Analytic Methods in Accident Research 2 (2014) 21–29

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

# Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model

Brendan J. Russo [a,1], Peter T. Savolainen [a,*], William H. Schneider IV [b,2], Panagiotis Ch. Anastasopoulos [c,3]

[a] Department of Civil and Environmental Engineering, Wayne State University, 5050 Anthony Wayne Drive, EDC 0504.01, Detroit, MI 48202, USA
[b] Department of Civil Engineering, The University of Akron, Auburn Science and Engineering Center, Room 210, Akron, Ohio 44325, USA
[c] Department of Civil, Structural and Environmental Engineering, Institute for Sustainable Transportation and Logistics, University at Buffalo, The State University of New York, 241 Ketter Hall, Buffalo, New York 14260, USA

ARTICLE INFO

*Article history:*
Received 30 November 2013
Received in revised form
10 March 2014
Accepted 10 March 2014
Available online 19 April 2014

*Keywords:*
Crash severity
Intersection crash
Fault status
Random parameter bivariate ordered probit

ABSTRACT

The extant traffic safety research literature includes numerous examples of studies that assess those factors affecting the degree of injury sustained by crash-involved motor vehicle occupants. One important methodological concern in such work is the potential correlation in injury outcomes among occupants involved in the same crash, which may be due to common unobserved factors affecting such occupants. A second concern is unobserved heterogeneity, which is reflective of parameter effects that vary across individuals and crashes. To address these concerns, a random parameters bivariate ordered probit model is estimated to examine factors affecting the degree of injury sustained by drivers involved in angle collisions. The modeling framework distinguishes between the effects of relevant factors on the injury outcomes of the at-fault and not-at-fault parties. The methodological approach allows for consideration of within-crash correlation, as well as unobserved heterogeneity, and results in significantly improved fit as compared to a series of independent models with fixed parameters. While the factors affecting injury severity are found to be similar for both drivers, the magnitudes of these effects vary between the at-fault and not-at-fault drivers. The results demonstrate that injury severity outcomes are correlated for drivers involved in the same crash. Further, the impacts of specific factors may be over- or under-estimated if such correlation is not accounted for explicitly as a part of the analysis. Various factors are found to affect driver injury severity and the random parameters framework shows these effects to vary across crashes and individuals. The analytical approach utilized provides a useful framework for injury severity analysis.



* Corresponding author. Tel.: +313 577 9950; fax: +313 577 8126.
*E-mail addresses:* brendan@wayne.edu (B.J. Russo), savolainen@wayne.edu (P.T. Savolainen), whs4@uakron.edu (W.H. Schneider IV), panastas@buffalo.edu (P.Ch. Anastasopoulos).
[1] Tel.: +313 577 8228; fax: +313 577 8126.
[2] Tel.: +330 972 2426.
[3] Tel.: +716 645 4362.

http://dx.doi.org/10.1016/j.amar.2014.03.001
2213-6657/

## 1. Introduction

Although traffic crashes, injuries, and fatalities in the United States have been trending downward in recent years, in 2010 alone there were 23,303 motor vehicle occupant fatalities and 2,027,000 injuries (NHTSA, 2012). According to the National Safety Council (2010), each crash-related fatality results in an average economic cost of $1,410,000 while each non-fatal disabling injury costs society $70,200. The true costs of traffic crash injuries and fatalities are even greater as these economic costs do include measures such as lost quality of life. Gaining a better understanding of the factors that affect the degree of injury sustained by crash-involved occupants is critical in developing methods for addressing this public health issue.

Angle collisions are among the most severe crashes in terms of injury outcomes. These crashes typically occur at intersections or driveways when one of the crash involved drivers fails to yield or disregards a traffic control device. In 2010, these types of collisions represented 36.1% of all motor vehicle crashes (NHTSA, 2012). During this same period, angle collisions represented 50.7% of fatal crashes and 41.1% of injury crashes, illustrating the clear need for a careful examination of the factors associated with such crashes. In angle crashes, one party is generally found to be at fault and the relationship between fault status and injury outcomes is of general interest to the traffic safety community. While some prior research has investigated factors associated with fault status designation (Haque et al., 2009; Jiang et al., 2012; Kim et al., 2008; Schneider et al., 2012; Ulfarsson et al., 2010), research on the effects of fault status on injury outcomes is quite limited. Savolainen and Mannering (2007) found that motorcyclists who were at-fault were more likely to be fatally injured; however, other studies have shown mixed results with respect to fault status and injury outcomes (Abdel-Aty, 2003; Helai et al., 2008). The analysis of injury severity data from angle crashes is complicated by several factors, including the potential correlation in injury outcomes among occupants involved in the same crash.

In this study, injury severity data are examined from angle crashes occurring at intersections in the state of Michigan. A random parameters bivariate ordered probit (RPBOP) model is developed to jointly examine the degree of injury sustained by each crash-involved driver. The use of the RPBOP model allows for consideration of within-crash correlation while the estimation of random parameters allows for the effects of variables to vary across observations. The injury severity levels of the at-fault and not-at-fault drivers are estimated separately, such that differential effects of each independent variable can be identified.

## 2. Literature review

Numerous studies have examined factors that affect crash injury severity. The broad body of existing research covers a wide array of topics and analytical techniques. Generally, such studies focus on a specific crash type, vehicle type, or roadway type. Findings from these studies can provide valuable insights into factors that affect injury outcomes, though these results may not be applicable to all traffic crash scenarios. Savolainen et al. (2011) provide a summary of methodological issues encountered when analyzing crash severity data, as well as an overview of statistical techniques that have been used in analyzing such data. Commonly applied statistical models include multinomial logit models, ordered logit and probit models, and random parameter (mixed) logit models.

