The Joint Analysis of Injury Severity of Drivers in Two-Vehicle Crashes Accommodating Seat Belt Use Endogeneity

Kibrom A. Abay
University of Copenhagen
Department of Economics
Øster Farimagsgade 5, Building 26
DK-1353 Copenhagen K, Denmark
E-mail: kibrom.araya.abay@econ.ku.dk

Rajesh Paleti
Parsons Brinckerhoff
One Penn Plaza, Suite 200
New York, NY 10119
Phone: 512-751-5341
Email: rajeshp@mail.utexas.edu

Chandra R. Bhat*
The University of Texas at Austin
Dept of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172
Phone: 512-471-4535, Fax: 512-475-874
E-mail: bhat@mail.utexas.edu

*corresponding author

Original: July 20, 2012
Revision 1: December 20, 2012
Revision 2: January 29, 2013

ABSTRACT

The current study contributes to the existing injury severity modeling literature by developing a multivariate probit model of injury severity and seat belt use decisions of both drivers involved in two-vehicle crashes. The modeling approach enables the joint modeling of the injury severity of multiple individuals involved in a crash, while also recognizing the endogeneity of seat belt use in predicting injury severity levels as well as accommodating unobserved heterogeneity in the effects of variables. The proposed model is applied to analyze the injury severity of drivers involved in two-vehicle road crashes in Denmark.

The empirical analysis provides strong support for the notion that people offset the restraint benefits of seat belt use by driving more aggressively. Also, men and those individuals driving heavy vehicles have a lower injury risk than women and those driving lighter vehicles, respectively. At the same time, men and individuals driving heavy vehicles pose more of a danger to other drivers on the roadway when involved in a crash. Other important determinants of injury severity include speed limit on roadways where crash occurs, the presence (or absence) of center dividers (median barriers), and whether the crash involves a head-on collision. These and other results are discussed, along with implications for countermeasures to reduce injury severities in crashes. The analysis also underscores the importance of considering injury severity at a crash level, while accommodating seat belt endogeneity effects and unobserved heterogeneity effects.

Keywords: Multivariate ordered-response probit, crash analysis, injury severity modeling, seat belt use endogeneity, offsetting behavior, maximum simulated likelihood.

1. INTRODUCTION

Improving the safety of road users remains a top priority of transport and safety planners. This is not surprising, since roadway crashes are the leading cause of death in countries around the world. According to the *Global Status Report on Road Safety*, published by the World Health Organization (2009), nearly 1.3 million people are killed annually and between 20 and 50 million people get injured every year around the globe in roadway crashes. The estimated cost of highway crashes to governments worldwide is estimated to be 518 billion US dollars. In the U.S., roadway crashes constitute the single most important cause of death among individuals 5-24 years of age (Murphy *et al*., 2012), and killed 32,885 individuals in 2010 (NHTSA, 2012). In Denmark, roadway crashes constitute the single most important cause of unintentional injury-caused death among individuals 15-19 years of age (EuroSafe, 2012), and killed 255 individuals in 2010.[1] In addition to the loss of life, traffic crashes impose a tremendous emotional, social, and psychological cost on non-fatally injured crash victims, their families, and society as a whole. This has led traffic safety researchers to explore the leading causes of crashes and the injury severities sustained in crashes as a precursor to developing countermeasures to reduce the occurrence of crashes as well as their severity.

In the literature, it is typical to adopt a two-step approach to model crash occurrence (frequency) and the injury severity sustained by those involved in a crash. Lord and Mannering (2010) provide a review of methods for crash frequency analysis, while Savolainen *et al*. (2011) present a corresponding review of methods for injury severity analysis conditional on a crash. In this paper, the objective is to contribute to the methods for injury severity analysis by proposing an approach to jointly model the injury severity of multiple individuals involved in a crash (see Section 1.1), while also recognizing the endogeneity of seat belt use in predicting injury severity levels (Section 1.2) as well as accommodating unobserved heterogeneity in the effects of variables (see Section 1.3). The proposed model is applied to two-vehicle road crashes in Denmark.

1.1. Joint Injury Severity Modeling of Multiple Individuals in a Crash

The injury severity of individuals involved in traffic crashes is a combined effect of a multitude of factors, including the characteristics of the drivers involved, the characteristics of the

---

[1] See http://statbank.dk/statbank5a/default.asp?w=1093.

vehicle(s) involved, environmental conditions, roadway characteristics, and crash characteristics. Ideally, then, the injury severities of all individuals involved in a crash should be considered at once, to acknowledge that the injuries sustained by these individuals are likely to be inter-related through their connection to the single crash event. In contrast, most crash-related injury severity studies in the safety literature either (a) pool all individuals across all crashes and estimate an individual-level injury severity model that completely severs the link between individuals involved in the same crash, or (b) model the injury severity of the most severely injured individual in a crash. The problem with the first approach, as alluded to before, is that it ignores potential shared unobserved crash-specific factors that impact the injury severity of those involved in the same crash. These unobserved factors may include variables such as the condition and maintenance record of the vehicles involved in the crash, vehicle speeds at the time of crash, condition and effectiveness of safety equipment installed in the vehicles involved, and mental and physical state of the vehicle occupants. Ignoring these unobserved crash-specific factors can, and in general will, lead to inefficiency in model parameter estimation. Further, if there are unobserved crash-specific factors that moderate the impact of explanatory variables on the injury severity levels sustained by all individuals involved in the crash (see Section 1.3), the result is heteroscedasticity in the error terms across individuals involved in different crashes. Unfortunately, this heteroscedasticity gets ignored when individuals are pooled together across crashes, leading not only to inefficient parameters but also, in general, to biased parameter estimation in the commonly used non-linear univariate injury severity models. The problem with the second approach of modeling the injury severity of only the most severely injured person in a crash is that it does not provide a comprehensive view of the nature and severity of all injuries sustained in the crash.

To be sure, there have been earlier studies focusing on jointly modeling the injury severity of multiple individuals in a crash. However, almost all of these earlier studies have examined the injury severities of multiple individuals in the same vehicle in a crash, but not individuals from different vehicles in the same crash. For example, Hutchinson (1986) and Yamamoto and Shankar (2004) adopted a bivariate ordered probit model to jointly examine the injury severity of the driver and the most severely injured passenger in the same vehicle. Eluru *et al*. (2010) employed a copula-based approach to model the injury severity of all occupants of the same vehicle. Rana *et al*. (2010) also used a copula-based model to jointly examine the injury

2

severity of the two drivers in two-vehicle crashes. This Rana *et al* study is the only study that we are aware of that jointly models the injury severity of occupants in different vehicles involved in a crash. Overall, earlier joint modeling studies have demonstrated the need to undertake a comprehensive and joint injury severity modeling approach considering all victims in a crash.

1.2. Seat Belt Use and Injury Severity

An important issue in safety analysis is to assess the benefit of safety measures, such as the effectiveness of seat belt use in reducing injury severity. Earlier studies have suggested that seat belt use reduces the risk of fatality by 45-60 percent when used by passenger car drivers (see, for example, Cummings *et al*., 2003, Cummings and Rivara, 2004, Evans 1986, and NHTSA, 2009). NHTSA (2009) further suggests that the seat belt use effectiveness rates from earlier studies may be underestimations because of the growing improvement in seat belt designs (and vehicle designs that work in tandem with seat belt designs to reduce injury severity in crashes).

