Accident Analysis and Prevention 39 (2007) 955–963

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes

Peter Savolainen[a], Fred Mannering[b,*]

[a] Department of Civil Engineering, Wayne State University, Detroit, MI 48202, United States
[b] Department of Civil Engineering, Purdue University, West Lafayette, IN 47907, United States

Received 4 August 2006; received in revised form 14 December 2006; accepted 22 December 2006

Abstract

Motorcycle fatalities have more than doubled in the United States since 1997—highlighting the need to better understand the many interrelated factors that determine motorcyclists' crash-injury severities. In this paper, using a detailed crash database from the state of Indiana, we estimate probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Nested logit (estimated with full information maximum likelihood) and standard multinomial logit model results show a wide-range of factors significantly influence injury-severity probabilities. Key findings show that increasing motorcyclist age is associated with more severe injuries and that collision type, roadway characteristics, alcohol consumption, helmet use, unsafe speed and other variables play significant roles in crash-injury outcomes.

Keywords: Crash injuries; Motorcycle crashes; Nested logit models; Motorcycle fatalities

1. Introduction

In the US, motorcycle crash fatalities have been on a surprising upward trend. From an historic low of 2116 fatalities in 1997, fatalities have increased each year since, reaching a projected 4315 in 2005 (National Highway Traffic Safety Administration, 2006). There are potentially many reasons for this observed increase in fatalities. Motorcycle registrations have climbed significantly during the 1997–2005 time period; many states have repealed their helmet laws; and the average age of motorcyclists has increased with reflex and skill degradation a concern. With many factors likely involved in this disturbing upward trend in fatalities, there is a clear need for research to provide some insight into the relative effect of various factors so that appropriate countermeasures can be implemented to save lives.

The intent of this study is to develop probabilistic models of motorcycle crash-injury severity in an effort to provide additional insight into the factors that determine injuries to motorcyclists. We use data from the state of Indiana, whose fatality trends have followed national trends closely. With a database of all crashes reported to Indiana police from January 1, 2003 to October 15, 2005, we develop detailed multivariate analyses to determine the factors that significantly influence crash-injury severity.

In terms of appropriate statistical methods for the multivariate analysis of crash-injury severity data, recent literature indicates that a variety of methods have been used with an emphasis on car and truck crashes. For some examples, Abdel-Aty et al. (1998) used log-linear models to examine the relationship between driver age and crash characteristics, including severity; Farmer et al. (1997) used a binomial regression model to investigate the impact of vehicle and crash characteristics on injury severity in two-vehicle, side-impact crashes; O'Donnell and Connor (1996) assessed the probabilities of four levels of injury severity as a function of driver attributes using ordered logit and ordered probit specifications; Kockelman and Kweon (2002) used ordered probability models to investigate separate datasets for all crashes, two-vehicle crashes and single-vehicle crashes; Carson and Mannering (2001) developed multinomial logit models to examine the effect of ice-warning signs on crash-injury severity for different roadway functional classes; and Ulfarsson and Mannering (2004) explored differences in injury severity between male and female drivers in single and two-vehicle crashes involving passenger cars, pickups, sport-utility vehicles, and minivans using multinomial logit models.

* Corresponding author.
E-mail addresses: savolainen@wayne.edu (P. Savolainen), flm@purdue.edu (F. Mannering).

0001-4575/$ – see front matter
doi:10.1016/j.aap.2006.12.016

Among the extant literature, at least two recent studies focused exclusively on motorcycle crash severity. Shankar and Mannering (1996) considered environmental, roadway, vehicular, and rider characteristics in their multinomial logit analysis of single-vehicle motorcycle crash-injury severity. And, Quddus et al. (2002) utilized ordered probit models to analyze motorcycle damage and injury severity resulting from crashes.

The review of the literature shows that two preferred approaches have emerged in the statistical modeling of crash-injury severity data, ordered probability models (ordered logit and probit) and unordered probability models (multinomial and nested logit). Because injury levels are typically progressive (ranging from no-injury to fatality), ordered probability models would seem to be a natural choice to account for the ordering of injury severities. O'Donnell and Connor (1996), Duncan et al. (1998), Renski et al. (1999), Khattak (2001), Kockelman and Kweon (2002), Khattak et al. (2002), Kweon and Kockelman (2003), Abdel-Aty (2003), Yamamoto and Shankar (2004) are some of the many that have used this technique. However, there are at least two potential problems with applying ordered probability models to injury-severity analysis. The first relates to the fact that non-injury crashes may be under-represented in police-reported crash data since lower injury levels make reporting to authorities less likely. The presence of underreporting in an ordered probability model can result in biased and inconsistent model coefficient estimates.[1] In contrast, the coefficient estimates of an unordered multinomial logit probability model are consistent except for the constant term (see McFadden, 1981; Washington et al., 2003).

The second problem is more difficult to correct and relates to the restriction that ordered probability models place on variable influences. To see this, we follow the example used in Washington et al. (2003). Consider the effect that air-bag deployment has on occupant injury with four possible injury outcomes: no injury, non-incapacitating injury, incapacitating injury and fatality. The ordered probability model constrains the airbag-deployment indicator variable to either increase the probability of fatality (and subsequently decrease the probability of no injury) or decrease the probability of fatality (and subsequently increase the probability of no injury). This precludes the possibility that airbag deployment may simultaneously increase (or decrease) the probabilities of a fatality and no-injury because the deployment may save lives but may also cause minor injuries (non-incapacitating injuries) in doing so. Unordered probability approaches do not impose this constraint and such models have been applied in crash-injury severity analysis by numerous researchers (Shankar et al., 1996; Chang and Mannering, 1999; Carson and Mannering, 2001; Lee and Mannering, 2002; Ulfarsson and Mannering, 2004; Khorashadi et al., 2005) and will be the approach adopted in this study.