O'Donnell and Conner (1996) estimated ordered probit and ordered logit models to examine factors affecting injury severity for all crash types and all locations occurring in New South Wales, Australia. Kockelman and Kweon (2002) investigated driver injury severity by estimating separate ordered probit models for single-vehicle crashes and two-vehicle crashes. Haleem and Abdel-Aty (2010) examined crash injury severity for all crash types at unsignalized intersections in Florida using ordered probit, binary probit, and nested logit models. One finding from this study of particular relevance is that young at-fault drivers were less likely to experience severe injuries. Abdel-Aty (2003) developed separate ordered probit models to examine driver injury severity of crashes occurring at signalized intersections, toll plazas, and road segments. Exclusive to the signalized intersection model, it was found that at-fault drivers were less likely to sustain injuries than not at-fault drivers. Helai et al. (2008) developed a Bayesian hierarchical binomial logistic model to examine driver injury severity at intersections and found that at-fault drivers were more likely to be severely injured. In general, prior research has not explicitly examined the differences in injury outcomes between the at-fault and not-at-fault parties.

To address this gap in the research literature, a random parameters bivariate ordered probit (RPBOP) model is estimated to analyze factors affecting the injury severity of each crash-involved driver. In the transportation research literature, bivariate ordered probit (BOP) models have been utilized to examine outcomes or decisions that may be correlated. As one example, Anastasopoulos et al. (2012a) estimated a BOP model to examine automobile and motorcycle ownership. However, research utilizing BOP models in the context of traffic safety is limited. Mannering and Bhat (2014) provide a summary of recent methodological approaches that have been used to analyze both crash frequency and crash injury severity, noting several examples of bivariate/multivariate ordered probit models. Yamamoto and Shankar (2004) used a BOP model to simultaneously examine the injury level of the driver and most severely injured passenger in single-vehicle collisions with fixed objects. de Lapparent (2008) used BOP models to jointly analyze seat belt use and crash-related injury severity. Chiou et al. (2013) used a bivariate generalized ordered probit (BGOP) model to examine driver injury severity in two-vehicle crashes. Various other statistical techniques have been used to account for potential correlation or endogeneity in injury severity outcomes. Examples include a copula-based approach to simultaneously examine the degree of injury sustained by drivers, front seat passengers, and rear-seat passengers (Eluru et al., 2010); a copula-based joint multinomial logit-ordered logit model (Rana et al., 2010); a multivariate tobit model of accident-injury-severity rates on multilane divided highways (Anastasopoulos et al., 2012c); and a multivariate probit model of injury severity and seat belt use (Abay et al., 2013).

## 3. Data description

Data for all two-vehicle angle-type collisions occurring in the state of Michigan during calendar year 2011 were obtained from the Michigan Office of Highway Safety Planning (2012). These crash data provide details of the roadway characteristics, environmental conditions at the time of the crash, vehicle information, and driver information. Data were excluded for those crashes in which the injury level for a driver was unknown or if neither driver was found to be at-fault. Driver injury severity is assessed on an ordinal scale, which classifies each driver's injury severity into one of five discrete categories:

- Fatality (results in the death of the driver)
- Incapacitating injury (any injury, other than a fatal injury, that prevents the injured driver from walking, driving or normally continuing the activities the person was capable of performing before the injury occurred.)
- Non-incapacitating injury (any injury not incapacitating but evident to observers at the scene of the crash in which the injury occurred.)
- Possible injury (any injury reported or claimed that is not a fatal injury, incapacitating injury or non-incapacitating injury.)
- No injury (driver reported as not receiving bodily harm from the motor vehicle crash.)

The final dataset included a total of 19,941 crashes involving 39,882 motor vehicle drivers. Fault was assigned to one driver in each crashed based upon the judgment of the investigating officer, who determined the at-fault driver to have performed one or more hazardous actions (e.g. fail to yield, disregard of traffic control) that contributed to the crash. When analyzing the resultant injury severity data, drivers were separated into two groups: at-fault and not-at-fault. Separating drivers by fault status allows for examination of potential differential influences of key factors within the two driver groups. It should be noted that due to the low number of fatal injuries, the fatal and incapacitating injury levels were combined for the purposes of this analysis.

A tabular summary of the joint injury severity data for the at-fault and not-at-fault drivers is shown in Table 1. Inspection of these summary statistics shows that not at-fault drivers tend to be more severely injured than at-fault drivers. Only 14.9% of at-fault drivers sustained possible or evident injuries, compared to 20.6% of not-at-fault drivers. This finding is consistent with prior research (Abdel-Aty, 2003; Chiou et al., 2013; Haleem and Abdel-Aty, 2010), which has theorized this difference may be due to the fact that the at-fault driver tends to be driving the striking vehicle and the (not-at-fault) driver of the struck vehicle would tend to experience more severe injuries (Abdel-Aty, 2003).

Numerous explanatory variables were examined to determine whether they influenced these injury severity outcomes. Table 2 provides summary statistics for those driver- and vehicle-specific variables that were found to significantly affect injury severity levels while Table 3 presents details of crash-specific variables that were significant determinants of injury severity. It is interesting to note some general trends between the at-fault and not-at-fault drivers. For example, belt use and alcohol/drug use were lower among the not-at-fault drivers. This is consistent with prior work in this area, which has shown at-fault drivers to exhibit other high-risk behaviors (Schneider et al., 2012). While prior work has identified substantive differences between the characteristics of the at-fault and not-at-fault drivers, investigation of potential differences in injury outcomes between the two parties warrants further investigation.

