Analytic Methods in Accident Research 12 (2016) 1–17

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models

Ali Behnood [a,*], Fred L. Mannering [b]

[a] Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall, West Lafayette, IN 47907-2051, USA
[b] Department of Civil and Environmental Engineering, University of South Florida, 4202 E Fowler Avenue, ENC 3300, Tampa, FL 33620, USA

ARTICLE INFO

Article history:
Received 9 December 2015
Received in revised form
7 July 2016
Accepted 12 July 2016

Keywords:
Economic recession
Pedestrian
Temporal stability
Transferability
Mixed logit model
Latent class logit model

ABSTRACT

This study explores the differences in pedestrian injury severity in three distinct economic time periods from the recent global recession (the Great Recession): pre-recession, recession, and post-recession. Using data from pedestrian crashes in Chicago, Illinois over an eight-year period, separate time-period models of pedestrian-injury severities (with possible outcomes of severe injury, moderate injury, and minor injury) were estimated using latent-class logit and mixed logit models. Likelihood ratio tests were conducted to examine the overall stability of model estimates across time periods and marginal effects of each explanatory variable were also considered to investigate the temporal stability of the effect of individual parameter estimates on pedestrian injury-severity probabilities. A wide range of variables potentially affecting injury severities was considered including time, location, and severity of crashes, as well as data on roadway and environmental conditions, pedestrian characteristics, and crash characteristics. Our findings show significant temporal instability, which likely results from a combination of the economic recession and the long-term evolution of the influence of factors that affect pedestrian-injury severity. Understanding and explicitly modeling the evolution of driver and pedestrian behavior is a promising direction for future research, but this would unfortunately require far more extensive data than is currently available in traditional safety databases.

1. Introduction

Pedestrians are among the most vulnerable road users because they are not protected by safety features, vehicle crash-absorbing features, body-armored clothing (such as those used by some motorcyclists and in motorcycle racing), or helmets. In 2013, there were 4735 pedestrian fatalities and about 66,000 pedestrian injuries in traffic crashes in the U.S. (NHTSA, 2015). Overall traffic-related fatalities in the U.S. have generally shown a decreasing trend over the years (particularly during the Great Recession, officially defined as occurring from December 2007 until June 2009), and pedestrian fatalities have, for the most part, followed this trend. However, pedestrian traffic fatalities have become an increasing percentage of total traffic fatalities rising from 10.91% of total fatalities in 2004 to 14.47% of total fatalities in 2013 (NHTSA, 2015; FARS, 2015). With pedestrian fatalities representing a larger and larger percentage of overall traffic fatalities, there is a clear need to study pedestrian-injury severity.

* Corresponding author.
E-mail addresses: abehnood@purdue.edu (A. Behnood), flm@usf.edu (F.L. Mannering).

http://dx.doi.org/10.1016/j.amar.2016.07.002

Pedestrian safety, and specifically the study of pedestrian-injury severity, has been the emphasis of a number of important research efforts over the years (Lee and Abdel-Aty, 2005; Kim et al., 2008, 2010; Alluri et al., 2013; Venkataraman et al., 2013; Haleem et al., 2015; Sasidharan et al., 2015). All of these studies have provided valuable insight into the factors that determine pedestrian-injury severity. However, the issue of temporal stability in pedestrian-injury severity models (whether the estimated model parameters are stable over time) has not really been addressed to date. There are at least two reasons to suspect that temporal instability may be an issue in pedestrian injury severity models. First, the fact that pedestrian fatalities are making up a larger and larger percentage of overall traffic fatalities in recent years may suggest the possibility of temporal instability in injury severities (as noted above). However, one could argue that the increasing percentages may be simply due to increases in pedestrian volume and exposure (distance walked), both of which are virtually impossible to measure with existing data sources. The second, and perhaps more compelling reason, is that there may be fundamental recession-induced shifts in driver and pedestrian behavior that result in temporally unstable model-parameter estimates. It is a well-established fact that recessions generally result in a decline in the total number of motor-vehicle fatalities as well as the number of motor vehicle fatalities per distance driven (Ruhm, 2000; Peterman, 2013). To understand why this is the case, recent work by Maheshri and Winston (2016) studied reasons why U.S. overall traffic fatality rates (vehicle occupant and pedestrian fatalities) declined in the most recent world-wide recession (the Great Recession). They found that the recession had differential impacts on road users, with more crash-prone drivers driving less and safer drivers driving more, which tended to result in lower overall crash rates during the recession. This finding is supported by the empirical work of Behnood and Mannering (2015) who found that crash-injury severity models exhibited significant temporal instability over the 2004–2012 time period (which included the time before and after the December 2007 to June 2009 Great Recession). Possible recessionary-induced effects, such as the effect of safer drivers driving more and less-safe drivers driving less as found in Maheshri and Winston (2016), would clearly influence pedestrian injury severities. In addition, there could be changes in pedestrian behavior and exposure during recessionary periods that would affect injury severity. While it is not possible to get access to detailed pedestrian exposure data (as Maheshri and Winston (2016) were able to do with vehicle odometer readings for drivers), we can access detailed pedestrian-injury crash reports and use these reports as a basis for studying possible temporal instability that may have been recession-induced.

