Analytic Methods in Accident Research 16 (2017) 75–89

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach

Sikai Chen [a,*], Tariq Usman Saeed [a,b], Samuel Labi [a]

[a] Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., W. Lafayette, IN 47907, USA
[b] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT), Cambridge 02139, MA, USA

ARTICLE INFO

Article history:
Received 6 August 2016
Received in revised form 8 September 2017
Accepted 8 September 2017

Keywords:
Rural highway
Unobserved heterogeneity
Accident severity
Peltzman effect
Risk compensation
Risk homeostasis

ABSTRACT

Recent studies have begun to shed more light on the crashes experienced on rural roads by examining the influence of a road's pavement surface condition. In a bid to contribute to this growing body of knowledge and to facilitate comprehensive evaluation of pavement maintenance projects, this paper explores the safety effects of the pavement condition of rural roads. The paper tests the hypotheses that pavement roughness generally has a non-trivial residual impact on safety outcomes and that the magnitude and direction of these impacts differ across road segments. To explore these hypotheses, the paper presents crash frequency models for three levels of crash severity and also across five levels of road surface condition. The developed models use the multivariate random parameters negative binomial specification to account for the unobserved heterogeneity and correlation among the different levels of crash severity. The model results suggest that for pavements in fair or good condition, the surface condition parameter has fixed effects on the crash frequency, irrespective of the crash severity level. However, for pavements in poor condition, the surface condition variable in the crash model has a significant random parameter that is normally distributed. The positive portions of the parameter density function suggest that higher roughness (poorer condition) generally increases the expected crash frequency, likely because drivers may lose control of their vehicles. The negative portions suggest that within that condition range, higher surface roughness is generally associated with a lower expected crash frequency, likely because drivers are generally likely to drive more carefully on pavements in very poor condition (a manifestation of risk-compensation behavior). The developed models can help highway engineers quantify not only the safety benefits of road resurfacing projects but also the safety consequences of worsening road surface conditions arising from delay of pavement maintenance.

1. Introduction

Highway safety research literature is replete with methodological developments and increasing model sophistication over the last several decades; these have spawned improvements in predictive capability and reliability of models (Lord and Mannering, 2010; Mannering and Bhat, 2014). Some of these advancements include estimation of multivariate/bivariate crash count models (Anastasopoulos, 2016; Heydari et al., 2016; Serhiyenko et al., 2016), incorporation of random parameters in crash models (Anastasopoulos and Mannering, 2009; Venkataraman et al., 2014; Coruh et al., 2015; Barua et al.,

* Corresponding author at: Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr, W. Lafayette, IN 47907, USA
E-mail addresses: chen1670@purdue.edu (S. Chen), tusaeed@purdue.edu, tusaeed@mit.edu (T.U. Saeed), labi@purdue.edu (S. Labi).

http://dx.doi.org/10.1016/j.amar.2017.09.001
2213-6657/

2016; Bhat et al., 2017); accounting for spatial/temporal correlation in crash models (Chiou et al., 2014; Chiou and Fu, 2015; Hong et al., 2016) and estimation of latent class, finite mixture, two-state Markov switching, mixed logit and neural network models (Malyshkina et al., 2009; Park and Lord, 2009; Behnood et al., 2014; Behnood and Mannering, 2016; Zeng et al., 2016). These improvements arose from concerted efforts to identify and address a number of complex issues associated with highway crash data and analysis, which includes unobserved heterogeneity, spatial and temporal correlation, and correlated collision types (Lord and Mannering, 2010; Aguero-Valverde, 2013; Mannering and Bhat, 2014; Chiou and Fu, 2015; Anastasopoulos et al., 2016; Hong et al., 2016; Mannering et al., 2016; Amoh-Gyimah et al., 2017; Behnood and Mannering, 2017a; Huang et al., 2017). For example, failure to address any unobserved heterogeneity in the data could result in biased, inefficient, or inconsistent parameter estimates when a traditional fixed-parameter model is specified, could consequently lead to incorrect inferences (Washington et al., 2011). Taking a similar position, Anastasopoulos and Mannering (2016) and Mannering et al. (2016) argued that neglecting any unobserved heterogeneity (thus, assuming implicitly that the effects of observable factors are the same across all observations) may cause model specification problems, biased and inefficient parameters, and ultimately, erroneous predictions.

In this vein, a number of recent research studies have used a variety of statistical tools to account for the unobserved heterogeneity across observations in crash data (Gkritza and Mannering, 2008; Milton et al., 2008; Anastasopoulos and Mannering, 2009, 2011; El-Basyouny and Sayed, 2009a; Dinu and Veeragavan, 2011; Garnowski and Manner, 2011; Ukkusuri et al., 2011; Venkataraman et al., 2011, 2013, 2014; Anastasopoulos et al., 2012a; Wu et al., 2013; Xiong and Mannering, 2013; Behnood et al., 2014; Chen and Tarko, 2014; Russo et al., 2014; El-Basyouny et al., 2014a; Anastasopoulos, 2016; Barua et al., 2016; Sarwar and Anastasopoulos, 2016; Mothafer et al., 2017; Gomes et al., 2017).

