A NOTE ON GENERALIZED ORDERED OUTCOME MODELS

Naveen Eluru*
Associate Professor
Department of Civil, Environmental & Construction Engineering
University of Central Florida
Tel: 1-407-823-4815, Fax: 1-407-823-3315
Email: naveen.eluru@ucf.edu

Shamsunnahar Yasmin
Department of Civil Engineering & Applied Mechanics
McGill University
Suite 483, 817 Sherbrooke St. W., Montréal
Ph: 514 398 6823, Fax: 514 398 7361
Email: shamsunnahar.yasmin@mail.mcgill.ca

* Corresponding author

ABSTRACT

While there is growing application of generalized ordered outcome model variants (widely known as Generalized Ordered Logit (GOL) model and Partial Proportional Odds Logit (PPO) model) in crash injury severity analysis, there are several aspects of these approaches that are not well documented in extant safety literature. The current research note presents the relationship between these two variants of generalized ordered outcome models and elaborates on model interpretation issues. While these variants arise from different mathematical approaches employed to enhance the traditional ordered outcome model, we establish that these are mathematically identical. We also discuss how one can facilitate estimation and interpretation while building on the ordered outcome model estimates – a useful process for practitioners considering upgrading their existing traditional ordered logit/probit injury severity models. Finally, the note presents the differences within GOL and PPO model frameworks, for accommodating the effect of unobserved heterogeneity, referred to as Mixed Generalized Ordered Logit (MGOL) and Mixed Partial Proportional Odds Logit (MPPO) models while also discussing the computational difficulties that may arise in estimating these models.

Keywords: Ordered discrete outcome models, transportation safety, ordinal discrete variables, generalized ordered logit, partial proportional odds model, unobserved heterogeneity

1 INTRODUCTION

Road traffic crash injury severity outcomes are often reported as an ordinal scale variable (such as no injury, minor injury, major injury, and fatal injury). Naturally, road safety researchers have widely employed different econometric approaches within ordered outcome frameworks to evaluate the influence of exogenous factors on ordinal-level crash injury severity outcomes[1] (for example O'Donnell and Connor, 1996; Renski et al., 1999; Yasmin and Eluru, 2013). The ordered outcome models explicitly recognize the inherent ordering within the outcome variable. These models represent the outcome process under consideration using a single latent propensity. Thus, the outcome probabilities are determined by partitioning the uni-dimensional propensity into as many categories as the dependent variable alternatives through a set of thresholds.

Traditional ordered outcome formulations (such as ordered logit/probit) are the primary tools to model the ordinal-level outcomes. But the traditional ordered outcome models impose a restrictive and monotonic impact – most widely referred to as proportional odds or parallel line regression assumption (McCullagh, 1980) – of the exogenous variables on the injury severity alternatives. Imposing such restriction can lead to inconsistent parameter estimation. The recent revival in the ordered regime has addressed this limitation by either allowing the analyst to estimate individual level thresholds as function of exogenous variables or allowing the impact of exogenous variables to vary across alternatives. In fact several generalized ordered frameworks (partial proportional odds model, proportional odds model with partial proportionality constraints and generalized ordered model) relaxing this restrictive assumption have been proposed and employed in extant econometric literature (Fullerton (2009)). More recent research efforts in safety literature following Wang and Abdel-Aty (2008) and Eluru et al. (2008), have encompassed two methodological approaches of generalized ordered outcome formulation that rely on logistic distribution[2] and relax the fixed threshold assumption. These approaches are widely referred to as the Generalized Ordered Logit (GOL) model and Partial Proportional Odds Logit (PPO) model. The generalization of traditional ordered logit (OL) model is achieved in GOL model by allowing the thresholds to be linear functions of observed exogenous variables (as proposed in Terza (1985)). On the other hand, PPO model allows a subset of the explanatory variables to vary across alternatives of interest in generalizing the tradition OL model (as proposed in Peterson and Harrell (1990))[3].

