Accident Analysis and Prevention 40 (2008) 1695–1702

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic logit analysis

Joon-Ki Kim[a,1], Gudmundur F. Ulfarsson[b,*], Venkataraman N. Shankar[c,2], Sungyop Kim[d,3]

[a] Korea Research Institute for Human Settlements (KRIHS), Transportation Research Division, 1591-6 Gwanyang-dong, Dongan-gu, Anyang-si, Gyeonggi-do, 431-712, Republic of Korea

[b] University of Iceland, Civil and Environmental Engineering, Hjardarhagi 6, IS-107 Reykjavik, Iceland

[c] Pennsylvania State University, Department of Civil and Environmental Engineering, 226C Sackett Building, University Park, PA 16802, USA

[d] University of Missouri – Kansas City, Department of Architecture, Urban Planning and Design, 5100 Rockhill Road, Kansas City, MO 64110-2499, USA

ARTICLE INFO

Article history:
Received 2 April 2008
Received in revised form 6 May 2008
Accepted 5 June 2008

Keywords:
Age
Aging
Pedestrian
Injury
Heteroskedasticity
Logit
Accident
Severity

ABSTRACT

This research explores the injury severity of pedestrians in motor-vehicle crashes. It is hypothesized that the variance of unobserved pedestrian characteristics increases with age. In response, a heteroskedastic generalized extreme value model is used. The analysis links explanatory factors with four injury outcomes: *fatal*, *incapacitating*, *non-incapacitating*, and *possible or no injury*. Police-reported crash data between 1997 and 2000 from North Carolina, USA, are used.

The results show that pedestrian age induces heteroskedasticity which affects the probability of fatal injury. The effect grows more pronounced with increasing age past 65. The heteroskedastic model provides a better fit than the multinomial logit model. Notable factors increasing the probability of fatal pedestrian injury: increasing pedestrian age, male driver, intoxicated driver (2.7 times greater probability of fatality), traffic sign, commercial area, darkness with or without streetlights (2–4 times greater probability of fatality), sport-utility vehicle, truck, freeway, two-way divided roadway, speeding-involved, off roadway, motorist turning or backing, both driver and pedestrian at fault, and pedestrian only at fault. Conversely, the probability of a fatal injury decreased: with increasing driver age, during the PM traffic peak, with traffic signal control, in inclement weather, on a curved roadway, at a crosswalk, and when walking along roadway.

1. Introduction

Walking is a simple and fundamental mode of transportation. In fact, without handicap, walking indispensably occurs in daily life. In the United States, there were 4641 pedestrians killed and 68,000 pedestrians injured in traffic crashes in 2004 (NHTSA, 2004). These fatalities account for 10.7% of total transport fatalities in the US by transportation mode (BTS, 2006).

Previous studies of pedestrian–vehicle crashes have considered characteristics describing the pedestrian, driver, vehicle, geometry, time, weather, etc. There is a relationship between pedestrian age and pedestrian–vehicle crashes. Older pedestrians suffer more serious injuries than other age groups (Sklar et al., 1989; Zeeger et al., 1996; Fontaine and Gourlet, 1997; Harruff et al., 1998; Al-Ghamdi, 2002), whereas older pedestrians are more cautious than other age groups (Harrell, 1991). Older pedestrian crashes may be increasingly caused by misjudgment of gaps in traffic (Oxley et al., 2005). Even though age by itself is not a causal factor, age is related with a decline of cognitive function (Maring and van Schagen, 1990; DeLucia et al., 2003) and a weakening of the body (Jensen, 1999).

Several studies have shown the influence of speed (Pasanen and Salmivaara, 1993; Anderson et al., 1997; Davis, 2001; Gårder, 2004; Leden et al., 2006) and vehicle types (Zajac and Ivan, 2003; Ballesteros et al., 2004; Lefler and Gabler, 2004; Matsui, 2005) on pedestrian–vehicle crashes. Alcohol related pedestrian–vehicle

* Corresponding author. Tel.: +354 525 4907; fax: +354 525 4632.
E-mail addresses: kimjoonki@krihs.re.kr (J.-K. Kim), gfu@hi.is (G.F. Ulfarsson), shankarv@engr.psu.edu (V.N. Shankar), kims@umkc.edu (S. Kim).
1 Tel.: +82 31 380 0285.
2 Tel.: +1 814 865 9434.
3 Tel.: +1 816 235 6898.

0001-4575/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2008.06.005

crashes have not shown the same reduction in pedestrian fatalities as alcohol related vehicle crashes (Öström and Eriksson, 2001). Intoxication (driver or pedestrian) affects the risk of pedestrian–vehicle crashes (daSilva et al., 2003), but also generally increases pedestrian injury severity when crashes occur (Miles-Doan, 1996; Jensen, 1999; Öström and Eriksson, 2001). The influences of other factors have been studied, such as traffic signal spacing (Shankar et al., 2003), crosswalks (Zeeger et al., 1996; Leden et al., 2006), intersections (Koepsell et al., 2002; Lee and Abdel-Aty, 2005), sidewalks (McMahon et al., 2002), time (Al-Ghamdi, 2002), culpability (Preusser et al., 2002), traffic volume (Davis et al., 2002).

Pedestrian–vehicle crash research continues to grow. Nevertheless, to date, heteroskedasticity across individual pedestrians has not been considered. To explain, a person's age is correlated with other conditions such as physical fragility and cognitive function, which differ across individuals. This leads to the hypothesis that variance increases across individuals with age. Typical models for injury severity assume individuals have equal variances, but when the variances differ heteroskedasticity arises. Heteroskedasticity can be important for all road users although it is likely to be most pronounced for vulnerable groups such as pedestrians. This is because vehicle safety features are designed to give vehicle occupants a broader margin of safety. Thus, the objective is to develop a heteroskedastic multivariate model of pedestrian injury severity. The goal is to improve the modeling and thereby understanding of pedestrian crashes in order to promote a safer pedestrian environment and further assist in prevention of serious injury.

