned} \tag{10}$$

The random parameters  $\theta_{2n}$  and  $\theta_{3n}$  in all the four scenarios (S1, S2, S3, and S5) were simulated from two normal distributions  $N(\theta_2, \sigma_2)$  and  $N(\theta_3, \sigma_3)$  with a correlation level of  $\rho_{23}$  (which was assumed to be  $-0.7$ ).

In the fourth scenario (S4), however, correlated random parameters were introduced on constants in the first and second thresholds (with a correlation parameter  $\rho_{12} = -0.7$ )<sup>6</sup>, while keeping the coefficients of the protective clothing variable to be fixed, as shown in Equation (11) below.

$$\begin{aligned}
\Psi_{1n} &= \alpha_{1n}, \\
\Psi_{2n} &= \alpha_{1n} + \exp(\alpha_{2n} + \theta_2 \times PC_n), \\
\Psi_{3n} &= \alpha_{1n} + \exp(\alpha_{2n} + \theta_2 \times PC_n) + \exp(\alpha_3 + \theta_3 \times PC_n).
\end{aligned} \tag{11}$$

Note from Table 2 that scenario S1 simulated a high percentage (50.3%) of fatal crashes (although empirical contexts with such a high percentage of fatal crashes are rare), S2 simulated a low percentage (6.6%) of fatal crashes, S3 simulated approximately equal shares for all injury-severity levels, S4 simulated a low percentage (7.6%) of fatal crashes, and S5 simulated a high percentage of fatal crashes. This allows us to examine the above-discussed issues in different data generation settings as defined by the percentage of different ordered outcomes. For the first four scenarios (S1, S2, S3, and S4), a sample size of 5,000 motorcyclists was generated from the assumed distributions for  $age_n$ ,  $male_n$ ,  $intersection_n$ , and  $PC_n$  variables. For the fifth scenario

---

<sup>6</sup> With regard to the equation 8, the variance of higher threshold can be less than the variance of higher threshold only when there is a negative covariance between those thresholds. We considered a higher negative correlation value of  $-0.7$  to achieve a higher negative covariance and make the variance reduction technique possible. Other correlation values of  $-0.5$  and  $-0.9$  were tried in the experimental design and found to produce similar findings.



(S5), the S1 scenario was simply repeated by increasing the sample size of motorcyclists from 5,000 to 10,000. This was done to evaluate the influence of sample size, while keeping all else same.

In each of the five scenarios, the outcome injury-severity level for each observation was obtained by mapping the propensity function with the thresholds as shown in Eq. (2). For each of the five scenarios, the data generation process was repeated 100 times to obtain 100 different datasets by drawing different values for the random components ( $\varepsilon_n$  and random parameters) from their corresponding distributions.

For each of the 100 datasets in each of the five scenarios, two MGORL models were estimated (a total of  $100 \times 5 \times 2 = 1,000$  MGORL model estimations). In the first model, correlation was allowed (estimated) between random parameters in the thresholds. In the second model, the correlation term was fixed to zero. To examine the recovery of random parameters and negative correlations in the thresholds, under different severity scenarios, propensity function and thresholds specifications are forced to be the same as the ones considered during the data simulation process. All models were estimated using the maximum simulated likelihood (MSL) approach with 400 Halton draws to simulate the distribution of random parameters (Bhat, 2003). Model estimations were carried out using codes written in the Gauss matrix programming language for the MGORL model with correlated random parameters.

#### *4.2 Model performance metrics*

To evaluate the performance of MGORL models estimated with and without correlated random parameters (on simulated data with correlations), the following criteria were considered:

**a.** The Log-likelihood improvement in the MGORL model after allowing correlated random parameters in thresholds was evaluated using a likelihood ratio test. Here, a model without correlation was a restricted version of a model with correlation, and the number of restrictions was equal to the difference in the number of parameter estimates in both the models. Chi-square value ( $\chi^2$ ) which is equal to  $-2 \times [LL_{mwoc} - LL_{mwc}]$  was computed for each dataset and was then compared with the critical chi-square value for a given number of restrictions at 95% confidence level.

**b.** The Absolute Percentage Bias (APB) for each parameter was calculated as the absolute percentage difference of the mean parameter estimate from the true parameter value. The mean estimate of each parameter was the average of all estimates across 100 datasets. This metric is a measure of accuracy of parameter estimates, expressed as given below:

$$APB = \left| \frac{\text{mean estimate} - \text{true value}}{\text{true value}} \right| \times 100 \quad (12)$$

**c.** The Finite Sample Standard Error (FSSE) for a given parameter was calculated as the standard deviation of that estimated parameters across 100 datasets. FSSE for a parameter may be interpreted as the empirical standard error of its estimate in finite samples. Comparison of this metric for each parameter in the two models (model with correlation and model without correlation) was used to assess the loss/gain in the precision of parameter estimate when the correlation between random parameters was allowed.

**d.** The Coverage Probability (CP) for each parameter was calculated using the formula:  $CP = 1/N \sum_{n=1}^N I[\hat{\beta}_X^n - t_\alpha \times se(\hat{\beta}_X^n) \leq \beta_X \leq \hat{\beta}_X^n + t_\alpha \times se(\hat{\beta}_X^n)]$ , where  $N$  is the total number of datasets (100),  $\hat{\beta}_X^n$  is the estimated parameter for a dataset  $n$ ,  $\beta_X$  is the true parameter value,  $se(\hat{\beta}_X^n)$  is the asymptotic standard error of  $\hat{\beta}_X^n$ , and  $I[\cdot]$  is an indicator function which takes a value 1 if the condition inside the bracket satisfies, otherwise equals to zero, and  $t_\alpha$  is the  $t$ -

statistic value for a given confidence level  $(1 - \alpha) \times 100$ . The confidence interval in this study was set to 95%. If the computed coverage probability is less than the nominal coverage probability (0.95), it suggests that the confidence intervals of the estimated parameter do not provide sufficient empirical coverage of the true parameter.

e. Predicted percentage shares were calculated for each injury-severity level using the estimated parameters from each dataset and were averaged across all datasets. The predicted percentage shares for both the models (model with correlation and model without correlation) were compared to each other to understand the impact of ignoring correlations on the predicted aggregate shares.

f. The marginal effects were calculated to understand the effect of each variable on the outcome response for the model with correlation and the model without correlation. For a continuous variable, marginal effects were calculated as the average change in the probability of injury severity levels for all individuals with a unit increase in the variable of interest from its current value. Marginal effects for an indicator variable were computed using the procedure presented by Eluru and Bhat (2007).

#### *4.3 Additional scenarios*

Apart from the five scenarios discussed in Section 4.1, additional scenarios were simulated. Recall that all the indicator variables (such as gender) were simulated as Bernoulli distributed with mean 0.5. However, the simulated data may not always represent the actual data. Therefore, scenario S1 was repeated with a gender split of 68% males 32% females, as observed in the motorcycle crashes reported in the 2016 crash data by Fatality Analysis Reporting System (FARS), keeping all other variables same. Such a scenario is labelled S6. The overall findings from this

scenario aligned with those from scenario S1 suggesting that the inferences from this study are applicable for simulated data based on real-world scenarios. Therefore, we retained the mean value of 0.5 for all the indicators in all other simulations.

To check the influence of number of draws on model estimation, we repeated scenario S1 by increasing the number of Halton draws from 400 to 1000 and the new scenario is termed S7. Results suggested that there is no notable change in the evaluation metrics, marginal effects, or percentage shares after increasing the number of Halton draws. Therefore, we used 400 Halton draws for all other estimations in this study.

Scenario S1 was repeated with a smaller sample size of 1000 (and labelled scenario S8). It was found that there is a decrease in the consistency of predictions with a decrease in the sample size. Moreover, there is a reduction in the consistency (and increase in the APB and FSSE) of parameter estimates with a decrease in the sample size. However, the comparison of models with and without correlation suggested that the new findings with regard to the evaluation of the variance-reduction technique are consistent with the S1's findings. Therefore, the results and findings of the scenario with a smaller sample size are not discussed further in the paper.

We repeated the scenario S4 with a higher proportion (43%) of fatal injuries and the new scenario is termed as S9. Table 1 contains the number of datasets with converged models and significantly improved log-likelihood for S9. Table 2, Table 3 and Figure 1 contains the evaluation metrics, marginal effects and predicted percentage shares respectively for S9. Similar to S4, S9 results suggest that the model with correlations is superior to that of a model without correlations in retrieving the parameter estimates in higher order thresholds. Moreover, there is no notable change in the evaluation metrics, marginal effects and percentage shares for S4 even after increasing the percentage of fatal crashes in the data.

Note that the results from the additional scenarios S6, S7, S8, and S9 are not reported in the paper to conserve space. Only the results from scenarios S1 through S5 are reported on the next section, because the overall findings from the additional scenarios are similar to those from the first set of scenarios.

## 5. Results

This section presents the results and findings of the simulation experiments to evaluate the efficacy of negative correlations as a variance reduction technique in threshold functions of MGORL models.