However, most of the previous estimates and studies on seat belt use effectiveness consider seat belt use as an exogenous variable in explaining injury severity. But the decision to fasten a seat belt may be endogenous to the injury severity level sustained, due to one or both of the following reasons: (a) unbelted drivers may be intrinsically unsafe drivers, who are likely to be involved in severe crashes due to their dangerous driving habits (sometimes referred to as the selective recruitment hypothesis; see Eluru and Bhat, 2007 and Evans, 1996), and (b) belted drivers may exhibit negligent and aggressive driving behavior due to the increased safety they perceive by belting up (sometimes referred to as the offsetting behavior hypothesis; see de Lapparent, 2008, and Adams, 1994).[2] If a selective recruitment hypothesis is at work, but the endogeneity of seat belt use is ignored when modeling injury severity, the result is an overestimation of the restraint effectiveness of seat belt use. On the other hand, if an offsetting behavior hypothesis is at work, but the endogeneity of seat belt use is ignored when modeling injury severity, the result is an underestimation of the restraint effectiveness of seat belt use. In either case, it behooves the analyst to consider seat belt use as being endogenous to the modeling of injury severity. While some earlier studies have done so, these studies are undertaken at the level of an individual driver. No earlier study that we are aware of has considered the

---

[2] The offsetting hypothesis has been empirically observed in other similar safety devices, such as the increased risky behavior in recreational boating because of wearing floatation devices (McCarthy and Talley, 1999) and the increased aggressive behavior due to the presence of airbags in vehicles (Winston *et al*., 2006).

endogeneity of seat belt use of a driver in a crash on the injury severity level of that driver, as well as on the injury severity level of other individuals involved in the crash (such as the injury severity level of the driver of a second vehicle involved in the crash).

Of course, our emphasis in the current study is limited to recognizing the endogeneity of seat-belt use on injury severity. Thus, in our analysis, we consider many other variables to be exogenous to injury severity, such as driving under intoxication (DUI), vehicle body type, whether the person is married, and even the road way conditions under which a person is driving. One could argue that these should also be considered endogenous to injury severity in that they reflect personality traits associated with risk avoidance or risk acceptance that also can affect injury severity in a crash. However, the idea of a predictive model is to abstract some from reality and make reasonable assumptions about the exogeneity and endogeneity of variables. In this regard, for a driver, the decision to wear seat belts or not is a regular and frequent choice when driving. So, one could surmise that considering seat belt use as being exogenous when it is endogenous would be more problematic than, for example, considering DUI as being exogenous when it is endogenous. Indeed, it is no surprise at all that endogeneity issues have been studied extensively in the context of seat belt use impact on injury severity (see de Lapparent, 2008, Eluru and Bhat, 2007, Cohen and Einav, 2003, Levitt and Porter, 2001), and not in the context of DUI or vehicle ownership decisions, or other driving decisions on injury severity. But, methods that accommodate endogeneity of a limited number of these other variables may provide additional benefits. In this context, all subsequent statements in this paper regarding the problems in considering seat belt use as a pure explanatory variable in injury severity modeling (rather than modeling seat belt use jointly with injury severity) should be viewed in the light that the proposed model in this paper itself may be further enhanced by modeling additional variables jointly with seat belt use and injury severity.

1.3. Unobserved Heterogeneity in the Effects of Variables

Injury severity studies to date typically make the *a priori* assumption that there are no variations in the effects of explanatory variables. However, it is possible that there are unobserved crash-specific factors that may moderate the impact of explanatory variables. For example, consider the effect of gender. Many earlier studies have indicated that, other things being equal, men are likely to be less severely injured in crashes relative to women, perhaps due to overall physical

build and weight considerations. However, this may not always be the case, since, in some crashes, the smaller structure and build of women may actually provide less surface body exposure to injuries. Such a possibility may be reflected by accommodating a random coefficient on the "male" dummy variable (with "female" being the base category) in the underlying injury risk propensity. Earlier studies that consider unobserved heterogeneity have invariably found statistically significant variations in the effects of variables. Further, these unobserved heterogeneity effects are not simply esoteric econometric enhancements, but can have very real implications for accurately assessing the overall effects of variables and to design countermeasures to reduce injury severity. This realization has resulted in many more studies in the past five years or so that consider unobserved heterogeneity effects in injury severity models (see Eluru and Bhat, 2007, Paleti *et al*., 2010, Christoforou *et al*., 2010, Milton *et al*., 2008, Anastasopoulos and Mannering, 2011, Chen and Chen, 2011, and Moore *et al*., 2011).

1.4. Current Study in Context

The current study contributes to the existing injury severity modeling literature by developing a multivariate probit model of seat belt use and injury severity of both drivers involved in two-vehicle crashes.[3] In doing so, it combines the three strands of relatively isolated literature discussed in the earlier three sub-sections. Indeed, to our knowledge, this is the first study in the injury severity research literature that models the injury outcome of both drivers in two-vehicle crashes and their safety belt use decisions, while also allowing unobserved heterogeneity effects. Two issues deserve particular attention regarding our specification of unobserved heterogeneity effects. First, we not only consider unobserved heterogeneity in the effects of exogenous variables on injury severity (as discussed in Section 1.3), but also consider unobserved heterogeneity in decisions to wear seat belts. Thus, using the same example as in Section 1.3, while men may, on average, be less likely to wear seat belts than women (as observed in several earlier studies; see Reinfurt *et al*., 1996, de Lapparent, 2008, and Eluru and Bhat, 2007), some men may indeed be much more defensive and safety-conscious than women. This can lead to a higher propensity to wear seat belts for some men relative to women, which can be captured by a

---

[3] Extending our model framework to include the seat belt use and injury severity of all individuals involved in a crash is left as a future exercise. However, as we note later in this section, about 70% of two-vehicle crashes in the U.S. had the driver as the sole vehicle occupant of each vehicle, while the corresponding rate for Denmark was about 82%.

random coefficient on the "male" dummy variable. Second, in the process of allowing unobserved heterogeneity effects of variables on injury severity, we allow such effects also on the seat belt use variable (in addition, of course, to recognizing the endogeneity of seat belt use). For instance, unobserved characteristics (such as physical frame or precise sitting posture) may moderate the effect of seat belt use on injury severity.

The joint model takes the form of ordered-response probit specifications for the two injury severity variables (corresponding to the injury severity levels sustained by the two drivers) and a binary probit specification for the two seat belt use variables (corresponding to the seat belt use of the two drivers). Our choice of the ordered-response specification for injury severity deserves some explanation here. Many earlier injury severity studies have used such a specification to appropriately recognize the ordinal nature of the injury severity levels (see, for example, Xie *et al.*, 2009, Christoforou *et al.*, 2010, Haleem and Abdel-Aty, 2010, Quddus *et al.*, 2010, Jung *et al.*, 2010, Paleti *et al.*, 2010, Zhu and Srinivasan, 2011). In recent years, unordered-response specifications have also been considered (see, for example, Rifaat *et al.*, 2011, Yan *et al.*, 2011, Malyshkina and Mannering, 2009, Huang *et al.*, 2008, Milton *et al.*, 2008, Kim *et al.*, 2010, and Moore *et al.*, 2011). These unordered-response specifications, while not recognizing the ordinal nature of injury severity levels, provide additional flexibility in capturing variable effects and are also less vulnerable to parameter inconsistency problems caused by varying under-reporting rates (across injury severity levels) in the data (see Ye and Lord, 2011). In addition, a few studies also use generalizations of the ordered-response specification that lie somewhere between the ordered- and unordered-response specifications in terms of capturing flexible variable effects, while also continuing to recognize the ordinal nature of injury severity levels (see, for example, Eluru *et al.*, 2008, and Castro *et al.*, 2012). Savolainen *et al.* (2011) and Castro *et al.* (2012) discuss in detail the advantages and limitations of each approach. In our multivariate model, the simple structure and relatively easy interpretation of the ordered-response specification has particular appeal, which is the reason for its use here.