[1] Correcting for the possibility of such underreporting in ordered probability models is more involved than a simple correction of the constant term, but can be done with a weighted likelihood function if true population portions are known (see Manski and Lerman, 1977).

2. Methodology

In applying an unordered probability model to assess motorcyclists' injury severity, we begin by defining a linear function that determines motorcyclist n's injury-severity outcome i as,

$$S_{in} = \boldsymbol{\beta}_i \boldsymbol{X}_{in} + \varepsilon_{in} \tag{1}$$

where $\boldsymbol{X}_{in}$ is a vector of measurable characteristics (motorcyclist characteristics, roadway characteristics, and so on) that determine the injury severity for motorcyclist n, $\boldsymbol{\beta}_i$ a vector of estimable coefficients, and ε_{in} is an error term accounting for unobserved effects influencing the injury severity of crash n. McFadden (1981) has shown that if ε_{in} are assumed to be generalized extreme value distributed the standard multinomial logit model results,

$$P_n(i) = \frac{\exp[\boldsymbol{\beta}_i \boldsymbol{X}_{in}]}{\sum_{\forall I} \exp(\boldsymbol{\beta}_I \boldsymbol{X}_{In})} \tag{2}$$

where $P_n(i)$ is the probability that crash n will result in motorcyclist injury outcome i and I is the set of possible crash-injury-severity outcomes. Eq. (2) is estimable by standard maximum likelihood techniques.

The generalized extreme value distribution can also be used to generate a family of models that includes the nested logit model (McFadden, 1981), which can overcome the restriction of the standard multinomial logit model that requires the assumption that the error terms (ε_{in}'s) are independently distributed across alternate outcomes This independence may not always be the case if some crash-injury severity levels share unobserved effects. For example, with the four injury categories we will consider in this paper (no injury, non-incapacitating injury, incapacitating injury and fatality), it is possible that no injury and non-incapacitating injuries may share unobserved effects that relate to lower-impact collisions, thus violating the assumption that the error terms are independently distributed across outcomes. The nested logit model resolves this by grouping alternatives that share unobserved effects into conditional nests. The outcome probabilities are determined by differences in the functions determining these probabilities with shared unobserved effects canceling out in each nest. The nested logit model has the following structure for motorcyclist n's crash resulting in injury outcome i

$$P_n(i) = \frac{\exp[\boldsymbol{\beta}_i \boldsymbol{X}_{in} + \phi_i LS_{in}]}{\sum_{\forall I} \exp[\boldsymbol{\beta}_I \boldsymbol{X}_{In} + \phi_I LS_{In}]} \tag{3}$$

$$P_n(j|i) = \frac{\exp[\boldsymbol{\beta}_{j|i} \boldsymbol{X}_n]}{\sum_{\forall J} \exp[\boldsymbol{\beta}_{J|i} \boldsymbol{X}_{Jn}]} \tag{4}$$

$$LS_{in} = LN\left[\sum\nolimits_{\forall J} \exp(\boldsymbol{\beta}_{J|i} \boldsymbol{X}_{Jn})\right], \tag{5}$$

where $P_n(i)$ is the unconditional probability of motorcyclist n having injury outcome i, $\boldsymbol{X}$'s vectors of measurable characteristics that determine the probability of injury outcomes, $\boldsymbol{\beta}$'s vectors of estimable coefficients, and $P_n(j|i)$ is the probability of motorcyclist n's crash having injury severity j conditioned on the injury severity being in injury-severity category i. For example, for a nested structure that assumes correlation among minimal injury outcomes (no-injury and non-incapacitating) the outcome category i would be “minor or no-injury” and $P_n(j|i)$ would be the binary logit model of the no-injury and non-incapacitating injury outcomes. Continuing, J is the conditional set of outcomes (conditioned on i), I is the unconditional set of outcome categories (minor or no-injury, incapacitating injury and fatality in the above example), LS_{in} is the inclusive value (logsum), and ϕ_i is an estimable parameter.

Estimation of a nested model is usually done in a sequential fashion where the procedure is first to estimate the conditional model (Eq. (4)) using only the observations in the sample that are observed having the subset of crash-injury outcomes J. Once these estimation results are obtained, the logsum is calculated (this is the denominator of one or more of the conditional models—see Eq. (5)) for all observations, both those resulting in injury-severity J and those not (for all crashes). Finally, these computed logsums are used as independent variables in the functions as shown in Eq. (3). However, this sequential estimation procedure has been shown to generate variance–covariance matrices that are too small and thus t-statistics are inflated (typically by about 10–15%). The alternative to this is to write a single likelihood function to estimate the entire model (all nests) simultaneously. While more computationally intensive, this full information maximum likelihood approach ensures that variance–covariance matrices are properly estimated. We use this full information maximum likelihood approach in our model estimations (see Greene, 2000 for additional details).