### *5.1 Order of variance in MGORL models with negatively correlated random parameters in thresholds*

We examined the order of threshold variances in simulated MGORL data with and without correlated random parameters. Since scenarios S1, S2, S3, and S5 include random parameters on a binary variable ( $PC_n$ ) entering the threshold functions, the thresholds are random only when the  $PC_n$  variable takes a value of 1. Therefore, in all these four scenarios, random parameters kick in only for 50% of the cases where the  $PC_n$  variable takes a value of 1. For all these cases, as discussed in Section 3, variances of the second and third thresholds ( $\psi_{n2}$  and  $\psi_{n3}$ ) in the data without correlated random thresholds were in an increasing order, with the variance of  $\psi_{n2}$  as 5.94 and that of  $\psi_{n3}$  as 8.14. This order reversed in data with correlated random thresholds, with the variance of  $\psi_{n2}$  as 5.94 and that of  $\psi_{n3}$  as 5.02, demonstrating the use of negative correlations for relaxing the assumption of non-decreasing variances. In the scenario S4, since random parameters are associated with the constants of the threshold functions, the thresholds are random for all cases

in any of the 100 datasets. Therefore, introducing a negative correlation of -0.7 rendered the order of thresholds to be decreasing for all cases.

### *5.2 Estimation issues of MGORL models with negatively correlated random parameters in thresholds*

During the estimation of MGORL models with correlated random thresholds (on the simulated data with correlated random thresholds), we encountered non-convergence issues for at least 10% of the datasets in each scenario. More specifically, the number of simulated datasets (out of 100) for which the MGORL models with correlated random thresholds converged in each scenario are reported in Table 3 (second column). For the remaining datasets in each scenario, the non-convergence issues arose due to the occurrence of maximum value for the log-likelihood ( $LL$ ) function at the boundary value (-1) of the correlation term, which was not a stationary point to satisfy the convergence criterion.

To examine this issue, we plotted the  $LL$  function profiles of un-converged models (in scenario S2) with respect to the correlation term while fixing the other parameter estimates. Figure 2 shows the variation of the  $LL$  function with respect to the correlation term for each of the 14 datasets on which the models did not converge in scenario S2. It can be observed that the  $LL$  function is monotonic in the range of the correlation term (-1 to +1). The maximum value of the log-likelihood function for all 14 profiles in Figure 2 is at the correlation level of -1 where the  $LL$  function is not stationary.

Interestingly, by changing the starting values of the parameters to be estimated, convergence was achieved for 3 out of the 14 datasets mentioned above, with correlation values different from -1. However, although the parameter estimates (including those of the correlation

parameter) obtained from both the converged and the corresponding un-converged models were different, the final  $LL$  values were not different between the converged models and the corresponding un-converged models at boundary values. This suggests a flat  $LL$  function profile, potentially due to identification problems in models with correlated random parameters in thresholds. Besides, the parameter estimates for the correlation term and the standard deviation of the second random parameter ( $\sigma_3$ , the random parameter in the higher threshold function) had high standard errors in some converged models, which again points to issues of parameter identifiability. Note also that for some of the converged models, the Hessian matrix could not be inverted at the final parameter estimates, and the  $t$ -statistics could not be computed using the sandwich estimator. So, the  $t$ -statistics were computed using the cross products of the first order derivatives. For subsequent analysis, we ignored the un-converged models and computed the evaluation measures only for the converged models in each scenario.

Interestingly, when we estimated models without correlation between random parameters in thresholds (again on simulated data with correlated random parameters), the standard deviation of the random parameter in the higher order random threshold (third threshold in S1, S2, S3, and S5 and second threshold in S4) was statistically insignificant at 95% confidence interval in almost all datasets for all five scenarios. This suggests difficulty in estimating random parameters in higher order thresholds of MGORL models, as discussed in Section 3.2. Therefore, the insignificant random parameter was replaced with a fixed parameter, and the model was re-estimated with only one random parameter, instead of two correlated random parameters<sup>7</sup>.

---

<sup>7</sup> Even in the model with correlated random parameters in second, third, and fourth scenarios, standard deviation of the second random parameter turned out to be statistically insignificant in at least 50% of the datasets. This could be due to the flat nature of the log-likelihood function and lower  $t$ -statistics resulting from the calculation of covariance matrix for the parameter estimates using the first order derivatives instead of using the sandwich estimator. Removing the insignificant random parameters in the model with correlation, and re-estimating the model, eliminates the need for the correlation term. We elected to keep the insignificant random parameter in the model and retain the correlation term to keep the results as general as possible.

### *5.3 Data fit of MGORL models with negatively correlated random parameters in thresholds vis-à-vis those without correlated random parameters*

For each dataset on which we could estimate a model with correlated random parameters without facing convergence issues (see column 3 of Table 3), a likelihood ratio test was conducted between the model without correlation and the model with correlation. This likelihood ratio test was associated with two degrees of freedom, since both the standard deviation of the random parameter in the higher order threshold and the correlation term were constrained to zero in the model without correlated random parameters. The results of the likelihood ratio tests are shown in Table 3 (last column) in the form of the number of datasets which show a statistically significant improvement in log-likelihood when correlated random thresholds were allowed. Interestingly, allowing correlated random parameters did not yield a significant improvement in data fit in a majority of the datasets in all five scenarios. Note also that increasing the sample size from 5,000 (in S1) to 10,000 (in S5) did not substantially change the results.

### *5.4 Recovery of parameters*

Table 4 reports, for all the five scenarios, the metrics of parameter recovery from the maximum simulated likelihood estimation technique for both the models (the model with correlated random parameters<sup>8</sup> and the model without correlated random parameters). Recall that these metrics (APB, FSSE, and CP) have already been defined in Section 4.

---

<sup>8</sup> For models with correlated random parameters, we ignored the un-converged models and computed the evaluation metrics only for the converged models in each scenario.



Comparing the metrics between scenario S5 and scenario S1, it can be observed that increasing the sample size did not change the APB, FSSE and CP values drastically for any of the parameters.

Comparison of mean estimates and APB values between the two models for scenarios S1, S2, S3, and S5 indicate that ignoring correlations lead to a greater rightward bias in the parameter estimate of the protective clothing variable in the third threshold. In scenario S4, ignoring correlations between random parameters resulted in a greater rightward bias in the estimates of mean values of constants for all three thresholds as well the standard deviations of random parameters. Further, coefficients of the gender and intersection dummy variables in the propensity function became biased to the left, where as the coefficient of the age variable became biased to the right.

In the context of precision (FSSE values) in parameter estimates, allowing or ignoring correlations did not have much influence in scenarios S1, S2, S3, and S5, except for the coefficient of the protective clothing variable in the third threshold. Although the parameter estimates in a model with correlation are typically expected to have better precision, probably due to the issues faced during estimation, the precision of parameters for the protective clothing variable was worse when correlations were allowed. In Scenario S4, parameters with greater bias were associated with greater FSSE values, particularly for the mean and standard deviation estimates of the first random threshold. In the context of the coverage probability (CP) values, there were not much differences between models with correlation and models without correlation.

It is important not to get carried away by the above reported loss in accuracy and/or precision of parameter estimates in models without correlation. This is because a single parameter estimate does not offer much interpretation by itself in ordered response models. That is, it is

difficult to use an individual parameter estimate to assess the change in probability of an outcome response when the corresponding independent variable changes. It is the combination of all parameters in the propensity and the threshold functions that determine outcome probabilities. Therefore, in the next section, we compare the predicted percentage shares and marginal effects between the model with correlation and the model without correlation.

### *5.5 Predicted percentage shares and marginal effects*

Figure 3 shows comparisons of predicted percentage shares for different injury severity levels by models with and without correlated random parameters for all the five scenarios. It can be observed that the predicted percentage shares by the model with correlated random parameters and the model without correlation are close in scenarios S1, S2, S3, and S5 but differ slightly for scenario S4. Also, there is no discernible difference between the predictions in scenarios S1 and S5, which have different sample sizes but control for all other factors.

In scenario S4, the model without correlated random parameters slightly overestimates the share of non-incapacitating injuries and underestimates the share of incapacitating injuries (when compared to the model with correlated random parameters). To further examine the differences between the two models in scenario S4, we computed the root mean square error (RMSE) between predicted and actual shares for each injury severity level (for all datasets with converged models) and then averaged across all injury severity levels. The average RMSE values were 1.675 and 1.574 for the model with correlations and model without correlations, respectively. While these RMSE values are close to each other, it is interesting to note that the model without correlation has a slightly lower RMSE than the model with correlation.

Table 5 compares the marginal effects of each variable on different injury severity levels for the two models (the model with correlated random parameters and the model without correlated

random parameters). For scenarios S1, S2, S3, and S5, the marginal effects of all the variables except the protective clothing (on which random parameters were estimated) are almost same for the two models. Even for the protective clothing variable, the marginal effects are not substantially different. In S4, the marginal effects of protective clothing variable differ slightly (but not drastically) for the incapacitating, non-incapacitating and fatal injuries.

Bringing together the findings in this section with those in the previous section, it appears that when correlations are ignored between random parameters in thresholds, the estimates of other parameters are adjusted in such a way that the marginal effects and predicted percentage shares are similar to those when correlation is considered.

## **6. Summary and conclusions**

This study highlights a potential limitation of MGOR models, as applied in most empirical research, that the variances of the random thresholds are implicitly assumed to be in a non-decreasing order. This restriction is not necessary and likely causes difficulty in estimating random parameters in higher order thresholds. To relax this restriction, we evaluated the use of negative correlations between the random parameters as a variance reduction technique. To do so, a simulation-based approach was used, where five different MGOR data scenarios were simulated with 100 datasets in each scenario using a known correlation structure between the random parameters. Two MGOR models were estimated on each simulated dataset – one allowing correlations between random parameters and the other not allowing correlations – and the performance of these two models was evaluated using various evaluation criteria.

Allowing negative correlations helped relax the non-decreasing variance property of MGOR models. However, when negative correlations were considered between random

parameters in thresholds, convergence issues and parameter identification problems were encountered. In addition, for a considerable number of simulated datasets, the correlation parameter estimate was associated with a high standard error. All these issues suggest the difficulty of the maximum simulated likelihood estimation and inference method for MGOR models with correlated random parameters in thresholds.

