Analytic Methods in Accident Research 20 (2018) 46–67

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

journal homepage: www.elsevier.com/locate/amar

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

Suryaprasanna Kumar Balusu[a], Abdul Rawoof Pinjari[b,*], Fred L. Mannering[c], Naveen Eluru[d]

[a] *Department of Civil and Environmental Engineering, University of South Florida, 4202 E. Fowler Avenue, ENB 118, Tampa, FL 33620, United States*
[b] *Department of Civil Engineering, Centre for Infrastructure, Sustainable Transportation, and Urban Planning (CiSTUP) , Indian Institute of Science (IISc), Bangalore 560012, India*
[c] *College of Engineering, Professor of Civil and Environmental Engineering, University of South Florida, 4202 E. Fowler Avenue, ENC 3506, Tampa, FL 33620, United States*
[d] *Department of Civil, Environmental and Construction Engineering, University of Central Florida, 12800 Pegasus Drive, Room 301D, Orlando, FL 32816, United States*

A R T I C L E I N F O

Keywords:
Generalized ordered response models
Random thresholds
Variance reduction
Negative correlation
Parameter identification

A B S T R A C T

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.

* Corresponding author.
E-mail addresses: balusu@mail.usf.edu (S.K. Balusu), abdul@iisc.ac.in (A.R. Pinjari), flm@usf.edu (F.L. Mannering), naveen.eluru@ucf.edu (N. Eluru).

https://doi.org/10.1016/j.amar.2018.09.003
Received 24 January 2018; Accepted 24 September 2018
2213-6657/

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. (2014b) 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.[1]

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 implement. The latter approach is a non-trivial[2] 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.

[1] 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).

[2] 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.