## 4. Statistical methodology

As discussed previously, various analytical techniques have been used to assess the degree of injury severity resulting from traffic crashes. One common limitation among prior work is that within-crash correlation is frequently not accounted for. Savolainen et al. (2011) note that unobserved elements relating to a specific crash (e.g., impact characteristics) may result in correlation among crash-injury observations. The current study utilizes a random parameters bivariate ordered probit (RPBOP) model to simultaneously examine injury severity of each driver involved in two-vehicle angle-type intersection collisions.

**Table 1**
Summary of injury severity for crash-involved drivers.

| Injury severity of not-at-fault driver | Injury severity of at-fault driver | | | | Total |
|---|---|---|---|---|---|
| | No injury | Possible injury | Non-incapacitating injury | Incapacitating or fatal injury | |
| **No injury** | 14,390 | 1033 | 306 | 102 | 15,831 (79.4%) |
| **Possible injury** | 1962 | 737 | 192 | 53 | 2944 (14.8%) |
| **Non-incapacitating injury** | 475 | 164 | 181 | 43 | 863 (4.3%) |
| **Incapacitating or fatal injury** | 140 | 57 | 37 | 69 | 303 (1.5%) |
| **Total** | 16,967 (85.1%) | 1991 (10.0%) | 716 (3.6%) | 267 (1.3%) | 19,941 (100.0%) |

**Table 2**
Summary statistics for driver- and vehicle-specific independent variables.

| Driver/vehicle variable | At-fault driver | | | | | Not-at-fault driver | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| | No injury | Possible injury | Non-incapacitating injury | Incapacitating/ fatal injury | Total | No injury | Possible injury | Non-incapacitating injury | Incapacitating/ fatal injury | |
| *Driver gender* | | | | | | | | | | |
| Male | 8895 | 807 | 337 | 123 | 10,162 | 8632 | 1204 | 406 | 168 | 10,410 |
| Female | 7989 | 1181 | 379 | 144 | 9693 | 7190 | 1740 | 456 | 134 | 9520 |
| Unknown | 83 | 3 | 0 | 0 | 86 | 9 | 0 | 1 | 1 | 11 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |
| *Driver seatbelt or helmet use* | | | | | | | | | | |
| Driver belted or helmeted | 16,107 | 1895 | 651 | 221 | 18,874 | 15,429 | 2867 | 827 | 283 | 19,406 |
| Driver unbelted or un-helmeted | 135 | 38 | 37 | 26 | 236 | 94 | 20 | 13 | 12 | 139 |
| Safety device use unknown | 725 | 58 | 28 | 20 | 831 | 308 | 57 | 23 | 8 | 396 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |
| *Driver ejected* | | | | | | | | | | |
| Driver not ejected | 16,958 | 1984 | 713 | 258 | 19,913 | 15,820 | 2937 | 853 | 287 | 19,897 |
| Driver ejected | 9 | 7 | 3 | 9 | 28 | 11 | 7 | 10 | 16 | 44 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |
| *Driver alcohol and drug use* | | | | | | | | | | |
| Driver did not use alcohol or drugs | 16,661 | 1913 | 673 | 242 | 19,489 | 15,798 | 2934 | 858 | 300 | 19,890 |
| Driver used alcohol or drugs | 306 | 78 | 43 | 25 | 452 | 33 | 10 | 5 | 3 | 51 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |
| *Driver age* | | | | | | | | | | |
| Age 24 and below | 5320 | 607 | 233 | 71 | 6231 | 3298 | 549 | 159 | 48 | 4054 |
| Age 25–39 | 3826 | 463 | 151 | 61 | 4501 | 4330 | 825 | 228 | 72 | 5455 |
| Age 40–59 | 4364 | 502 | 157 | 61 | 5084 | 5549 | 1046 | 298 | 121 | 7014 |
| Age 60 and above | 3216 | 412 | 171 | 74 | 3873 | 2629 | 522 | 178 | 62 | 3391 |
| Unknown | 241 | 7 | 4 | 0 | 252 | 25 | 2 | 0 | 0 | 27 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |

**Table 3**
Summary statistics for crash-specific independent variables.

| Crash variable | At-fault driver | | | | | Not-at-fault driver | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| | No injury | Possible injury | Non-incapacitating injury | Incapacitating/ fatal injury | Total | No injury | Possible injury | Non-incapacitating injury | Incapacitating/ fatal injury | |
| *Crash season* | | | | | | | | | | |
| Winter (Dec-March) | 6327 | 616 | 202 | 69 | 7214 | 6027 | 898 | 220 | 69 | 7214 |
| Non-Winter (April-Nov) | 10,640 | 1375 | 514 | 198 | 12,727 | 9804 | 2046 | 643 | 234 | 12,727 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |
| *Crash season* | | | | | | | | | | |
| Weekend | 3630 | 483 | 185 | 67 | 4365 | 3330 | 710 | 235 | 90 | 4365 |
| Weekday | 13,337 | 1508 | 531 | 200 | 15,576 | 12,501 | 2234 | 628 | 213 | 15,576 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |
| *Crash site speed limit (mph)* | | | | | | | | | | |
| 25 or lower | 4438 | 299 | 108 | 20 | 4865 | 4175 | 541 | 121 | 28 | 4865 |
| 30 to 50 | 10,295 | 1250 | 390 | 115 | 12,050 | 9599 | 1793 | 510 | 148 | 12,050 |
| 55 or higher | 2134 | 426 | 214 | 131 | 2905 | 1965 | 589 | 229 | 122 | 2,905 |
| Unknown | 100 | 16 | 4 | 1 | 121 | 92 | 21 | 3 | 5 | 121 |
| Total | 16,967 | 1991 | 716 | 267 | 19,941 | 15,831 | 2944 | 863 | 303 | 19,941 |

This method accounts for possible within-crash correlation due to the presence of such common unobserved factors. The dependent variables in this model are the injury severity levels sustained by each crash-involved driver, which can take one of four discrete values, ranging from property damage only to incapacitating/fatal injury. As these severity data follow an inherent ordering structure from least severe to most severe, ordered probit models present a useful analytical framework.