In comparing nested and un-nested logit models, it is important to note that if the estimated value of ϕ_i is not significantly different from 1, the assumed shared unobserved effects in the lower-nest are not significant and the nested model reduces to a simple multinomial logit model (see Eqs. (3)–(5) with ϕ_i's = 1). We will test for this possibility in our forthcoming estimations.

Finally, to assess the vector of estimated coefficients ($\boldsymbol{\beta}_i$), we calculate elasticities which measure the magnitude of the impact of specific variables on the injury-outcome probabilities. The elasticity is computed for each motorcyclist n (n subscripting omitted) as

$$E_{x_{ki}}^{P(i)} = [1 - P(i)]\beta_{ki}x_{ki}, \tag{6}$$

where β_{ki} is the estimated coefficient associated with variable x_{ki}. Elasticity values can be roughly interpreted as the percent effect that a 1% change in x_{ki} has on the injury-severity-outcome probability $P(i)$.

Note that elasticities are not applicable to indicator variables (those variables taking on values of 0 or 1). In these cases, a pseudo-elasticity can be calculated in percent as

$$E_{x_{ki}}^{P(i)} \left[\frac{\exp[\Delta(\boldsymbol{\beta}_i \boldsymbol{X}_i)]\sum_{\forall I}\exp(\boldsymbol{\beta}_{kI}x_{kI})}{\exp[\Delta(\boldsymbol{\beta}_i \boldsymbol{X}_i)]\sum_{\forall I_n}\exp(\boldsymbol{\beta}_{kI}x_{kI}) + \sum_{\forall I \neq I_n}\exp(\boldsymbol{\beta}_{kI}x_{kI})} - 1 \right] \times 100, \tag{7}$$

where I_n is the set of alternate injury-severity outcomes with x_k in the function determining the outcome, and I is the set of all possible injury-severity outcomes. The pseudo-elasticity of a variable with respect to a motorcyclist injury-severity category represents the percent change in the probability of that injury-severity category when the variable is changed from zero to one. Thus, a pseudo-elasticity of 95% for a variable in the no-injury category means that when the values of the variable in the subset of observations where $x_k = 0$ are changed from 0 to 1, the probabilities of the no-injury outcome for these observations increased, on average, by 95%. See Washington et al. (2003) for a complete explanation of elasticities.

3. Data

Our data is drawn from all police-reported motorcycle crashes in the state of Indiana between January 1, 2003 and October 15, 2005. We also combined the Indiana State Police crash database with rider training records obtained from the American Bikers Aimed Toward Education (ABATE) of Indiana to provide a more comprehensive dataset for analysis. Although police provide estimates of crash-injury severity at the scene of the crash in one of five categories (no-injury, possible injury, non-incapacitating injury, incapacitating injury and fatality), we went through individual crash reports to note the location and type of injury for each person involved in the crash (as noted by the reporting officer). From this analysis, it was apparent that, due to variations in officer reporting, we could not adequately distinguish between no-injury and possible injury categories and, consequently, both were combined into a single no-injury category for the subsequent analyses (leaving four possible outcomes: no-injury, non-incapacitating injury, incapacitating injury and fatality).[2]

The data include a rich set of information including: an environmental record containing details related to the circumstances leading up to and resulting from the crash; a person record containing demographic information on all involved persons, as well as details on injuries suffered and protective equipment worn; and a vehicle record that includes details on the vehicle type, make, and model.

Finally, for our empirical analysis, we separate single- and multi-vehicle because of the substantially different causality mechanisms and factors involved in these two crash types.

4. Single-vehicle crash-injury severity

A total of 2273 single-vehicle motorcycle crashes (occurring between January 1, 2003 and October 15, 2005 in the state

[2] This was tested by estimating a model with five of the alternate outcome categories (no-injury, possible injury, non-incapacitating injury, incapacitating injury and fatality) and one with the no-injury and possible injury categories combined. To determine the most appropriate of these two models, the test statistic used was the likelihood ratio test with $X = -2[\mathrm{LL}(0) - \mathrm{LL}(\beta)]$, where LL(0) is the log-likelihood with all parameters at zero and $\mathrm{LL}(\beta)$ is the log-likelihood at convergence. This statistic is χ^2 distributed with the degrees of freedom equal to the number of estimated parameters. The model with the no-injury and possible-injury categories combined produced the χ^2 with the highest p-value. Also, with the full five-alternative outcome model, we were unable to find any statistically significant variables to distinguish between the no injury and possible-injury categories.

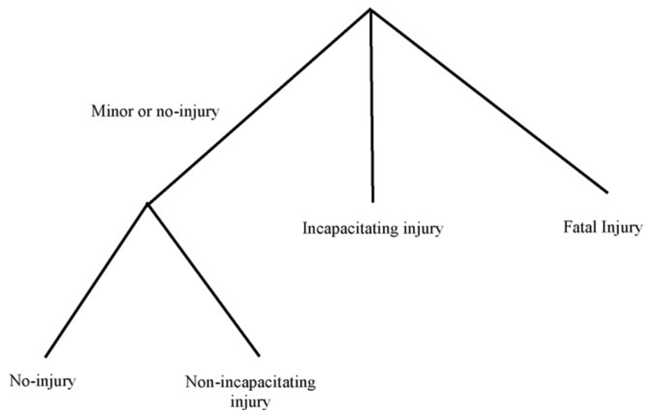

Fig. 1. Nested logit structure of the single-vehicle crash-injury severity model.