A list of earlier research on crash injury severity analysis that employed these variants of generalized ordered outcome approaches is provided in Table 1. While there is growing application of GOL and PPO models in severity analysis (as evident from table 1), there are still several aspects of these approaches that are not well documented in extant safety literature. It would be beneficial to discuss these variants of generalized ordered outcome models so that researchers and practitioners that consider their application are fully aware of the theoretical and practical similarities and differences between GOL and PPO models. Towards this end, the current research note presents the relationship between these two variants of generalized ordered outcome models

[1] To be sure, researchers have also employed unordered discrete outcome frameworks to study the influence of exogenous variables (see Yasmin and Eluru (2013) and Savolainen et al. (2011) for a detailed description of studies employing different econometric approaches).

[2] In current study, we focus on the logistic error term as it is the most commonly employed model in safety literature; however the same discussion will hold for normal error term assumption too.

[3] In several studies PPO model is also referred to as GOL model (for instance Kaplan and Prato, 2012). However, for ease of discussion we will refer threshold generalization as GOL model and generalization of estimates in propensity as PPO model throughout the note.

and elaborates on model interpretation issues. While these variants arise from different mathematical approaches employed to enhance the traditional ordered outcome model, we establish that these are mathematically identical. To illustrate this we derive the GOL/PPO models from the traditional OL model and show how one can facilitate estimation and interpretation while building on the OL model estimates – a useful process for practitioners considering upgrading their existing traditional ordered logit/probit injury severity models. Finally, the note presents the differences within GOL and PPO model frameworks, referred to as mixed generalized ordered logit (MGOL) and mixed partial proportional odds logit (MPPO) models while also discussing the computational difficulties that may arise in estimating these models.

2 Methodological Framework

In discussing the econometric details of the GOL and PPO models, we begin our discussion with the traditional OL model and build upon the OL framework to arrive at the GOL and PPO models.

2.1 Ordered Logit Model

In the traditional OL model, the discrete injury severity levels (y_i) are assumed to be associated with an underlying continuous, latent variable (y_i^*). This latent variable is typically specified as a linear function as follows

$$y_i^* = \boldsymbol{X}_i\boldsymbol{\beta} + \varepsilon_i, for\ i = 1,2, \dots \dots \dots, N \qquad (1)$$

where,

i $(i = 1,2, \dots \dots \dots, N)$ represents the individual

$\boldsymbol{X}_\text{i}$ is a vector of exogenous variables (excluding a constant)

$\boldsymbol{\beta}$ is a vector of unknown parameters to be estimated

ε is the random disturbance term assumed to be standard logistic

Let j $(j = 1,2, \dots \dots \dots, J)$ and τ_j denote the injury severity levels and the thresholds associated with these severity levels, respectively. These unknown thresholds are assumed to partition the propensity into $J - 1$ intervals. The unobservable latent variable y_i^* is related to the observable ordinal variable y_i by the τ_s with a response mechanism of the following form:

$$y_i = j,\ if\ \tau_{j-1} < \ y_i^* < \tau_j, for\ j = 1,2, \dots \dots \dots, J \qquad (2)$$

In order to ensure the well-defined intervals and natural ordering of observed severity, the thresholds are assumed to be ascending in order, such that $\tau_0 < \tau_1 < \ \dots \dots \dots < \tau_J$ where $\tau_0 = -\infty$ and $\tau_J = +\infty$. The probability expressions take the form:

$$\pi_{ij} = Pr(y_i = j|X_i) = \Lambda\left(\tau_j - \boldsymbol{X_i\beta}\right) - \Lambda\left(\tau_{j-1} - \boldsymbol{X_i\beta}\right) \qquad (3)$$

where $\Lambda(\cdot)$ represents the standard logistic cumulative distribution function and π_{ij} is the probability that individual i sustains an injury severity level j. The standard logistic cumulative distribution function (cdf), $\Lambda(t) = \frac{1}{1+e^{-t}}$; applying the transformation in equation 3, the probability takes the following form:

$$\pi_{ij} = Pr(y_i = j|X_i) = \frac{exp(\tau_j - \boldsymbol{X}_i\boldsymbol{\beta})}{\left(1 + exp(\tau_j - \boldsymbol{X_i}\boldsymbol{\beta})\right)} - \frac{exp(\tau_{j-1} - \boldsymbol{X_i}\boldsymbol{\beta})}{\left(1 + exp(\tau_{j-1} - \boldsymbol{X_i}\boldsymbol{\beta})\right)} \quad (4)$$

In equation 4, the parameter $\boldsymbol{\beta}$ are constrained to be the same across all alternatives – thus resulting in a monotonic impact of the exogenous variables on probability levels. Any enhancement to the systematic component in the ordered outcome system will require addressing the assumption of restricting $\boldsymbol{\beta}$ parameters.

