Accident Analysis and Prevention 42 (2010) 1805–1813

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Analyzing fault in pedestrian–motor vehicle crashes in North Carolina

Gudmundur F. Ulfarsson[a,*], Sungyop Kim[b,1], Kathleen M. Booth[c,2]

[a] University of Iceland, Civil and Environmental Engineering, Hjardarhagi 2-6, IS-107 Reykjavik, Iceland
[b] University of Missouri–Kansas City, Department of Architecture, Urban Planning and Design, 5100 Rockhill Road, Kansas City, MO 64110-2499, USA
[c] Clark Construction, 110 Tampico, Suite 100, Walnut Creek, CA 94598, USA

ARTICLE INFO

Article history:
Received 9 December 2009
Received in revised form 16 April 2010
Accepted 3 May 2010

Keywords:
Fault
Crash
Pedestrian
Driver
Motor vehicle

ABSTRACT

Crashes between pedestrians and motor vehicles are an important traffic safety concern. This paper explores the assignment of fault in such crashes, where observed factors are associated with pedestrian at fault, driver at fault, or both at fault. The analysis is based on police reported crash data for 1997 through 2000 in North Carolina, U.S.A.

The results show that pedestrians are found at fault in 59% of the crashes, drivers in 32%, and both are found at fault in 9%. The results indicate drivers need to take greater notice of pedestrians when drivers are turning, merging, and backing up as these are some of the prime factors associated with the driver being found at fault in a crash.

Pedestrians must apply greater caution when crossing streets, waiting to cross, and when walking along roads, as these are correlated with pedestrians being found at fault. The results suggest a need for campaigns focused on positively affecting pedestrian street-crossing behavior in combination with added jaywalking enforcement. The results also indicate that campaigns to increase the use of pedestrian visibility improvements at night can have a significant positive impact on traffic safety. Intoxication is a concern and the results show that it is not only driver intoxication that is affecting safety, but also pedestrian intoxication.

The findings show in combination with other research in the field, that results from traffic safety studies are not necessarily transferable between distant geographic locations, and that location-specific safety research needs to take place. It is also important to further study the specific effects of the design of the pedestrian environment on safety, e.g. crosswalk spacing, signal timings, etc., which together may affect pedestrian safety and pedestrian behavior.

1. Introduction

Walking, for exercise or as an alternative mode of transportation, is encouraged in today's society. Walking provides health benefits and decreases the use of motor vehicles. However, pedestrians share the world with motor vehicles and the resulting opportunities for conflict lead to a crash risk. The National Highway Traffic Safety Administration (NHTSA) reported a total of 68,000 pedestrian injuries and 4,641 pedestrian fatalities in the U.S.A. in 2004 (NHTSA, 2005). While these numbers account for only 2% of people injured in traffic crashes in the U.S.A., they represent 11% of traffic crash fatalities (NHTSA, 2005).

Fault is assigned to the party (pedestrian, driver, both, or in some cases neither) who acted negligently or is in other ways found to have caused the crash. Some studies indicate that the pedestrian, not the driver, is more commonly found at fault (Lee and Abdel-Aty, 2005; Preusser et al., 2002) but in a recent study drivers were more likely to be found at fault (Kim et al., 2008a,b).

Fault has been explored in non-pedestrian–motor vehicle crashes. For example, a young or an elderly driver is more likely to be found at fault than a middle-aged driver (McGwin and Brown, 1999). Young males are more likely to be at fault than young females, whereas the opposite holds for elderly females compared to elderly males (McGwin and Brown, 1999). However, relatively few studies have focused on fault and the factors correlated with the assignment of fault in pedestrian crashes, aside from a study in Hawaii (Kim et al., 2008b).

Research has explored a variety of aspects relating to pedestrian safety, including pedestrian injury severity (Kim et al., 2010, 2008a; Kim, 2007), effects of gender (Kim et al., 2010; Clifton et al., 2004), vehicle speed (Gårder, 2004), U.S. interstate crashes (Johnson, 1997), pedestrian age (Kim et al., 2010, 2008a; Oxley

* Corresponding author. Tel.: +354 525 4907; fax: +354 525 4632.
E-mail addresses: gfu@hi.is (G.F. Ulfarsson), kims@umkc.edu (S. Kim), katie.booth@clarkconstruction.com (K.M. Booth).
[1] Tel.: +1 816 235 1725; fax: +1 816 235 5226.
[2] Tel.: +1 925 279 4000; fax +1 925 279 4030.

0001-4575/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2010.05.001

et al., 1997, 2005; Zegeer et al., 1993), fault (Kim et al., 2008b), intersections (Lee and Abdel-Aty, 2005), truck crashes (Lefler and Gabler, 2004), pedestrian behavior (McMahon et al., 1999), crash types (Stutts et al., 1996), and crash frequency (Shankar et al., 2003) to name just a few of the areas that have been studied.

Multiple studies have examined the pedestrian actions that most commonly contribute to crash occurrence. Pedestrians failing to yield the right of way, disregarding traffic signals, running into the street, stepping from between parked cars, walking while intoxicated, or walking with traffic rather than against it are noted (Preusser et al., 2002; Oxley et al., 1997; Stutts et al., 1996; Baltes, 1998). Such behavior can be correlated with the assignment of fault.

Pedestrians are naturally not the only ones who cause crashes. Failure to yield, excessive speed, improper backing, moving violations, distraction, reckless driving, and intoxication have been identified as some of the most frequent contributing factors involving drivers (Stutts et al., 1996). A study of fatal hit-and-run pedestrian crashes showed that drivers are less likely to leave the scene when the pedestrian is a child or an elderly person (Solnick and Hemenway, 1995). When the pedestrian is female or a child, it is more likely that the driver will be identified. Drivers often run from crashes where they would be considered to be at fault (Solnick and Hemenway, 1995).