The bivariate ordered probit (BOP) model is derived first by defining the ordinal data $y$ for each observation; for example, for two outcomes (Greene and Hensher, 2009; Anastasopoulos et al., 2012a):

$$\begin{aligned} y_{i,1} &= \boldsymbol{\beta}'_1 \boldsymbol{X}_{i,1} + \varepsilon_{i,1}, \quad y_{i,1} = j \text{ if } \boldsymbol{\mu}_{j-1} < y_{i,1} < \boldsymbol{\mu}_j,\ j = 0, \ldots J_1, \\ y_{i,2} &= \boldsymbol{\beta}'_2 \boldsymbol{X}_{i,2} + \varepsilon_{i,2}, \quad y_{i,2} = j \text{ if } \boldsymbol{\theta}_{j-1} < y_{i,2} < \boldsymbol{\theta}_j,\ j = 0, \ldots J_2, \end{aligned} \tag{1}$$

where

$y$ corresponds to integer ordering of injury severity for each crash involved driver,
$\boldsymbol{\beta}$ vectors of estimable parameters,
$\boldsymbol{X}$ vectors of explanatory variables that affect driver injury severity,
$\boldsymbol{\mu}$,$\boldsymbol{\theta}$ estimable threshold parameters that define $y$,
$j$ integer ordered severity levels, and
$\varepsilon$ random error terms [assumed to be normally distributed ($N$) with zero mean and variance of one].

The cross-equation correlated error terms are

$$\begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right] \tag{2}$$

where $\rho$ cross-equation correlation coefficient of the error terms.

The BOP model with ordered selection joint probability for $y_{i,1} = j$ and $y_{i,2} = k$ is then defined as

$$P(y_{i,1} = j, y_{i,2} = k | \boldsymbol{X}_{i,1}, \boldsymbol{X}_{i,2}) = \begin{pmatrix} \Phi_2[(\boldsymbol{\mu}_j - \boldsymbol{\beta}'_1 \boldsymbol{X}_{i,1}), (\boldsymbol{\theta}_k - \boldsymbol{\beta}'_2 \boldsymbol{X}_{i,2}), \rho] \\ -\Phi_2[(\boldsymbol{\mu}_{j-1} - \boldsymbol{\beta}'_1 \boldsymbol{X}_{i,1}), (\boldsymbol{\theta}_k - \boldsymbol{\beta}'_2 \boldsymbol{X}_{i,2}), \rho] \end{pmatrix} - \begin{pmatrix} \Phi_2[(\boldsymbol{\mu}_j - \boldsymbol{\beta}'_1 \boldsymbol{X}_{i,1}), (\boldsymbol{\theta}_{k-1} - \boldsymbol{\beta}'_2 \boldsymbol{X}_{i,2}), \rho] \\ -\Phi_2[(\boldsymbol{\mu}_{j-1} - \boldsymbol{\beta}'_1 \boldsymbol{X}_{i,1}), (\boldsymbol{\theta}_k - \boldsymbol{\beta}'_2 \boldsymbol{X}_{i,2}), \rho] \end{pmatrix} \tag{3}$$

where $\Phi$ standard normal cumulative distribution function

After model estimation, the signs of the parameter estimates are of particular interest. A positive sign indicates an increase in the probability of the most severe outcome (incapacitating/fatal injury) and a decrease in the probability of the least severe outcome (property damage only). The converse is true for negative parameter estimates. In order to interpret the effects on the intermediate categories (possible and non-incapacitating injury), marginal effects are computed at the sample mean for each category (Greene, 2007; Washington et al., 2011; Anastasopoulos et al., 2012a):

$$\frac{P(y=j)}{\partial \boldsymbol{X}} = \left[\phi(\omega_{j-1} - \boldsymbol{\beta} X) - \phi(\omega_j - \boldsymbol{\beta} X)\right]\boldsymbol{\beta} \tag{4}$$

where

$P(y=j)$ probability of driver experiencing $j$ level injury
$\omega$ thresholds
$j$ integer ordered severity levels
$\phi(.)$ probability mass function of the standard normal distribution

While BOP models can account for within-crash correlation, another methodological concern is that the effects of certain parameters may vary across observations due to unobserved heterogeneity. Constraining the model parameters to be constant across observations may lead to inconsistent and biased parameter estimates (Washington et al., 2011). To address this issue, random parameters can be estimated, allowing for the effects of parameters to vary across observations (Eluru et al., 2008; Anastasopoulos and Mannering, 2009, 2011; Anastasopoulos et al., 2012a, 2012b). Such variability can be incorporated into the OP model by allowing parameters to vary as follows (Greene, 2007):

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \mu_i \tag{5}$$

where

$\boldsymbol{\beta}_i$ vector of driver-specific parameters and
$\mu_i$ randomly distributed term [normally distributed with mean zero and variance $\sigma^2$].

Random parameter bivariate ordered probit (RPBOP) models can be estimated by simulated maximum likelihood estimation. To improve the efficiency of estimation, 200 Halton draws are utilized as recommended through other research in this area (Halton, 1960; Bhat, 2003; Train, 2003; Washington et al., 2011).

## 5. Model estimation and discussion of results

Initially, two independent ordered probit (OP) models were estimated, one for at-fault drivers and another for not-at-fault drivers. All variables that were significant at a 95% confidence level were retained for the subsequent analyses. Next, a series of independent random parameters ordered probit (RPOP) model was developed. Conceptually, the RPOP allows for greater flexibility as it is able to account for the possibility that parameters may vary across crash-involved individuals (Washington et al., 2011). While the parameter estimates for the OP and the RPOP models were similar in direction, the RPOP model demonstrated substantial variability in parameter effects across individuals. A likelihood ratio test showed the RPOP to provide an improved fit compared to the OP model ($\alpha=0.05$).

