Analytic Methods in Accident Research 11 (2016) 33–47

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

Analysis of occupant injury severity in winter weather crashes: A fully Bayesian multivariate approach

Mohammad Saad Shaheed [a], Konstantina Gkritza [b,*], Alicia L. Carriquiry [c], Shauna L. Hallmark [d]

[a] Traffic Safety Specialist, Arizona Department of Transportation, 1615 West Jackson Street, Phoenix, AZ 85002, United States
[b] Associate Professor, Lyles School of Civil Engineering, G167B, Purdue University, West Lafayette, IN 47907, United States
[c] Distinguished Professor, Department of Statistics, 1121 C Snedecor Hall, Iowa State University, Ames, IA 50011, United States
[d] Professor, Civil, Construction, & Environmental Engineering, Iowa State University Ames, IA 50011, United States

ARTICLE INFO

Article history:
Received 10 February 2015
Received in revised form
17 June 2016
Accepted 17 June 2016
Available online 12 July 2016

Keywords:
Multivariate discrete data modeling
Full Bayesian (FB) estimation
Hierarchical analysis
Multivariate error term correlation
Winter weather
Disaggregate crash data
Occupant injury severity

ABSTRACT

Multivariate injury severity models that consider the cross-group heterogeneity in the crash data where individuals or occupants are nested within vehicles and vehicles are nested within crashes are limited in the literature. Most previous studies on crash injury severity were conducted at the crash level ignoring the potential correlation in severity for the vehicles involved in the same crashes or individuals involved in the same vehicles. Ignoring these correlation and dependence effects might result in underestimation of standard errors and erroneous inferences. The objective of this paper is to correctly determine the factors affecting occupant injury severity in winter seasons by addressing the within-crash and between-crash correlation of injury severity. To achieve this, fully Bayesian hierarchical multinomial logit models were developed for estimating occupant injury severity in weather-related crashes, non weather-related crashes, and all crashes. These models were developed using disaggregate crash data with occupants nested within crashes for four winter seasons in Iowa. Significant factors affecting occupant injury severity included factors related to occupants (gender, seating position, occupant trapped status, ejection status, and occupant protection used), as well as crash-level factors (road junction type, first harmful event and major cause of crash). Weather-related variables, such as visibility, pavement and air temperature, were also significant factors in winter weather crashes. Interaction effects involving crash-level variables and occupant-level variables were also found significant. Overall, the model diagnostics suggested significant within-crash correlation in the study dataset justifying the use of a multivariate model specification that addresses multivariate error term correlation issues.

1. Introduction and background

Most of the crash data used for road safety research are of hierarchical nature and belong to structures with several hierarchically-ordered levels. These hierarchical structures could be attributed to the spatial and temporal spread of the data

* Corresponding author.
E-mail addresses: mshaheed@azdot.gov (M.S. Shaheed), nadia@purdue.edu (K. Gkritza), alicia@iastate.edu (A.L. Carriquiry), shallmar@iastate.edu (S.L. Hallmark).

http://dx.doi.org/10.1016/j.amar.2016.06.002
2213-6657/

Table 1.
Descriptive statistics of the variables considered in the models.