Comparison of the models that did converge suggests that ignoring correlations leads to an estimation of fewer random parameters in higher order thresholds and results in bias and/or loss of precision for a few parameter estimates. However, when the converged models with correlated random parameters were compared with the corresponding models without correlations, we did not observe significant benefits of accounting for correlations. Neither did the data fit (as measured by likelihood ratio test) improve significantly nor did the predicted shares of different severity levels or the marginal effects differ substantially from those of the models that ignored correlations. In our experimental setup (in all five different scenarios), ignoring correlations lead to an adjustment of other parameter estimates such that overall likelihood values, predicted percentage shares, and the marginal effects were similar to those from the models with correlations. This again suggests potential identifiability issues of MGOR models with correlated random parameters in thresholds.

In summary, the technique of using negative correlation as a variance reduction technique was not effective in our experimental setup, in part due to convergence and identification issues associated with estimating MGORL models that have correlated parameters in thresholds. Therefore, more research is needed for an advanced model structure that can relax the assumption on the order of variance of thresholds in MGOR models (see Paleti and Pinjari, 2018). A relevant question in this context is whether (and to what extent) such assumption is an egregious restriction

to be concerned with. Finally, the issues explored with regard to the MGOR models in this paper add to the discussion of using ordered versus unordered models in the analysis of accident-injury data. Specifically, the tradeoff between the high-degree of model flexibility that an unordered model analysis can provide (such as the standard mixed logit and its various extensions) versus that ability to account for the ordering of alternatives that an ordered response model allows (see Eluru, 2013; Yasmin and Eluru 2013; Mannering and Bhat 2014).

### **Acknowledgements**

The authors wish to thank Rajesh Paleti, a discussion with whom was the genesis of ideas in this paper. The authors acknowledge two anonymous reviewers for their constructive suggestions, which helped in clarifying several points made in the original manuscript. The authors gratefully acknowledge support provided by the Center for Teaching Old Models New Tricks (TOMNET), a University Transportation Center sponsored by the US Department of Transportation through Grant No. 69A3551747116. A faculty startup grant to the second author from the Indian Institute of Science (IISc) is also gratefully acknowledged.

### **References**

- Aitchison, J., Silvey, S., 1957. The generalization of probit analysis to the case of multiple responses. *Biometrika* 44(1-2), 131-140.
- Anarkooli, A., Hosseinpour, M., Kardar, A., 2017. Investigation of factors affecting the injury severity of single-vehicle rollover crashes: a random-effects generalized ordered probit model. *Accident Analysis and Prevention* 106, 399-410.

- Baba, N., 2009. Increased presence of foreign investors and dividend policy of Japanese firms. *Pacific-Basin Finance Journal* 17(2), 163-174.
- Boes, S., Winkelmann, R., 2006. Ordered response models. *Allgemeines Statistisches Archiv* 90(1), 167-181.
- Boes, S., Winkelmann, R., 2010. The effect of income on general life satisfaction and dissatisfaction. *Social Indicators Research* 95(1), 111.
- Castro, M., Paleti, R., Bhat, C., 2013. A spatial generalized ordered response model to examine highway crash injury severity. *Accident Analysis and Prevention* 52, 188-203.
- Chiou, Y.-C., Hwang, C.-C., Chang, C.C., Fu, C., 2013. Modeling two-vehicle crash severity by a bivariate generalized ordered probit approach. *Accident Analysis and Prevention* 51, 175-184.
- Clifton, K., Burnier, C., Akar, G., 2009. Severity of injury resulting from pedestrian-vehicle crashes: What can we learn from examining the built environment?. *Transportation Research Part D* 14(6), 425-436.
- Eluru, N., 2013. Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. *Accident Analysis and Prevention* 5, 1-11.
- Eluru, N., Bhat, C., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. *Accident Analysis and Prevention* 39(5), 1037-1049.
- Eluru, N., Bhat, C., Hensher, D., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention* 40(3), 1033-1054.

- Erdogan, M., Sogut, O., Colak, S., Ayhan, H., Afacan, M., Satilmis, D., 2013. Roles of motorcycle type and protective clothing in motorcycle crash injuries. *Emergency Medicine International* 2013, Article 760205.
- Forbes, J., Habib, M., 2015. Pedestrian injury severity levels in the Halifax regional municipality, Nova Scotia, Canada: Hierarchical ordered probit modeling approach. *Transportation Research Record: Journal of the Transportation Research Board* 2519, 172-178.
- Fountas, G., Anastasopoulos, P Ch., 2017. A random thresholds random parameters hierarchical ordered probit analysis of highway accident injury-severities. *Analytic Methods in Accident Research* 15, 1-16.
- Greene, W., Harris, M., Hollingworth, B., Maitra, P., 2008. A bivariate latent class correlated generalized ordered probit model with an application to modeling observed obesity levels. New York University, New York, NY. Working paper.
- Greene, W., Harris, M., Hollingsworth, B., Weterings, T., 2014. Heterogeneity in ordered choice models: A review with applications to self-assessed health. *Journal of Economic Surveys* 28(1), 109-133.
- Greene, W., Hensher, D., 2010a. *Modeling ordered choices: A primer*: Cambridge University Press, Cambridge, UK.
- Greene, W., Hensher, D., 2010b. Ordered choices and heterogeneity in attribute processing. *Journal of Transport Economics and Policy* 44(3), 331-364.

- Habib, M., Forbes, J., 2014. Modeling bicyclists' injury severity levels in the province of Nova Scotia, Canada using a generalized ordered probit structure. In 93rd Annual Meeting of Transportation Research Board.
- Hosseinpour, M., Yahaya, A., Sadullah, A., 2014. Exploring the effects of roadway characteristics on the frequency and severity of head-on crashes: Case studies from Malaysian Federal Roads. *Accident Analysis and Prevention* 62, 209-222.
- Ierza, J., 1985. Ordinal probit: a generalization. *Communications in Statistics-Theory and Methods* 14(1), 1-11.
- Maddala, G., 1986. *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press, Cambridge, UK.
- Mannering, F., Bhat, C., 2014. Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research* 1, 1-22.
- Mannering, F., Shankar, V., Bhat, C., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research* 11, 1-16.
- McKelvey, R., Zavoina, W., 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology* 4(1), 103-120.
- Mentzakis, E., Moro, M., 2009. The poor, the rich and the happy: Exploring the link between income and subjective well-being. *The Journal of Socio-Economics* 38(1), 147-158.
- Paleti, R., Pinjari, A.R., 2018. A new class of ordered response models with stochastic thresholds. Pennsylvania State University, University Park, PA. Working paper.



- Pudney, S., Shields, M., 2000. Gender, race, pay and promotion in the British nursing profession: estimation of a generalized ordered probit model. *Journal of Applied Econometrics* 15(4), 367-399.
- Quddus, M.A., Noland, R.B., Chin, H.C., 2002. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. *Journal of Safety Research* 33(4), 445-462.
- Rifaat, S.M., Tay, R., De Barros, A., 2012. Severity of motorcycle crashes in Calgary. *Accident Analysis and Prevention* 49, 44-49.
- Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single-and multi-vehicle crashes. *Accident Analysis and Prevention*, 39(5), 955-963.
- Shabanpour, R., Golshani, N., Auld, J., Mohammadian, A., 2017. Willingness-to-pay for automated vehicles: A random parameters and random thresholds HOPIT model. In *International Choice Modelling Conference 2017*.
- Srinivasan, K., 2002. Injury severity analysis with variable and correlated thresholds: ordered mixed logit formulation. *Transportation Research Record: Journal of the Transportation Research Board* 1784, 132-141.
- Stanley, J., Hensher, D., Stanley, J., Currie, G., Greene, W., Vella-Brodrick, D., 2011. Social exclusion and the value of mobility. *Journal of Transport Economics and Policy* 45(2), 197-222.
- Washington, S., Karlaftis, M., Mannering, F., 2011. *Statistical and econometric methods for transportation data analysis*. Second Edition, Chapman and Hall/CRC, Boca Raton, FL.

- Xin, C., Guo, R., Wang, Z., Lu, Q., Lin, P. S., 2017. The effects of neighborhood characteristics and the built environment on pedestrian injury severity: A random parameters generalized ordered probability model with heterogeneity in means and variances. *Analytic Methods in Accident Research* 16, 117-132.
- Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. *Accident Analysis and Prevention* 59, 506-521.
- Yasmin, S., Eluru, N., Bhat, C., Tay, R., 2014. A latent segmentation based generalized ordered logit model to examine factors influencing driver injury severity. *Analytic Methods in Accident Research* 1, 23-38.
- Yasmin, S., Eluru, N., Pinjari, A., 2015a. Analyzing the continuum of fatal crashes: A generalized ordered approach. *Analytic Methods in Accident Research* 7, 1-15.
- Yasmin, S., Eluru, N., Pinjari, A., 2015b. Pooling data from fatality analysis reporting system (FARS) and generalized estimates system (GES) to explore the continuum of injury severity spectrum. *Accident Analysis and Prevention* 84, 112-127.
- Yasmin, S., Eluru, N., Ukkusuri, S., 2014. Alternative ordered response frameworks for examining pedestrian injury severity in New York City. *Journal of Transportation Safety and Security* 6(4), 275-300.
- Zou, W., Wang, X., Zhang, D., 2017. Truck crash severity in New York city: an investigation of the spatial and the time of day effects. *Accident Analysis and Prevention* 99, 249-261.