To undertake this study of the temporal stability of pedestrian injury severities, data were gathered from reported vehicle-pedestrian crashes in the city of Chicago, Illinois. To study the effect of the 2007–2009 economic recession on pedestrian injury severities, three distinct time periods will be defined: pre-recession, recession, and post-recession. The assessment of these data allows studying the temporal stability of injury severity models and their estimated parameters over time. In addition, two alternate methodological approaches were used to capture unobserved heterogeneity (random parameters and latent class approaches) to study the effect that different methodological approaches may have on temporal instability findings. As will be shown below, we find significant temporal instabilities with regard to pedestrian-injury severities.

2. Methodology

When assessing the potential temporal instability of pedestrian injury severities, it is important that the findings be robust across methodological alternatives. Over the years, crash-related injury severities have been studied by a variety of ordered and unordered discrete outcome approaches including ordered logit/probit models, multinomial logit models, dual-state multinomial logit models, nested logit models, latent-class logit models, mixed (random parameters) logit models, Markov-switching models, and others (Savolainen et al., 2011; Mannering and Bhat, 2014; Mannering et al., 2016). To account for possible unobserved heterogeneity in the data, more recent research has focused on random parameter approaches (Anastasopoulos and Mannering, 2011; Behnood and Mannering, 2015; Eluru et al., 2008; Kim et al., 2013; Milton et al., 2008; Morgan and Mannering, 2011; Venkataraman et al., 2013), latent class models (Behnood et al., 2014; Cerwick et al., 2014; Shaheed and Gkritza, 2014; Yasmin et al., 2014) or combination of both (Xiong and Mannering, 2013) to model the injury severities.

Heterogeneity models, such as those using random parameters and latent classes, are among the most advanced methodological approaches currently used in crash-injury severity analysis.[1] However, random parameters models with mixing distributions (such as the mixed logit) and models with latent classes both have their own drawbacks. Traditional random parameters models require that an assumption be made for the functional form of the mixing distribution to capture unobserved heterogeneity.[2] Latent class models (also referred to as finite mixture models) potentially overcome this limitation by identifying homogenous subgroups of data where distributional assumptions are not made. However, conventional latent class models do not account for the possibility of variation within the identified latent classes (homogenous parameter effects are assumed within each class to simplify model estimation). The within-class fixed parameter assumption may not be valid since the number of identified classes is usually small, thus leaving open the possibility of residual heterogeneity within each class.[3] The possibility of having latent classes with parameters distributed across

[1] Please see Mannering et al. (2016) for a complete discussion of unobserved heterogeneity in accident data and an overview of statistical approaches.

[2] Many econometric software packages provide a wide variety of functional forms for mixing distribution. However, the analyst must still make a decision regarding the appropriate functional form. In many cases, in order to simplify the modeling and interpretation of the estimation results, a commonly used distribution such as the normal distribution is assumed.

[3] In another attempt to find the different classes of data that share common features, several researchers have considered sub-groups of data based on