Analytic Methods in Accident Research 1 (2014) 86–95

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

Modeling safety of highway work zones with random parameters and random effects models

Erdong Chen *, Andrew P. Tarko [1]

Center for Road Safety, School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907-2051, USA

A R T I C L E I N F O

Article history:
Received 30 May 2013
Received in revised form
30 October 2013
Accepted 30 October 2013
Available online 2 December 2013

Keywords:
Highway safety
Work zone safety
Crash frequency modeling
Random parameters
Negative binomial model

A B S T R A C T

This paper presents an investigation of traffic safety in highway work zones using detailed data obtained from the results of a survey of project engineers and existing datasets. The observations were organized in monthly clusters that correspond to individual work zones; and a two-level random parameters negative binomial model that reflected the structure of the observations was estimated. The safety effects of various work zone design and traffic management features were identified, including lane shift, lane split, and detour, whose safety effects have not been evaluated in past research. This new insight into highway work zone safety was accomplished thanks to the better data acquired and the improved statistical model. A fixed parameters negative binomial model with random effects then was estimated to check its viability as an alternative to the random parameters model when the sample's large size makes estimation of the latter challenging. From a practical standpoint, the marginal effects on crash frequency computed from the model with random effects were quite similar to those computed from the random parameters model. This result indicates that, at least in some cases, convenient fixed parameters models may be a practical alternative to random parameters models. Utilization of an entire sample to estimate these conventional models may further compensate a less advanced model specification. The obtained negative binomial model with random effects has become useful for programming police enforcement in highway work zones in Indiana.

1. Introduction

Statistical models play an important role in acquiring knowledge and attaining a better understanding of highway safety, its factors, and effective countermeasures. Mannering and Bhat (this issue) point out the recent fast advancement in the analytic methodology of safety analysis. They also note the widening gap between the "frontier" models and modeling practice in highway safety. One point must be emphasized though. Not all applications of early models are missed opportunities for applying a better methodology if the selection of a model type considers not only the state-of-the art but also the purpose of developing the model.

An example of safety management comes to mind here. Safety management includes identification of roads that exhibit excessive frequency and severity of crashes; and, typically, only traffic volumes and key road characteristics are available for

* Corresponding author. Tel.: +1 765 494 9821; fax: +1 765 494 0395.
E-mail addresses: chen608@purdue.edu, chenerdong007@gmail.com (E. Chen), tarko@purdue.edu (A.P. Tarko).
[1] Tel.: +1 765 494 5027; fax: +1 765 494 0395.

2213-6657/$ - see front matter
http://dx.doi.org/10.1016/j.amar.2013.10.003

all the roads in large road networks. Thus, a model with a limited number of variables is fitted to data that represent the population of roads, and then the model is used to estimate the expected number of crashes on these roads. A road with a large positive residual (excessive crash frequency) is a good candidate for a detailed site investigation. Even if the model parameters reflect other variables omitted from the model (e.g., the traffic volume parameter may reflect better winter maintenance of busy roads), the model may still produce reasonable estimates of the crash frequency if applied to the same population. The parsimony of the model can sometimes be based on necessity and, more importantly, purpose. Nevertheless, applicable modeling advancements, such as accounting for a spatial correlation, should be utilized even in this case. On the other hand, if individual safety effects are to be estimated and this knowledge is to be applied outside of the studied population or sample, then the model must be carefully specified and the benefits offered by the state-of-the-art models utilized as discussed by Mannering and Bhat (this issue).

This paper discusses two statistical models developed to analyze work zone safety: a random parameters model and a random effects model. The benefit of using the random parameters model in providing additional insight into the safety factors of work zones will be emphasized. Furthermore, the justification of using a model with random effects, which might be considered by some researchers as less advanced among the two discussed, will be shown based on its performance and intended use.

First, the existing literature relevant to this paper's topic is summarized by pointing out representative examples of published work on the matter. Then, the data collected and assembled for the study is briefly described, and the two considered models are introduced and their estimation described. Next, the obtained models are discussed and compared, followed by a summary of the findings to conclude the paper.

1.1. Highway work zone safety

There is strong evidence that work zone conditions involve heightened risk of crash, which justifies the ongoing efforts to better understand the safety factors in work zones and to propose effective countermeasures. Early research on this subject compared the crash rates between construction and non-construction conditions (Rouphail, et al., 1988; Ha and Nemeth, 1995; Pal and Sinha, 1996; Khattak et al., 2002). Their research confirmed that crash rates during construction periods tend to be higher than in non-construction periods. More recent research used statistical modeling to examine the link between work zone characteristics and safety measured with crash frequency and severity. Lane closures and construction intensity were found to be significant factors of crash frequency in work zones (Pal and Sinha, 1996; Venugopal and Tarko, 2000) as well as the length of the work zone and the duration of the construction period (Khattak et al., 2002; Venugopal and Tarko, 2000). Different sections of the work zone were found to experience different prevailing crashes types (Garber and Zhao, 2002). Finally, traffic conditions and driver behavior were identified as major contributing factors that increase the frequency of crashes (Harb et al., 2008; Daniel et al., 2000; Wang et al., 1996).

A number of studies investigated the injury severity of work zone crashes (Khattak et al., 2002; Venugopal and Tarko, 2000; Li and Bai, 2008; Khattak and Targa, 2004). Venugopal and Tarko (2000) estimated the relationships between the work zone characteristics and the frequencies of crashes of different severity. The obtained relationships were similar for exposure variables across injury categories, while injury and fatal crashes were more affected by the construction intensity but not by the type of work. Li and Bai (2008) used logistic models to estimate crash severity and developed a Crash Severity Index, in which a set of factors of crash severity in work zones were identified. Khattak and Targa (2004) estimated work zone truck crash severity with an ordered logit model. Lane crossovers and the proximity of construction activities were found to significantly increase crash severity.

The relatively small number of studies on highway work zone crashes could be attributed to both the typically small sample of crashes (most work zones last for a short period) and the limited scope of data about work zones that cannot be easily increased afterwards when the work zone no longer exists. Traditionally, a work zone crash is identified through the police crash reports. Ullman and Scriba (2004) found that in 1992, only 14 states had explicit entries of data indicating work zone presence in their crash reports; 21 states had some work zone information entered, but it was not reported explicitly; and 16 states had no work zone information included in their crash reports. Although this situation has improved since the above study was conducted, missing work zone information still poses a major issue for work zone safety research.

Wang et al. (1996) also reviewed the current state of crash reporting issues along with federal efforts undertaken to improve the situation. They pointed out that work zone crashes are seriously under-represented due to (1) lack of explicit information, (2) non-inclusion of crashes outside the work zones but nonetheless caused by the construction activities, and (3) failure to identify a crash that happened in a work zone. The authors recommended that work zone information and status should be better reported in police crash reports. They also concluded that the existence of a work zone inventory database is essential for gaining a better understanding of work zone safety.

Another issue is the limited data about work zone characteristics. State transportation agencies preserve basic work zone information, but detailed road cross-section data and traffic management plans are not readily available. As a result, researchers either must use the existing limited data or make a great effort to collect detailed work zone characteristics. This situation is responsible for the limited number of variables included in past studies on work zone safety.