2. Methodology

When considering statistical models of injury severity in motor-vehicle crashes, we look towards models that are conditional on a crash having occurred. Such models hypothesize a function of observable (e.g. recorded in police or hospital records) and unobservable (e.g. a person's physical and mental state) factors that affect the probability of a particular injury severity category. Commonly employed models in the study of injury severity are the: logistic regression (e.g. Ballesteros et al., 2004), multinomial logit model (MNL) (e.g. Ulfarsson and Mannering, 2004), nested logit model (NL) (e.g. Shankar et al., 1996), and ordered probit model (e.g. Kockelman and Kweon, 2002).

These models differ in strengths and weaknesses but share a common issue when unobserved factors (such as a person's physical and mental health condition) are likely to vary differently across people leading to heteroskedasticity, and we hypothesize, especially for older people. If this differing variance is ignored as in the mentioned models, the estimated coefficients may be biased, inconsistent, and inefficient (Zeng, 2000; Greene, 2003).

Generalized extreme value (GEV) models (McFadden, 1981), of which MNL and NL are special cases, are a broad group of models that can be developed to handle heteroskedasticity. There are two fundamental types of heteroskedastic GEV models: the individual-specific (Zeng, 2000) and the alternative-specific (Bhat, 1995). Another way to overcome a variety of such restrictions is to employ a mixed logit framework (McFadden and Train, 2000). In this study, the emphasis is on accounting for individual-specific variance where a specialized heteroskedastic model can be useful.

2.1. Model structure

This research uses the Zeng (2000) model of heteroskedastic individual-specific variances to explore the effect of pedestrian age, and in fact gender, on the variance of unobserved factors. Conditional on a crash having occurred, a linear-in-parameters propensity function affecting the probability of injury severity i for pedestrian n is written:

$$U_{ni} = \boldsymbol{\beta}_i \mathbf{x}_n + \varepsilon_{ni}. \tag{1}$$

The vector $\mathbf{x}_n$ contains the observed variables; the vector $\boldsymbol{\beta}_i$ contains the estimable coefficients, and ε_{ni} denotes the error which is stochastic and unobserved.

To formalize the model, Zeng (2000) lets the error term in (1) be independently distributed extreme value type 1 with location parameter 0 and a positive individual-specific scaling parameter, θ_n:

$$\varepsilon_{ni} \sim \exp(-\exp(-\varepsilon_{ni}\theta_n)). \tag{2}$$

In the MNL model the scaling parameter is taken as 1. This leads to an individual-specific variance, i.e. heteroskedasticity:

$$\mathrm{var}(\varepsilon_{ni}) = \frac{\pi^2}{6\theta_n^2}. \tag{3}$$

Deriving the heteroskedastic logit (HET) model from this point is straightforward (see Zeng, 2000):

$$P_{ni} = \frac{e^{\boldsymbol{\beta}_i \mathbf{x}_n \theta_n}}{\sum_{j=1}^{I} e^{\boldsymbol{\beta}_j \mathbf{x}_n \theta_n}}, \tag{4}$$

where I is the number of alternatives. The positive individual-specific scaling parameter is parameterized (Zeng, 2000):

$$\theta_n = e^{\boldsymbol{\gamma} \mathbf{z}_n}, \tag{5}$$

where $\mathbf{z}_n$ is a vector of observed individual-specific variables (believed to generate heteroskedasticity, they can also be in $\mathbf{x}_n$) and $\boldsymbol{\gamma}$ is a vector of estimable coefficients. As shown in (3) and (4), the utility function, $\boldsymbol{\beta}_i \mathbf{x}_n$, is weighted with θ_n, which affects the variance of the unobserved terms. If $\boldsymbol{\gamma}$ is not significantly different from zero, (4) reduces to MNL. The model (4) predicts the conditional probability of injury severity i for pedestrian n (conditional on a crash having occurred) as a function of observable characteristics.

2.2. Model estimation

The model in (4) is estimated using full information maximum likelihood using code adapted from Zeng (2000) who makes the code available. The log-likelihood function to be maximized is

$$L = \sum_{n=1}^{N} \sum_{i=1}^{I} y_{ni} \ln P_{ni}, \tag{6}$$

where I is the number of injury severities and N is the number of pedestrians. P_{ni} is given by (4) and y_{ni} is defined:

$$y_{ni} = \begin{cases} 1 & \text{if the } n\text{th pedestrian suffers injury } i, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

2.3. Elasticity

To simplify the interpretation of the variables we explore their impacts on the probability by using elasticity as a measure of marginal effects. A direct elasticity, $E_{x_{nk}}^{P_{ni}}$, refers to the percentage change in the probability of injury severity i, P_{ni}, due to a percent change in the k-th variable for person n, x_{nk}.

In case of continuous variables a direct elasticity for each pedestrian n is computed by (Zeng, 2000):

$$E_{x_{nk}}^{P_{ni}} = \begin{cases} x_{nk}\theta_n \left(\beta_{ik} - \sum_{j=1}^{I} P_{nj}\beta_{jk} \right) & \text{if } x_{nk} \text{ is not in } \theta_n, \\ x_{nk}\theta_n \left(\gamma_k \left[\beta_{ik}x_{nk} - \sum_{j=1}^{I} P_{nj}\beta_{jk}x_{nk} \right] \right. \\ \qquad \left. + \left[\beta_{ik} - \sum_{j=1}^{I} P_{nj}\beta_{jk} \right] \right) & \text{otherwise}, \end{cases} \quad (8)$$

where P_{ni} is defined by (4), θ_n is defined by (5), β_{ik} is the estimable coefficient of the k-th independent variable in (1), γ_k is the estimable coefficient of the k-th independent variable in (5), and I is the number of injury severity categories.

For an indicator variable that has the value 0 or 1 the direct pseudo-elasticity is calculated (Shankar and Mannering, 1996). It measures the change in estimated probability of injury severity due to a switch in the values of the indicator variable between 0 and 1:

$$E_{x_{nk}}^{P_{ni}} = \frac{P_{ni}[\text{given } x_{nk}=1] - P_{ni}[\text{given } x_{nk}=0]}{P_{ni}[\text{given } x_{nk}=0]}. \quad (9)$$

When interpreting pseudo-elasticities a value of 1, for example, means a 100% change or a doubling of the probability when the variable is switched.

Underreporting, especially of no injury crashes, leads to biased injury probabilities which affect the elasticities. This is less of an issue for pedestrian crashes due to pedestrian injury concerns, even in relatively low impact crashes. Pedestrians have motivation to report in order to assert rights against the vehicle driver and insurer. Drivers have motivation to report to defend against pedestrian claims and to indicate that the pedestrian was at fault.