The binary assignment of fault to either driver or pedestrian in Hawaii has been explored (Kim et al., 2008b). Fault and the factors affecting it have also been studied for bicycle–motor vehicle crashes (Kim and Li, 1996). The present study explores how observable factors associated with pedestrian–motor vehicle crashes correlate with which party is found at fault in a multinomial fashion, i.e. driver at fault, pedestrian at fault, or both at fault. The study omits hit-and-run crashes, which are about 12% of the complete dataset. In many such cases, the driver and vehicle information is not known and in the majority of those cases fault has been assigned to the driver. This may lead to an underrepresentation of intoxicated drivers, since it can be speculated that intoxicated drivers may be more likely to leave the scene.

The results of this study provide new information on behavior and factors associated with being found at fault. This information can be used to increase awareness of pedestrian and driver behavior that contributes to crash occurrence.

2. Methodology

Numerous models exist for exploring data, predicting behavior, and forecasting outcomes. Kim et al. (2008b) studied fault in pedestrian–motor vehicle crashes using logistic regressions for each crash type. In their study there were two outcomes, pedestrian at fault or driver at fault. For binary outcomes the logistic regression or odds-ratio analysis are natural methods, but other methods could be used, such as discriminate analysis and analysis of variance (see, e.g. the textbook by Agresti, 2002).

In the present study, there are three discrete outcomes, *pedestrian at fault*, *driver at fault*, and *both at fault*. The fundamental limitation of the binary methods is a lack of statistical efficiency when applied in a multinomial context, since these methods do not use all the available data simultaneously as they consider only two alternatives at once. A more statistically efficient method when dealing with three or more outcomes is the multinomial logit (MNL) model (McFadden, 1974, 1981). The MNL model has been commonly used for modeling outcomes of events which are partly random and partly systematic, such as crash injury severity outcomes (e.g. Ulfarsson and Mannering, 2004). The MNL approach is able to simultaneously estimate the effect of numerous observed variables on the probability of multiple outcomes. Whereas, the binary methods loose statistical efficiency by not considering all outcomes simultaneously, e.g. odds-ratios consider the ratio between one outcome and one other outcome at a time.

The MNL model represents the probability of one outcome occurring over another based upon each outcome's propensity, S_{ni}, as shown in

$$P_{ni} = P(\max_{i' \neq i} S_{ni'} \leq S_{ni}), \tag{1}$$

where i, $i' \in I$, and I represents the set of possible outcomes. Also, $n \in N$, where N stands for all observed crashes. The propensity itself can be represented with the linear-in-parameters form:

$$S_{ni} = \boldsymbol{\beta}_i \mathbf{x}_n + \varepsilon_{ni}, \tag{2}$$

where $\boldsymbol{\beta}_i$ is a vector of estimable coefficients specific to outcome i, and $\mathbf{x}_n$ is a vector of observed variables specific to crash n. The random component, ε_{ni}, is assumed to be identically and independently distributed type I extreme value. This leads to the MNL model of the probability of fault outcome i conditional on crash n having occurred (McFadden, 1974, 1981):

$$P_{ni} = \frac{e^{\boldsymbol{\beta}_i \mathbf{x}_n}}{\sum_{\forall i' \in I} e^{\boldsymbol{\beta}_{i'} \mathbf{x}_n}}. \tag{3}$$

The coefficients, $\boldsymbol{\beta}_i$, are estimated with the method of maximum likelihood (see, e.g. Greene, 2007).

It should be reiterated that the MNL model has a strict assumption that the unobserved random component, ε_{ni}, is identically and independently distributed. This leads to the independence of irrelevant alternatives (IIA) property of the MNL. IIA means that adding or deleting alternative outcomes should not affect the odds among the remaining outcomes. If this property is violated the MNL model results will become biased (McFadden, 1981). A statistical test of the violation of IIA is performed using the Small–Hsiao test (Small and Hsiao, 1985). The test statistic is χ^2 distributed and if the statistic is larger than a χ^2 table value at a particular level of significance it is evidence to reject the null hypothesis of the IIA property holding. In this study, the 0.05 level of significance is chosen.

To enable focus on the most significant coefficients, coefficients that are not found significantly different from zero at the 0.05 level of significance using a t-test (see, e.g. Greene, 2007) are removed. To avoid artificial accuracy, the likelihood ratio statistic (see, e.g. Greene, 2007) is used to test whether two or more coefficients on related variables (or on the same variable in different outcome equations) are significantly different from each other at the 0.05 level of significance and if not they are constrained to be equal.

To judge the overall model fit it is possible to compare the model's log-likelihood at convergence with the log-likelihood of a naïve model, e.g. a model with all coefficients set to zero (equivalent to assigning equal probability to all outcomes), or a model with only alternative-specific constants (equivalent to assigning probability to outcomes equal to the observed share of the outcomes in the dataset). The latter test is stronger. This is performed by calculating the statistic (see, e.g. the textbook by Washington et al., 2003):

$$\rho^2 = 1 - \frac{\mathrm{LL}(\boldsymbol{\beta})}{\mathrm{LL}(\mathbf{c})}, \tag{4}$$

where $\mathrm{LL}(\boldsymbol{\beta})$ represents the log-likelihood at model convergence, $\mathrm{LL}(\mathbf{c})$ represents the log-likelihood of a model with alternative-specific constants only. The ρ^2 goes from 0 (for no improvement in log-likelihood) to 1 for a perfect fit. A value for ρ^2 larger than 0.1 already indicates meaningful improvement.

To facilitate the interpretation of results, the percentage change in probability of a fault category is calculated when each variable is switched between 0 and 1, since all the variables are such binary indicators. This has been termed the direct pseudo-elasticity