## Appendix A

### Computation of variance of thresholds in MGOR models

Let  $VAR(\cdot)$ ,  $COV(\cdot)$  and  $E(\cdot)$  represent the variance, covariance and expected value of random variables. Let the first 3 thresholds be

$$\Psi_{n1} = 0,$$

$$\Psi_{n2} = \exp(\alpha_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2}),$$

$$\Psi_{n3} = \exp(\alpha_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2}) + \exp(\alpha_3 \mathbf{U}_{n3} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3}),$$

In the presence of normally distributed random parameters, variance of the third threshold is

$$VAR(\Psi_{n3}) = VAR(\exp(\alpha_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2})) + VAR(\exp(\alpha_3 \mathbf{U}_{n3} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})) + C_{23},$$

where,

$$\begin{aligned} C_{23} &= 2COV(\exp(\alpha_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2}), \exp(\alpha_3 \mathbf{U}_{n3} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})) \\ &= 2[E(\exp(\alpha_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2}) \times \exp(\alpha_3 \mathbf{U}_{n3} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3}))] - E(\exp(\alpha_2 \mathbf{U}_{n2} + \boldsymbol{\theta}_{n2} \mathbf{V}_{n2})) \\ &\quad \times E[\exp(\alpha_3 \mathbf{U}_{n3} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})] \\ &= 2 \times \exp(\alpha_2 \mathbf{U}_{n2} + \alpha_3 \mathbf{U}_{n3}) [E(\exp(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})) - E(\exp(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2})) \times \\ &\quad E(\exp(\boldsymbol{\theta}_{n3} \mathbf{V}_{n3}))] \end{aligned}$$

If the mean and variance of normally distributed random variable  $X$  are  $\mu$  and  $\sigma^2$ , then the expected value of  $\exp(X)$  is  $\exp(\mu + \frac{\sigma^2}{2})$ .

Therefore,

$$\begin{aligned} C_{23} &= 2 \exp(\alpha_2 \mathbf{U}_{n2} + \alpha_3 \mathbf{U}_{n3} + E(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})) \times \\ &\quad \left[ \exp\left(\frac{VAR(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})}{2}\right) - \exp\left(\frac{VAR(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2}) + VAR(\boldsymbol{\theta}_{n3} \mathbf{V}_{n3})}{2}\right) \right] \\ &= 2 \exp(\alpha_2 \mathbf{U}_{n2} + \alpha_3 \mathbf{U}_{n3} + E(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2} + \boldsymbol{\theta}_{n3} \mathbf{V}_{n3}) + \frac{(Var(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2}) + Var(\boldsymbol{\theta}_{n3} \mathbf{V}_{n3}))}{2}) \times \\ &\quad [\exp(COV(\boldsymbol{\theta}_{n2} \mathbf{V}_{n2}, \boldsymbol{\theta}_{n3} \mathbf{V}_{n3})) - 1] \end{aligned}$$

Table 1: Summary of empirical studies that used generalized ordered response models in accident research.

Study	Abbreviation(s) of considered model(s)	Ordered outcome response representation	Estimation of random parameters (RP)		Major findings/ contributions from methodological stand point
			Propensity function	Thresholds	
Srinivasan (2002)	Standard Ordered Response Logit (SORL) and Ordered Mixed Logit (OML)	Traffic crash injury severity – four category variable	<ul style="list-style-type: none"> <li>Propensity function specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>Each threshold in the OML model was expressed as a linear function of explanatory variables, and RP with correlations between them were allowed over the constants.</li> <li>All the correlated random parameters were found to be statistically significant.</li> <li>Interestingly, the variances of thresholds were 0.013, 0.937, and 0.0026 and did not follow any order.</li> </ul>	<ul style="list-style-type: none"> <li>OML model provided a better fit for the observed crash data than SORL model.</li> <li>Prediction capability of OML model was significantly better than the SORL model.</li> </ul>
Eluru et al. (2008)	SORL and Mixed Generalized Ordered Response Logit (MGORL)	Pedestrian and bicyclist injury severity - four category variable	<ul style="list-style-type: none"> <li>Propensity function specification allowed for the estimation of RP in MGORL model.</li> <li>No RP were found to be statistically significant.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification allowed for the estimation of RP in MGORL model.</li> <li>No RP were found to be statistically significant.</li> </ul>	<ul style="list-style-type: none"> <li>SORL model estimation resulted in inconsistent estimates for several variables.</li> <li>MGORL model provided better statistical fit over SORL model.</li> </ul>
Clifton et al. (2009)	Ordered Mixed Probit (OMP)	Pedestrian injury severity - three category variable	<ul style="list-style-type: none"> <li>Propensity function specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>Incorporating built environment characteristics and environmental conditions significantly improved the explanatory power of OMP model.</li> </ul>
Chiou et al. (2013)	Bivariate Generalized Ordered Response Probit (BGORP) and Bivariate Standard Ordered Response Probit (BSORP)	Injury severity of the drivers in a two vehicle crash - four category variable	<ul style="list-style-type: none"> <li>Propensity function specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>BGORP model performed significantly better than the BSORP model in terms of goodness-of-fit indices.</li> <li>BGORP model had better prediction accuracy than the BSORP model.</li> </ul>
Castro et al. (2013)	Spatial Random Parameters Generalized Ordered Response Probit (SRP-GORP) and Standard Ordered Response Probit (SORP)	Injury severity of highway crashes - four category variable	<ul style="list-style-type: none"> <li>Propensity function specification allowed for the estimation of RP in SRP-GORP model.</li> <li>No RP were found to be statistically significant.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification did not allow for the estimation of random parameters.</li> </ul>	<ul style="list-style-type: none"> <li>SRP-GORP model provided statistically better data fit than the Standard Ordered Response Probit (SORP).</li> <li>Predicted shares of different injury severity levels from SRP-GORP model were closer to the actual shares as compared to SORP.</li> </ul>

Table 1: (Continued) Summary of empirical studies that used generalized ordered response models in accident research.

Study	Abbreviation(s) of considered model(s)	Ordered outcome response representation	Estimation of random parameters (RP)		Major findings/ contributions from methodological stand point
			Propensity function	Thresholds	
Eluru (2013)	GORL, SORL, and Multinomial Logit (MNL)	Four alternatives ordered variable	<ul style="list-style-type: none"> <li>Propensity function specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>SORL model was found to be restrictive as compared to MNL model for analyzing an ordered response outcome.</li> <li>GORL model can act as a true ordered equivalent of MNL model.</li> <li>Irrespective of aggregate shares, GORL model performed well as compared to the MNL model.</li> </ul>
Yasmin and Eluru (2013)	SORL, GORL, MGORL, MNL, Nested Logit (NL), and Mixed Multinomial Logit (MMNL)	Passenger vehicle injury severity - four category variable	<ul style="list-style-type: none"> <li>Propensity function specification allowed for the estimation of RP in MGORL model.</li> <li>Three random parameters were found to be statistically significant, and corresponding variables were                             <ol style="list-style-type: none"> <li>1. Restrained system use – Unrestrained (base: restrained)</li> <li>2. Airbag deployment – deployed (base: not deployed)</li> <li>3. Collision location: Intersection (base: non-intersection).</li> </ol> </li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification allowed for the estimation of random parameters in MGORL model.</li> <li>Two random parameters were found to be statistically significant and were in the thresholds demarcating                             <ol style="list-style-type: none"> <li>1. second and third injury severity levels                                     <ol style="list-style-type: none"> <li>a. Vehicle rolled over</li> </ol> </li> <li>2. third and fourth injury severity levels                                     <ol style="list-style-type: none"> <li>a. Collision with stationary object (base: another moving object).</li> </ol> </li> </ol> </li> </ul>	<ul style="list-style-type: none"> <li>Elasticities obtained using the under-reported sample were incorrect in both the MGORL and MMNL models.</li> <li>Both the MMNL and MGORL models had similar prediction results at the aggregate and disaggregate levels.</li> </ul>
Yasmin et al. (2014)	SORL, GORL, and Latent Segmentation based Standard Ordered Response Logit (LS-SORL)	Pedestrian injury severity - three category variable	<ul style="list-style-type: none"> <li>Propensity function specification did not allow for the estimation of RP.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification did not allow for the estimation of random parameters.</li> </ul>	<ul style="list-style-type: none"> <li>GORL and LS-SORL models provided better data fit as compared to the SORL model.</li> <li>Also, LS-SORL model provided better data fit than GORL.</li> <li>In the model validation, GORL model performed marginally better than LS-SORL.</li> </ul>

Table 1: (Continued) Summary of empirical studies that used generalized ordered response models in accident research.