3. Empirical setting

This study analyzed police-reported pedestrian–vehicle crashes from the State of North Carolina, USA, from 1997 to 2000, involving only one pedestrian and one vehicle, excluding data if pedestrian age is below 18, and omitting observations with missing values. This yields a total of 5808 observations. Table 1 shows the key variables tested in this study and their observed distribution.

The most dominant category in reported pedestrian–vehicle crashes was possible or no injury (37.0%), being followed by non-incapacitating injury (34.0%), incapacitating injury (18.8%), and fatal injury (10.2%). Contrast this with the share of fatal driver injuries in motor-vehicle crashes, which e.g. range from 0.5% to 4.1% (based on gender, vehicle type, single-vehicle or two-vehicle crashes) in the study by Ulfarsson and Mannering (2004), and the share of fatal bicyclist injuries in motor-vehicle crashes which was 3.5% in North Carolina (Kim et al., 2007).

As age increased, the proportion of fatal injury increased: age 18–24 (4.9%), age 25–54 (9.9%), 55–64 (13.6%), 65–74 (15.6%), and 75+ (21.8%). About one fifth of 75+ year-old pedestrian victims is fatally injured. In terms of gender, 64.4% were male. Intoxicated pedestrians accounted for 19.0% in fatal injury and 24.1% in incapacitating injury, whereas non-intoxicated pedestrians accounted for 7.8% and 17.4%, respectively.

About 44.2% of pedestrian–vehicle crashes occurred when a pedestrian was crossing at a crosswalk; followed by a pedestrian walking off the roadway (13.3%), and a pedestrian waiting to cross the road (10.3%). Most pedestrian–vehicle crashes were related to a low estimated vehicle speed, less than 32.2 km/h (20 mph). Most of

Table 1
Descriptive statistics for key variables

Variable			Total
Pedestrian characteristics			
	Age	18–24	986 (17.0%)
		25–54	3,766 (64.8%)
		55–64	470 (8.1%)
		65–74	320 (5.5%)
		75+	266 (4.6%)
	Average*		40.9 (16.1)
	Gender	Male	3,743 (64.4%)
		Female	2,065 (35.6%)
Driver characteristics			
	Age	<25	1,413 (24.3%)
		25–54	3,243 (55.8%)
		55–64	531 (9.1%)
		65–74	334 (5.8%)
		75+	287 (4.9%)
	Average*		39.5 (17.6)
	Gender	Male	3,467 (59.7%)
		Female	2,341 (40.3%)
	Intoxicated	No	5,566 (95.8%)
		Yes	242 (4.2%)
Temporal characteristics			
Time		6:00 a.m.–9:59 a.m.	738 (12.7%)
		10:00 a.m.–2:59 p.m.	1,285 (22.1%)
		3:00 p.m.–5:59 p.m.	1,152 (19.8%)
		6:00 p.m.–8:59 p.m.	1,211 (20.9%)
		9:00 p.m.–5:59 a.m.	1,422 (24.5%)
Control characteristics			
	Control type	No traffic control	4,529 (78.0%)
		Traffic signal	738 (12.7%)
		Traffic sign	383 (6.6%)
		Human control	94 (1.6%)
		Other sign (rail road etc.)	64 (1.1%)
Land development characteristics			
	Land use	Residential area	1,775 (30.6%)
		Commercial area	3,017 (51.9%)
		Institutional area	183 (3.2%)
		Industrial area	58 (1.0%)
		Farm/woods/pastures	775 (13.3%)
Environmental characteristics			
	Weather	Clear	4,235 (72.9%)
		Cloudy	996 (17.1%)
		Fog/smog/smoke	44 (0.8%)
		Rain	504 (8.7%)
		Snow	20 (0.3%)
		Other weather	9 (0.2%)
	Light	Daylight	3,127 (53.8%)
		Dawn or dusk	222 (3.8%)
		Dark-Lighted	1,336 (23.0%)
		Dark-Unlighted	1,123 (19.3%)
Vehicle characteristics			
	Vehicle type	Car	3,945 (67.9%)
		Pickup	849 (14.6%)
		Minivan	147 (2.5%)
		Sport-utility vehicle	229 (3.9%)
		Van	337 (5.8%)
		Bus	–
		Truck	194 (3.3%)
		Motorcycle/Moped	29 (0.5%)
		Public (police etc.)	3 (0.1%)
		Other vehicle types	75 (1.3%)

Table 1 (*Continued*)

Variable		Total
Geometry characteristics		
Road class type	Freeway	132 (2.3%)
	US route	439 (7.6%)
	North Carolina state route	401 (6.9%)
	N. Carolina secondary state route	716 (12.3%)
	Local city street	2,902 (50.0%)
	Public vehicular area	1,218 (21.0%)
Road geometry	Curved	355 (6.1%)
	Straight level	4,295 (73.9%)
	Straight grade	1,158 (19.9%)
Road type	One-way	405 (7.0%)
	Two-way not divided	4,448 (76.6%)
	Two-way divided	955 (16.4%)
Crash characteristics		
Speeding-involved	No	5,734 (98.7%)
	Yes	74 (1.3%)
Pedestrian behavior		
	Pedestrian in crosswalk	2,568 (44.2)
	Off roadway	775 (13.3)
	Pedestrian lying on roadway	61 (1.1)
	Waiting to cross	596 (10.3)
	Walking along roadway	499 (8.6)
	Working in roadway	142 (2.4)
	Other	1,167 (20.1)
Motorist maneuver		
	Motorist turning/merging	449 (7.7)
	Motorist backing	566 (9.7)
Fault	Both driver and pedestrian	566 (9.7%)
	None or can not be determined	565 (9.7%)
	Driver only at fault	1,953 (33.6%)
	Pedestrian only at fault	2,724 (46.9%)

This table displays frequency (percentage). Percentages are calculated categories of each variable (e.g. two rows for No–Yes variables) for the total frequency. Averages marked with * display mean (standard deviation).

the crashes involved a passenger car (67.9%), followed by a pickup truck (14.6%). Pedestrians were more often believed responsible for the crashes than drivers: pedestrian fault (46.9%) and driver fault (33.6%).