Next, a bivariate ordered probit (BOP) model was estimated. The BOP model showed strong correlation in injury outcomes between drivers who were involved in the same crash. Capturing this correlation resulted in more efficient parameter estimates (i.e., reduced standard errors) and the goodness-of-fit was markedly improved compared to the OP models, as well as the RPOP models (all independent variables were statistically significant at a 99% confidence level).

Finally, a random parameter bivariate ordered probit (RPBOP) model was estimated, which simultaneously addressed the issues of cross-equation error correlation within injury outcomes, and unobserved heterogeneity (unobserved factors varying systematically across the observations). The RPBOP provided substantive improvements compared to the other three model formulations (all independent variables were statistically significant at $\alpha=0.001$). This provides compelling evidence that: (a) there is correlation among the degree of injury severity sustained by drivers involved in the same crash; and (b) the parameter effects vary across both crashes and drivers due to unobserved heterogeneity. For comparison purposes, Table 4 presents the results of the fixed and random parameters bivariate ordered probit models (results of the separate univariate models have been excluded). These results include parameter estimates, standard errors, and $p$-values for each variable. Explanatory variables were generally coded as a series of binary (0/1) indicators, with the exception of speed limit. Speed limit was treated as a continuous variable for model estimation as this parameterization resulted in significantly improved model fit.

For the RPBOP model, further details are provided as to the percentage of drivers who exhibit parameters above or below zero (in the case of the random parameters). These results suggest that the effects of some factors may vary across crashes and drivers. Specific parameters may be positive for some crashes/drivers and negative for others, which may reflect the effects of unobserved factors that are specific to these crashes/drivers. If this heterogeneity is not explicitly accounted for, biased parameter estimates may result as shown by the fixed and random parameter model results in Table 4.

Some variables, such as belt use or ejection from the vehicle, were estimated as fixed parameters as their effects were found to be consistent across the sample of crash-involved drivers. Other factors, such as speed limit, showed some variability, but were generally consistent in one direction. In this case, higher speeds tended to consistently increase injury severity, a finding that is in line with past research (Elvik, 2006). Conversely, crashes that occurred on the weekend showed significant variability in injury outcomes across the sample. While on average, crashes tended to be less severe on the weekend, this effect tended to be highly variable on a crash-to-crash or driver-to-driver basis. This variability may be due to unique weekend travel patterns that are not captured by the crash data.

While the parameter estimates in Table 4 provide a general sense of the direction of impacts of specific factors on injury outcomes, Table 5 presents marginal effects for these variables. These marginal effects represent the average change in the probability of a specific injury outcome that corresponds to a one-unit change in an independent variable. For example, when a driver was unbelted, the probabilities of incapacitating or fatal injuries were found to increase by 11.49 percentage points and 2.63 percentage points for at-fault and not at-fault drivers, respectively. The subsequent discussion provides further details of the analysis results, with specific emphasis on the RPBOP model.

In addition to addressing concerns as to cross-equation error correlation and unobserved heterogeneity, another advantage of the RPBOP model is that it allows for the estimation of the tetrachoric correlation parameter ($\rho$), which quantifies the correlation in the error terms between the ordered probit models for each pair of crash-involved drivers. This parameter, in effect, reflects the correlation in unobserved factors affecting injury severity in each crash event. The RPBOP model results, shown in Table 4, reflect a correlation parameter of 0.259 ($p$-value $<0.0001$). This suggests that these unobserved factors tend to jointly increase (or decrease) the degree of injury sustained by drivers involved in the same crash. This correlation may relate to factors such as impact speed, which are not typically captured in the police crash report forms, but would tend to increase (or decrease) the severity level for each driver involved in the same crash.

Returning to the results from Table 4, the final RPBOP model includes nine covariates for both the at-fault and not-at-fault drivers. Of these, six parameters exhibited significant variability among the sample of at-fault drivers (as evidenced by the standard deviation estimate for these random parameters) and seven parameters were found to vary among not at-fault drivers. It is interesting to note that the variability in these parameters for not at-fault drivers tended to be larger than those of at-fault drivers. This indicates significantly greater heterogeneity among the injury outcomes of not-at-fault drivers. This is not surprising as the not-at-fault driver generally becomes aware of an impending collision much later than an at-fault driver. Consequently, there are likely to be significant differences between drivers due to unobserved factors, such as driving experience and physiological differences. In any case, this is an area where further research may provide important insights into the myriad of factors involved in such collisions.

Turning to the other parameters of interest, crashes occurring during the winter months (December through March) tended to result in less severe injuries to both drivers. This may be attributable to drivers being more cautious and driving at