The use of random-parameters models to account for unobserved heterogeneity has been prevalent in recent studies (Gkritza and Mannering, 2008; Milton et al., 2008; Anastasopoulos and Mannering, 2009, 2016; Anastasopoulos et al., 2012a; Chen and Tarko, 2014; Russo et al., 2014; Venkataraman et al., 2014; Coruh et al., 2015; Behnood and Mannering, 2015, 2016, 2017b; Sarwar et al., 2017a,b; Seraneeprakarn et al., 2017; Anderson and Hernandez, 2017; Fountas and Anastasopoulos, 2017; Bogue et al., 2017). This is because random parameter models can explicitly account for heterogeneity across observations that are mainly due to unobserved explanatory traffic, environmental, roadway design, vehicle characteristics, driver behavior, and other related factors. This technique allows every estimated parameter in the model to vary across each individual observation according to an analyst-defined continuous distribution and thus requires a parametric distribution. Since it is quite possible that the individual coefficients in the model will follow different distributions, using a parametric approach could potentially cause statistical problems (Mannering et al., 2016). Analysts find that it is often impractical to consider all these factors simultaneously because highway crashes is often the outcome of a combination of factors such as those related to the driver, vehicle, roadway, and natural environment and the road (engineering) and also of the complex interactions among them (Sinha and Labi, 2007; Mannering et al., 2016; Sarwar and Anastasopoulos, 2017).

Furthermore, crash counts are correlated in nature. Therefore, modeling crashes within each crash severity level separately and not accounting for possible correlations between the crashes across the severity levels can be problematic. Using a univariate modeling approach for the correlated crash counts can lead to less precise estimates of the risk factors associated with the different severity levels. In this regard, using multivariate approaches to model crash frequency jointly across the different crash severity levels has gained attention in recent years (Bijleveld, 2005; Ma and Kockelman, 2006; Song et al., 2006; Park and Lord, 2007; Ma et al., 2008; Aguero-Valverde and Jovanis, 2009; Wang et al., 2011; Anastasopoulos et al., 2012b; El-Basyouny et al., 2014a,b; Lee et al., 2015; Abay and Mannering, 2016; Barua et al., 2016; Serhiyenko et al., 2016; Zeng et al., 2016; Heydari et al., 2016, 2017; Sarwar et al., 2017a,b). Multivariate models have been found to adequately account for correlations among the different levels of crash severity; such correlations seem to emerge from a variety of sources including data on traffic collisions that involve multiple occupant injuries from the same crash; see Mannering et al., 2016 for more detailed information. In this example, the occupants would likely suffer different levels of injury (crash severity), but the unobserved factors that influence the injury severity levels would be correlated (Eluru et al., 2010; Abay et al., 2013; Yasmin et al., 2014; Russo et al., 2014; Ma et al., 2017). The unobserved factors may impact the multiple crash counts of different severity levels simultaneously for each roadway segment under consideration (Mannering et al., 2016); therefore, estimating separate univariate models could cause statistical problems. To avoid these problems, this paper uses the concept of random parameters in a multivariate modeling framework to accommodate the unobserved heterogeneity in the correlated crash data.

A number of recent studies found significant variation in the effects of roughness or overall pavement condition on crash frequency and severity (Al-Masaeid et al., 1993; Lamptey, 2004; Labi, 2006; Mayora and Piña, 2009; Anastasopoulos et al., 2012a,b; Li et al., 2013; Buddhavarapu et al., 2013; Anastasopoulos, 2016). A school of thought contends that crash propensity increases as pavement condition improves because a very good pavement condition is generally associated with higher speeds and subsequently higher crash counts. Other researchers maintain that such effects have yet to be proven beyond a reasonable doubt (Harwood et al., 2003), a position that was seemingly corroborated by Agent et al. (2004) whose analysis using Kentucky highway data found no difference in crash frequency before and after pavement resurfacing. Al-Masaeid (1997), on the other hand, indicated that pavement condition, expressed in terms of the international roughness index (IRI) had significant but rather opposite effects on single- and multiple-vehicle crashes. NCHRP Project 17-9(2) reported that resurfacing had a negative effect on safety in certain states and a positive effect in others. Earlier work by Cleveland (1987) seemed rather prescient: the researcher had suggested that the direction of impact was rather ambiguous and that additional research was needed. Drivers tend to behave less carefully on roads with adequate levels of service, where they