4. Results

The injury severity model was first estimated with MNL and then with the HET model in (4). Before MNL was selected as the model for comparison, several NL structures were tested. Statistical tests rejected each in favor of the MNL (Kim, 2007). A number of explanatory variables were examined. We avoided pairs of variables that were strongly correlated to each other. Coefficients that were not statistically significantly different from zero at the 0.05 level of significance were restricted to zero to reach a fairly parsimonious model. When the coefficients on a particular variable were not found statistically significantly different across injury severities, those coefficients were constrained to be equal. The final estimation results are shown in Table 2.

To test a hypothesis of no heteroskedasticity, a likelihood ratio test was performed to compare MNL to HET. The likelihood-ratio test statistic is 7.13 with one degree of freedom, which is greater than the χ^2 table value, 6.63, at the 0.01 level of significance, thereby rejecting the null hypothesis. This implies that HET is favored over MNL. The adjusted likelihood ratio index (Ben-Akiva and Lerman, 1985) also favors HET over MNL in this study.

The signs of all coefficients on variables in the MNL and HET models are consistent. To facilitate the interpretation of the results,

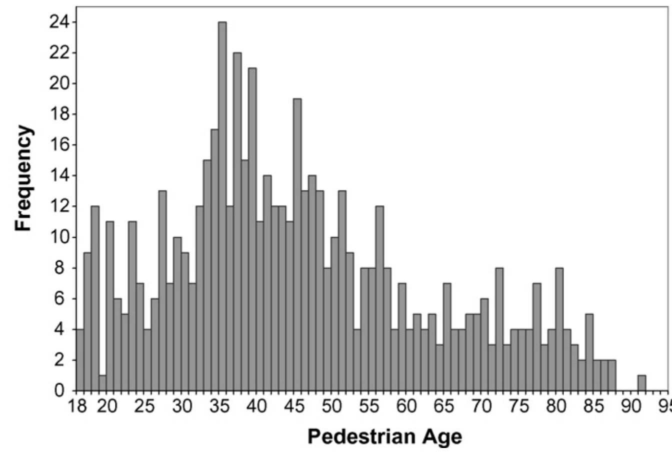


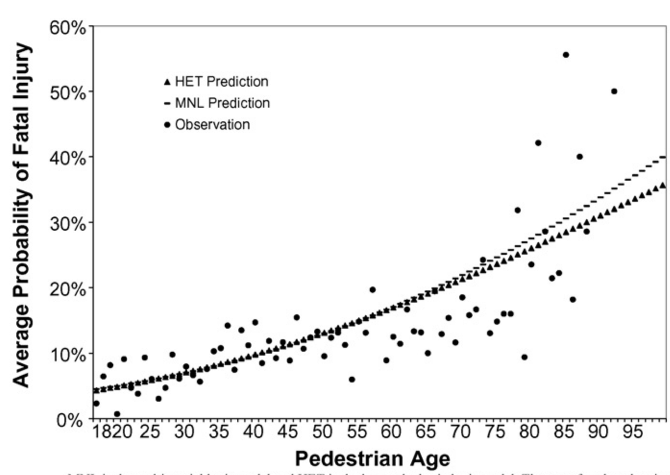
Fig. 1. Number of fatally injured pedestrians by age in the dataset.

direct elasticity for continuous variables and average direct pseudo-elasticity for indicator variables were calculated and are presented in Table 3.

4.1. Pedestrian characteristics

We first compare MNL and HET in terms of their predictions and the effect of pedestrian age. Fig. 1 shows the observed frequency of fatal injury for pedestrians by age. The MNL and HET models are used to predict the probability of fatal injury using simulated age as each pedestrian's age is changed from 18 to 100. Fig. 2 shows the average predicted probability of fatal injury for each age for both models; the dots show the fraction of pedestrians at each age that were observed to suffer fatal injury. The share of pedestrians that is fatally injured grows with age and both MNL and HET capture that. The MNL is more sensitive to the larger share of older adults with fatal injuries, despite there being few observations for those ages (Fig. 1). Fig. 2 shows that the models diverge at around age 65. There is little age-specific heteroskedastic effect for younger pedestrians. Fig. 2 supports the paper's initial hypothesis that heteroskedasticity is important for older adults.

Fig. 3 shows the average predicted probability of fatal injury for the data. The MNL and HET are nearly identical for younger ages



MNL is the multinomial logit model and HET is the heteroskedastic logit model. The age of each pedestrian is varied from 18 to 100 in this simulation, but otherwise each observation is left unchanged. The predicted probability of fatal injury is then average across all pedestrians for each simulated age to result in these smooth overall trend curves.

Fig. 2. Prediction of fatal injury probabilities using simulated pedestrian age.

Table 2
Pedestrian injury severity model estimation results

	Fatal injury		Incapacitating injury		Non-incapacitating injury	
	MNL	HET	MNL	HET	MNL	HET
Heteroskedastic variable ($\boldsymbol{\gamma}\mathbf{z}$)						
Pedestrian age		−0.005		−0.005		−0.005
Explanatory variables ($\boldsymbol{\beta}\mathbf{x}$)						
Alternative specific constant	−6.165†	−7.303†	−1.449†	−1.627†	−0.457†	−0.425†
Pedestrian						
Age	0.052†	0.059†	0.013†	0.013†	0.013†	0.013†
Driver						
Age	−0.008†	−0.010†	−0.008†	−0.010†	−0.008†	−0.010†
Male	0.472†	0.559†	0.266†	0.327†	0.266†	0.327†
Intoxicated driver	1.199†	1.467†	0.452†	0.551†		
Time						
PM peak (15:00–17:59)	−0.550†	−0.688†				
Control						
Traffic signal	−0.473	−0.597				
Traffic sign			−0.390†	−0.497		
Land development						
Commercial area			−0.246†	−0.294†	−0.246†	−0.294†
Environment						
Inclement weather	−0.485†	−0.571				
Dark-lighted	1.035†	1.285†	0.258†	0.324†		
Dark-unlighted	1.719†	2.090†	0.435†	0.531†		
Vehicles						
Sport-utility vehicle	0.542	0.671	−0.858†	−0.999†		
Truck	1.458†	1.789†				
Geometry						
Freeway	1.997†	2.382†	1.997†	2.382†	1.008†	1.169†
US route	1.376†	1.684†	0.448†	0.528†		
State route	0.965†	1.205†				
Curved			0.295	0.346		
Straight grade	0.371†	0.451†				
Two-way divided	0.524†	0.675†				
Crash						
Speeding-involved	1.857†	2.248†	0.952†	1.146†		
Pedestrian in crosswalk			0.530†	0.641†	0.206†	0.253†
Off roadway			−0.646†	−0.802†	−0.442†	−0.551†
Walking along roadway	−0.574†	−0.702†				
Motorist turning/merging			−1.221†	−1.524†	−0.426†	−0.536†
Motorist backing			−0.912†	−1.156†	−0.426†	−0.536†
Both driver and pedestrian fault	0.842†	1.030†				
Pedestrian fault	2.039†	2.479†	0.753†	0.893†	0.548†	0.645†
Number of observations				5808		
Log-likelihood at zero				−8051.6		
Log-likelihood at convergence		−6567.03 (MNL)			−6563.47 (HET)	
Adjusted Log-likelihood ratio index		0.1789 (MNL)			0.1792 (HET)	