2.2 Generalized Ordered Outcome Approach

The restrictive fixed threshold assumption of traditional ordered outcome models can be relaxed by modifying equation 1: (1) either for τ_j – will result in GOL model (2) or for $\boldsymbol{\beta}$ – will result in PPO model. The mathematical formulations of these models are presented in the following sections.

2.2.1 Generalized Ordered Logit Model

The basic idea of the GOL approach is to represent the threshold parameters as a linear function of exogenous variables (Terza, 1985; Srinivasan, 2002; Eluru et al., 2008). We can employ the following parametric form:

$$\tau_{i,j} = \alpha_j + \boldsymbol{\delta}_j\boldsymbol{Z}_{ij} \quad (5)$$

where, $\boldsymbol{Z}_i$ is a set of exogenous variable (without a constant).
$\boldsymbol{\delta}_j$ is a vector of parameters to be estimated.
With the modification the probability expression of equation 4 takes the following form:

$$\pi_{ij} = \frac{exp(\alpha_j + \boldsymbol{\delta}_j\boldsymbol{Z}_i - \boldsymbol{X}_i\boldsymbol{\beta})}{\left(1 + exp(\alpha_j + \boldsymbol{\delta}_j\boldsymbol{Z}_i - \boldsymbol{X}_i\boldsymbol{\beta})\right)} - \frac{exp(\alpha_{j-1} + \boldsymbol{\delta}_{j-1}\boldsymbol{Z}_i - \boldsymbol{X}_i\boldsymbol{\beta})}{\left(1 + exp(\alpha_{j-1} + \boldsymbol{\delta}_{j-1}\boldsymbol{Z}_i - \boldsymbol{X}_i\boldsymbol{\beta})\right)} \quad (6)$$

It is important to note that the $\boldsymbol{\beta}$ vector is still restricted to be the same in the above model.

2.2.2 Partial Proportional Odds Model

The PPO model is generated from the idea that some of the explanatory variables may meet the proportional odds assumption, while a subset of explanatory variables may not (Peterson and Harrell, 1990). Thus, in PPO model the vector of exogenous variables ($\boldsymbol{X}_i$) in equation 1 is partitioned into two groups – coefficients of variables not-varying across alternatives ($\boldsymbol{X}_{i1}$) and coefficients of variables varying across alternatives ($\boldsymbol{X}_{i2}$). $\boldsymbol{X}_{i1}$ and $\boldsymbol{X}_{i2}$ have no common elements. Thus, the probability expression for PPO model can be expressed as:

$$\pi_{ij} = \frac{exp(\tau_j - \boldsymbol{X}_{i1}\boldsymbol{\beta}_1 - \boldsymbol{X}_{i2}\boldsymbol{\beta}_{2j})}{(1 + exp(\tau_j - \boldsymbol{X}_{i1}\boldsymbol{\beta}_1 - \boldsymbol{X}_{i2}\boldsymbol{\beta}_{2j}))} - \frac{exp(\tau_{j-1} - \boldsymbol{X}_{i1}\boldsymbol{\beta}_1 - \boldsymbol{X}_{i2}\boldsymbol{\beta}_{2(j-1)})}{(1 + exp(\tau_{j-1} - \boldsymbol{X}_{i1}\boldsymbol{\beta}_1 - \boldsymbol{X}_{i2}\boldsymbol{\beta}_{2(j-1)}))} \quad (7)$$

where, $\boldsymbol{\beta}_1$ is the vector of coefficients associated with $\boldsymbol{X}_{i1}$ (the subset of independent variables for which the parallel regression assumption is not violated)

$\boldsymbol{\beta}_{2j}$ is the vector of coefficients associated with $\boldsymbol{X}_{i2}$ (the subset of independent variables for which the parallel regression assumption is violated)