Table 1
Single-vehicle crash summary statistics

Variable	Value
Rider characteristics and behavior	
Average rider age in years (standard deviation)	38.32 (13.37)
Percent of female riders	9.1
Percent of motorcycles less than 5 years old	49.2
Percent of riders taking the Basic Rider Course over 2 years ago	3.0
Percent of crashes with alcohol involved	5.5
Percent of riders wearing helmet	41.0
Percent of crashes with passenger	16.6
Percent of motorcyclists cited for speeding	14.8
Roadway characteristics	
Percent of crashes on wet pavement	3.0
Percent of crashes at intersections	21.3
Percent of crashes on roads with speed limit over 50 mph	32.7
Percent of crashes occurring on a horizontal curve	39.3
Crash characteristics	
Percent of crashes occurring in April	9.3
Percent of crashes occurring in July	16.2
Percent that are collisions in darkness	25.6
Percent of crashes that are run-off-road crashes	20.7
Percent that are collisions with a tree	1.8
Percent that are collisions with a pole	3.0
Percent that are collisions with a curb	5.3
Percent that are collisions with a culvert	1.1
Percent that are collisions with guardrail	3.9

Table 2
Full-information maximum likelihood estimation results for single-vehicle rider crash-injury severity

Severity level	Variable	Coefficient estimate	*t*-Ratio
No-injury (lower nest)	Constant	−1.368	−11.94
	Wet pavement	0.880	2.86
	Helmet used	0.529	4.50
	Crash at intersection	0.380	2.82
	Motorcycle age less than 5 years	0.278	2.38
Non-incapacitating injury (lower nest)	Speeding	0.674	3.49
	Alcohol	1.445	3.33
	Passenger	0.658	3.80
	Female	0.773	3.50
Incapacitating injury	Constant	−2.749	−10.90
Fatality	Constant	−5.702	−15.10
	Speeding	1.301	3.96
	April	0.959	2.61
	July	0.909	2.98
	Over 2 years since Basic Riding Course	1.262	2.32
	Run-off-road	1.231	3.95
	Collision with tree	1.487	3.16
	Collision with pole	1.388	3.13
	Darkness	0.700	2.55
Minor or no-injury	Collision with curb	−0.645	−2.63
	Collision with pole	−0.820	−2.53
	Collision with culvert	−1.425	−3.17
	Horizontal curvature	−0.324	−2.62
	Age	−0.021	−4.38
	Alcohol	−1.032	−2.49
	Speeding	−0.676	−2.96
	Helmet	0.618	4.55
	Darkness	−0.347	−2.34
	Collision with guardrail	−0.738	−2.82
	Collision with tree	−1.462	−3.59
	Speed limit over 50 mph	−0.423	−3.30
	Inclusive value	0.421	1.70
Number of observations		2273	
Log Likelihood at zero		−3481.50	
Log Likelihood at convergence		−1967.66	
ρ^2		0.43	

of Indiana) were used for model estimation. We consider only motorcycle-operator injury levels (not passengers) and, of these crashes, 20% resulted in no-injury, 62% in non-incapacitating injury, 15% in capacitating injury and 3% in fatality. Using these data, we develop models to estimate the probability of the four discrete driver-injury severity outcomes conditioned on a crash having occurred and having been reported to police.

We ran extensive tests on all possible logit-model structures including all nesting possibilities.[3] Our findings indicate that the appropriate model structure for single-vehicle crash-injury severity is a nested model with no-injury and non-incapacitating injury sharing unobserved effects as illustrated in Fig. 1.[4] Summary statistics for all variables included in the model are shown

[3] For model estimation, LIMDEP Version 7.0 was used. Likelihood ratio tests were conducted to determine the best model structure, based on the highest χ^2 statistics and associated confidence levels (see Washington et al., 2003).

[4] There is a possibility that the outcome data (no-injury, possible injury, non-incapacitating injury, incapacitating injury and fatality) may be incorrectly recorded by the officer at the scene. For example, an injury may not be detectable until later. Hausman et al. (1998) developed a correction procedure for this possibility for discrete models. A modified version of this procedure was applied to police-reported crash data by Winston et al. (2006) and they found that such misclassification was not statistically significant, indicating that standard model-estimation methods are appropriate.

Table 3
Single-vehicle elasticities in percent

Variable	No-injury	Non-incapacitating injury	Incapacitating injury	Fatality	Minor or no-injury
Rider characteristics and behavior					
Age					−1.15
Female		20			
Motorcycle age less than 5 years	20				
Over 2 years since Basic Riding Course				171	
Alcohol		12			−10
Helmet used	50				11
Passenger		18			
Speeding		−2		212	−14
Roadway characteristics					
Wet pavement	77				
Crash at Intersection	29				
Speed limit over 50 mph					−10
Horizontal curve					−8
Crash characteristics					
April				111	
July				98	
Darkness				95	−9
Run-off-road				137	
Collision with tree				525	−43
Collision with pole				344	−25
Collision with curb					−15
Collision with culvert					−35
Collision with guardrail					−17

in Table 1 and full-information maximum likelihood model estimation results are presented in Table 2 with corresponding elasticities in Table 3. The estimation results in Table 2 show that a wide variety of variables were found to be statistically significant, and that the overall model fit is quite good with a ρ^2 of 0.43.[5]