MNL is the multinomial logit model and HET is the heteroskedastic logit model. Pedestrian age is both a regular explanatory variable and in the heteroskedasticity. Level of significance: all with $p < 0.05$ and † with $p < 0.01$. Coefficients that were not significant at the 0.05 level were restricted to zero and omitted from the table. Possible or no injury is the base case with coefficients restricted at zero.

but the models start diverging for ages past 65; with the HET yielding a smaller variance in predictions than the MNL. The variance in prediction does not behave like the variance in the unobserved effect (which increases as noted below). The visual difference in prediction probabilities is small between models, although it is statistically significant as indicated at the start of Section 4. Prediction by itself is not the objective; rather it is the analysis of the process, where the HET provides a richer, more realistic interpretation. The HET model indicates that the differences due to age-specific heteroskedasticity grow increasingly important past age 65.

Unlike MNL, the HET model can have different elasticities (Table 3) for injury categories that have identical coefficients (Table 2). Table 3 shows that the elasticity of age towards fatal injury is slightly greater for the HET model than the MNL model. This can appear surprising as the MNL model predicts a greater probability of fatal injury, however, elasticity is the proportional change in probability over the proportional change in age, and this value is larger for the HET model.

As pedestrian age goes up, pedestrians are more likely to be involved in serious injury when in a pedestrian–vehicle crash, in fact, age is elastic (Table 3). One reason might be physical deterioration in the aging process. Increased fragility plays a significant and important role in exacerbating injury severity with age. Alternatively, older pedestrians have been observed to cross streets under unsafe traffic conditions more frequently than other age groups (Oxley et al., 1997). Due to these and other reasons, e.g. older pedestrian's slower and perhaps unexpected movements, elders can also be hit harder by a vehicle.

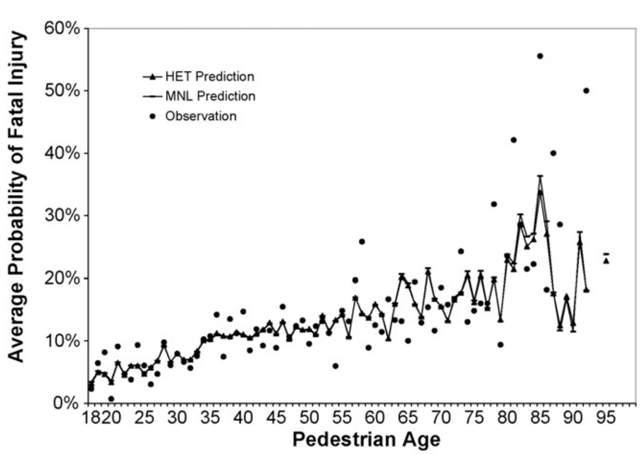
The sign of the coefficient, γ in (5), on the heteroskedastic variable (a pedestrian's age) is negative (Table 2). This implies that

Table 3
Comparison of average elasticities

	Fatal Injury		Incapacitating Injury		Non-Incapacitating Injury		Possible or No Injury	
	MNL	HET	MNL	HET	MNL	HET	MNL	HET
Pedestrian								
Age*	1.604	1.853	0.001	0.111	0.001	-0.020	-0.532	-0.448
Driver								
Age*	-0.121	-0.126	-0.121	-0.126	-0.121	-0.126	0.194	0.200
Male	33.3%	32.3%	8.5%	9.0%	8.5%	9.0%	-16.8%	-16.9%
Intoxicated driver	160.8%	167.8%	23.6%	24.4%	-21.4%	-21.4%	-21.4%	-21.4%
Time								
PM peak (15:00-17:59)	-39.3%	-40.5%	5.2%	5.2%	5.2%	5.2%	5.2%	5.2%
Control								
Traffic signal	-34.9%	-36.3%	4.5%	4.6%	4.5%	4.6%	4.5%	4.6%
Traffic sign	6.7%	7.1%	-27.8%	-29.2%	6.7%	7.1%	6.7%	7.1%
Land development								
Commercial area	14.0%	13.9%	-10.9%	-10.8%	-10.9%	-10.8%	14.0%	13.9%
Environment								
Inclement weather	-35.6%	-35.0%	4.6%	4.4%	4.6%	4.4%	4.6%	4.4%
Dark-lighted	139.3%	148.0%	10.0%	11.0%	-15.0%	-15.3%	-15.0%	-15.3%
Dark-unlighted	322.6%	338.1%	17.0%	18.4%	-24.2%	-24.1%	-24.2%	-24.1%
Vehicle								
Sports utility vehicle	80.8%	83.4%	-55.4%	-54.3%	5.1%	4.8%	5.1%	4.8%
Truck	251.0%	265.3%	-18.4%	-18.4%	-18.4%	-18.4%	-18.4%	-18.4%
Geometry								
Freeway	139.8%	143.2%	139.8%	143.2%	-10.8%	-11.8%	-67.4%	-66.6%
US route	205.3%	216.4%	20.7%	20.1%	-22.9%	-22.7%	-22.9%	-22.7%
State route	136.5%	146.3%	-9.9%	-10.0%	-9.9%	-10.0%	-9.9%	-10.0%
Curved	-5.9%	-5.8%	26.4%	25.7%	-5.9%	-5.8%	-5.9%	-5.8%
Straight grade	39.4%	40.1%	-3.8%	-3.8%	-3.8%	-3.8%	-3.8%	-3.8%
Two-way divided	59.7%	65.5%	-5.4%	-5.7%	-5.4%	-5.7%	-5.4%	-5.7%
Crash								
Speeding-involved	297.2%	309.1%	60.7%	62.2%	-38.0%	-37.8%	-38.0%	-37.8%
Pedestrian in crosswalk	-15.6%	-15.8%	43.4%	43.4%	3.6%	3.9%	-15.6%	-15.8%
Off roadway	28.4%	29.4%	-32.7%	-33.5%	-17.5%	-18.1%	28.4%	29.4%
Walking along roadway	-40.6%	-41.1%	5.5%	5.4%	5.5%	5.4%	5.5%	5.4%
Motorist turning/merging	36.1%	37.6%	-59.9%	-61.2%	-11.1%	-11.9%	36.1%	37.6%
Motorist backing	31.9%	33.4%	-47.0%	-48.9%	-13.9%	-14.6%	31.9%	33.4%
Both driver and ped. fault	109.9%	113.9%	-9.6%	-9.5%	-9.6%	-9.5%	-9.6%	-9.5%
Pedestrian fault	363.6%	386.4%	28.2%	28.4%	4.4%	4.3%	-39.6%	-39.1%
Mean Probability	10.2%	10.2%	18.8%	18.8%	34.0%	34.0%	37.0%	37.0%