Study	Abbreviation(s) of considered model(s)	Ordered outcome response representation	Estimation of random parameters (RP)		Major findings/ contributions from methodological stand point
			Propensity function	Thresholds	
Yasmin et al. (2014)	Latent Segmentation based Generalized Ordered Response Logit(LS-GORL) and LS-SORL	Driver injury severity – three category variable	• Propensity function specification did not allow the estimation of RP.	• Threshold specification did not allow for the estimation of RP.	• At an aggregate level, LS-GORL model performed well as compared to LS-SORL on validation sample.
Hosseinpour et al. (2014)	Random Effects Ordered Mixed Probit (REOMP), Ordered Mixed Probit (OMP) and SORP	Head on crash severity injury severity - four category variable	• Propensity function specification in REOMP model allowed for the estimation of random effects parameter on the constant and was found to be statistically significant.	• Threshold specification did not allow for the estimation of random parameters.	• REOMP model was found to be statistically better than the OMP and SORP models in terms of data fit.
Habib and Forbes (2014)	OMP with Non-Linear Thresholds specification <sup>9</sup> (OMPnLT) and SORP	Bicyclist injury severity - five category variable	• Propensity function specification did not allow for the estimation of RP.	• Threshold specification did not allow for the estimation of random parameters.	• OMPnLT model with neighborhood and land use characteristics was found to be statistically better than OMPnLT model without such characteristics and SORP model.
Yasmin et al. (2015a)	Mixed Generalized Ordered Logit Response Model (MGORL)	Severity of fatal injury - seven category variable obtained using the survival time in a fatal crash	• Propensity function specification allowed for the estimation of RP.  • Two random parameters were found to be statistically significant, and corresponding variables were  1. Previous record of other harmful motor vehicle convictions  2. Speed limit above 50 mph (base: speed limit < 26mph).	• Threshold specification allowed for the estimation of random parameters.  • No random parameters were found to be statistically significant.	• Endogeneity on the outcome variable due to emergency medical service (EMS) response time variable was addressed using a 2 stage model comprising MGORL for the fatality timeline and regression equation for the EMS response time.
Yasmin et al. (2015b)	Generalized Ordered Response Logit Model (GORL)	Passenger vehicle driver injury severity - eleven category variable	• Propensity function specification did not allow for the estimation of RP.	• Threshold specification did not allow for the estimation of random parameters.	• A simple approach was developed to combine information from Fatality Analysis Reporting System (FARS) data and Generalized Estimates System (GES) data.
Forbes and Habib (2015)	OMPnLT and SORP	Pedestrian injury severity - five category variable	• Propensity function specification did not allow for the estimation of RP.	• Threshold specification did not allow for the estimation of random parameters.	• OMPnLT model performed well as compared to SORP model in terms of model fit.

<sup>9</sup> Similar to ordered mixed response model structure, each threshold is expressed as an exponential function of linear function of covariates.

Table 1: (Continued) Summary of empirical studies that used generalized ordered response models in accident research.

Study	Abbreviation(s) of considered model(s)	Ordered outcome response representation	Estimation of random parameters (RP)		Major findings/ contributions from methodological stand point
			Propensity function	Thresholds	
Fountas and Anastasopoulos (2017)	Mixed Generalized Ordered Response Probit (MGORP), Generalized Ordered Response Probit (GORP), SORP, and Random Parameters SORP (RPSORP)	Single vehicle crash injury severity - four category variable	<ul style="list-style-type: none"> <li>Propensity function specification allowed for the estimation of RP in MGORP and RPSORP models.</li> <li>Seven random parameters were found to be statistically significant in both MGORP and RPSORP models, and corresponding indicator variables were                             <ol style="list-style-type: none"> <li>Presence of vertical curve with the curve length greater than 400 feet</li> <li>Average annual daily traffic per lane greater than 8500 vehicles.</li> <li>Driving under the influence of alcohol or drugs</li> <li>Vehicle crashed due to out of control</li> <li>Vehicle exceeded a reasonable safe level or speed limit</li> <li>Vehicle was travelling straight at the time of crash</li> <li>Pedestrian was involved in the crash.</li> </ol> </li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification allowed for the estimation of random parameters on the constants in the thresholds in MGORP model.</li> <li>Both constants in the thresholds were statistically significant at 95% confidence level.</li> </ul>	<ul style="list-style-type: none"> <li>MGORP model was found to be statistically better than the GORP, SORP, and RPSORP models in terms of data fit.</li> <li>MGORP models had better forecasting accuracy as compared to its model counterparts.</li> </ul>
Anarkooli et al. (2017)	REOMP, OMP, Random Effects Ordered Response Probit (REORP), SORP, MNL, and MMNL	Single vehicle rollover crash severity - three category variable	<ul style="list-style-type: none"> <li>Propensity function specification allowed for the estimation of random effects parameter on the constant in REOMP model and was found to be statistically significant.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification did not allow for the estimation of random parameters.</li> </ul>	<ul style="list-style-type: none"> <li>REOMP model was found to be statistically better than the OMP, REORP, SORP, MNL and MMNL models in terms of model fit.</li> </ul>
Zou et al. (2017)	SRP-GORP and RPSORP	Single-vehicle and multi-vehicle truck crash injury severity - four category variable	<ul style="list-style-type: none"> <li>Propensity function specification allowed for the estimation of RP in both SRP-GORP and RPSORP models.</li> <li>One random parameter was found to be statistically significant in RPSORP model for single vehicle crash severity and corresponding variable is                             <ol style="list-style-type: none"> <li>Truck registered weight.</li> </ol> </li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification did not allow for the estimation of random parameters.</li> </ul>	<ul style="list-style-type: none"> <li>Spatial dependency and temporal effects have significant effect on the single vehicle and multi-vehicle truck crash severity.</li> </ul>

Table 1: (Continued) Summary of empirical studies that used generalized ordered response models in accident research.

Study	Abbreviation(s) of considered model(s)	Ordered outcome response representation	Estimation of random parameters (RP)		Major findings/ contributions from methodological stand point
			Propensity function	Thresholds	
Xin et al. (2017)	MGORP with Heterogeneity in Means and Variances (MGORPHMV), MGORP, GORP and ORP	Pedestrian injury severity - four category variable	<ul style="list-style-type: none"> <li>Propensity function specification allowed for the estimation of RP in MGORPHMV and MGORP models.</li> <li>Two random parameters were found to be statistically significant in both MGORPHMV and MGORP models, and corresponding variables were                             <ol style="list-style-type: none"> <li>Elderly pedestrian indicator</li> <li>Very elderly pedestrian indicator.</li> </ol> </li> <li>Moreover, the random parameter on elderly pedestrian indicator had significant heterogeneity in both means and variance.</li> </ul>	<ul style="list-style-type: none"> <li>Threshold specification allowed for the estimation of random parameters in both MGORPHMV and MGORP models.</li> <li>No random parameters were found to be statistically significant.</li> </ul>	<ul style="list-style-type: none"> <li>The order of statistical superiority (high to low) of models in terms of data fit is MGORPHMV, MGORP, GORP and ORP model.</li> </ul>



Table 2: Summary of different scenarios<sup>a</sup> simulated for motorcyclist injury severity.

Scenario number <sup>b</sup>	Scenario detail	Scenario description	Sample size	Average percentage shares across simulated datasets			
				No injury	Non-incapacitating injury	Incapacitating injury	Fatal injury
S1	$y_n^* = 0.1 \times age_n - 0.3 \times male_n - 0.75 \times intersection_n + \varepsilon_n$ $\Psi_{1n} = 0.2$ $\Psi_{2n} = 0.2 + \exp(0.25 + \theta_{n2} \times PC_n)$ $\Psi_{3n} = 0.2 + \exp(0.25 + \theta_{n2} \times PC_n) + \exp(0.75 + \theta_{n3} \times PC_n)$ $\theta_{n2} = N(0.5, 0.75), \theta_{n3} = N(-0.5, 0.75), \text{ and } \rho_{23} = -0.7$	Greater share for higher ordered outcome	5,000	4.4	18.2	27.1	50.3
S2	$y_n^* = 0.1 \times age_n - 0.3 \times male_n - 0.75 \times intersection_n + \varepsilon_n$ $\Psi_{1n} = 3.5$ $\Psi_{2n} = 3.5 + \exp(0.25 + \theta_{n2} \times PC_n)$ $\Psi_{3n} = 3.5 + \exp(0.25 + \theta_{n2} \times PC_n) + \exp(0.75 + \theta_{n3} \times PC_n)$ $\theta_{n2} = N(0.5, 0.75), \theta_{n3} = N(-0.5, 0.75), \text{ and } \rho_{23} = -0.7$	Greater share for lower ordered outcome	5,000	48.2	27.4	17.8	6.6
S3	$y_n^* = 0.1 \times age_n - 0.3 \times male_n - 0.75 \times intersection_n + \varepsilon_n$ $\Psi_{1n} = 2.1$ $\Psi_{2n} = 2.1 + \exp(0.06 + \theta_{n2} \times PC_n)$ $\Psi_{3n} = 2.1 + \exp(0.06 + \theta_{n2} \times PC_n) + \exp(0.56 + \theta_{n3} \times PC_n)$ $\theta_{n2} = N(0.5, 0.75), \theta_{n3} = N(-0.5, 0.75), \text{ and } \rho_{23} = -0.7$	Approximately equal shares for all outcomes	5,000	24.7	25.9	24.8	24.6
S4	$y_n^* = 0.1 \times age_n - 0.3 \times male_n - 0.75 \times intersection_n + \varepsilon_n$ $\Psi_{1n} = \alpha_{1n}$ $\Psi_{2n} = \alpha_{1n} + \exp(\alpha_{2n} + 0.5 \times PC_n)$ $\Psi_{3n} = \alpha_{1n} + \exp(\alpha_{2n} + 0.5 \times PC_n) + \exp(0.75 - 0.5 \times PC_n)$ $\alpha_{1n} = N(3.5, 1.75), \alpha_{2n} = N(0.25, 0.75), \text{ and } \rho_{12} = -0.7$	Greater share for lower ordered outcome	5,000	48.2	28.4	15.8	7.6
S5	$y_n^* = 0.1 \times age_n - 0.3 \times male_n - 0.75 \times intersection_n + \varepsilon_n$ $\Psi_{1n} = 0.2$ $\Psi_{2n} = 0.2 + \exp(0.25 + \theta_{n2} \times PC_n)$ $\Psi_{3n} = 0.2 + \exp(0.25 + \theta_{n2} \times PC_n) + \exp(0.75 + \theta_{n3} \times PC_n)$ $\theta_{n2} = N(0.5, 0.75), \theta_{n3} = N(-0.5, 0.75), \text{ and } \rho_{23} = -0.7$	Greater share for higher ordered outcome	10,000	4.4	18.2	27.1	50.3

<sup>a</sup> For each of the five scenarios, a total of 100 datasets were simulated.