In effect, then, our joint system takes a multivariate ordered-response probit model form. A particularly attractive feature of the model form is its flexible covariance structure that allows for a manifestation of both the selective recruitment hypothesis as well as the offsetting behavior hypothesis through the unobserved correlations between the seat belt use propensity equation for a driver and the injury severity level risk propensity for that driver and the other driver. Thus, if

there is a negative correlation, this suggests that, in the net, the selective recruitment hypothesis is at work (that is, drivers wearing seat belts are intrinsically more cautious and responsible drivers who drive defensively, leading to lower injury severity levels of both drivers involved in the crash). On the other hand, if there is a positive correlation between the seat belt use propensity error term for driver A and the injury severity level error terms for driver A and driver B, this suggests that, in the net, the offsetting behavior hypothesis dominates over the selective recruitment hypothesis (that is, drivers wearing seat belts drive more aggressively and in a risky fashion, leading to higher injury severity levels of both drivers involved in the crash). Of course, empirically speaking, it is possible that the directions of correlation are different between driver A's seat belt use and driver A's injury severity level, and driver A's seat belt use and driver B's injury severity level. If such an empirical situation manifests itself, it would be difficult to intuitively explain things and may suggest a constrained inference procedure that disallows such a result.

To summarize, the current paper contributes methodologically by formulating and estimating a specific type of multivariate ordered probit system. The proposed system is different from the one by Bhat and Srinivasan (2005) in two important ways. First, Bhat and Srinivasan do not consider unobserved heterogeneity in the effects of variables, which, as we have discussed in Sections 1.3 and 1.4, can be critical for the accurate estimation of the effects of variables and the design countermeasures to reduce injury severity. While such unobserved heterogeneity can theoretically be introduced in Bhat and Srinivasan's model framework, it becomes very cumbersome because Bhat and Srinivasan mix a multivariate normal error component vector (to capture correlations in the overall propensities across dependent variables) with an independent and identically distributed kernel logistic value term (for the propensity underlying each ordered-response dependent variable). In this setting, the addition of normal unobserved heterogeneity in response to each variable will increase the integration dimensionality in the unconditional probability expression for the ordinal outcomes by one. This creates problems when testing unobserved heterogeneity for many variables at the same time. However, the conjugate nature of the sum of normal distributions enables the inclusion of unobserved heterogeneity on multiple variables at the same time in our proposed approach (because we use a kernel normal term for the injury severity propensity) without increasing the dimensionality of integration (which remains fixed at four, corresponding to the four dependent variables of seat-belt use and injury

severity of each of the two drivers involved in the crash). Thus, our proposed approach leads to substantial computational efficiency, and also is a more natural way to accommodate unobserved heterogeneity rather than assuming a multivariate normal distribution for the coefficient heterogeneity but a logistic distribution for the kernel error terms. Second, our proposed formulation constitutes a specific version of the generalized Roy model that is referred to as the treatment effects model (see Heckman and Vytlacil, 2005 and Bhat and Eluru, 2009). In our context, "seat belt use" of each driver is the "treatment", and the "outcomes" correspond to the injury severities of the two drivers. Further, unlike earlier treatment effects models that have a single treatment and a single outcome, the proposed model considers the case of two treatments and two outcomes (with both treatments considered endogenous and each treatment affecting both outcomes). That is, in addition to error correlations in the multivariate model system, the seat belt use of each driver appears as a determinant of injury severity propensity of that driver as well as the injury severity propensity of the other driver. Thus, while the resulting system is a type of multivariate ordered-response probit system, there is much more structure in our proposed formulation than the simple reduced form formulation of Bhat and Srinivasan. This is also evidenced in our discussion later in Section 3.1.

In addition to making a methodological contribution to the injury severity literature, the current paper also makes a substantive empirical contribution by studying seat belt use and injury severity of both drivers in two-vehicle crashes (a "vehicle" as used in the current paper refers to a motorized transport form in which drivers sit in an enclosed space; examples include passenger cars, sports-utility vehicles, and minivans, but, for this paper, we exclude crashes involving buses and commercial vehicles). The reader will note that two-vehicle crashes constitute the most common type of crash scenario in most parts of the developed world. For instance, in 2010, 51% of all roadway crashes in the U.S. involved two vehicles, while the corresponding figure for Denmark was 61%.[4] Further, for the same year, within the pool of two-vehicle crashes, about 70% of these crashes in the U.S. had the driver as the sole vehicle occupant of each vehicle, while the corresponding rate for Denmark was about 82%. In examining seat belt use and injury severity levels of the two drivers in such two-vehicle crashes,

---

[4] These figures are computed using the 2010 NASS data (available at: ftp://ftp.nhtsa.dot.gov/NASS/2010/Formatted%20Data/) and the Statistics Denmark data base (available at: http://statbank.dk/statbank5a/default.asp?w=1093), respectively.

we also consider a comprehensive set of driver attributes that have seldom been considered in the past, including marital status and crime history of drivers.

The remaining sections of this paper are organized as follows. Section 2 provides an overview of the data source and sample statistics. Section 3 describes the specification of our model and presents the estimation approach. Section 4 discusses the estimation results. Finally, Section 5 offers concluding thoughts.

2. THE DATA

2.1. Data Source

The data used in this research is based on police-reported road crash data from Denmark. This nationwide crash data base provides comprehensive information on crash location, road users involved in the crash, driver and other vehicle occupant characteristics for each vehicle involved in the crash, crash type characteristics, and the injury severity of each individual involved in the crash.[5] From this database, we culled out two-vehicle, two-driver (only the driver as the occupant in each vehicle) crashes during the years between (and including) 2002 and 2008. The final sample for analysis included 5,077 crashes involving 10,154 drivers.

2.2. Sample Description

The police-reported crash data used in this paper records the injury severity of drivers in four ordinal levels: (1) no injury, (2) minor injury, (3) serious injury, and (4) fatal injury. In our sample, there were very few crash records that resulted in a fatal injury (only about 1% of the 10,154 drivers sustained fatal injury), and so we collapsed the serious injury and fatal injury categories into a single "serious" injury category. Also, the police-reported data labels the drivers involved in two-vehicle crashes as first party and second party based on the number of driving violations committed by the drivers, with the first party being the driver with a higher number of driving violations. However, this police designation is not an exact or scientific assignment. Thus, we do not attach any significance to the designation as the "first party" or "second party" in our analysis, and treat drivers as being exchangeable in the econometric analysis (thus, for example, the effects of driver characteristics on the driver's seat belt use and injury severity are held to be the same regardless of whether the driver is designated as the "first party" or "second

---

[5] Further details of the database are available at http://statbank.dk/statbank5a/default.asp?w=1093.

party"; similarly, the covariance generated by unobserved driver characteristics in a driver's seat belt use propensity and injury severity risk is constrained to be the same regardless of "first party" or "second party" designation).