2.2.3 Mathematical Equivalency

If one compares the probability expressions in equations 6 and 7, it is evident that both approaches relaxing the traditional OL model yield exactly the same mathematical model. Specifically, if we set $\alpha_j = \tau_j, \boldsymbol{\beta} = \boldsymbol{\beta}_1$ and $\boldsymbol{\delta}_j = -\boldsymbol{\beta}_{2j}$, identical mathematical structures for both formulations are arrived at. The only *difference* is that parameters corresponding to varying group might offer opposite signs in the two models because in one structure (GOL model) these parameters enter the thresholds and in the other (PPO model) the parameters enter the propensity. Hence, one can establish that the GOL and PPO models are mathematically equivalent and thus the results of one model can be converted into the estimates of the other one.

2.2.4 Model Estimation Procedure

In both the GOL and PPO formulation, the objective is to identify variables for which the parallel line assumption is violated and consider additional parameters for this purpose. The identification process requires careful additional analysis. We outline the procedure for GOL and PPO models for a single exogenous variable.

In the GOL structure, the analyst would estimate a model with only one coefficient in the propensity (a simple ordered model) and another model with the variable appearing in the propensity and thresholds (J-1 parameters). The analyst then would conduct a Wald test at a specific confidence level (95% is most commonly used confidence level) based on the t-statistic to see if all the parameters (single estimate in the simple ordered model or the multiple estimates of the GOL) are statistically significant. If a subset of the parameters are statistically insignificant, the analyst would drop the insignificant parameters and re-estimate the model. After obtaining the best specification between the simple ordered and GOL structure, the analyst can compare model performance using the Log-likelihood ratio (LR) test[4]. For the GOL model, if the propensity parameter and additional parameters are significant then the LR test will definitely outperform the simple ordered model. The LR test is particularly useful if the propensity variable in the GOL is insignificant and only threshold parameters are significant. In this case, a Wald test is not adequate and a LR ratio test is required to identify the superior model.

In the PPO structure, the analyst would employ a similar approach of estimating a simple ordered model and the PPO model with J-1 parameters. A combination of Wald test and LR test will allow the analyst to identify if the parallel line assumption is violated. For PPO model, another

[4] The LR test statistic can be computed as $2[LL_U - LL_R]$, where LL_U and LL_R are the log-likelihood of the unrestricted and the restricted models, respectively.

diagnostic tool, proposed by Brant (1990), is also commonly used for identifying the set of $\boldsymbol{\beta}$s varying across alternatives. This method assesses the non-proportionality not only for the whole model, but also on a detailed variable by variable basis using Wald test. However, LR test is a universal approach and is widely used for testing if the addition of significant variable in threshold (for GOL) and across alternative specific equation (for PPO) has any significant impact on the corresponding log-likelihood value at convergence.

The above procedure needs to be repeated for every exogenous variable. While the approach might seem very burdensome, once the analysts starts model estimation, the testing process is relatively straight-forward and is not different from an unordered multinomial logit model estimation.

2.2.5 Parameter Interpretation

In GOL model, $\boldsymbol{\beta}$ retains the same interpretation as the traditional OL model. However, the δ_j parameters represent shifting of thresholds depending on decision unit specific exogenous variables. Thus, in GOL model when the threshold parameter is positive (negative) the result implies that the threshold is bound to increase (decrease) thus resulting in increase (decrease) in the probability of the alternative to the left of the threshold and decrease (increase) in the probability of the alternative to the right of the threshold.

In PPO model formulation, $\boldsymbol{\beta}_1$ retains the same interpretation as the traditional OL model. However, the $\boldsymbol{\beta}_{2j}$ parameters represent varying impact of exogenous variables across alternatives. The interpretation of $\boldsymbol{\beta}_{2j}$ is similar to unordered logistic regressions *i.e.* a positive coefficients indicate higher likelihood of being in a higher category of the outcome, whereas negative coefficients indicate higher likelihood of being in the current or a lower category of the outcome.

In both mathematical formulations, the analyst can easily interpret the impact of each coefficient. However, when all the possible coefficients for a particular exogenous variable are statistically significant in GOL or PPO structure the net impact of these variables on the ordered outcome variable is generally not straight forward and would require an elasticity or marginal effect computation.

