Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

Evaluating temporal variability of exogenous variable impacts over 25 years: An application of scaled generalized ordered logit model for driver injury severity

Robert Marcoux, Shamsunnahar Yasmin*, Naveen Eluru, Moshiur Rahman

Department of Civil, Environmental and Construction Engineering, University of Central Florida, United States

HIGHLIGHTS

- We examine changes in the impact of exogenous variables on driver injury severity.
- We consider scaled and mixed versions of generalized ordered logit model.
- We demonstrate our model by using General Estimates System (GES) crash database.
- We consider data for a span of 25-year from 1989 through 2014 in 5-year increments.

ARTICLE INFO

Keywords:
Driver injury severity
Temporal instability
Pseudo-panel
Generalized ordered logit
Mixed model
Scaled model

ABSTRACT

The current study undertakes a unique research effort to quantify the impact of various exogenous factors on crash severity over time. Specifically, we examine if over time, the impact of exogenous variables has changed and if so what is the magnitude of the change. The research contributes to driver injury severity analysis both methodologically and empirically by proposing a framework that addresses the challenges associated with pooled (or pseudo-panel) data. For our analysis, we draw data from the General Estimates System (GES) over a span of twenty-five years. The data is compiled for driver injury severity in single or two vehicle crashes from 1989 through 2014 in 5-year increments (1989, 1994, 1999, 2004, 2009 and 2014). The alternative econometric frameworks considered for the analysis include ordered logit, generalized ordered logit, scaled generalized ordered logit and mixed generalized ordered logit models. A host of comparison metrics are computed to evaluate the performance of these alternative models in examining the pooled data. The model development exercise is conducted with a host of exogenous variables including driver characteristics, vehicle characteristics, roadway attributes, environmental factors, crash characteristics and temporal attributes. The model estimation results are further augmented by performing a detailed policy scenario analysis, probability profile representations and elasticity effects for different driving and situational conditions across different years.

* Corresponding author.
E-mail addresses: BobbyMrcx@Knights.ucf.edu (R. Marcoux), shamsunnahar.yasmin@ucf.edu (S. Yasmin), naveen.eluru@ucf.edu (N. Eluru), moshiur@knights.ucf.edu (M. Rahman).

https://doi.org/10.1016/j.amar.2018.09.001
Received 17 March 2018; Received in revised form 1 September 2018; Accepted 1 September 2018
2213-6657/

1. Introduction

The negative consequences of road traffic crashes have a significant impact on the emotional and financial well-being of the society. In the United States in 2015, roadway crashes were responsible for over 35,000 fatalities, while about 2.35 million individuals were injured or disabled. Financially, road traffic crashes cost the US nearly 230 billion annually. Given the magnitude of consequences of road traffic crashes, it is not surprising that road safety is a well-researched field. Earlier research has examined factors affecting crash occurrence using crash frequency models (see Yasmin and Eluru, 2016 for a review) and crash consequence in the event of a crash using crash severity models (see Yasmin and Eluru, 2013 for a review). The current study contributes to road safety literature by focusing on crash severity models. Earlier literature on severity modeling has identified several factors that significantly influence severity of vehicle occupants involved in traffic crashes including vehicle occupant age, restraint system use, driving under the influence of alcohol or drugs, vehicle age and type (such as sedan, van or pickup truck) of the vehicles involved, and collision type (such as head-on, rear-end and angle).

To be sure, road safety in the US has improved over the years. The number of traffic fatalities have reduced from about 47,000 in 1965 to about 35,000 in 2015 (NHTSA, 2016). The fatality rate per 100 million vehicle miles travelled has dropped from 5.3 to 1.12 during the same period. The improvement in traffic crash associated fatalities can chiefly be attributed to: (1) design and enforcement of several policies such as mandatory seat belt use, vehicle regulations requiring airbags, child rear facing seats, (2) advances in vehicle technology to improve occupant safety and (3) concerted effort dedicated to educational awareness campaigns for different driver age groups to encourage safe driving behavior. However, in recent years (since 2000) instances of increase in the number of traffic fatalities compared to the previous year have occurred multiple times. Thus, in spite of the significant progress made over the years, there is further scope for improving road safety.