<sup>b</sup> See text for complete scenario-number definitions.

Table 3: Summary of results from simulation experiments.

Scenario number <sup>a</sup>	Scenario description	Number of datasets with converged models	Number of datasets with significantly improved log-likelihood <sup>b</sup>
S1	Greater share for higher ordered outcome	91	6
S2	Greater share for lower ordered outcome	86	4
S3	Approximately equal shares for all outcomes	94	3
S4	Greater share for lower ordered outcome	82	0

<sup>a</sup> See text for complete scenario-number definitions.

<sup>b</sup> A likelihood ratio test was carried out between the models with and without correlation between the random parameters at 95% confidence level, with 2 degrees of freedom.

Table 4: Evaluation of estimated parameters in the presence and absence of correlations between random parameters.<sup>a</sup>

Parameters		Performance metrics <sup>b</sup>	Scenario-S1		Scenario-S2		Scenario-S3		Scenario-S4		Scenario-S5	
			With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations
Propensity function	Age	True value	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
		Mean estimate	0.1	0.101	0.1	0.1	0.1	0.1	0.112	0.28	0.1	0.101
		APB	0	1	0	0	0	0	12	180	0	1
		FSSE	0.004	0.004	0.002	0.002	0.003	0.003	0.053	0.549	0.003	0.003
		CP	0.967	0.912	0.976	0.988	0.957	0.936	0.931	1	0.966	0.933
	Gender	True value	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
		Mean estimate	-0.312	-0.313	-0.31	-0.309	-0.308	-0.308	-0.345	-0.774	-0.309	-0.311
		APB	4	4.33	3.33	3	2.67	2.67	15	158	3	3.67
		FSSE	0.065	0.066	0.057	0.055	0.059	0.059	0.193	1.219	0.046	0.047
		CP	0.978	0.978	0.964	0.952	0.947	0.957	0.958	1	0.944	0.944
	Intersection indicator	True value	-0.75	-0.75	-0.75	-0.75	-0.75	-0.75	-0.75	-0.75	-0.75	-0.75
		Mean estimate	-0.756	-0.758	-0.751	-0.749	-0.761	-0.76	-0.846	-2.122	-0.752	-0.755
		APB	0.8	1.07	0.13	0.13	1.47	1.33	12.8	182.93	0.27	0.67
		FSSE	0.067	0.067	0.05	0.05	0.054	0.053	0.397	4.219	0.045	0.045
		CP	0.967	0.978	0.988	0.988	0.957	0.947	0.972	1	0.978	0.966
First threshold	Constant	True value	0.2	0.2	3.5	3.5	2.1	2.1	3.5 (1.75)	3.5 (1.75)	0.2	0.2
		Mean estimate	0.191	0.183	3.512	3.504	2.09	2.086	3.904 (2.028)	9.846 (6.016)	0.193	0.179
		APB	4.5	8.5	0.34	0.11	0.48	0.67	11.54 (15.89)	181.31 (243.77)	3.5	10.5
		FSSE	0.132	0.129	0.108	0.106	0.104	0.104	1.811 (1.454)	19.53 (13.01)	0.104	0.102
		CP	0.934	0.956	0.976	0.964	0.968	0.957	0.944 (0.958)	1 (1)	0.933	0.944

<sup>a</sup> Values in parentheses represent the standard deviation of random parameters.

<sup>b</sup> APB = Absolute Percentage Bias; FSSE = Finite Sample Standard Error; CP = Coverage Probability.

<sup>c</sup> Parameter estimates are statistically insignificant (in at least 50% of datasets) at 95% confidence level.

Table 4: (Continued) Evaluation of estimated parameters in the presence and absence of correlations between random parameters.<sup>a</sup>

Parameters		Performance metrics <sup>b</sup>	Scenario-S1		Scenario-S2		Scenario-S3		Scenario-S4		Scenario-S5	
			With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations
Second threshold	Constant	True value	0.25	0.25	0.25	0.25	0.06	0.06	0.25 (0.75)	0.25(0.75)	0.25	0.25
		Mean estimate	0.247	0.247	0.255	0.254	0.071	0.07	0.277 (0.78) <sup>c</sup>	1.053 (-)	0.251	0.251
		APB	1.2	1.2	2	1.6	18.33	16.67	10.8 (4)	321.2 (-)	0.4	0.4
		FSSE	0.055	0.055	0.037	0.037	0.039	0.038	0.375 (1.605)	1.147 (-)	0.038	0.038
		CP	0.945	0.945	0.94	0.929	0.968	0.968	0.986 (0.856)	1 (-)	0.978	0.978
	Protective clothing	True value	0.5 (0.75)	0.5 (0.75)	0.5 (0.75)	0.5 (0.75)	0.5 (0.75)	0.5 (0.75)	0.5	0.5	0.5 (0.75)	0.5 (0.75)
		Mean estimate	0.487 (0.77)	0.488 (0.77)	0.503 (0.79)	0.502 (0.77)	0.486 (0.77)	0.488 (0.77)	0.526	0.437	0.495 (0.75)	0.494 (0.75)
		APB	2.6 (3.2)	2.4 (2.93)	0.6 (5.33)	0.4 (2.8)	2.8 (3.07)	2.4 (2.4)	5.2	12.6	1 (0.53)	1.2 (0.67)
		FSSE	0.07 (0.082)	0.069 (0.08)	0.06 (0.197)	0.06 (0.189)	0.05 (0.096)	0.05 (0.093)	0.12	0.075	0.05 (0.061)	0.05 (0.062)
		CP	0.96 (0.989)	0.96 (0.978)	0.94 (0.952)	0.94 (0.94)	0.97 (0.989)	0.97 (0.979)	0.875	0.653	0.97 (0.955)	0.98 (0.944)
Third threshold	Constant	True value	0.75	0.75	0.75	0.75	0.56	0.56	0.75	0.75	0.75	0.75
		Mean estimate	0.755	0.755	0.756	0.755	0.561	0.561	0.807	1.282	0.753	0.754
		APB	0.67	0.67	0.8	0.67	0.18	0.18	7.6	70.93	0.4	0.53
		FSSE	0.029	0.029	0.038	0.038	0.03	0.03	0.298	0.958	0.022	0.022
		CP	0.967	0.967	0.976	0.976	0.957	0.957	0.917	1	0.966	0.978
	Protective clothing	True value	-0.5(0.75)	-0.5(0.75)	-0.5(0.75)	-0.5(0.75)	-0.5(0.75)	-0.5(0.75)	-0.5	-0.5	-0.5(0.75)	-0.5(0.75)
		Mean estimate	-0.53 (0.78)	-0.175 (-)	-0.43 <sup>c</sup> (0.85) <sup>c</sup>	-0.17 (-)	-0.48 <sup>c</sup> (0.75) <sup>c</sup>	-0.17 (-)	-0.501	-0.514	-0.42 <sup>c</sup> (0.56) <sup>c</sup>	-0.18 (-)
		APB	5.2 (4.13)	65 (-)	14.2 (12.8)	66.2 (-)	3 (0.27)	65.6 (-)	0.2	2.8	15 (24.93)	64.2 (-)
		FSSE	0.19 (0.30)	0.33 (-)	0.17 (0.25)	0.341 (-)	0.124 (0.18)	0.337 (-)	0.083	0.083	0.19 (0.41)	0.328 (-)
		CP	0.956 (0.98)	0.978 (-)	1 (1)	0.94 (-)	1 (1)	0.979 (-)	0.903	1	0.843 (0.86)	0.944 (-)

<sup>a</sup> Values in parentheses represent the standard deviation of random parameters.

<sup>b</sup> APB = Absolute Percentage Bias; FSSE = Finite Sample Standard Error; CP = Coverage Probability.

<sup>c</sup> Parameter estimates are statistically insignificant (in at least 50% of datasets) at 95% confidence level.