3 UNOBSERVED HETEROGENEITY

In crash injury severity analysis missing or unobserved information is a very common issue. The conventional police/hospital reported crash databases may not include individual specific behavioural or physiological characteristics and vehicle safety equipment specifications for crashes. Due to the possibility of such critical missing information, it is important to incorporate the effect of unobserved attributes within the modeling approach (see for example Srinivasan, 2002; Eluru et al., 2008; Kim et al., 2013). In non-linear models, neglecting the effect of such unobserved heterogeneity can result in inconsistent estimates (Chamberlain, 1980; Bhat, 2001). Hence, it is also important to discuss the variants of generalized ordered outcome models in the context of accommodating unobserved heterogeneity. In the following section, we discuss the potential structure of GOL and PPO model frameworks, referred to as mixed generalized ordered logit (MGOL) and mixed partial proportional odds logit (MPPO) models, in accommodating the effect of unobserved heterogeneity. Further, we also discuss the computational difficulties that may arise in estimating these mixed models.

3.1 Mixed Generalized Ordered Logit Model

The MGOL model accommodates unobserved heterogeneity in the effect of exogenous variable on outcome levels in both the latent propensity function and the threshold functions (Srinivasan 2002; Eluru et al., 2008). Let us assume that $\boldsymbol{\mu}_i$ and $\boldsymbol{\gamma}_{ij}$ are two column vectors representing the unobserved factors specific to individual i in equation 1 and 5, respectively. Thus, conditional on $\boldsymbol{\mu}_i$ and $\boldsymbol{\gamma}_{ij}$, the probability expression for individual i and alternative j in MGOL model take the following form:

$$\begin{aligned}\pi_{ij} &= Pr\left(y_i = j \middle| \boldsymbol{\mu}_i, \boldsymbol{\gamma}_{ij}\right) \\ &= \Lambda[(\tau_{i,j} - (\boldsymbol{\beta} + \boldsymbol{\mu}_i)\boldsymbol{X}_i] - \Lambda[\tau_{i,j-1} - (\boldsymbol{\beta} + \boldsymbol{\mu}_i)\boldsymbol{X}_i]\end{aligned} \tag{8}$$

where $\Lambda(.)$ represents the standard logistic cumulative distribution function and

$$\tau_{i,j} = \alpha_j + (\boldsymbol{\delta}_j + \boldsymbol{\gamma}_{ij})\boldsymbol{Z}_{ij}.$$

The unconditional probability can subsequently be obtained as:

$$P_{ij} = \int_{\boldsymbol{\alpha}_i, \boldsymbol{\gamma}_{ij}} [Pr\left(y_i = j \middle| \boldsymbol{\alpha}_i, \boldsymbol{\gamma}_{ij}\right)] * \boldsymbol{dF}\left(\boldsymbol{\mu}_i, \boldsymbol{\gamma}_{ij}\right)\boldsymbol{d}(\boldsymbol{\mu}_i, \boldsymbol{\gamma}_{ij}) \tag{9}$$

3.2 Mixed Partial Proportional Odds Logit Model

The MPPO model allows the parameters for exogenous variables to vary across individual by accommodating unobserved heterogeneity on the propensity functions for different outcome levels. Let us assume that $\boldsymbol{\vartheta}_i$ and $\boldsymbol{\omega}_{ij}$ are two column vectors representing the unobserved factors specific to individual i for $\boldsymbol{X}_{i1}$ and $\boldsymbol{X}_{i2}$, respectively, in equation 7. Thus, conditional on $\boldsymbol{\vartheta}_i$ and $\boldsymbol{\omega}_{ij}$, the probability expression for individual i and alternative j in MPPO model takes the following form:

$$\begin{aligned}\pi_{ij} &= Pr\left(y_i = j \middle| \boldsymbol{\vartheta}_i, \boldsymbol{\omega}_{ij}\right) \\ &= \Lambda[\tau_j - (\boldsymbol{\beta}_1 + \boldsymbol{\vartheta}_i)\,\boldsymbol{X}_{i1} - \left(\boldsymbol{\beta}_{2,j} + \boldsymbol{\omega}_{ij}\right)\boldsymbol{X}_{i2}] - \Lambda[\tau_{j-1} - (\boldsymbol{\beta}_1 + \boldsymbol{\vartheta}_i)\,\boldsymbol{X}_{i1} \\ &\qquad - \left(\boldsymbol{\beta}_{2,j-1} + \boldsymbol{\omega}_{i,j-1}\right)\boldsymbol{X}_{i2}]\end{aligned} \tag{10}$$