**Table 4**
Results of fixed and random parameter bivariate ordered probit models.

| **At-fault driver** | **Fixed parameter model** | | | **Random parameter model** | | | **Percent observations** | |
|---|---|---|---|---|---|---|---|---|
| | **Parameter** | **Std. error** | ***p*-value** | **Parameter** | **Std. error** | ***p*-value** | **above 0** | **below 0** |
| Constant | −2.048 | 0.048 | < 0.001 | −2.301 | 0.045 | < 0.001 | 2.66% | 97.34% |
| *Standard deviation* | | | | *1.190* | *0.114* | *< 0.001* | | |
| Speed limit | 0.021 | 0.001 | < 0.001 | 0.024 | 0.001 | < 0.001 | 98.54% | 1.46% |
| *Standard deviation* | | | | *0.011* | *0.001* | *< 0.001* | | |
| Winter crash | −0.152 | 0.024 | < 0.001 | −0.190 | 0.025 | < 0.001 | 14.56% | 85.44% |
| *Standard deviation* | | | | *0.180* | *0.023* | *< 0.001* | | |
| Weekend crash | 0.069 | 0.026 | 0.009 | 0.084 | 0.029 | 0.003 | 56.74% | 43.26% |
| *Standard deviation* | | | | *0.495* | *0.086* | *< 0.001* | | |
| Driver unbelted | 0.738 | 0.061 | < 0.001 | 0.937 | 0.089 | < 0.001 | | |
| Driver ejected | 1.299 | 0.223 | < 0.001 | 1.364 | 0.244 | < 0.001 | | |
| Driver alcohol/drug use | 0.372 | 0.062 | < 0.001 | 0.643 | 0.075 | < 0.001 | | |
| Driver age 15–24 | −0.018 | 0.025 | 0.478 | −0.018 | 0.003 | < 0.001 | 25.25% | 74.75% |
| *Standard deviation* | | | | *0.027* | *0.001* | *< 0.001* | | |
| Driver age 60+ | 0.121 | 0.027 | < 0.001 | 0.038 | 0.006 | < 0.001 | 79.56% | 20.44% |
| *Standard deviation* | | | | *0.046* | *0.003* | *< 0.001* | | |
| Driver female | 0.259 | 0.021 | < 0.001 | 0.288 | 0.024 | < 0.001 | 77.75% | 22.25% |
| *Standard deviation* | | | | *0.377* | *0.071* | *< 0.001* | | |
| $\mu_1$ | 0.689 | 0.015 | < 0.001 | 0.742 | 0.018 | < 0.001 | | |
| $\mu_2$ | 1.456 | 0.032 | < 0.001 | 1.391 | 0.029 | < 0.001 | | |
| **Not-at-fault driver** | **Fixed parameter model** | | | **Random parameter model** | | | **Percent observations** | |
| | **Parameter** | **Std. error** | ***p*-value** | **Parameter** | **Std. error** | ***p*-value** | **above 0** | **below 0** |
| Constant | −1.698 | 0.043 | < 0.001 | −5.524 | 0.093 | < 0.001 | 3.54% | 96.46% |
| *Standard deviation* | | | | *3.057* | *0.047* | *< 0.001* | | |
| Speed limit | 0.019 | 0.001 | 0.001 | 0.065 | 0.002 | < 0.001 | 94.79% | 5.21% |
| *Standard deviation* | | | | *0.040* | *< 0.001* | *< 0.001* | | |
| Winter crash | −0.223 | 0.022 | < 0.001 | −0.737 | 0.042 | < 0.001 | 8.34% | 91.66% |
| *Standard deviation* | | | | *0.533* | *0.029* | *< 0.001* | | |
| Weekend crash | 0.134 | 0.024 | < 0.001 | 0.378 | 0.037 | < 0.001 | 69.95% | 30.05% |
| *Standard deviation* | | | | *0.723* | *0.032* | *< 0.001* | | |
| Driver unbelted | 0.433 | 0.097 | 0.001 | 1.494 | 0.167 | < 0.001 | | |
| Driver ejected | 1.726 | 0.142 | < 0.001 | 5.879 | 0.263 | < 0.001 | | |
| Driver alcohol/drug use | 0.209 | 0.154 | 0.175 | 0.446 | 0.173 | 0.010 | 87.36% | 12.64% |
| *Standard deviation* | | | | *0.390* | *0.057* | *< 0.001* | | |
| Driver age 15–24 | −0.089 | 0.025 | < 0.001 | −0.302 | 0.041 | < 0.001 | 45.22% | 54.78% |
| *Standard deviation* | | | | *2.514* | *0.293* | *< 0.001* | | |
| Driver age 60+ | 0.106 | 0.026 | < 0.001 | 0.290 | 0.043 | < 0.001 | 82.29% | 17.71% |
| *Standard deviation* | | | | *0.313* | *0.043* | *< 0.001* | | |
| Driver Female | 0.265 | 0.02 | < 0.001 | 0.835 | 0.033 | < 0.001 | 80.19% | 19.81% |
| *Standard deviation* | | | | *0.984* | *0.059* | *< 0.001* | | |
| $\mu_1$ | 0.780 | 0.014 | < 0.001 | 0.869 | 0.015 | < 0.001 | | |
| $\mu_2$ | 1.481 | 0.027 | < 0.001 | 1.525 | 0.027 | < 0.001 | | |
| $\rho$ | 0.233 | | < 0.001 | 0.259 | | < 0.001 | | |
| *Final log-likelihood* | −20,695.23 | | | −19,869.25 | | | | |

lower speeds under adverse winter conditions. It should be noted that this finding may be specific to northern states that experience more adverse weather conditions (e.g., snow and ice) during the winter months.

Younger drivers (age 15–24) were generally less likely to be injured, which may be reflective of physiological differences as compared to older drivers. It is interesting to note that there was significant variability in injury outcomes for younger drivers, particularly those who were not at-fault. This may be another indication of varying degrees of skill and experience among this youngest age cohort.

Female drivers and drivers aged 60 and above tended to sustain more severe injuries whether they were at fault or not. The more severe injuries sustained by older drivers are consistent with past studies and may be attributable to physiological differences or differences in driver behavior (Abay et al., 2013; Chiou et al., 2013; Kockelman and Kweon, 2002; O'Donnell and Conner, 1996; Rana et al., 2010). However, the effect of gender on injury severity is a subject that warrants further research as the results are mixed in comparison to prior work, which has found males to be more at risk for injuries and fatalities (Abdel-Aty, 2003; Chen et al., 2012).