Considering the elasticities presented in Table 3 and turning first to rider characteristics and behavior, our results show that increasing age is less likely to result in minor or no-injury crashes (a 1% increase in age results in a 1.15% decline in the probability of these injury severity outcomes). This would seem to provide some support for the theory that recent increases in motorcycle injuries are due in some significant part to older riders. The reasons for this may relate to different behavior before the crash (for example, the late application of brakes) or may be due to physiological differences with older riders simply being more susceptible to injury. A recent study by Islam and Mannering (2006) found a similar trend in their analysis of single-vehicle car crashes, with significant differences in injury outcomes as drivers aged. In other analyses of single-vehicle motorcycle crashes, Shankar and Mannering (1996) found a similar result with older riders more likely sustain a disabling injury or be killed; and Quddus et al. (2002) also found that older motorcyclists are more likely to be severely injured.

Female riders were 20% more likely to be involved in non-capacitating injury crashes and this finding could reflect riding style and behavior relative to males. Similar gender differences have also been found in single-vehicle car crashes (Islam and Mannering, 2006). Also, owning a motorcycle less than 5 years resulted in a 20% higher probability of no injury. This finding could be an indication of the types of riders attracted to newer motorcycles or the better safety features of newer bikes (for example, improved braking which would reduce the speed at collision).

Riders that had taken the Motorcycle Safety Foundation's Beginning Rider Course over 2 years ago were 171% more likely to be fatally injured. This could be the result of a number of factors including: (1) unobserved heterogeneity in our data in that people that are taking the course are a self-selected group of less-skilled riders, (2) some degradation in learned material, and/or (3) the courses, as taught, are giving riders a "safety-skill" set that they are using to ride more aggressively instead of more safely (this is the well-known safety-compensation effect). As some support for these possibilities, Savolainen and Mannering (2006) have argued that individuals taking the Basic Riding Course may be inherently less skilled and, as learned material is forgotten, they may return to pre-course crash rates. In terms of safety-compensation, numerous researchers have shown this to be a major concern with the introduction of automobile safety features (Peltzman, 1975; Winston et al., 2006). This compensating behavior is also likely to be present when motorcyclists expand their skill-set in a safety course—that is they will extract some (and possibly all) of the benefits of such a course with increased mobility (speed) at the expense of increased safety.

5 ρ^2 is defined as $1 - LL(\beta)/LL(0)$, where $LL(\beta)$ is the log-likelihood at convergence with coefficient vector β and LL(0) is the initial log-likelihood (with all coefficients set to zero).

As an example, if learning proper application of front and rear brakes enables motorcyclists to stop in shorter distances, motorcyclists may ride faster or apply their brakes later to maintain roughly the same crash-risk level they had before their braking skills were improved.

Alcohol involvement resulted in a 10% reduction in minor or no-injury and a 12% increase in non-incapacitating injury. Helmet use resulted in a 50% increase in no-injury and an 11% increase in minor or no injury. Another notable finding is that speeding (motorcyclists cited for unsafe speed) resulted in a 212% increase in the probability of a fatality. These results are broadly similar to the earlier findings of Shankar and Mannering (1996), although they used a simple multinomial logit model structure.

With regard to roadway characteristics, the notable findings were that wet pavement and intersection crashes were more likely to result in no injury. This could be a result of lower speeds being maintained by riders in these situations as they adjust for the perceived higher risk (safety compensation along the lines discussed above).

Crash characteristics strongly influenced injury-severity outcomes as indicated by the calculated elasticities. Crashes occurring in the months of April and July had 111% and 98% greater probability of being fatal. The April finding may be the result of riders "brushing the rust" off their motorcycling skills after the winter months. The July finding may reflect, among other factors, more reckless behavior during ideal riding conditions. Also, crashes occurring in darkness were 95% more likely to be fatal.

Finally, as expected, the elasticities in Table 3 show that crashes that result in running off the road and collisions with roadside objects are much more likely to be severe, with collisions with trees and poles most likely to produce a fatality.

5. Multi-vehicle crash-injury severity

Multi-vehicle crashes involving motorcyclists have been identified as major safety concern with crash characteristics that vary considerably from those found in single-vehicle motorcycle crashes (Motorcycle Safety Foundation, 2000). Because motorcycles are less conspicuous than passenger cars or trucks, they are often more difficult to detect and their approaching speed is more difficult to determine, and many have argued that this is a major contributing factor to the high crash rate of motorcycles (Thomson, 1980; Wulf et al., 1989).

A total of 2213 multi-vehicle crashes involving motorcycles (occurring between January 1, 2003 and October 15, 2005 in the state of Indiana) were used in our injury-severity-model estimations. We consider crashes only involving exactly two vehicles to better isolate specific vehicle effects. Of the crashes considered, 53% resulted in no-injury, 26% in non-incapacitating injury, 17% in capacitating injury and 4% in fatality. It is also noteworthy that most of the crashes were related to intersection traffic movements with 75.7% classified as either right-angle, rear-end or left-turn collisions (see Table 4).

Using these data, we again develop models to estimate the probability of the four discrete driver-injury severity outcomes conditioned on a crash having occurred and having been reported to police. Unlike the single-vehicle crash case, our statistical tests show that a nested logit model formulation is not justified and we thus estimate a standard multinomial logit model as shown in Eq. (2).[6] Model estimation results are presented in Table 5 with corresponding elasticities presented in Table 6.