Across both drivers, the percentage of drivers who use a seat belt is about 89%, while 11% do not use a seat belt (this sample seat belt use rate is close to the 92% nationwide Denmark seat belt use rate for personal vehicles).[6] The distribution of injury severity levels across both drivers is as follows: no injury (47.7%), minor injury (33.1%), and serious injury (19.2%). Table 1a provides a cross-tabulation of seat belt use by a specific driver and the injury severity level sustained by that same driver (across both drivers involved in each crash and across all crashes; so, the sample size for this cross-tabulation is $5,077 \times 2 = 10,154$ observations). The cross-tabulation indicates the clear negative association between seat belt use and injury severity. Thus, among those drivers using a seat belt, 50.6% sustain no injury, while 17.4% sustain a severe injury. However, among those drivers not using a seat belt, the corresponding percentages are 23.5% and 34.3%. Table 1b provides a similar cross-tabulation as Table 1a, but now for the relationship between seat belt use by a specific driver and the injury severity sustained by the other driver. Interestingly, the cross-tabulation indicates a positive relation between seat belt use of one driver and the injury severity of the other driver. To be specific, among those drivers involved in a crash with other belted drivers, 46.4% sustained no injury and 19.5% are seriously injured while the corresponding rates for those drivers involved in a crash with other unbelted drivers are 57.7% and 17.3%. So, the preliminary descriptive analysis seems to suggest that the effect of the offsetting behavior of a belted driver is proving to be dangerous to the other driver involved in the crash.

Exogenous variables

Several explanatory variables were considered in our multivariate analysis. For each variable, a range of possible functional form specifications were attempted. Thus, for continuous variables such as the age of the driver, we considered a linear effect as well as a non-linear effect (by specifying piece-wise linear effects as well as dummy variables). For categorical variables, we initially specified dummy indicators at the most disaggregate discrete levels available for the variable (including interaction effects among variables at this disaggregate level), and

---

[6] See http://www.sikkertrafik.dk/I-bil/Sikkerhedssele/~/media/Files/Rapporter/selerapport-2010.ashx.

progressively worked toward more efficient specifications based on intuitive considerations as well as statistical considerations (for example, combining the effects of two discrete categories if the effects were not statistically different). Further, we also tested a whole host of interaction variables in both the seat belt use and injury severity propensity equations (for example, such as driver age-gender interaction effects and vehicle type-environmental conditions interaction effects). But these interactions effects turned out to be statistically insignificant.

The explanatory variables may be grouped into the following categories: driver characteristics (own characteristics) and other driver characteristics (characteristics of the other driver involved in the crash), vehicle characteristics and other driver's vehicle characteristics, environmental attributes, roadway characteristics, and crash characteristics. Driver characteristics included gender, age, alcohol use, whether with legal driving license or not, marital status, and the crime history of drivers (an indicator variable that takes a value of one if a driver has committed one of more registered crimes in the three years before the crash).[7] The vehicle characteristics included the body type and weight of the vehicles involved in the crash. Specifically, the vehicle type may be a private car, a van weighing less than 2000 kg, or a van weighing above 2000 kg.[8] As in earlier safety studies (see, for example, Zhu and Srinivasan, 2011, Rana *et al.*, 2010, Jung *et al.*, 2010, and Huang *et al.*, 2008), we do not consider vehicle make/model in our analysis because it leads to an explosion in the number of model parameters. Besides, this make/model data was not available in the otherwise rich Danish crash data set used in the current analysis. Environmental attributes included weather conditions (dry weather, raining, or other types of adverse weather conditions) and lighting conditions (daylight, night-lighted and night-unlighted). Roadway attributes included the type of the road (two lane with no center divider, two-lane with center divider, four lane, and motorway),road geometry (intersection with four legs, intersection with three legs, roadway segment with horizontal curvature, straight road, or other type), and the posted speed limit.[9] Finally, the crash

---

[7] Variables such as marital status and crime history are considered in our analysis as proxy variables for aggressiveness/reckless tendencies that may impact seat belt use and injury severity levels.

[8] The private car body type includes all motorized passenger vehicles except vans.

[9] Most crash data sets that are based on police reports and/or based on immediate reporting of collision characteristics do not have the speed at which the vehicles were traveling just prior to the crash (because the observation is made in the aftermath of a collision). So, we use the roadway speed limit as a proxy for the speed of vehicles at the time of the crash, as has been done in most earlier crash analysis studies (see, for example, Anastasopoulos and Mannering, 2011, Haleem and Abdel-Aty, 2010, Rana *et al.*, 2010, Malyshkina and Mannering, 2009, and Eluru *et al.*, 2008).

characteristics considered were collision type (frontal, rear-end, or angle), driver's trajectory of motion at the time of the crash and that of the other driver's trajectory of motion (turning left, turning right, heading straight or reversing), and the role of each vehicle (overtaking, striking, or struck). Overall, we consider a wide range of crash characteristics in the current study.

3. THE MODEL

3. 1. Multivariate Ordered-Response Probit Model

The model system in this paper includes two ordered-response variables (corresponding to the two drivers' injury severity level in the crash) and two binary-response variables (corresponding to the seat belt use decisions of the two drivers). However, since a binary-response variable is a special case of an ordered-response variable with only two discrete states, we will refer to our model system as a multivariate ordered-response model system. Further, to generate a flexible (but also theoretically appropriate) correlation pattern among the four decision variables, we adopt a multivariate normal distribution for the four error terms, which leads to the proposed multivariate ordered-response probit (MORP) model system. To write this system, let q be an index for crashes ($q = 1, 2, \ldots, Q$), d be an index for drivers ($d = 1, 2$), g_{qd} be the actual observed seat belt use decision for driver d in crash q (g_{qd} may take a value of '0' indicating seat belt non-use or '1' indicating seat belt use), and m_{qd} be the actual observed injury severity level for driver d in crash q (this observed injury severity level may take values between 1 to 3, with the value of '1' corresponding to "no injury", the value of '2' corresponding to "minor injury", and the value of '3' being "serious injury". Then, the MORP model system takes the following form:

$$s_{qd}^* = \boldsymbol{\theta}_q' \mathbf{z}_{qd} + \xi_{qd}, s_{qd} = g_{qd} \quad \text{if } \mu^{g_{qd}-1} < y_{qd}^* < \mu^{g_{qd}}$$

$$y_{qd}^* = \boldsymbol{\beta}_q' \mathbf{x}_{qd} + \varepsilon_{qd}, y_{qd} = m_{qd} \quad \text{if } \quad \psi^{m_{qd}-1} < y_{qd}^* < \psi^{m_{qd}}$$

(1)

where s_{qd}^* is the latent seat belt use propensity of driver d in crash q, $\mathbf{z}_{qd}$ is an ($L{\times}1$) vector of exogenous variable characteristics that explains driver d's seat belt use ($\mathbf{z}_{qd}$ excludes a constant), $\boldsymbol{\theta}_q$ is the corresponding vector of crash specific coefficients (this coefficient vector is held to be the same across both drivers to recognize the exchangeability of the drivers), ξ_{qd} is an idiosyncratic standard normally distributed error term assumed to be identically and

independently distributed across crashes q (with additional specifications given later), $\mu^{g_{qd}}$ is the upper threshold corresponding to the seat belt use decision g_{qd} (with $\mu^{-1} < \mu^0 < \mu^1$; $\mu^{-1} = -\infty$, $\mu^1 = +\infty$), y_{qd}^* is the latent injury risk propensity associated with driver d in crash q, x_{qd} is ($K{\times}1$) vector of variables (excluding a constant, but including exogenous variables affecting injury severity as well as the endogenous seat belt use decisions of the driver d (s_{qd}) and that of the other driver d' ($s_{qd'}$)), $\boldsymbol{\beta}_q$ is a corresponding crash-specific vector of coefficients to be estimated (this coefficient vector is also held to be the same across both drivers, to recognize the exchangeability of the drivers), ε_{qd} is an idiosyncratic standard normal error term assumed to be identically and independently distributed across crashes q (with additional specifications given later), and $\psi^{m_{qd}}$ is the upper threshold corresponding to injury severity outcome m_{qd} (with $\psi^0 < \psi^1 < \psi^2 < \psi^3$; $\psi^0 = -\infty$, $\psi^3 = +\infty$). In Equation system (1), the latent propensities y_{qd}^* ($d = 1, 2$) are mapped to the injury severity level y_{qd} by the ψ thresholds in the usual ordered-response fashion (again, to ensure exchangeability, the thresholds are specified to be fixed across drivers).[10]