where $\Lambda(.)$ represents the standard logistic cumulative distribution function. The unconditional probability can subsequently be obtained as:

$$P_{ij} = \int_{\boldsymbol{\vartheta}_i, \boldsymbol{\omega}_{ij}} [Pr\left(y_i = j \middle| \boldsymbol{\vartheta}_i, \boldsymbol{\omega}_{ij}\right)] * \boldsymbol{dF}\left(\boldsymbol{\vartheta}_i, \boldsymbol{\omega}_{ij}\right)\boldsymbol{d}(\boldsymbol{\vartheta}_i, \boldsymbol{\omega}_{ij}) \tag{11}$$

The reader would note that the formulation presented here has never been documented in existing literature.

3.3 Computational Difficulties of Mixed Models

In ordered outcome framework, a necessary condition for non-negative probability predictions is that the thresholds remain ordered. However, in the generalized ordered outcome models this requirement is modified. Specifically, to maintain the ordering conditions and thus to ensure the non-negative probability, $\left(\alpha_j + \boldsymbol{\delta}_j \boldsymbol{Z}_i - \boldsymbol{X}_i \boldsymbol{\beta}\right) > \left(\alpha_{j-1} + \boldsymbol{\delta}_{j-1} \boldsymbol{Z}_i - \boldsymbol{X}_i \boldsymbol{\beta}\right)$ condition should maintain in equation 6 of GOL model framework, while $\left(\tau_j - \boldsymbol{X}_{i1}\boldsymbol{\beta}_1 - \boldsymbol{X}_{i2}\boldsymbol{\beta}_{2j}\right) > \left(\tau_{j-1} - \boldsymbol{X}_{i1}\boldsymbol{\beta}_1 - \boldsymbol{X}_{i2}\boldsymbol{\beta}_{2(j-1)}\right)$ condition should maintain in equation 7 of PPO model framework. For, these generalized ordered outcome models with fixed parameters *i.e.* when we ignore the presence of unobserved heterogeneity, the convergence estimates will *rarely* violate the above conditions (theoretically possible). However, if we need to incorporate unobserved heterogeneity within these structures the possibility of the error becomes very critical and might occur often (see Srinivasan, 2002 and Eluru et al., 2008 for a discussion).

These two mathematical formulations of generalized ordered outcome approach employed in literature differ in this aspect. Within GOL model framework, a possible way around to theoretically avoid such potential negative probability issues is to adopt the following non-linear parameterization of the thresholds as a function of exogenous variables, as proposed in Eluru et al. (2008):

$$\tau_{i,j} = \tau_{i,j-1} + exp((\boldsymbol{\delta}_j + \boldsymbol{\gamma}_{ij})\, \boldsymbol{Z}_{ij}) \tag{12}$$

The above formulation capitalizes on the fact that the thresholds are parameterized and hence ensuring they are ordered will ensure that the probabilities remain positive. Thus, it is computationally feasible to estimate the model as presented in equation 9 while employing the parameterization of equation 12. In fact, several previous studies in existing safety literature have employed this approach in accommodating the effect of unobserved heterogeneity within GOL framework (Yasmin and Eluru, 2013).