Alcohol or drug use by either driver tended to increase injury severity, highlighting the importance of continuing campaigns aimed at reducing impaired driving. Drivers who were unbelted were at increased risk for severe injury, with injuries being particularly severe if the driver was ejected from the vehicle. Crash-involved motorcyclists who were not helmeted also tended to experience significantly more severe injuries. These findings are consistent with past research

**Table 5**
Marginal effects for random parameter bivariate ordered probit model.

| At-fault driver | No injury | Possible injury | Non-incapacitating injury | Incapacitating or fatal injury |
|---|---|---|---|---|
| Speed limit | −0.0057 | 0.0035 | 0.0012 | 0.0010 |
| Winter crash | 0.0363 | −0.0248 | −0.0091 | −0.0024 |
| Weekend crash | −0.0161 | 0.0110 | 0.0038 | 0.0013 |
| Driver unbelted | −0.2986 | 0.1523 | 0.1007 | 0.0456 |
| Driver ejected | −0.4782 | 0.1925 | 0.1708 | 0.1149 |
| Driver alcohol/drug use | −0.1629 | 0.0954 | 0.0492 | 0.0183 |
| Driver age 15–24 | 0.0031 | −0.0019 | −0.0008 | −0.0004 |
| Driver age 60+ | −0.0088 | 0.0054 | 0.0021 | 0.0013 |
| Driver female | −0.0550 | 0.0374 | 0.0135 | 0.0041 |
| **Not-at-fault driver** | **No injury** | **Possible injury** | **Non-incapacitating injury** | **Incapacitating or fatal injury** |
| Speed limit | −0.0047 | 0.0032 | 0.0009 | 0.0006 |
| Winter crash | 0.0054 | −0.0033 | −0.0012 | −0.0009 |
| Weekend crash | −0.0038 | 0.0031 | 0.0004 | 0.0003 |
| Driver unbelted | −0.0688 | 0.0432 | 0.0184 | 0.0072 |
| Driver ejected | −0.9939 | 0.4811 | 0.4865 | 0.0263 |
| Driver alcohol/drug use | −0.1607 | 0.0916 | 0.0452 | 0.0239 |
| Driver age 15–24 | 0.0017 | −0.0014 | −0.0002 | −0.0001 |
| Driver age 60+ | −0.0028 | 0.0023 | 0.0003 | 0.0002 |
| Driver female | −0.0074 | 0.0069 | 0.0003 | 0.0002 |

(Abay et al., 2013; Abdel-Aty, 2003; Chen et al., 2012; O'Donnell and Conner, 1996; Rana et al., 2010) and are reflective of the importance of occupant protection systems. Higher speed limits also increased injury severity, another indication of the effects of increased impact forces.

Some notable variables that were not found to significantly affect driver injury severity were time of day, number of traffic lanes, type of traffic control, weather conditions, lighting conditions, and vehicle type. Several of these variables had limited variability in the analysis dataset, while other variables (such as whether the driver was trapped in the vehicle) were excluded due to concerns of endogeneity.

## 6. Conclusions

This study assessed the degree of injury sustained by drivers involved in angle collisions in consideration of fault status. Injury severity was examined through the estimation of a random parameters bivariate ordered probit (RPBOP) model. The use of the RPBOP model allowed for consideration of the within-crash correlation between each pair of crash-involved drivers, as well as the varying effects of certain variables across observations. The model results showed the RPBOP to provide superior fit to a series of univariate ordered probit models (with both fixed and random effects), as well as to bivariate models that assumed fixed parameter effects. In addition, the RPBOP model demonstrated correlation in unobserved factors among drivers involved in the same crash. The results suggest that these factors tend to consist jointly increase (or decrease) the degree of injury sustained by those drivers involved in the same crash. Furthermore, comparing the parameter estimates shows that it is important to account for both the within-crash correlation and the heterogeneity in parameter effects in order to avoid potentially over- or understating the effects of specific variables.

All other factors being equal, those drivers who were not-at-fault tended to experience more severe injuries. This is significant as it suggests drivers who are not-at fault are paying a higher price in terms of crash outcomes than those drivers who were responsible for the crash. The results of the RPBOP model showed various factors to affect injury severity, including time of year, speed limit, age, gender, restraint/helmet use, and alcohol/drug use. The results demonstrated that these effects tended to vary significantly across the sample of crash-involved drivers, particularly among those who were not at fault. It is important to note that such nuances in the data could not be accounted for under a fixed parameters modeling framework.

Ultimately, the results of this study reveal important differences in injury outcomes between at-fault and not at-fault drivers. The results also showed that the random parameters bivariate ordered probit models provided significant flexibility that allowed for a more careful assessment of the effects of those factors that affect injury outcomes. The methodological approach explored as a part of this study presents a promising framework for subsequent injury severity research. There are several promising avenues for extending this framework. While this study assessed correlation among drivers involved in the same crash, a similar approach could be utilized to assess spatial or temporal correlation in injury outcomes or correlation in injuries among occupants of the same vehicle. It would also be interesting to examine the conditions under which the correlation in injury outcomes varies (e.g., by examining different types of collisions or examining different sets of variables). Another natural extension of this research would be to examine alternate means of handling unobserved heterogeneity, such as through the estimation of latent class ordered choice models. A recent example of latent class models, within the framework of a multinomial logit model, is provided by Shaheed and Gkritza (2014). Under this framework, crash-involved drivers can be aggregated into similar classes, with separate injury outcome models being estimated for each class.

## References


Abay, K., Paleti, R., Bhat, C., 2013. The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity. Transportation Research Part B 50, 74–89.

Abdel-Aty, M., 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models. Journal of Safety Research 34 (5), 597–603.