Table 4: (Continued) Evaluation of estimated parameters in the presence and absence of correlations between random parameters.<sup>a</sup>

Parameters	Performance metrics <sup>b</sup>	Scenario-S1		Scenario-S2		Scenario-S3		Scenario-S4		Scenario-S5	
		With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations
Correlation term	True value	-0.7	-	-0.7	-	-0.7	-	-0.7	-	-0.7	-
	Mean estimate	-0.689	-	-0.632 <sup>c</sup>	-	-0.652 <sup>c</sup>	-	-0.76 <sup>c</sup>	-	-0.566	-
	APB	1.57	-	9.71	-	6.86	-	8.57	-	19.14	-
	FSSE	0.198	-	0.265	-	0.208	-	0.178	-	0.351	-
	CP	0.956	-	1	-	1	-	0.931	-	0.933	-

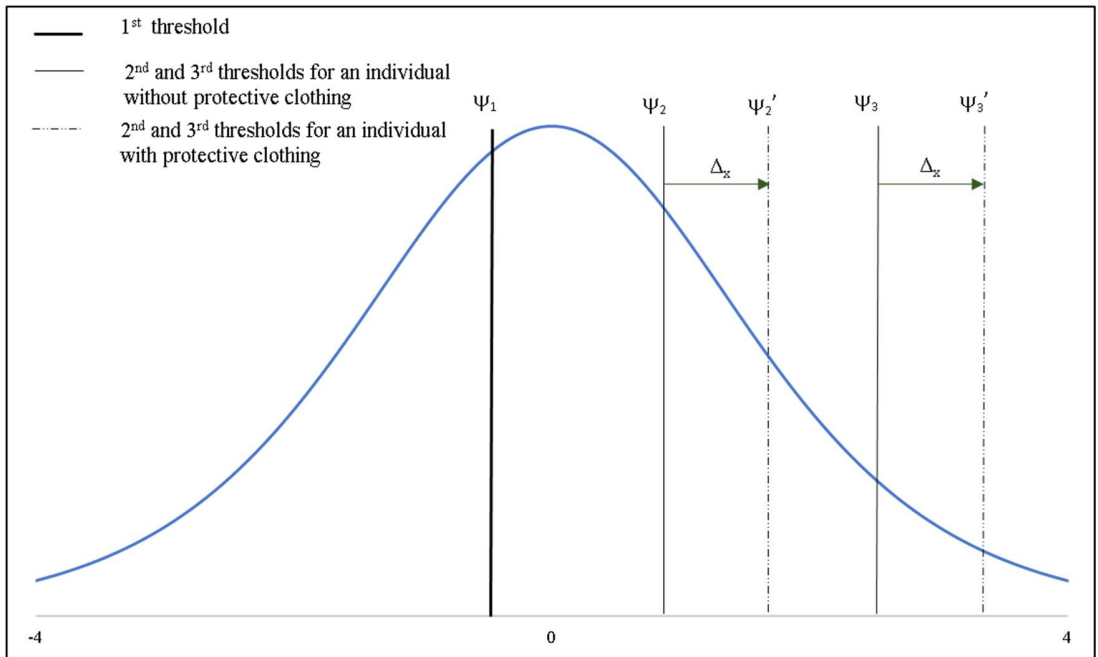
<sup>a</sup> Values in parentheses represent the standard deviation of random parameters.

<sup>b</sup> APB = Absolute Percentage Bias; FSSE = Finite Sample Standard Error; CP = Coverage Probability.

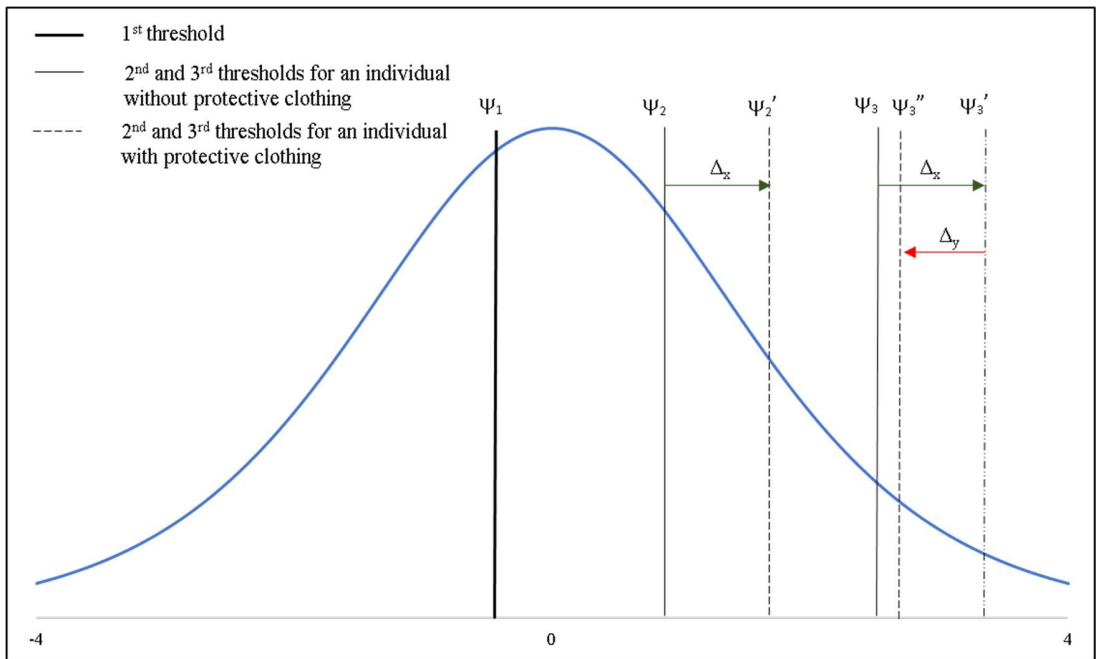
<sup>c</sup> Parameter estimates are statistically insignificant (in at least 50% of datasets) at 95% confidence level.

Table 5: Comparison of marginal effects for the estimated models with and without correlation between random parameters.

Scenario	Variable description	Age		Gender		Intersection indicator		Protective clothing indicator	
	Model	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations	With correlations	Without correlations
S1	No injury	-0.004	-0.004	0.024	0.024	0.058	0.058	0	0
	Non-incapacitating injury	-0.007	-0.007	0.041	0.041	0.101	0.101	0.397	0.396
	Incapacitating injury	-0.006	-0.006	0.035	0.034	0.087	0.085	-0.145	-0.13
	Fatal Injury	0.016	0.016	-0.1	-0.1	-0.245	-0.244	-0.252	-0.266
S2	No injury	-0.018	-0.018	0.112	0.112	0.273	0.273	0	0
	Non-incapacitating injury	0.005	0.005	-0.033	-0.033	-0.08	-0.08	0.215	0.215
	Incapacitating injury	0.008	0.008	-0.047	-0.047	-0.116	-0.116	-0.19	-0.187
	Fatal Injury	0.005	0.005	-0.032	-0.032	-0.077	-0.076	-0.024	-0.028
S3	No injury	-0.014	-0.014	0.089	0.089	0.221	0.221	0	0
	Non-incapacitating injury	-0.002	-0.002	0.011	0.011	0.028	0.029	0.329	0.33
	Incapacitating injury	0.003	0.003	-0.019	-0.02	-0.049	-0.049	-0.213	-0.204
	Fatal Injury	0.013	0.013	-0.081	-0.081	-0.201	-0.2	-0.117	-0.125
S4	No injury	-0.0183	-0.0205	0.1137	0.1274	0.2803	0.3138	0	0
	Non-incapacitating injury	0.0047	0.0093	-0.0308	-0.0596	-0.0762	-0.1486	0.2297	0.2102
	Incapacitating injury	0.0077	0.0082	-0.0475	-0.0503	-0.1178	-0.1243	-0.228	-0.2064
	Fatal Injury	0.006	0.003	-0.0354	-0.0175	-0.0863	-0.0409	-0.0018	-0.0038
S5	No injury	-0.004	-0.004	0.024	0.024	0.057	0.057	0	0
	Non-incapacitating injury	-0.007	-0.007	0.041	0.041	0.101	0.101	0.4	0.399
	Incapacitating injury	-0.006	-0.006	0.036	0.035	0.086	0.089	-0.146	-0.131
	Fatal Injury	0.016	0.016	-0.1	-0.1	-0.244	-0.246	-0.254	-0.267



a) Shift in the thresholds when protective clothing variable enters only the second threshold directly and takes value 1



b) Shift in the thresholds when protective clothing variable enters both the second and third thresholds directly and takes value 1

Figure 1: Influence of protective clothing variable on the shift of thresholds.

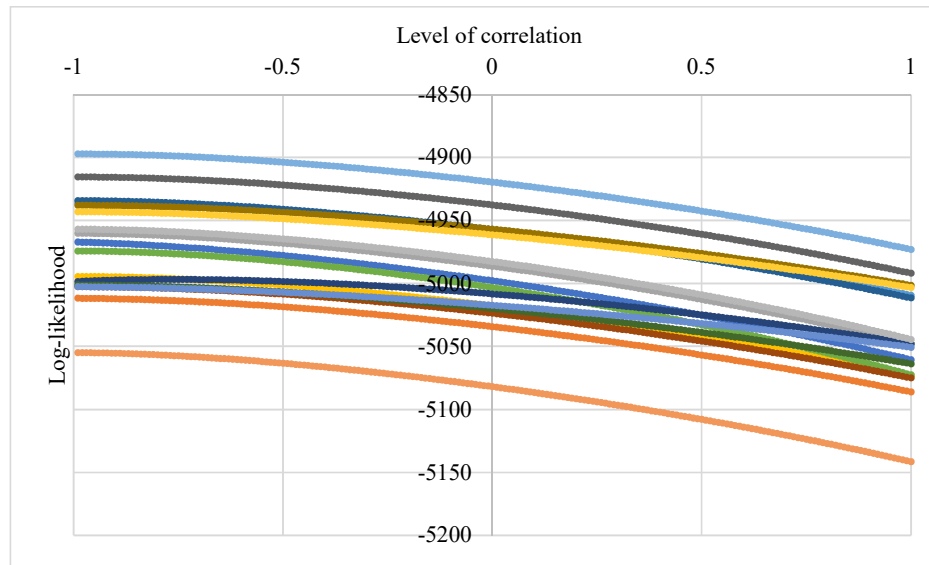


Figure 2: Log-likelihood profiles of unconverged models with respect to the correlation term.