Table 4
Multi-vehicle crash summary statistics

Variable	Value
Rider/motorist characteristics and behavior	
Average motorcyclist age in years (standard deviation)	38.46 (13.48)
Percent of motorcyclists ages 60 years or over	5.3
Average motorcycle age in years (standard deviation)	10.11 (8.93)
Percent helmet used and sport utility vehicle collision	3.2
Percent helmet used and pickup truck collision	5.4
Percent helmet used and right-angle collision	10
Percent motorcyclist at fault	37.4
Percent other motorist at fault	58.4
Percent alcohol use (other motorist)	1.1
Percent involving sportbikes	11.3
Percent failure to yield (other motorist)	32.4
Percent of motorcyclists cited for speeding	6.0
Roadway characteristics	
Percent with speed limit over 30 mph	89.1
Percent with speed limit over 40 mph	43.6
Percent with speed limit over 50 mph	18.7
Percent with speed limit over 60 mph	1.3
Percent of crashes on interstates	3.1
Percent occurring on a vertical curve	2.5
Percent occurring on a horizontal curve	8.3
Crash characteristics	
Percent occurring in April	10.9
Percent that are head-on	6.7
Percent that are rear-end	29.4
Percent that are right-angle	30.9
Percent occurring in darkness	18.5
Percent that are collisions with motorcycle	5.6
Percent that are collisions with pickup truck	14.1
Percent that are collisions with sport utility vehicle	9.6
Percent that are collisions with tractor-trailer	3.2

Turning to Table 6 and looking first at some of the more interesting findings relating rider and motorist characteristics and their behavior, we find that increasing motorcyclists' age results in a much higher likelihood of an incapacitating injury (a 1% increase in age increases the likelihood of incapacitating injury by 4.2%). As with single-vehicle crashes, this points again to the fact that older riders are more likely to be involved in severe collisions even when controlling all other factors (crash type, motorcycle type, etc.).

Similar to the single-vehicle case, we find that newer motorcycles are less likely to result in severe crashes. In this case, each 1% increase in motorcycle age results in a 0.06% decrease in the likelihood of a no-injury crash (and thus a higher likelihood of other crash-injury-severity types).

[6] In this case, the nested logit model resulted in an estimated value of ϕ_i that was not significantly different from 1 (see Eqs. (3)–(5)). Thus, the nested model reduces to a simple multinomial logit model (see Eqs. (3)–(5) with ϕ_i's = 1).

Table 5
Multi-vehicle multinomial logit estimation results for rider crash-injury severity

Severity level	Variable	Coefficient estimate	t-Ratio
No-injury	Head-on	−0.535	−2.61
	Motorcycle age	−1.700	−3.40
	Alcohol use (other motorist)	−0.013	−2.50
	Speed limit over 30 mph	−0.389	−2.46
	Speed limit over 40 mph	−0.611	−4.44
	Speed limit over 50 mph	−0.372	−2.51
	Speed limit over 60 mph	1.597	3.37
Non-incapacitating injury	Constant	−1.253	−7.19
	Sportbike	0.379	2.44
	Collision with sport utility vehicle	−1.264	−4.22
	Collision with motorcycle	0.513	2.41
	Failure to yield (other motorist)	0.680	4.62
	Other motorist at fault	−0.532	−3.94
	Motorcyclist 60 years old or older	0.719	3.37
	Speed limit over 40 mph	−0.722	−5.18
	Interstate	0.913	2.91
	Helmet use in sport utility vehicle collisions	0.948	2.24
	Helmet use in pickup collisions	0.463	2.12
Incapacitating injury	Constant	−2.600	−9.96
	Rear-end	−0.864	−5.45
	Vertical curve	0.743	2.32
	Collision with pickup	0.560	3.55
	Failure to yield (other motorist)	0.334	2.37
	Motorcyclist age	0.013	2.87
	Horizontal curve	0.449	2.32
	Motorcyclist speeding	0.552	2.16
Fatality	Constant	−4.714	−14.59
	Right-angle	1.255	4.27
	Head-on	1.785	5.59
	Collision with tractor-trailer	1.289	3.29
	Darkness	0.727	2.89
	Motorcyclist at fault	0.854	3.44
	Speed limit over 50 mph	0.681	2.68
	April	−1.053	−1.97
	Motorcyclist speeding	0.919	2.51
	Helmet use in right-angle collisions	−0.972	−2.04
Number of observations		2092	
Log Likelihood at zero		−2900.13	
Log Likelihood at convergence		−2174.92	
ρ^2		0.25	

Helmet use is shown to be significant in increasing the probability of non-incapacitating injuries in collisions with sport utility vehicles and pickup trucks, and reducing the likelihood of fatality in right-angle collisions. Other findings show alcohol use by the other motorist increases injury severity by reducing the likelihood of no-injury by 65%; when the motorcyclist is identified as being at fault, the likelihood of fatality is 126% higher; and when the motorcyclist is cited for speeding the likelihood of fatality is 116% higher.

With regard to roadway characteristics, high-speed roads are, as expected, associated with higher injury-severity levels. Roads with speed limits exceeding 50 mph have a 132% higher likelihood of a fatal injury. Also, crashes that occur on vertical or horizontal curves are significantly more likely to result in incapacitating injury.