Anastasopoulos, P., Karlaftis, M., Haddock, J., Mannering, F., 2012a. Household automobile and motorcycle ownership analyzed with random parameters bivariate ordered probit model. Transportation Research Record 2279, 12–20.

Anastasopoulos, P., Mannering, F., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accident Analysis and Prevention 41 (1), 153–159.

Anastasopoulos, P., Mannering, F., 2011. An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data. Accident Analysis and Prevention 43 (3), 1140–1147.

Anastasopoulos, P., Mannering, F., Shankar, V., Haddock, J., 2012b. A study of factors affecting highway accident rates using the random-parameters tobit model. Accident Analysis and Prevention 45 (1), 628–633.

Anastasopoulos, P., Shankar, V., Haddock, J., Mannering, F., 2012c. A multivariate tobit analysis of highway accident-injury-severity rates. Accident Analysis and Prevention 45 (1), 110–119.

Bhat, C., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transportation Research Part B 37 (9), 837–855.

Chen, H., Cao, L., Logan, D., 2012. Analysis of risk factors affecting the severity of intersection crashes by logistic regression. Traffic Injury Prevention 13 (3), 300–307.

Chiou, Y., Hwang, C., Chang, C., Fu, C., 2013. Modeling two-vehicle crash severity by a bivariate generalized ordered probit approach. Accident Analysis and Prevention 51, 175–184.

de Lapparent, M., 2008. Willingness to use safety belts and levels of injury in car accidents. Accident Analysis and Prevention 40 (3), 1023–1032.

Eluru, N., Bhat, C., Heshner, D., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis and Prevention 40 (3), 1033–1054.

Eluru, N., Paleti, R., Pendyala, R., Bhat, C., 2010. Modeling multiple vehicle occupant injury severity: a copula based multivariate approach. Transportation Research Record 2165, 1–11.

Elvik, R., 2006. Speed and road safety: synthesis of evidence from evaluation studies. Transportation Research Record 1908, 59–69.

Greene, W., 2007. Limdep, Version 9.0. Econometric Software, Inc, Plainview, NY.

Greene, W., Hensher, D., 2009. Modeling Ordered Choices: a Primer, first ed. Cambridge University Press, Cambridge, United Kingdom.

Haleem, K., Abdel-Aty, M., 2010. Examining traffic crash injury severity at unsignalized intersections. Journal of Safety Research 41 (4), 347–357.

Halton, J., 1960. On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numerische Mathematik 2 (1), 84–90.

Haque, M., Chin, H., Huang, H., 2009. Modeling fault among motorcyclists involved in crashes. Accident Analysis and Prevention 41 (2), 327–335.

Helai, H., Chor, C., Haque, M., 2008. Severity of driver injury and vehicle damage in traffic crashes at intersections: a Bayesian hierarchical analysis. Accident Analysis and Prevention 40 (1), 45–54.

Jiang, X., Qui, Y., Lyles, R., Zhang, H., 2012. Issues with using police citations to assign responsibility in quasi-induced exposure. Safety Science 50 (4), 1133–1140.

Kim, K., Brunner, I., Yamashita, E., 2008. Modeling fault among accident-involved pedestrians and motorists in Hawaii. Accident Analysis and Prevention 40 (6), 2043–2049.

Kockelman, K., Kweon, Y., 2002. Driver injury severity: an application of ordered probit models. Accident Analysis and Prevention 34 (3), 313–321.

Mannering, F., Bhat, C., 2014. Analytic methods in accident research: methodological frontier and future direction. Analytic Methods in Accident Research 1, 1–22.

Michigan Office of Highway Safety Planning - Michigan Traffic Crash Facts Data Query Tool, 2012. 〈http://dx.doi.org/www.michigantrafficcrashfacts.org〉. Accessed October 26, 2012.

National Highway Traffic Safety Administration (NHTSA), 2012. Traffic Safety Facts 2010, DOT HS 811 659, Washington DC.

National Safety Council (NSC), 2010. Estimating the Costs of Unintentional Injuries, 〈http://www.nsc.org/news_resources/injury_and_death_statistics/Pages/EstimatingtheCostsofUnintentionalInjuries.aspx〉. (accessed 9. 12. 2012.

O'Donnell, C., Conner, D., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Accident Analysis and Prevention 28 (6), 739–753.

Rana, T., Sikdar, S., Pinjari, A., 2010. Copula-based method for addressing endogeneity in models of severity of traffic crash injuries. Transportation Research Record 2147, 75–87.

Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accident Analysis and Prevention 39 (5), 955–963.

Savolainen, P., Mannering, F., Lord, D., Quddus, M., 2011. The statistical analysis of crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention 43 (5), 1666–1676.

Schneider, W., Savolainen, P., Van Boxel, D., Beverley, R., 2012. Examination of factors determining fault in two-vehicle motorcycle crashes. Accident Analysis and Prevention 45, 669–676.

Shaheed, M.S., Gkritza, K., 2014. A latent class analysis of single-vehicle motorcycle crash severity outcomes. Analytic Methods in Accident Research 2, 30–38, http://dx.doi.org/10.1016/j.amar.2014.03.002.

Train, K., 2003. Discrete Choice Methods with Simulation, first ed. Cambridge University Press, Cambridge, United Kingdom.

Ulfarsson, G., Kim, S., Booth, K., 2010. Analyzing fault in pedestrian-motor vehicle crashes in North Carolina. Accident Analysis and Prevention 42 (6), 1805–1813.

Washington, S., Karlaftis, M., Mannering, F., 2011. Statistical and Econometric Methods for Transportation Data Analysis, second ed. Chapman and Hall/CRC, Boca Raton, FL.

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