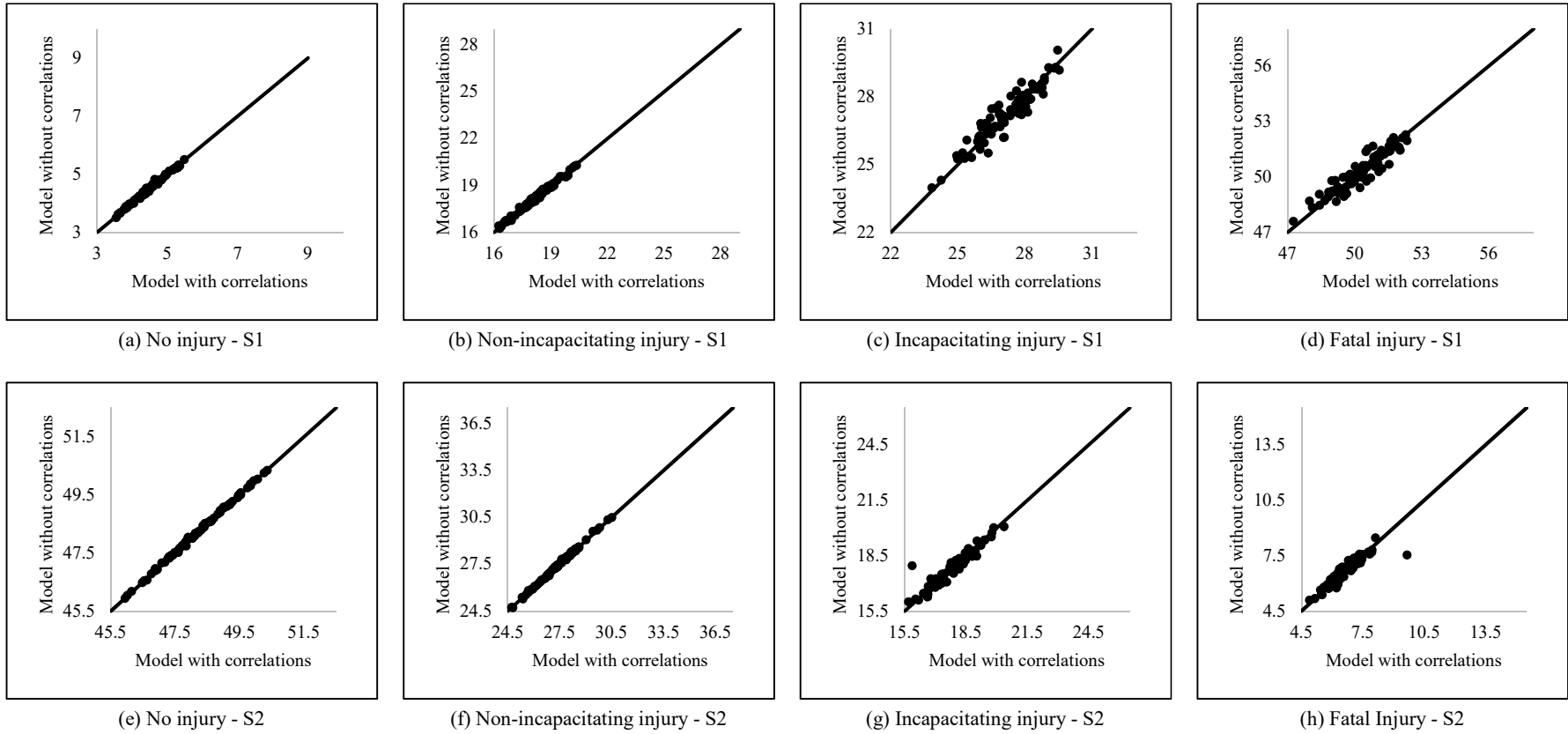


Figure 3: Predicted percentage shares of different injury severity levels by the models with and without correlated random parameters.

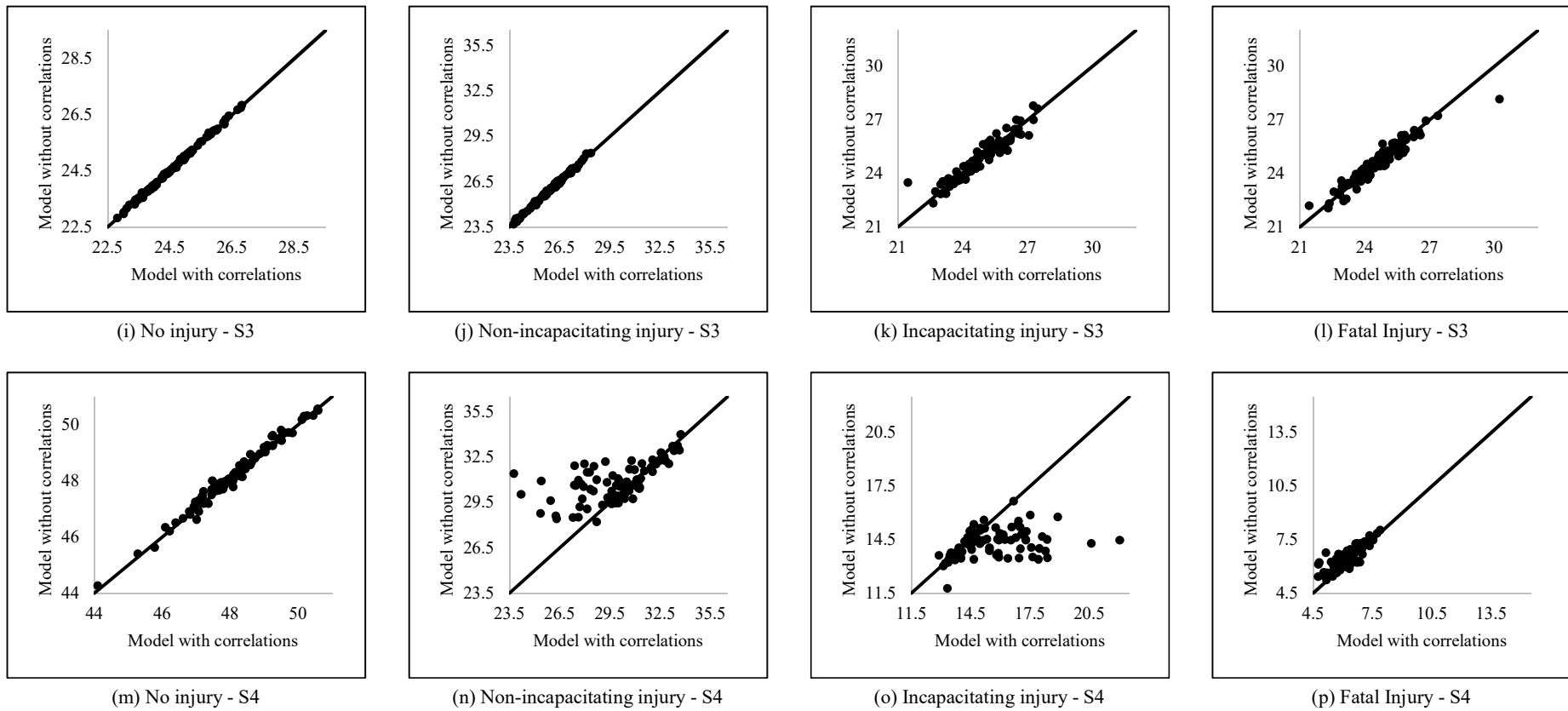
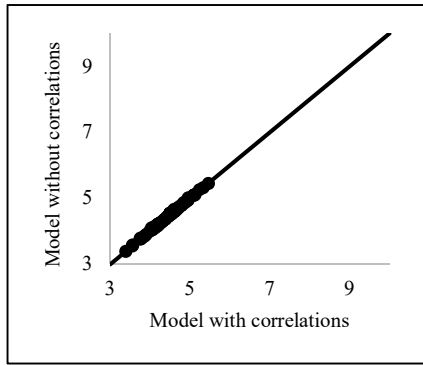
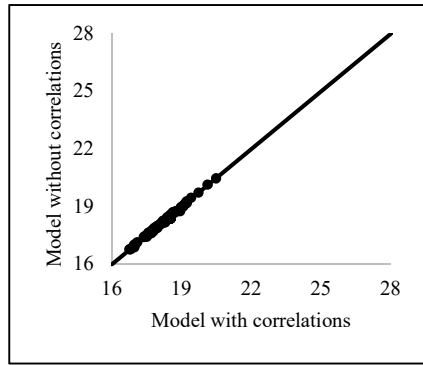


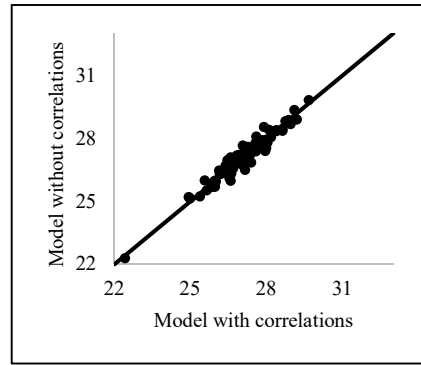
Figure 3: (Continued) Predicted percentage shares of different injury severity levels by the models with and without correlated random parameters.



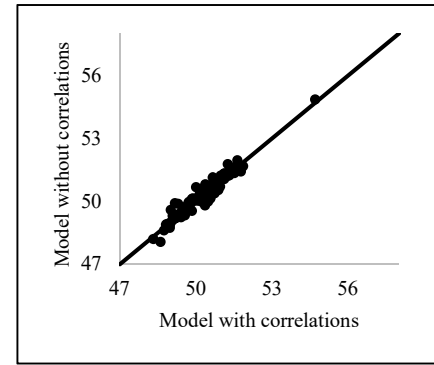
(q) No injury – S5



(r) Non-incapacitating injury – S5



(s) Incapacitating injury – S5



(t) Fatal Injury – S5

Figure 3: (Continued) Predicted percentage shares of different injury severity levels by the models with and without correlated random parameters.