MNL is the multinomial logit model and HET is the heteroskedastic logit model. Age variables marked with * are continuous variables and therefore have true direct elasticity values (8). The other variables are indicator variables and have pseudo-elasticities (9). Shading indicates an elastic variable (elasticity greater than 1) or a percentage change greater than 100% for pseudo-elasticities.

as pedestrian age increases, the variance of the unobserved term increases, as interpreted by (3). This result corresponds with our hypothesis that unobserved effects of pedestrians vary more widely as age increases. We also tested pedestrian gender as a factor for heteroskedasticity but it remained insignificant.



MNL is the multinomial logit model and HET is the heteroskedastic logit model. The observation bullets note the percentage share of observed pedestrians at each age that were fatally injured. The lines were drawn to better separate the model predictions from the observations but they do not imply a continuum, predictions were calculated only for integer ages.

Fig. 3. Prediction of fatal injury probabilities for observed pedestrians.

4.2. Driver characteristics

As driver age increased, the propensities of injuries decreased (Table 3). As a group, older drivers are more cautious than other age groups, and they drive more on lower speed limit roads. Drivers aged 65–74 and 75+ had lower shares of fatal injury and incapacitating injury than the other age groups of drivers (Table 1). Male drivers were more likely to be involved in serious pedestrian–vehicle crashes. Intoxicated drivers increased the probability of fatal injury 167.8% (nearly a tripling of the probability of fatality) and incapacitating injury by 24.4%. Pedestrian intoxication was not statistically significant in this study.

4.3. Temporal characteristics

During the PM peak (15:00–17:59), the probability of fatal injury was decreased by about 39% (shown in Table 3). During the peak traffic period, greater congestion may lead to slower speeds, or exposure effects such as walking occurs more on minor roads or places where there are less severe conflicts between pedestrians and vehicles (e.g., parking lot, driveway). Previous research (Campbell et al., 2004) has shown that fatal pedestrian–vehicle crashes tend to occur during night time. This trend was also shown in this data. The effects of night were captured by light conditions (Section 4.6).

4.4. Control characteristics

Traffic signals are linked with a decrease in the probability of fatal injury (−36.3%), whereas traffic signs show an increase (7.1%), see Table 3. Both pedestrians and drivers have a clearer indication of their right of way at locations with traffic signals. However, the right of a pedestrian at a traffic sign is perhaps not as obvious, or honored, as at a traffic signal. Pedestrians and drivers should therefore be more cautious at intersections governed by traffic signs.

4.5. Land development characteristics

Commercial areas had a high frequency of pedestrian–vehicle crashes (51.9% in Table 1) and are linked with slightly increased probability of fatal injury and possible or no injury (13.9% and 13.9% in Table 3, respectively). This suggests a need to develop a more pedestrian friendly environment in commercial areas. Also, the result indicates a complex relationship, which needs to be tackled in future research into the interplay between the built environment and pedestrian–vehicle crashes. Such research requires more detailed information about land development than is available here.

4.6. Environmental characteristics

Inclement weather decreased the probability of fatal injury (−35.0%), in Table 3. Inclement weather includes rain, snow, fog, and smog. The reduction in the probability of fatal injury indicates that drivers and pedestrians may adjust their behavior to the conditions (i.e. they compensate and show appropriately more caution in the rain).

Darkness, with or without streetlights, leads to a significant increase in the probability of fatal injury (148.0%) and incapacitating injury (338.1%) for pedestrians. About 42% of total pedestrian–vehicle crashes occurred during darkness with or without streetlights (Table 1). Lighting conditions affect crash risk but this shows visibility also affects injury severity since a driver's ability to notice pedestrians is reduced by darkness, which can lead to drivers braking later or taking less effective avoidance maneuvers, leading to greater severity if a crash occurs.

4.7. Vehicle characteristics

Being struck by a sport-utility vehicle increased the pedestrian's probability of fatal injury (83.4% in Table 3) compared with an otherwise identical crash with a passenger car. Trucks had a larger effect on fatal injury probabilities (265.3% in Table 3). With a larger mass and a longer stopping distance, sport-utility vehicles and, in particular, trucks have greater momentum, which leads to greater impacts and more serious injury. Matsui (2005) pointed out that bumper height was associated with type of injury which can affect severity. Ballesteros et al. (2004) emphasized that not only larger vehicle mass but also vehicle design contributed to different injury patterns. Previous research and our findings confirm that when considering vehicle safety, we cannot look only to vehicle occupants but need to consider vulnerable crash opponents such as pedestrians.

4.8. Geometry characteristics

The various roadway types were associated with pedestrian injury severity. Freeway, US route, and state route, including secondary state route, increased the probability of fatal injury, compared with local city streets (Table 3). This result indicates quite naturally that higher speed roadways are associated with more severe pedestrian injury.