Crash characteristics show that, in contrast to single-vehicle crashes, the month of April is associated with a 64% lower probability of fatality. The cause of these contradictory findings cannot be determined with any reasonable certainty. It is possible that these findings may be an outgrowth of motorcyclists' pre-crash behavior in the month with crashes mostly resulting in loss of control (single-vehicle) being more likely to produce a fatality and those resulting from collisions with other vehicles (multi-vehicle) less likely to result in fatality. A better understanding of motorcyclists' skill set and the evolution of this skill set after several months of not riding (winter months) is a fruitful area for future research.

Several crash types were found to significantly influence injury severity. Head-on, right-angle, those occurring in darkness, and collisions with tractor-trailers all greatly increased the probability of fatality—with head-on collisions resulting in a 566% higher likelihood of fatality.

Finally, for this model we also test for the possibility that motorcyclists determined to be at-fault in the crash may have fundamentally different parameters generating their severity probabilities relative to motorcyclists that were not at fault. The hypothesis being that these two types of riders may have significant differences in their risk-taking behavior. In fact, as Tables 5 and 6 show, we find a number of at-fault variables to be significant including "other motorist at fault" which indicates a decrease in the likelihood of a non-incapacitating injury by 33% and "motorcyclist at fault" which indicates a 126% increase in the likelihood of a fatality. Related variables are also statistically significant such as "alcohol use by other motorist" (65% decrease in no-injury), "failure to yield by other motorist" (56% increase in non-incapacitating injury, 11% increase in incapacitating injury), and "speeding by motorcyclist" (50% increase in incapacitating injury, 116% increase in fatality). However, when we estimate separate injury-severity models for crashes when the motorcyclist is at fault and for crashes when the motorcyclist is not at fault, we do not find statistically significant differences.[7] This suggests that possible differences

[7] The appropriate test statistic is $-2[\mathrm{LL}(\beta_T) - \mathrm{LL}(\beta_f) - \mathrm{LL}(\beta_{nf})]$, where $\mathrm{LL}(\beta_T)$ is the log-likelihood at convergence of the model estimated with all data, $\mathrm{LL}(\beta_f)$ is the log-likelihood at convergence of the model using crashes where motorcyclists are at fault, and $\mathrm{LL}(\beta_{nf})$ is the log-likelihood at convergence of the model using crashes where motorcyclists are not at fault. This statistic is χ^2 distributed with degrees of freedom equal to the summation of the number of estimated parameters in the fault and non-fault models minus the number of estimated parameters in the overall model. The resulting χ^2 statistic indicates that the hypothesis that the at-fault/non-at-fault models are the same cannot be rejected at the 90% confidence level.

Table 6
Multi-vehicle crash elasticities in percent

Variable	No-injury	Non-incapacitating injury	Incapacitating injury	Fatality
Rider/motorist characteristics and behavior				
Motorcyclist age			4.2	
Motorcyclist 60 years old or older		65		
Motorcycle age	−0.06			
Helmet use (sport utility vehicle)		88		
Helmet use (pickup truck)		39		
Helmet use (right-angle)				−61
Motorcyclist at fault				126
Other motorist at fault		−33		
Alcohol use (other motorist)	−65			
Sportbike		32		
Failure to yield (other motorist)		56	11	
Motorcyclist speeding			50	116
Roadway characteristics				
Speed limit over 30 mph	−15			
Speed limit over 40 mph	−9	−19		
Speed limit over 50 mph	−19			132
Speed limit over 60 mph	62			
Interstate		84		
Vertical curve			81	
Horizontal curve			45	
Crash characteristics				
April				−64
Head-on	−35			566
Rear-end			−53	
Right-angle				227
Darkness				100
Collision with motorcycle		44		
Collision with pickup truck			59	
Collision with sport utility vehicle		−65		
Collision with tractor-trailer				232

between at-fault motorcyclists and those that are not at fault are being appropriately captured with the individual variables that we have included in the model.

6. Conclusions

Motorcycle safety in Indiana is plagued by many of the same problems faced throughout the United States. The crash-injury severity analysis presented in this paper revealed several problem areas leading to more severe injuries: poor visibility (horizontal curvature, vertical curvature, darkness); unsafe speed (citations for speeding); alcohol use; not wearing a helmet; right-angle and head-on collisions; and collisions with fixed objects. There were some findings that motorcyclists may be managing risks. Crashes were found to be less severe under wet pavement conditions, near intersections, and when passengers were on the motorcycle. This may indicate that riders may be riding in a more cautious manner under such situations—resulting in less severe crashes once they occur.

There were also some disturbing findings. Perhaps key among these is the finding that older motorcyclists were more likely to be involved severe-injury crashes—even when controlling for all other elements of the crash in a multivariate analysis. This supports some recent aggregate data trends and would seem to contradict the long-held belief that young and inexperienced motorcyclists are more likely to be involved in severe crashes.

In terms of the possible safety-oriented strategies that our findings may support, it would seem that rider education would make the most sense. Making riders aware of the factors found to increase crash-injury severity, such as age, and the more obvious factors such as helmet use, alcohol consumption and speeding. An important note is that considerable caution must be taken with educational efforts aimed at expanding motorcyclists' skill set. Our finding indicating that motorcyclists taking the Basic Rider Course were more likely to be fatally injured in single-vehicle accidents could be the result of a number of factors (as explained in the paper) including risk compensating behavior that could negate a significant portion of the safety benefits of such courses.