On the other hand, within PPO model framework accommodating such parameterization is far from straight forward because any parameter in $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_{2j}$ of equation 10 could potentially affect the nature of the probability expression thus not allowing for non-negative probabilities for all realisations of $\boldsymbol{\vartheta}_i$ and $\boldsymbol{\omega}_{ij}$ *i.e.* for some realization of $\boldsymbol{\vartheta}_i$ and $\boldsymbol{\omega}_{ij}$, it is theoretically possible that $\tau_j - (\boldsymbol{\beta}_1 + \boldsymbol{\vartheta}_i)\, \boldsymbol{X}_{i1} - \left(\boldsymbol{\beta}_{2,j} + \boldsymbol{\omega}_{ij}\right)\boldsymbol{X}_{i2} < \tau_{j-1} - (\boldsymbol{\beta}_1 + \boldsymbol{\vartheta}_i)\, \boldsymbol{X}_{i1} - \left(\boldsymbol{\beta}_{2,j-1} + \boldsymbol{\omega}_{i,j-1}\right)\boldsymbol{X}_{i2}$. This would lead to a negative probability value. To be sure, this does not mean that we cannot accommodate unobserved heterogeneity in the MPPO model. To ensure non-negative probability values, while generating the realisations of $\boldsymbol{\vartheta}_i$ and $\boldsymbol{\omega}_{ij}$, an initial screening procedure to ensure that $\tau_j - (\boldsymbol{\beta}_1 + \boldsymbol{\vartheta}_i)\, \boldsymbol{X}_{i1} - \left(\boldsymbol{\beta}_{2,j} + \boldsymbol{\omega}_{ij}\right)\boldsymbol{X}_{i2} < \tau_{j-1} - (\boldsymbol{\beta}_1 + \boldsymbol{\vartheta}_i)\, \boldsymbol{X}_{i1} - \left(\boldsymbol{\beta}_{2,j-1} + \boldsymbol{\omega}_{i,j-1}\right)\boldsymbol{X}_{i2}$ is not violated can be added to the simulation procedure. Of course, this would require access to a larger number of random draws compared to the MGOL model estimation process. Given that the screening process will add substantial computation burden for each iteration of the maximum simulated likelihood process, this could yield to substantial increase in model convergence run times. In fact, a similar procedure was considered in Srinivasan, 2002 for a variant of the MGOL model. In summary, the MPPO model requires additional computation to estimate the model whereas the MGOL approach offers a direct parameterization approach while offering the same flexibility. In fact, this could potentially be a reason why no study in literature has employed the MPPO model.

4 CONCLUSIONS

With increasing application of generalized ordered outcome model variants (such as generalized ordered logit (GOL) and partial proportional odds logit (PPO) models in severity analysis, there are several aspects of these model approaches that are not well documented in extant safety literature. The research note discussed the equivalency between the two variants of generalized ordered outcome models. We also presented how one can facilitate enhanced interpretations while building on the ordered outcome model estimates – a useful process for practitioners considering upgrading their existing traditional ordered logit/probit injury severity models. While the variants offer equivalent mathematical and estimation approaches for traditional specification, in the presence of unobserved heterogeneity the GOL model variant offers enhanced model specification and estimation framework due to its parameterization structure. To be sure, the mixed PPO model can also be accommodated to account for unobserved heterogeneity but it would require additional steps to ensure consistent model estimation.

ACKNOWLEDEGMENTS

The corresponding author would like to acknowledge Prof. Chandra Bhat's contribution with whom the original Generalized Ordered Logit formulation discussed was developed.

REFERENCES

Abegaz, T., Berhane, Y., Worku, A., Assrat, A., Assefa, A., 2014. Effects of excessive speeding and falling asleep while driving on crash injury severity in Ethiopia: A generalized ordered logit model analysis. Accident Analysis and Prevention 71, 15-21.

Anowar, S., Yasmin, S., Tay, R., 2014. Factors influencing the severity of intersection crashes in Bangladesh. Asian transport studies 3 (2), 143-154.

Bhat, C. R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transportation Research Part B 35 (7), 677-693.

Brant, R., 1990. Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics 46 (4), 1171-1178.

Castro, M., Paleti, R., Bhat, C. R., 2013. A spatial generalized ordered response model to examine highway crash injury severity. Accident Analysis and Prevention 52, 188-203.

Chamberlain, G., 1980. Analysis of Covariance with Qualitative Data. The Review of Economic Studies 47 (1), 1980, 225-238.

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.

Eluru, N., 2013. Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. Accident Analysis and Prevention 55, 1-11.

Eluru, N., Bhat, C. R., Hensher, D. A., 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.

Fullerton A. S., 2009. A conceptual framework for ordered logistic regression models. Sociological methods and research 38 (2), 306-47.

Kaplan, S., Prato, C. G., 2012. Risk factors associated with bus accident severity in the United States: A generalized ordered logit model. Journal of Safety Research 43 (3), 171-180.