The current study undertakes a unique research effort to quantify the impact of various exogenous factors on crash severity over time (see Mannering, 2018 for discussion on temporal instability). Specifically, we examine if over time, the impact of exogenous variables has changed and if so what is the magnitude of the change. For example, the injury severity of a vehicle occupant in a crash that occurred in 1995 was a function of the vehicle safety features available at the time. An observationally identical crash occurring in 2015 is likely to result in either the same level of injury severity or lower. This is an example of how improvement in vehicle technology has affected injury severity. In our study, we systematically attempt to identify exogenous variables that offer time-varying effects and quantify the change in their impact. For this purpose, we draw data from the General Estimates System (GES), over a span of twenty-five years. The data is compiled for driver injury severity in single or two passenger vehicle involved crashes from 1989 through 2014 in 5-year increments (1989, 1994, 1999, 2004, 2009 and 2014). In our analysis, injury severity is classified in four levels as follows: no injury, possible injury, non-incapacitating injury, and incapacitating/fatal injury.

The data compiled is a pooled dataset obtained from stitching together 6 cross-sectional datasets providing us with a pseudo-panel data. Such data pooling of different observations across multiple years offers unique methodological challenges. The modeling methodology should recognize the differences across multiple time points adequately since the outcome process for the observations in a year might be influenced by various observed and unobserved attributes (Anowar et al., 2016; Train, 2009). For illustrative purposes, consider the possibility that airbags were made mandatory in vehicles only in 2000. The data compiled after this time period would possibly experience lower injury severity for crashes relative to the years prior due to the installation of airbags. This is an example of an observed attribute affecting either one or more cross-sections of data in the overall pooled dataset. On the other hand, consider the possibility of a cultural phenomenon that encourages good driving behavior – such as wearing a seatbelt – happened between 1995 and 2000. As a result, more drivers wear seatbelts after 2000, thus potentially reducing injury severity consequences for cross-sections compiled later. However, parsing the true impact of the cultural phenomenon in the model is quite challenging and usually remains hidden or unobserved for a long time. This is an example of how an unobserved effect specific to one-time period or multiple time periods can affect severity outcomes. In our study, we implement modeling approaches that simultaneously accommodate for the influence of observed and unobserved attributes on driver injury severity across multiple time points.

Given the inherent ordering of the data, we adopt a generalized ordered logit (GOL) model kernel structure for severity analysis. The GOL framework relaxes the restrictive assumption of the traditional ordered outcome (Ordered logit/probit) models (monotonic effect of exogenous variables) while simultaneously recognizing the inherent ordering of the injury severity variable – information that unordered model alternatives fail to consider (see Eluru et al., 2008; Eluru and Yasmin, 2015; Yasmin and Eluru, 2013). Further, to incorporate the effect of observed and unobserved temporal effects, we specifically consider two versions of the GOL model – the mixed GOL model and the scaled GOL model. The two variants differ in the way they incorporate the influence of unobserved attributes within the outcome process. The mixed GOL model captures unobserved heterogeneity by allowing the variable effects (including the constant) influencing injury severity to be distributed across the observations, while the scaled model accommodates for common unobserved heterogeneity by allowing the variance of the unobserved component to vary by time period. The two approaches also vary in how they accommodate for heteroscedasticity across observations. We estimate both models and employ data fit comparison metrics to determine the appropriate model structure. The model specification is undertaken to quantify how the impact of exogenous variables has altered over the 25-year period on driver injury severity. The model development exercise is conducted with a host of exogenous variables including driver characteristics, vehicle characteristics, roadway attributes, environmental factors, crash characteristics and temporal attributes. In summary, the current research effort contributes empirically by identifying if the impact of any exogenous factors has varied over time, and if so, quantifying the change in magnitude. Methodologically, the research contributes to driver injury severity analysis by proposing a framework that addresses the challenges associated with pooled (or pseudo-panel) data.

The remainder of the paper is organized as follows. Section 2 provides a brief literature review. Section 3 briefly outlines the