Crashes occurring on a curve were found positively correlated with incapacitating injury (25.7% in Table 3). Pedestrian–vehicle crashes at curves are however relatively infrequent in the data (6.1% in Table 3), lack of observations can be the reason that fatal injuries do not show up as significant. Straight roadway sections with an up/downgrade increased fatal injury and decreased the other injuries (Table 3). This can be related to different driving behavior on sloped roads, and to different angles between the pedestrian and vehicle in a crash. Downgrades can be associated with higher speed, but these data do not allow separation between up- and downgrades.

Generally two-way divided roads have better geometric design but if a pedestrian–vehicle crash happens, a pedestrian is more likely to suffer from fatal injury (65.5% increase in Table 3) most likely due to higher speed on two-way divided roads.

4.9. Crash characteristics

Intuitively, injury severity is strongly related with speed. However, we were forced to exclude estimated vehicle speed and posted speed limit because the distribution of missing values in those variables was biased in the data. Instead, an indicator variable termed speeding-involved, which tells whether or not the police indicated that speeding was a factor in the crash, was employed. The roadway classes correspond to speeds to some degree. When speeding was involved in a pedestrian–vehicle crash, the likelihoods of fatal injury and incapacitating injury were increased: 309.1% and 62.2%, respectively (Table 3). This captures some of the speed effect, but surely not all. We have discussed potential correlations with high-speed limit in other variables that can be capturing this effect, such as two-way divided roadways, and the roadway classes.

When pedestrians were crossing streets at crosswalks, there was an increase of incapacitating injury and non-incapacitating injury and a decrease of fatal injury and possible or no injury. This result is reasonable because drivers are alerted to the potential presence of pedestrians at crosswalks with signs and other indicators. Crosswalks appear effective in reducing the probability of fatal injuries.

An interesting finding is that the probability of fatal injury increased when a pedestrian–vehicle crash occurred off roadway (29.4% in Table 3). The frequency of such crashes was not negligible (13.3%). There is a need for further research that classifies off roadway locations in detail, such as into parking lots, or driveways, etc.

Vehicle maneuvers such as turning or backing had a U-shape effect: the probabilities of fatal injury and possible or no injury increased, while those of incapacitating injury and non-incapacitating injury decreased (Table 3). In the data, there were few observations for fatal injury when a vehicle turned or backed which likely explains the lack of significance of the coefficient for fatal injury, which made that category inseparable from the base category and resulted in the uptick.

In terms of fault, the model results, and the percentage changes in probability from Table 3, indicate that pedestrians found solely at fault were associated more strongly with fatal injury (386.4%) than drivers found solely at fault. Fatally injured pedestrians cannot disagree with the driver's description of events, which may explain some of this result.

5. Conclusions

Walking is one of the most basic and necessary modes of transportation. Hence, it is important to contribute to a safer pedestrian environment which would encourage people to walk more often. This study analyzed factors associated with pedestrian injury severity in pedestrian–vehicle crashes. The results show that a pedestrian's age was a significant contributor to heteroskedasticity by increasing the variance in the error terms across pedestrians with age, but that gender did not affect the variation. The results showed (Fig. 2) that the age-specific heteroskedasticity becomes increasingly important past age 65.

The results also identify important factors that significantly increase the probability of fatal injury for pedestrians: intoxicated driver, darkness with or without streetlights, greater pedestrian age, sport-utility vehicle, truck, freeway, US route, state route, and speeding-involved. Important variables that decrease the probability of fatal injury for pedestrians are the PM peak (15:00–17:59), traffic signal control, and inclement weather.

This study provides direct policy recommendations to promote pedestrian safety. Darkness is associated with greater severity. Reflectors are therefore not only helpful to reduce the probability of a crash in the first place but also to reduce severity. Successful campaigns to increase pedestrian use of reflectors are likely to have significant benefits for safety. Drunk drivers are an important issue. Not only from a crash frequency standpoint, but as shown here, intoxicated drivers are strongly associated with greater severity in pedestrian crashes. Commercial areas need to be targeted to improve pedestrian safety. Also, pedestrian–vehicle crashes off roadways (such as in parking lots, driveways, etc.) need to receive more attention. Separation between pedestrian traffic and truck corridors is important due to the significantly greater severity in pedestrian–truck crashes, compared to pedestrian–car crashes. The greater fatality probabilities associated with sport-utility vehicles indicate that passenger vehicle safety should in part be determined by the effect the vehicle has on pedestrians, not only vehicle occupants.

Acknowledgments

The authors gratefully acknowledge the assistance of the Highway Safety Research Center at the University of North Carolina, which provided the data for this study. The study was supported in part by the Department of Civil Engineering, Washington University in St. Louis.