Kim, J.-K., Ulfarsson, G. F., Kim, S., Shankar, V. N., 2013. Driver-injury severity in single-vehicle crashes in California: a mixed logit analysis of heterogeneity due to age and gender. Accident Analysis and Prevention 50, 1073-1081.

Kweon, Y. J., Lee, J., 2010. Potential risk of using General Estimates System: Bicycle safety. Accident Analysis and Prevention 42 (6), 1712-1717.

McCullagh, P., 1980. Regression models for ordinal data. Journal of the Royal Statistical Society, 42 (2), 109-142.

Mooradian, J., Ivan, J. N., Ravishanker, N., Hu, S., 2013. Analysis of driver and passenger crash injury severity using partial proportional odds models. Accident Analysis and Prevention 58, 53-58.

O'Donnell, C. J., Connor, D. H., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Accident Analysis and Prevention 28 (6), 739-753.

Peterson, B., arrell, F. E., Jr., 1990. Partial Proportional Odds Models for Ordinal Response Variables. Journal of the Royal Statistical Society 39 (2), 205-217.

Quddus, M. A., Wang, C., Ison, S. G., 2010. Road traffic congestion and crash severity: Econometric analysis using ordered response models. Journal of Transportation Engineering 136 (5), 424-435.

Renski, H., Khattak, A., Council, F., 1999. Effect of Speed Limit Increases on Crash Injury Severity: Analysis of Single-Vehicle Crashes on North Carolina Interstate Highways. Transportation Research Record 1665, 100-108.

Rifaat, S. M., Tay, R., de Barros, A., 2012a. Urban Street Pattern and Pedestrian Traffic Safety. Journal of Urban Design 17 (3), 337-352.

Rifaat, S. M., Tay, R., De Barros, A., 2012b. Severity of motorcycle crashes in Calgary. Accident Analysis and Prevention 49, 44-49.

Sasidharan, L., Menéndez, M., 2014. Partial proportional odds model - An alternate choice for analyzing pedestrian crash injury severities. Accident Analysis and Prevention 72, 330-340.

Savolainen, P. T., Mannering, F. L., Lord, D., Quddus, M. A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention 43 (5), 1666-1676.

Srinivasan, K., 2002. Injury Severity Analysis with Variable and Correlated Thresholds: Ordered Mixed Logit Formulation. Transportation Research Record 1784, 132-141.

Terza, J. V., 1985. Ordinal probit: a generalization. Communications in Statistics-Theory and Methods 14 (1), 1-11.

Wang, X., Abdel-Aty, M., 2008. Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models. Accident Analysis and Prevention 40 (5), 1674-1682.

Yasmin, S., Anowar, S., Tay, R., 2012. Effects of Drivers' Actions on Severity of Emergency Vehicle Collisions. Transportation Research Record 2318, 90-97.

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., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Accident Analysis and Prevention 59, 506-521.

Yasmin, S., Eluru, N., Ukkusuri, S. V., 2014a. Alternative Ordered Response Frameworks for Examining Pedestrian Injury Severity in New York City. Journal of Transportation Safety and Security 6 (4), 275-300.

Yasmin, S., Eluru, N., Bhat, C. R., and Tay, R., 2014b. A latent segmentation based generalized ordered logit model to examine factors influencing driver injury severity. Analytic Methods in Accident Research 1, 23-38.

TABLE 1 Existing Injury Severity Studies Employing Generalized Ordered Outcome Approaches

Alternate Approaches	Previous Studies
Partial Proportional Odds Model	Wang and Abdel-Aty (2008); Kweon and Lee (2010); Quddus et al. (2010); Kaplan and Prato (2012); Rifaat et al. (2012a); Rifaat et al. (2012b); Yasmin et al. (2012); Mooradian et al. (2013); Abegaz et al. (2014); Anowar et al. (2014); Sasidharan and Menéndez (2014)
Generalized Ordered Models	Srinivasan (2002); Eluru et al. (2008); Chiou et al. (2013); Eluru (2013); Castro et al. (2013); Yasmin and Eluru (2013); Yasmin et al. (2014a); Yasmin et al. (2014b)