Accident Analysis and Prevention 109 (2017) 10–17

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Full length article

Multilevel analysis of the role of human factors in regional disparities in crash outcomes

Emmanuel Kofi Adanu[a,*], Randy Smith[b], Lars Powell[c], Steven Jones[a]

[a] Department of Civil, Construction and Environmental Engineering, The University of Alabama Tuscaloosa, AL, United States
[b] Department of Computer Science, The University of Alabama Tuscaloosa, AL, United States
[c] Alabama Center for Insurance Information and Research, The University of Alabama Tuscaloosa, AL, United States

ARTICLE INFO

Keywords:
Multilevel analysis
Human factors
Regional factors
Serious injury crashes

ABSTRACT

A growing body of research has examined the disparities in road traffic safety among population groups and geographic regions. These studies reveal disparities in crash outcomes between people and regions with different socioeconomic characteristics. A critical aspect of the road traffic crash epidemic that has received limited attention is the influence of local characteristics on human elements that increase the risk of getting into a crash. This paper applies multilevel logistic regression modeling techniques to investigate the influence of driver residential factors on driver behaviors in an attempt to explain the area-based differences in the severity of road crashes across the State of Alabama. Specifically, the paper reports the effects of characteristics attributable to drivers and the geographic regions they reside on the likelihood of a crash resulting in serious injuries. Model estimation revealed that driver residence (postal code or region) accounted for about 7.3% of the variability in the probability of a driver getting into a serious injury crash, regardless of driver characteristics. The results also reveal disparities in serious injury crash rate as well as significant proportions of serious injury crashes involving no seatbelt usage, driving under influence (DUI), unemployed drivers, young drivers, distracted driving, and African American drivers among some regions. The average credit scores, average commute times, and populations of driver postal codes are shown to be significant predictors for risk of severe injury crashes. This approach to traffic crash analysis presented can serve as the foundation for evidence-based policies and also guide the implementation of targeted countermeasures.

1. Introduction

Road safety is both a public health and socioeconomic concern. The World Health Organization (WHO) estimates that about 1.25 million deaths occur annually through road traffic crashes, with millions of people sustaining various degrees of injury (World Health Organization, 2013). Globally, road traffic crashes are the main cause of death among those aged 15–29 years (World Health Organization, 2015). To be able to improve traffic safety, there is the need to understand the prevalence and underlying contributing factors of crashes. Treat et al. (1979) cited human factors as the primary contributor to roughly 93% of crashes when compared to roadway and vehicle-related factors. A growing body of research has examined the disparities in road traffic safety among population groups and geographic regions (e.g., Abdalla et al., 1997; Ameratunga et al., 2006; Factor et al., 2008; Anderson, 2010; Sehat et al., 2012). These studies reveal disparities in crash outcomes that exist between people and regions with different socioeconomic status and have overwhelmingly observed the disproportionately high fatalities in low income regions (e.g., Nantulya and Reich, 2003; Traynor 2009; Chen et al., 2010; Harper et al., 2015; World Health Organization, 2015).

Human factors-based traffic safety analyses, however, often focus on issues that directly contribute to crashes such as driver error (e.g., failure to yield) and risky actions (e.g., speeding or overtaking) discerned during the crash reporting process. There are, however, complex relationships between traffic safety and the broader social, economic, and environmental context of the individuals involved (Tillman and Hobbs, 1949; Abdalla et al., 1997; Lu et al., 2000; Kim et al., 2006; Choudhry et al., 2007; Factor et al., 2008; Traynor 2009; Anderson, 2010; Lee et al., 2014). Indeed, considerable work has been done to define and explore how safety behavior relates to deeper cultural currents (American Automobile Association, 2007; Rakauskas et al., 2009; Lund and Rundmo, 2009; Albrecht et al., 2013; Edwards et al., 2014; Atchley et al., 2014; Nordfjaern et al., 2014; American Automobile Association, 2015). Understanding the influence of both direct and indirect crash contributing factors is particularly relevant to crash studies

* Corresponding author.
E-mail addresses: ekadanu@crimson.ua.edu (E.K. Adanu), rsmith@cs.ua.edu (R. Smith), lars.powell@cba.ua.edu (L. Powell), sjones@eng.ua.edu (S. Jones).

http://dx.doi.org/10.1016/j.aap.2017.09.022
Received 19 April 2017; Received in revised form 2 August 2017; Accepted 26 September 2017

involving human elements as it has long been established that people are influenced by the sociocultural and economic conditions of where they live; people sharing the same context are more likely to be similar (Hox, 2010).

Quantitative crash studies are typically conducted to uncover hidden patterns in crash data or to predict the safety performance of specific transportation facilities (e.g., intersections, two-lane roads). Crash prediction models (e.g., Safety Performance Functions) are popular techniques to predict crash rates for a particular facility or location type (e.g. Jones et al., 1991; Miaou and Lum, 1993; American Association of State Highway and Transportation Officials, 2009; Brimley et al., 2012; Mehta et al., 2015). For other purposes, crash studies may be concerned with identifying factors that influence crash severities (e.g. Shankar and Mannering, 1996; Al-Ghamdi, 2002; Quddus et al., 2002; Abdel-Aty and Keller, 2005). Crash models have evolved with the development of sophisticated statistical methods to improve the accuracy of crash prediction. Crash prediction and analytical studies typically involve exploring multiple years of crash records. These studies generally lead to understanding the immediate dynamics of crashes, usually limited to the factors gathered at the time of the crash or factors believed to be directly linked to the crash occurrence (i.e., information contained in crash reports). If it is established, for instance, that a high proportion of a certain crash type occurs at a particular location, it can be inferred that the contributing factors may be attributable to the characteristics of the people involved or to the location characteristics, or some interaction between both. Similarly, if crash analyses reveal an overrepresentation of certain population groups in crashes, then further investigations may be required to understand the underlying influences.