The parameter vector $\boldsymbol{\theta}_q$ in Equation (1) is assumed to be a realization from a multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Gamma}$.[11] Then, we can write $\boldsymbol{\theta}_q = \boldsymbol{\theta} + \widetilde{\boldsymbol{\theta}}_q$, where $\widetilde{\boldsymbol{\theta}}_q \sim MVN_L(0, \boldsymbol{\Gamma})$ (MVN_L represents the multivariate normal distribution of dimension L). Similarly, the parameter vector $\boldsymbol{\beta}_q$ is defined as a realization from a multivariate normal distribution with mean vector $\boldsymbol{b}$ and covariance matrix $\boldsymbol{\Omega}$, so that we can write $\boldsymbol{\beta}_q = \boldsymbol{b} + \widetilde{\boldsymbol{\beta}}_q$, $\widetilde{\boldsymbol{\beta}}_q \sim MVN_K(0, \boldsymbol{\Omega})$. These specifications accommodate unobserved

---

[10] In the model in Equation (1), we do not employ the seat belt use indicator for one driver in the z_{qd} vector of the other driver to explain the other driver's seat belt use propensity. This is because of two reasons. First, there is no reason to expect that one particular driver's seat belt use should determine the disposition of the other driver's seat belt use. Second, doing so implies that we should allow each driver's seat belt use to affect the other driver's seat belt use (because of exchangeability). However, the seat belt use of both drivers are being jointly modeled, and there will be a breakdown in estimation if each driver's seat belt use is introduced as an explanatory variable in the other driver's seat belt use propensity. Fundamentally, there is a logical inconsistency problem that arises, as explained by Maddala (1983, pages 205-240).

[11] For ease in presentation, we will treat all elements of $\boldsymbol{\theta}_q$ as random, but this is not necessary; the researcher can fix some elements of $\boldsymbol{\theta}_q$ and let the remaining elements be random.

heterogeneity effects in the impact of variables on seat belt use as well as injury severity, as discussed earlier in Section 1.3 and 1.4.

Next, consider the following stacked (4×1) vector of error terms: $\varsigma_q = (\xi_{q1}, \xi_{q2}, \varepsilon_{q1}, \varepsilon_{q2})'$. This vector is multivariate normally distributed with a correlation matrix given below:

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{bmatrix} \tag{2}$$

However, such a free correlation matrix will violate exchangeability considerations. In this regard, the correlation parameter ρ_{13} captures the endogeneity of seat belt use of one driver on that driver's injury severity level, while ρ_{24} captures the endogeneity of seat belt use of the other driver in the crash on that other driver's injury severity level. For exchangeability reasons, we should have $\rho_{13} = \rho_{24} = \rho_{self}$, where the notation ρ_{self} refers to the self-endogeneity seat belt use effect. As discussed in Section 1.4, if this correlation parameter comes out to be positive (negative), it implies that offsetting behavior dominates over (is subordinate to) self-selection. Similarly, we also should have $\rho_{14} = \rho_{23} = \rho_{cross}$, which represents the cross-endogeneity seat belt use effect of one driver on the injury severity of the other driver. Of course, we expect the self- and cross-endogeneity seat belt use effects to be in the same direction, and also the self-effect to be larger in magnitude than the more indirect cross-effect. In addition, we expect that ρ_{12} in the correlation matrix of Equation (2), which captures correlation effects in the seat belt use decisions of the two drivers involved in the crash, should be zero, since the two drivers in a crash are unlikely to be related in any way. So, we constrain this parameter to a value of zero.[12] The final covariance matrix has three parameters for estimation, as follows:

$$\Sigma = \begin{bmatrix} 1 & 0 & \rho_{self} & \rho_{cross} \\ 0 & 1 & \rho_{cross} & \rho_{self} \\ \rho_{self} & \rho_{cross} & 1 & \rho_{IS} \\ \rho_{cross} & \rho_{self} & \rho_{IS} & 1 \end{bmatrix}, \tag{3}$$

---

[12] We also had this parameter free in our empirical estimations as part of our specification testing, but it came out to be consistently very small in magnitude and not statistical significantly different from zero at even the 0.50 level of significance.

where ρ_{IS} $(=\rho_{34})$ represents crash-specific unobserved correlation effects in the injury severity risk of the two drivers. Obviously, if we set all the off-diagonal terms to zero in the above correlation matrix, the model collapses to an independently formulated system of four equations.[13]

3.2. Estimation Approach

The parameter vector to be estimated in the model system is $\boldsymbol{\delta} = (\boldsymbol{\theta}',\boldsymbol{b}',\mu^0,\psi^1,\psi^2,\overline{\boldsymbol{\Gamma}}',\overline{\boldsymbol{\Omega}}',\overline{\boldsymbol{\Sigma}}')'$, where $\overline{\boldsymbol{\Gamma}},\overline{\boldsymbol{\Omega}}$, and $\overline{\boldsymbol{\Sigma}}$ represent column vectors that vertically stack the upper diagonal elements of the matrices $\boldsymbol{\Gamma},\boldsymbol{\Omega}$, and $\boldsymbol{\Sigma}$, respectively. To write the likelihood function, we define additional vectors and matrices. Let $\boldsymbol{z}_q = (\boldsymbol{z}_{q1},\boldsymbol{z}_{q2})'$ ($2 \times L$ matrix), $\boldsymbol{x}_q = (\boldsymbol{x}_{q1},\boldsymbol{x}_{q2})'$ ($2 \times K$ matrix), $\boldsymbol{\lambda} = (\boldsymbol{\theta}',\boldsymbol{b}')'$ [$(L+K) \times 1$ vector], $\widetilde{\boldsymbol{\lambda}}_q = (\widetilde{\boldsymbol{\theta}}_q,\widetilde{\boldsymbol{\beta}}_q)'$ [$(L+K) \times 1$ vector], $\varphi_q^{\text{UP}} = \left(\mu^{g_{q1}},\mu^{g_{q2}},\psi^{m_{q1}},\psi^{m_{q2}}\right)'$ (4×1 vector), $\varphi_q^{\text{DOWN}} = \left(\mu^{g_{q1}-1},\mu^{g_{q2}-1},\psi^{m_{q1}-1},\psi^{m_{q2}-1}\right)'$ (4×1 vector), $\boldsymbol{sy}_q^* = \left(s_{q1}^*,s_{q2}^*,y_{q1}^*,y_{q2}^*\right)'$ (4×1 vector), $\boldsymbol{sy}_q = \left(s_{q1},s_{q2},y_{q1},y_{q2}\right)$ (4×1 vector), and $\boldsymbol{gm}_q = \left(g_{q1},g_{q2},m_{q1},m_{q2}\right)'$ (4×1 vector). Define $\boldsymbol{w}_q$ as a block diagonal matrix of size $4 \times (L+K)$, with the first two rows and the first L columns occupied by the matrix $\boldsymbol{z}_q$, and the last two rows and the last K columns occupied by the matrix $\boldsymbol{x}_q$. Also, let $\boldsymbol{\Lambda}$ be another block diagonal matrix of size $(L+K) \times (L+K)$ with the first block occupied by $\boldsymbol{\Gamma}$ and the second by $\boldsymbol{\Omega}$. Then, in matrix form, we may write Equation (1) compactly as:

$$\boldsymbol{sy}_q^* = \boldsymbol{w}_q\boldsymbol{\lambda} + \boldsymbol{w}_q\widetilde{\boldsymbol{\lambda}} + \varsigma_q, \ \boldsymbol{sy}_q = \boldsymbol{gm}_q \ \text{ if } \ \varphi_q^{\text{DOWN}} < \boldsymbol{sy}_q^* < \varphi_q^{\text{UP}} \tag{4}$$

where $\boldsymbol{sy}_q^* \sim MVN_4(\boldsymbol{B}, \boldsymbol{\Delta})$, $\boldsymbol{B} = \boldsymbol{w}_q\boldsymbol{\lambda}$ and $\boldsymbol{\Delta} = \boldsymbol{w}_q\boldsymbol{\Lambda}\boldsymbol{w}_q' + \boldsymbol{\Sigma}$,

The log-likelihood function then is given by:

$$L_q(\boldsymbol{\delta}) = \text{Prob}[\boldsymbol{sy}_q = \boldsymbol{gm}_q] = \Phi_4\left[(\boldsymbol{\omega}_\Delta)^{-1}(\varphi_q^{\text{UP}} - \boldsymbol{B}),\boldsymbol{\Delta}^*\right] - \Phi_4\left[(\boldsymbol{\omega}_\Delta)^{-1}(\varphi_q^{\text{DOWN}} - \boldsymbol{B}),\boldsymbol{\Delta}^*\right], \tag{5}$$

---

[13] As an aside, we also estimated models that allowed a completely free covariance matrix $\boldsymbol{\Sigma}$ to examine the data fit offered by such a general, but unintuitive correlation structure, and to check the robustness of the constrained covariance matrix estimated in this paper. Interestingly, the estimated free covariance matrix strongly supported the *a priori* restrictions we imposed. Interested readers may obtain the free covariance matrix estimation results from the authors, but we will not be presenting these results in this paper to conserve on space.

where $\Delta^* = \boldsymbol{\omega}_\Delta^{-1} \boldsymbol{\Lambda} \boldsymbol{\omega}_\Delta^{-1}$, $\boldsymbol{\omega}_\Delta$ is the diagonal matrix of the standard deviations of $\boldsymbol{\Delta}$, and $\Phi_4(.,.)$ represents the 4-variate standard multivariate cumulative normal distribution (MVNCD) function.

Several simulation and analytic approximations have been proposed for evaluating the MVNCD function. In this paper, we adopt the most commonly used GHK (Geweke-Hajivassiliou-Keane) simulator (see Geweke, 1991, Hajivassiliou and McFadden, 1998, and Keane, 1994), but by enhancing it through the use of Halton draws rather than pseudo-random draws (see Bhat, 2011 and Bhat *et al.*, 2010 for extensive discussions of alternative approaches to evaluate the MVNCD function, including a detailed discussion of the Halton draw-enhanced GHK simulator). We used 150 Halton draws per individual within the GHK simulator, but also extensively tested for the stability of the GHK-based estimates using varying number of Halton draws (see Bhat, 2001, and Bhat, 2003 for discussions of the Halton draw technique). The estimation results were stable even for a small number of draws, with literally no change in the parameter estimates and standard errors beyond even 100 Halton draws.

A few final points regarding estimation. First, we need to ensure the positive definiteness of the covariance matrix $\boldsymbol{\Delta}$, which is assured if the covariance matrices $\boldsymbol{\Gamma}$, $\boldsymbol{\Omega}$, and $\boldsymbol{\Sigma}$ are positive definite. To ensure the positive definiteness of $\boldsymbol{\Gamma}$ and $\boldsymbol{\Omega}$, all that needs to be done is to parameterize the likelihood function in terms of the Cholesky matrices of $\boldsymbol{\Gamma}$ and $\boldsymbol{\Omega}$. A similar procedure can be adopted for the matrix $\boldsymbol{\Sigma}$, except that $\boldsymbol{\Sigma}$ is also a correlation matrix. This is maintained by adopting the procedure proposed by Bhat and Srinivasan (2005) for assuring a positive definite correlation matrix. Second, due to simulation noise when using finite number of draws, standard errors for our parameters are computed using the inverse of the sandwich information matrix (sometimes referred to as the Huber-White or robust standard errors; see McFadden and Train, 2000). Third, technically speaking, there is no need for exclusion restrictions in the vectors of explanatory variables $\boldsymbol{x}_{qd}$ and $\boldsymbol{z}_{qd}$, because the non-linearities in the model system allow identification even without such exclusion restrictions. In our model system, however, many variables that impacted seat belt use did not significantly impact injury severity. Such variables help in stabilizing the overall model system and aid the estimation process. For example, the Danish Council for Safe Driving introduced a Point Demerit System (PDS) in the year 2005 that began penalizing drivers who failed to restrain children below the age of 14. The penalization was in the form of a demerit point (a citation) on the driving record for each

16

violation, with three demerit points in a span of three years resulting in the suspension of the driving privilege (license). The introduction of the PDS program has been shown to increase the overall national seat belt use rate in Denmark, not just for children but also for all vehicle occupants including drivers.[14] So, we introduced a dummy variable corresponding to whether or not a crash occurred after 2005 as a dummy variable in the seat belt use equation (as part of the z_{qd} vector) to proxy the effect of the PDS program on seat belt use. However, this dummy variable did not statistically significantly affect the injury severity level of the drivers involved in a crash.

4. ESTIMATION RESULTS AND DISCUSSION

In this section, we discuss the key empirical results, their implications, and model fit statistics. To assess the comparative performance of our preferred model, we estimated three different models: (1) Independent ordered-response injury severity probit specifications for the two drivers, and independent seat belt use binary probits for the two drivers, which we will refer to as the independent probit (IP) model, (2) a multivariate (ordered-response) probit model for injury severity and seat belt use that accounts for seat belt use endogeneity and cross-driver injury severity dependence due to shared unobserved factors (but ignores parameter heterogeneity), which we will refer to as the multivariate probit (MVP) model, and (3) a multivariate probit model for injury severity and seat belt use, as well as unobserved heterogeneity effects in the injury severity and seat belt use equations of both drivers, which we will refer to as the random coefficients MVP (or RCMVP) model. In all the three models, the final specification is obtained through a systematic process of removing statistically insignificant variables and combining variable effects if appropriate and not statistically different. The specification process was also guided by prior research and parsimony considerations. Conveniently, we obtained the same variable specification as the best specification for all the three models, and thus we are able to use standard likelihood ratio tests to assess data fit among the three models. In particular, if the off-diagonal elements in Equation (3) in the MVP model are constrained to be all zero (*i.e.*, if $\rho_{self} = 0$, $\rho_{cross} = 0$, and $\rho_{IS} = 0$), the MVP model collapses to the IP model. Further, if there are

---

17

no statistically significant random coefficients in the RCMVP model (*i.e.,* if all elements of $\boldsymbol{\Gamma}$ and $\boldsymbol{\Omega}$ are identically zero), the RCMVP model collapses to the MVP model.