Variable	Weather-related crashes	Non weather-related crashes	All crashes
	Percentage (Frequency)	Percentage (Frequency)	Percentage (Frequency)
Occupant injury severity (fatal, incapacitating or non-incapacitating injury/possible injury/no injury/unknown)	8.8/11.5/79.2/0.5	8.0/10.8/80.7/0.6	8.4/11.1/80.0/0.5
Gender (if male = 1, otherwise = 0)	66.1 (2,459)	63.0 (2,591)	64.5 (5,050)
Seating position (if driver = 1, otherwise = 0)	92.7 (3,444)	94.3 (3,878)	93.5 (7,322)
Occupant protection (Used/not used/not reported or unknown)	87.5 (3,252)/2.2 (83)/ 10.3 (382)	71.2 (2,928)/3.1 (128)/ 25.7 (1,056)	78.9 (6,180)/2.7 (211)/18.4 (1,860)
Ejection status (not ejected/ejected/unknown, not reported)	96.8 (3,598)/0.5 (20)/ 2.7 (99)	95.6 (3,932)/0.3 (12)/4.1 (168)	96.2 (7,530)/0.3 (32)/ 3.5 (267)
Occupant trapped status (not trapped/trapped/unknown or not reported)	93.5 (3,475)/3.7 (137)/ 2.8 (105)	92.9 (3,821)/3.0 (125)/ 4.0 (166)	93.2 (7,296)/3.3 (262)/3.4 (271)
Type of roadway junction (if intersection = 1, otherwise = 0)	12.0 (446)	26.6 (1,094)	19.7 (1,540)
Road surface condition (surface condition dry/surface condition icy, wet, snowy or slushy/surface conditions others and not reported)	0.6 (21)/99.2 (3,688)/0.2 (8)	65.4 (2,688)/20.0 (826)/ 14.5 (598)	34.6 (2,709)/57.6 (4,514)/ 7.7 (606)
Air temperature (if below 32 °F = 1, otherwise = 0)	85.0	33.1	57.7
First harmful event (non-collision including overturn, rollover, jackknife/collision with vehicles/collision with non-vehicles including animal, debris, work zone equipment, etc.)	23.2 (864)/51.2 (1,904)/ 25.5 (949)	9.9 (406)/75.8 (3,115)/ 14.4 (591)	16.2 (1,270)/64.1 (5,019)/ 19.6 (1,540)
Major cause (if run-off road = 1, otherwise = 0)	19.3 (717)	11.9 (489)	15.4 (1,206)
Occupant protection used and crashed occurred at an intersection (if yes = 1, otherwise = 0)	10.3 (383)	28.7 (876)	19.7 (1,259)
Road surface condition dry and air temperature greater than 32 °F (if yes = 1, otherwise = 0)	20.8 (773)	23.8 (979)	22.4 (1,752)
Road surface condition dry and pavement temperature greater than 32 °F (if yes = 1, otherwise = 0)	11.9 (397)	46.8 (1,923)	29.6 (2,320)
Weather condition clear and major cause reported too fast for the condition (if yes = 1, otherwise = 0)	2.7 (101)	1.33 (55)	2.00 (156)
Weather condition rain, sleet, hail, mist, snow, fog, wind and major cause reported swerving or evasive action (if yes = 1, otherwise = 0)	12.4 (462)	1.92 (79)	6.9 (541)
Visibility greater than 3 miles and clear weather (if yes = 1, otherwise = 0)	23.2 (862)	27 (1,111)	25.2 (1,973)
Road surface condition wet, icy, snowy, or slushy and visibility up to 6 miles (if yes = 1, otherwise = 0)	19.7 (732)	6.5 (267)	12.8 (999)

or the hierarchical nature of the crash data itself. These two types of hierarchies (spatial and crash) are associated with aggregate and disaggregate crash data, respectively and can be distinguished as geographical and crash hierarchies (Huang and Abdel-Aty, 2010). The analysis of aggregate crash data mainly focuses on the geographical part of the hierarchy, where traffic crashes are nested within sites and sites are nested within geographic region. As shown in Barua et al. (2015), the analysis of aggregate crash data can account for spatial dependence through spatial analyses; temporal effects can be included as well, when considering traffic entity levels along a time horizon. On the other hand, the analysis of disaggregate crash data mainly focuses on individual occupants or vehicles involved in a crash, where individuals or occupants are nested within vehicles and vehicles are nested within crashes. Information regarding vehicles, drivers, or occupants is clustered within the crashes as each vehicle, driver, or occupant observation pertains to one crash only. These two complementary hierarchies can be incorporated into a single framework, as shown in past work (Dupont et al., 2013).

From a methodological standpoint, the two structures described above necessitate the consideration of multivariate models (where multiple dependent variables that are interrelated with each other are modeled at the same time) that accommodate the hierarchy of the crash data and spatial and/or temporal dependencies. Ignoring such dependencies will, in general, result in inefficient and inconsistent parameter estimates (Mannering and Bhat, 2014). Multivariate modeling efforts addressing spatial and temporal correlation have mainly concentrated on the analysis of aggregate crash data and the estimation of crash frequency models. Examples of past studies can be found in Mannering and Bhat (2014) (please refer to Table 1 under spatial and temporal correlation models) and Barua et al. (2014, 2016). Accommodating spatial and/or temporal dependence effects in injury severity levels in crashes has not received much attention to date, but rather the literature is limited to Castro et al. (2013), and two other studies that incorporated spatial dependence in simultaneously modeling crash frequency and severity (Chiou et al., 2014; Chiou and Fu, 2015).

Turning to the analysis of disaggregate crash data for modeling injury severity, it is reasonable to assume that the characteristics of the vehicle in which occupants are traveling affect the probability of injuries of occupants. In this case, injuries within the same vehicle might be similar compared to injuries in different vehicles. Similarly, vehicles/drivers involved in a multi-vehicle crash event might sustain damages/injuries with similar severity or vice versa. Whether positive