Our study data were limited to Indiana's police-reported crashes. An obvious extension would be to include a larger database of additional states that would allow one to better understand the effects of weather, geography, and perhaps different state licensing requirements. Also, data that are more detailed than police crash reports (a more in-depth assessment of the accident scene) would open up additional analysis possibilities and allow more precise model estimation.

References

Abdel-Aty, M.A., Chen, C., Schott, J.R., 1998. An assessment of the effect of driver age on traffic accident involvement using log-linear models. Accid. Anal. Prev. 30 (6), 851–861.

Abdel-Aty, M.A., 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models. J. Saf. Res. 34 (5), 597–603.

Carson, J., Mannering, F., 2001. The effect of ice warning signs on accident frequencies and severities. Accid. Anal. Prev. 33 (1), 99–109.

Chang, L.-Y., Mannering, F., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. Accid. Anal. Prev. 31 (5), 579–592.

Duncan, C., Khattak, A., Council, F., 1998. Applying the ordered probit model to injury severity in truck-passenger car rear-end collisions. Transport. Res. Rec. 1635, 63–71.

Farmer, C., Braver, E., Mitter, E., 1997. Two-vehicle side impact crashes: the relationship of vehicle and crash characteristics to injury severity. Accid. Anal. Prev. 29 (3), 399–406.

Greene, W., 2000. Econometric Analysis, fourth ed. Prentice Hall, Upper Saddle River, NJ.

Hausman, J., Abrevaya, J., Scott-Morton, F., 1998. Misclassification of the dependent variable in a discrete-response setting. J. Economet. 87, 239–269.

Islam, S., Mannering, F., 2006. Driver aging and its effect on male and female single-vehicle accident injuries: some additional evidence. J. Saf. Res. 37 (3), 267–276.

Khattak, A., 2001. Injury severity in multi-vehicle rear-end crashes. Transport. Res. Rec. 1746, 59–68.

Khattak, A., Pawlovich, D., Souleyrette, R., Hallmark, S., 2002. Factors related to more severe older driver traffic crash injuries. J. Transport. Eng. 128 (3), 243–249.

Khorashadi, A., Niemeier, D., Shankar, V., Mannering, F., 2005. Differences in rural and urban driver-injury severities in accidents involving large trucks: an exploratory analysis. Accid. Anal. Prev. 37 (5), 910–921.

Kockelman, K., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. Accid. Anal. Prev. 34 (4), 313–321.

Kweon, Y.-J., Kockelman, K., 2003. Overall injury risk to different drivers: combining exposure, frequency, and severity models. Accid. Anal. Prev. 35 (3), 414–450.

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Accid. Anal. Prev. 34 (2), 149–161.

Manski, C., Lerman, S., 1977. The estimation of choice probabilities from choice based samples. Econometrica 45 (8), 1977–1988.

McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C.F., McFadden, D. (Eds.), Structure Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge MA.

Motorcycle Safety Foundation, 2000. National Agenda for Motorcycle Safety, Irvine, CA, http://www.ahainc.com/nams/.

National Highway Traffic Safety Administration, 2006. Recent Trends in Fatal Motorcycle Crashes: An update, National Center for Statistics and Analysis, DOT HS 810 606, June 2006, Washington, D.C.

O'Donnell, C., Connor, D., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Accid. Anal. Prev. 28 (6), 739–753.

Peltzman, S., 1975. The effects of automobile safety regulation. J. Polit. Econ. 83, 677–725.

Quddus, M., Noland, R., Chin, H., 2002. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. J. Saf. Res. 33 (4), 445–462.

Renski, H., Khattak, A., Council, F., 1999. Impact of speed limit increases on crash injury severity: analysis of single-vehicle crashes on North Carolina interstate highways. Transport. Res. Rec. 1665, 100–108.

Savolainen, P., Mannering F., 2006. Additional evidence on the effectiveness of motorcycle training and motorcyclists' risk-taking behavior. Working paper.

Shankar, V., Mannering, F., 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. J. Saf. Res. 27 (3), 183–194.

Shankar, V., Mannering, F., Barfield, W., 1996. Statistical analysis of accident severity on rural freeways. Accid. Anal. Prev. 28 (3), 391–401.

Thomson, G., 1980. The role frontal motorcycle conspicuity has in road accidents. Accid. Anal. Prev. 12 (3), 165–178.

Ulfarsson, G., Mannering, F., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accid. Anal. Prev. 36 (2), 135–147.

Washington, S., Karlaftis, M., Mannering, F., 2003. Statistical and Econometric Methods for Transportation Data Analysis. Chapman and Hall/CRC, Boca Raton, FL, p. 425.

Winston, C., Maheshri, V., Mannering, F., 2006. An exploration of the offset hypothesis using disaggregate data: the case of airbags and antilock brakes. J. Risk Uncertainty 32 (2), 83–99.

Wulf, G., Hancock, P., Rahimi, M., 1989. Motorcycle conspicuity: and evaluation and synthesis of influential factors. J. Saf. Res. 20 (4), 153–175.

Yamamoto, T., Shankar, V., 2004. Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. Accid. Anal. Prev. 36 (5), 869–876.