In our estimations, the log-likelihood values at convergence for the three models were as follows: IP model (-12915), MVP model (-12857), and the RCMVP model (-12827). In the MVP model, only the ρ_{self} and ρ_{IS} coefficients turned out to be statistically significantly different from zero (with ρ_{cross} not being statistically different from zero, and hence being removed from the specification), as we will discuss later in more detail. Thus, the MVP model has two more parameters than the IP model. For the RCMVP model, the best specification included three random coefficients (with a diagonal covariance specification) in the injury severity equation, but no statistically significant random coefficients in the seat belt use equation, leading to three additional parameters relative to the MVP model. The likelihood ratio test statistic for comparing the MVP and IP models returns a value of 116, which is higher than the corresponding chi-squared table value with two degrees of freedom at any reasonable level of significance. Similarly, the likelihood ratio test statistic for comparing the RCMVP and MVP models is 60, which is higher than the corresponding chi-squared value with three degrees of freedom at even the 0.001 significance level). Clearly, the RCMVP model decisively rejects the other two models, indicating the presence of seat belt use endogeneity, cross-driver injury severity correlation due to unobserved factors, as well as unobserved heterogeneity in the effects of a few variables.

In the rest of this section, we discuss only the results of the RCMVP model for presentation focus and ease, though we will comment as appropriate on the other two models when discussing specific coefficient estimates. Table 2 presents the RCMVP model results for both the injury severity and seat belt use equations.

4.1. Drivers' Injury Severity Component

Several categories of variables turned out to be important determinants of the injury severity of the drivers. In the subsequent presentation of results, we separate out driver characteristics from the remaining characteristics to streamline the presentation.

Driver Characteristics

Among driver characteristics, drivers who do not wear their seat belt and who drive under the influence of alcohol have higher injury risk propensity than their peers who wear their seat belt and who driver sober, respectively. These results indicate the effectiveness of seat belt usage as a safety restraint device, and suggest that driving under the influence (DUI) of alcohol may lead to aggressive driving as well as an inability to react in ways to reduce the impact of a crash as it starts to develop. In the context of the seat belt use effect, note that we are controlling seat belt endogeneity effects (discussed later), so the parameter on seat belt use represents a "pure" engineering safety effect through restraining the driver in a crash and distributing impact forces. The results also show that male drivers, on average, are less likely to suffer serious injuries relative to female drivers, perhaps due to physiological gender disparities in absorbing impact forces. This is consistent with earlier safety research (see, for example, Eluru and Bhat, 2007, Ulfarsson and Mannering, 2004). However, the gender effect is heterogeneous, as can be observed from the highly statistically significant standard deviation estimate on the "male" variable (Paleti *et al*., 2010, and Eluru and Bhat, 2007 also note a similar finding). That is, there are crashes where, other observed characteristic being equal, women do come out less severely injured than men. The mean and standard deviation estimates indicate that, other characteristics being equal, men have a lower injury risk propensity than women in about 90% of crashes.

Several different linear and non-linear specifications were considered for the age effect, but the best specification was obtained using dummy variables for four categories: less than 25 years of age, 25-30 years of age, 31-60 years of age (base category in Table 2), and over 60 years of age. The results indicate a higher injury risk propensity, on average, for older drivers (over 60 years of age) relative to younger drivers, possibly attributable to lower bone mass density and less flexibility as individuals age. However, once again, there is substantial heterogeneity in this effect, revealing that, while 72% of drivers aged above 60 years face a higher injury risk than their younger peers, 28% of drivers aged above 60 years have a lower injury risk than their younger peers (this last result may be because of more driving experience and ability to take evasive action as a crash starts to develop).

The seat belt use (or not) of the driver of the other vehicle involved in the crash is also an important determinant of injury severity of the driver, as can be observed from the statistically significant positive effect in the injury severity component in Table 2 on seat belt use of the other

driver (under "other driver's characteristics"). This result is consistent with the preliminary descriptive statistics reported in Section 2 and lends support to the offsetting behavior of seat belt restrained drivers. Specifically, seat belt restrained drivers could exhibit negligent driving behavior due to the increased safety they perceive by belting up (see the discussion in Section 1.2). The positive coefficient on the male variable suggests that drivers hit by or colliding with vehicles driven by male drivers are likely to sustain more serious injuries compared to drivers hit by or colliding with vehicles driven by female drivers (we also interacted the gender of the driver with that of the other driver to see if there were any gender interaction influences, but did not find any such statistically significant effects). A reason for the gender effect may be the higher likelihood of aggressive driving behavior among men relative to women, as documented by Paleti *et al.* (2010), Shinar and Compton (2004), and Dahlen and White (2006). Taken in combination with the effect of a driver being male, the results suggest an interesting situation. Crashes with a male driver are likely to lead to higher injury severity for the other vehicle's driver, but also the men involved in the crash suffer less severe injuries than the women involved in the crash. Overall, women seem to bear the brunt of the effects of a crash. One possible countermeasure may be to stress in information campaigns that men, if they drive aggressively, are not putting themselves as much at risk as they are putting others on the road at risk, especially women drivers. An appeal to civic responsibility and the uneven distribution of the impacts of crashes may be an effective educational campaign tool to reduce aggressive driving among males (though it is perhaps a lack of societal concern that underlies, at least to some extent, aggressive driving in the first place).[15]

Other Characteristics

The results for vehicle characteristics in Table 2 reveal that drivers in the relatively lighter sedan class of vehicles tend to suffer serious injuries more so than drivers in other vehicle types (SUVs, pick-up trucks, minivans, and vans). The safest individuals, from a post-crash injury severity point of view, are those driving heavy mini-vans and vans. Not surprisingly, the effects are almost exactly reversed for drivers who are involved in crashes with other drivers driving a

---

[15] In another study, Rana *et al.* (2010) noted that alcohol consumption of one of the drivers involved in the accident increases the injury severity risk of the other driver, although we did not find a similar effect in our analysis. Also, we tested extensively for interactions between gender and age of both drivers involved in the crash, but did not observe any such statistically significant effects.

sedan or a heavy minivan/van. These results are obviously attributable to differences in the mass and body size of different vehicle types. However, in addition to the direct effects of mass/size differences of the vehicles involved, it is also quite possible (as found by Paleti *et al.*, 2010) that drivers of large vehicles drive more aggressively (than drivers of small vehicles) because of the powerful engine capability of large vehicles and the knowledge that they are not likely to be severely injured in crashes. As with the gender effect discussed earlier, the sensitive issue here is that drivers who are least likely to be injured severely are most likely to injure other users on the roadway. Focused information campaigns to underscore this result may help drivers of smaller vehicles make better driving decisions when around larger vehicles, while also appealing to the civic responsibility of drivers of larger vehicles to be sensitive to the vulnerabilities of drivers in smaller vehicles. Further, the results point to the continued need for vehicle manufacturers to invest in safety technology for both small and large vehicles to minimize crash-related injury severity.

Among the roadway characteristics considered in our analysis, the speed limit posted on the roadway (a proxy for actual driving speeds on the roadway) and the type of roadway on which the crash occurs are both important determinants of injury severity. Specifically, drivers experience progressively higher injury severity risk with progressively higher speed limits (with the base speed category being roads with speed limits not exceeding 50 kilometers per hour. Further, crashes on roadways with two lanes (across both directions of travel) and a central divider (median barrier) are associated with lower injury risk compared to crashes on roadways with two-lanes without a central divider, a reflection of the physical benefits of central dividers in separating traffic traveling in opposite directions and providing drivers with more physical safety margins (see also Anastopoulos and Mannering, 2011 for a similar result). Also, two-lane roads with central dividers are usually located in urban and crowded areas where drivers may be more cautious and drive more defensively (de Lapparent, 2008).