This paper explores human-related crash factors and is premised on the common assumption that regional (or sub-regional) factors interact with individual driver characteristics to influence the occurrence and severity of crashes. The clustering of crashes within regions introduces multilevel correlation among observations and can have implications for crash model parameter estimates. The primary objective of this paper, then, is to apply multilevel regression analysis to investigate the effects of characteristics attributable to drivers and the geographic regions they reside on the likelihood of a crash resulting in serious injuries.

2. Previous work

Understanding the human-centered elements that lead to disparities in crash outcomes among regions requires investigating the relationship between individuals and the segment(s) of society in which they live. Individuals are influenced by the social groups to which they belong and the groups are in turn influenced by the individuals who make up that group (Jencks and Mayer, 1990; Jones and Duncan, 1995; Kreft and De Leeuw, 1998; Wilkinson, 1999; Snijders and Bosker, 1999; Raudenbush and Bryk, 2002; O'Connell and McCoach, 2008). Social groups may be categorized based on population characteristics common to the group. A category of social groups can be defined based on similarities (e.g. risk taking behaviors, attitudes towards law enforcement, socioeconomic characteristics) among individuals comprising the groups. It is therefore possible to define a category to contain social groups that are widely separated. Individuals and their societies may be viewed as a hierarchical system of individuals nested within societies (Hox, 2010). Since humans are in some way responsible for over 90% of road traffic crashes, it is possible that disparities in crash frequencies and consequent severities between regions may be due to the driving characteristics of the people (direct human factors) in those regions. On the other hand, clusters of crashes and outcomes may arise for reasons less strongly associated with the individuals who live in the regions (indirect human factors). This implies that regions and their residents can exert influences on each other in factors contributing to crash occurrence and outcome. Due to this nested structure, the odds of an individual getting into a crash are not truly independent because individuals who share common regional characteristics (e.g. driving regulations, land use patterns, social networks, socioeconomic characteristics, roadway conditions) may be similar in their risk of involvement in crashes of varying severities. Local (i.e., regional) factors are known to indirectly influence driver behaviors that impact crash occurrence and severity. For instance, regional socioeconomic characteristics often reflect local investment in the development, operation, and maintenance of transportation infrastructure. Similarly, enforcement of driving regulations in addition to availability and quality of emergency response services are linked to regional socioeconomic characteristics. Lack of enforcement of driving regulations can result in the emergence of poor safety culture (i.e., prevalence of risky driving behaviors) in a region and the absence of rapid emergency response service and well-equipped trauma centers in some regions can impact post-crash injury severity. In effect, the characteristics of crashes (type, severity, and contributing factors) involving drivers from the same region could be correlated. This means that instead of viewing each driver as an independent unit, there is value in exploring similarities and possible dependencies based on the social groups to which they belong. This presents crash data in a hierarchical structure where drivers are nested within regions (i.e., social contexts).

The general idea that there exists bidirectional influential effects between individuals and the social contexts or groups to which they belong, and that the individuals and the social groups are conceptualized as a hierarchical system of individuals nested within groups is the sine qua non of multilevel research. Individuals from the same geographical area are seen to be more similar to each other than are individuals from different geographical areas as this spatial proximity tends to influence or reflect social grouping (Hox, 2010). Samples of individuals from different geographical areas are therefore not completely independent. The average correlation, expressed as intraclass correlation (ICC), between variables measured on individuals from the same geographical area would therefore be expected to be higher than the average correlation between variables measured on individuals from different geographical areas (Hox, 2010). ICC is an indication of the proportion of the variance explained by the grouping structure in the population. The partition of variance at different levels of the hierarchical structure improves statistical estimation (Goldstein et al., 2002; Merlo, 2003). Standard statistical tests are based on the assumption of independence of the observations. If this assumption is violated, which is always the case for hierarchical data, the estimates of the standard errors of conventional statistical tests may be wrong and possibly lead to an overstatement of statistical significance.

The use of multilevel analysis allows for the exploration of causal heterogeneity (Western, 1998). Specifying cross-level interactions makes it possible to determine whether causal effect of lower-level predictors is influenced by higher-level predictors (Steenbergen and Jones, 2002). Multilevel modeling allows simultaneous study of ecological (or regional) and individual-level risk factors, which is particularly useful in understanding how regional factors translate into differences in individual-level risk (Bryk and Raudenbush, 1992; DiPrete and Forristal, 1994; Huttner and Eeden, 1995; O'Campo et al., 1997; Gelman and Hill, 2007). Multilevel analysis eliminates potential confounding of individual-level explanatory models resulting from the omission of higher-level factors. Conducting an analysis at any of these levels while ignoring the lower levels (e.g., individuals) or contextual levels (e.g., regions) can lead to erroneous conclusions. Studies have shown that ignoring a level of nesting in data can impact estimated variances and power to detect covariate effects (Julian, 2001; Shadish et al., 2002; Moerbeek, 2004), can inflate Type I error rates (Wampold and Serlin, 2000), and may lead to significant errors among regression estimates (Rodriguez and Goldman, 1995; Goldstein, 2003) and consequently in the interpretation of results (Nich and Carroll, 1997; Snijders and Bosker, 1999). Multilevel models have been developed to properly account for the hierarchical (correlated) nesting of data (Heck