References

Al-Ghamdi, A.S., 2002. Pedestrian–vehicle crashes and analytical techniques for stratified contingency tables. Accident Analysis and Prevention 34 (2), 205–214.
Anderson, R.W.G., McLean, A.J., Farmer, M.J.B., Lee, B.H., Brooks, C.G., 1997. Vehicle travel speeds and the incidence of fatal pedestrian crashes. Accident Analysis and Prevention 29 (5), 667–674.
Ben-Akiva, M.E., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge MA.
Bhat, C.R., 1995. A heteroscedastic extreme value model of intercity travel mode choice. Transportation Research Part B 29 (6), 471–483.
Ballesteros, M.F., Dischinger, P.C., Langenberg, P., 2004. Pedestrian injuries and vehicle type in Maryland 1995–1999. Accident Analysis and Prevention 36 (1), 73–81.
BTS, 2006. Distribution of Transportation Fatalities by Mode. Bureau of Transportation Statistics, Department of Transportation, Washington, DC.
Campbell, B.J., Zegeer, C.V., Huang, H.H., Cynecki, M.J., 2004. A Review of Pedestrian Safety Research in the United States and Abroad. Publication FHWA-RD-03-042, FHWA, U.S. Department of Transportation.
daSilva, M.P., Smith, J.D., Najm, W.G., 2003. Analysis and Pedestrian Crashes. HS-809 585. U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC.
Davis, G.A., 2001. Relating severity of pedestrian injury to impact speed in vehicle–pedestrian crashes: simple threshold model. Transportation Research Record 1773, 108–113.
Davis, G.A., Sanderson, K., Davuluri, S., 2002. Development and Testing of a Vehicle/Pedestrian Collision Model for Neighborhood Traffic Control. Minnesota Department of Transportation, St. Paul, Final Report 2002-23.
DeLucia, P.R., Bleckley, M.K., Meyer, L.E., Bush, J.M., 2003. Judgment about collision in younger and older drivers. Transportation Research Part F 6 (1), 63–80.
Fontaine, H., Gourlet, Y., 1997. Fatal pedestrian accidents in France: a typological analysis. Accident Analysis and Prevention 29 (3), 303–312.
Gårder, P.E., 2004. The impact of speed and other variables on pedestrian safety in Maine. Accident Analysis and Prevention 36 (4), 533–542.
Greene, W., 2003. Econometric Analysis, 5th ed. Prentice Hall.
Harrell, W.A., 1991. Precautionary street crossing by elderly pedestrians. International Journal of Aging and Human Development 32 (1), 65–81.
Harruff, R., Avery, A., Alter-Pandya, A., 1998. Analysis of circumstances and injuries in 217 pedestrian traffic fatalities. Accident Analysis and Prevention 30 (1), 11–20.
Jensen, S.U., 1999. Pedestrian safety in Denmark. Transportation Research Record 1674, 61–69.
Kim, J.-K., 2007. Statistical Methods in Transportation: Application of Behavioral Models for Safety Analysis of Vulnerable Groups. D.Sc. dissertation, Washington University in St. Louis, UMI Dissertation Publishing, p. 182.
Kim, J.-K., Kim, S., Ulfarsson, G.F., Porrello, L.A., 2007. Bicyclist injury severities in bicycle-motor vehicle accidents. Accident Analysis and Prevention 39 (2), 238–251.
Kockelman, K.M., Kweon, Y., 2002. Driver injury severity: an application of ordered probit models. Accident Analysis and Prevention 34 (3), 313–321.
Koepsell, T., McCloskey, L., Wolf, M., Moudon, A., Buchner, D., Kraus, J., Patterson, M., 2002. Crosswalk markings and the risk of pedestrian–motor vehicle collisions in older pedestrians. Journal of the American Medical Association 288 (17), 2136–2143.
Leden, L., Gårder, P., Johansson, C., 2006. Safe pedestrian crossings for children and elderly. Accident Analysis and Prevention 38 (2), 289–294.
Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accident Analysis and Prevention 37 (4), 775–786.
Lefler, D.E., Gabler, H.C., 2004. The fatality and injury risk of light truck impacts with pedestrians in the United States. Accident Analysis and Prevention 36 (2), 295–304.
McFadden, D., 1981. In: Manski, C., McFadden, D. (Eds.), Econometric Models of Probabilistic Choice. Structural Analysis of Discrete Data Using Econometric Applications. MIT Press, Cambridge, MA.
McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15, 447–470.
Maring, W., van Schagen, I., 1990. Age dependence of attitudes and knowledge in cyclists. Accident Analysis and Prevention 22 (2), 127–136.
Matsui, Y., 2005. Effects of vehicle bumper height and impact velocity on type of lower extremity injury in vehicle–pedestrian accidents. Accident Analysis and Prevention 37 (4), 633–640.
McMahon, P.J., Zegeer, C.V., Duncan, C., Knoblauch, R.L., Stewart, R., Khattak, A.J., 2002. An Analysis of Factors Contributing to "Walking Along Roadway" Crashes: Research Study and Guidelines for Sidewalks and Walkways. Department of Transportation, Federal Highway Administration, VA.
Miles-Doan, R., 1996. Alcohol use among pedestrians and the odds of surviving an injury: evidence from Florida law enforcement data. Accident Analysis and Prevention 28 (1), 23–31.
NHTSA, 2004. Traffic Safety Facts 2004 Data: Pedestrians. National Highway Traffic Safety Administration, Washington, DC.
Öström, M., Eriksson, A., 2001. Pedestrian fatalities and alcohol. Accident Analysis and Prevention 33 (2), 173–180.
Oxley, J., Fildes, B., Ihsen, E., Charlton, J., Day, R., 1997. Difference in traffic judgements between young and old adult pedestrians. Accident Analysis and Prevention 29 (6), 839–847.
Oxley, J.A., Ihsen, E., Fildes, B.N., Charlton, J.L., Day, R.H., 2005. Crossing roads safely: an experimental study of age differences in gap selection by pedestrians. Accident Analysis and Prevention 37 (5), 962–971.
Pasanen, E., Salmivaara, H., 1993. Driving speeds and pedestrian safety in the city of Helsinki. Traffic Engineering and Control 34 (6), 308–310.
Preusser, D.F., Wells, J.K., Williams, A.F., Weinstein, H.B., 2002. Pedestrian crashes in Washington, DC and Baltimore. Accident Analysis and Prevention 34 (5), 703–710.
Shankar, V., Mannering, F., 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. Journal of Safety Research 27 (3), 183–194.
Shankar, V., Mannering, F., Barfield, W., 1996. Statistical analysis of accident severity on rural freeways. Accident Analysis and Prevention 28 (3), 391–401.
Shankar, V.N., Ulfarsson, G.F., Pendyala, R.M., Nebergall, M.B., 2003. Modeling crashes involving pedestrians and motorized traffic. Safety Science 41, 627–640.
Sklar, D., Demarest, G., McFeeley, P., 1989. Increased pedestrian mortality among the elderly. The American Journal of Emergency Medicine 7 (4), 387–390.
Ulfarsson, G.F., Mannering, F.L., 2004. Difference in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accident Analysis and Prevention 36 (2), 135–147.
Zajac, S.S., Ivan, J.N., 2003. Factors influencing injury severity of motor vehicle-crossing pedestrian crashes in rural Connecticut. Accident Analysis and Prevention 35 (3), 369–379.
Zeeger, C.V., Stutts, J.C., Rodgman, E., 1996. Analysis of elderly pedestrian accidents and recommended countermeasures. Journal of Safety Research 27 (2), 128.
Zeng, L.A., 2000. Heteroscedastic generalized extreme value discrete choice model. Sociological Methods and Research 29 (1), 119–145.