The final set of exogenous variables in Table 2 is related to crash characteristics. While we attempted many different combinations of collision type, trajectory of motion of the two vehicles at the time of the crash, and the role of each vehicle in the crash (see Section 2.2), the final specification in this variable category comprised only three dummy variables as shown in Table 2, with the base category being all crashes not involving a frontal impact of any kind. Overall, according to the results, frontal collisions are, in general, more likely to result in high

injury severity risk compared to other types of collisions. The risk is particularly high when a driver is struck by another vehicle moving frontally rather than when the driver strikes another vehicle frontally. This is to be expected because vehicles moving directly forward are likely to be traveling at high speeds just prior to the point of impact. Crashes that are head-on (both vehicles moving frontally) are, on average, the most dangerous due to the high speeds involved in such crashes. However, there is also substantial heterogeneity, with 23% of such head-on crashes actually resulting in lower injury severity risk for the drivers involved than other crash types (attributable perhaps to last-minute injury minimizing actions through body positioning and protective postures).

4.2. Seat Belt Use Component

The empirical findings in Table 2 with regard to seat belt use are consistent with the findings from earlier safety research. Among the driver characteristics, men and younger adults are less likely to fasten their seat belts relative to women and older drivers, respectively. Two other driver characteristics, seldom used in earlier safety research on seat belt use, also turned out to be important determinants of seat belt use. Specifically, married drivers are more likely to fasten their seat belts relative to unmarried drivers, presumably a reflection of more familial and financial responsibilities that make married drivers adopt less risk-taking driving behavior than other drivers.[16] On the other hand, drivers with a registered crime history are less likely to buckle their seat belts compared to their peers, probably due to their aggressive personalities in general (Abay, 2012). These results suggest targeted information and education campaigns to increase seat belt use, emphasizing the fact that wearing seat belt use is not only a matter of self-protection, but also an important citizenship act to reduce crash-related injury severity in general.

The roadway and crash characteristics indicate lower seat belt use when drivers travel on two lane roads with central dividers (mostly located in urban areas) and higher seat belt use when drivers travel on four lane roadways (usually located in rural highways), relative to drivers on two lane roads with no central dividers (see Gras *et al.*, 2007 for a similar result). Finally, the parameter estimate on the indicator variable for whether the crash occurred after 2005, the year when PDS was introduced in Denmark, suggests that the program has been successful in increasing seat belt use.

---

[16] Winston *et al.* (2006) reported similar evidence for ownership of cars with airbags.

4.3. Estimated Error-Correlation Structure

The error correlation parameter estimates ρ_{self} and ρ_{IS} are both statistically significant (see Table 2). The interpretation of the positive ρ_{self} parameter is that, in the net, the offsetting behavior hypothesis dominates over the selective recruitment hypothesis. A similar result was also reported in another study by de Lapparent (2008). That is, it appears that belted drivers offset the safety benefits that accrue from using a seat belt by driving more aggressively. This implies that if seat belt endogeneity is ignored, it would artificially deflate the "true" engineering restraint safety benefit of seat belt use. In our estimations, we noted this effect by comparing the coefficients on the seat belt use variable of the MVP model and the IP model (both of these models have the same overall error scale for injury severity, and so the coefficient estimates from these two models can be directly compared). The seat belt coefficient estimate from the MVP model was -1.124 (*t*-statistic of -7.91), while that from the IBO model was -0.720 (*t*-statistic of -19.18), reflecting the statistically significant underestimation of seat belt use effectiveness as a restraint device if seat belt use endogeneity is ignored.[17]

The estimate for ρ_{IS} indicates the presence of common crash-specific unobserved factors that simultaneously influence the injury severity of both drivers involved in the crash. Intuitively, one may expect that drivers involved in the same crash would share common crash-specific unobserved factors that affect both drivers' injury severity in the same direction. However, this may not be the case for a variety of reasons. For instance, if the crash is primarily instigated by one of the drivers, as seems to be the case in our sample from the police subjective reports, that driver may be more aware of the impending crash and can take evasive actions to minimize injury severity to her/him, while the other driver may not be aware of the impending danger and may experience severe injuries. This may be one explanation for the highly significant negative coefficient on ρ_{IS} in Table 2.

---

[17] The RCMVP model coefficient on seat belt use (the first coefficient in the injury severity component of Table 2) is not directly comparable to the MVP or IBO model seat belt coefficients due to the presence of random coefficients in the RCMVP model that lead to the normalization of the model coefficients with respect to a smaller overall scale.

4.4. Variable Effects and Their Implications

The parameter estimates in Table 2 do not directly provide a sense of the magnitude of impact of factors considered in our analysis. These effects vary across crashes based on the crash context. Our estimated model can be used to predict injury severity of each of the two drivers for any crash context by first predicting the multivariate probability of the combination of each state of the seat belt use variable (use or non-use) and each possible level of injury severity level (no injury, minor injury, and severe injury) for each of the two drivers for each crash in the estimation sample (using Equation (5) in Section 3.2). Next, these multivariate probabilities may be appropriately added up to obtain the marginal bivariate crash-level probabilities for each combination of injury severity for the two drivers involved in the crash (for ease in discussion, we will refer to these probabilities simply as bivariate injury severity (BIS) probabilities in the rest of this section).

The reader will note that a crash context is essentially represented by a specific combination of exogenous variables in the model system. Of course, since there would be a huge number of possible crash context combinations based on the determinant variables in Table 2, it is impossible to provide the effects of variables for each combination in this paper. But, to obtain a sense of the magnitude of variable effects from the model, we first develop a synthetic profile of a crash context with the following attributes (that correspond to the combination of the base categories for all variables in Table 2): (a) both drivers do not use seat belts, are sober, are single women in the age group of 31-60 with no crime history and driving minivans weighing less than 2000 kgs, and (b) the crash occurs on an undivided roadway with a speed limit of 50 kmph or less, and (c) the crash does not involve a frontal collision of any kind and occurs before 2005. For this synthetic crash context, we compute the BIS probabilities for each of 200 draws from the sampling distributions of the parameters, and average these probabilities. Then, we change each determinant variable from the base condition to an altered state (such as seat belt not used to seat belt used for one of the drivers) without changing any other variable to obtain the new averaged (across the same 200 draws) BIS probabilities for the two drivers in this altered state (corresponding to one driver wearing a seat belt and the other not). The percentage difference in the averaged BIS probabilities between the altered state and the base state provides a "pseudo-elasticity" effect (which we will simply label as the "elasticity" effect) for the altered state relative to the base state.

The above procedure can be undertaken to obtain the elasticity effects for a whole host of altered states. In Table 3, we present the elasticity effects of only selected altered states and for the injury severity combination category that represents serious injury for both the drivers in the crash (label this as the SIBD injury level for "serious injury for both drivers" injury level). Further, to keep the discussion focused and the presentation uncluttered, only the elasticity effects for the IP and RCMVP models are presented. The numbers in the table may be interpreted as the percentage change in the probability of the SIBD injury level due to a change in the crash scenario as identified in the first column (relative to the synthetic crash scenario). For example, the first entry in the table indicates that, according to the IP model, the probability of the SIBD injury level is 55.16% lower when one driver wears a seat belt and the other does not, compared to when both drivers do not wear seat belts (other characteristics being the same as those in the synthetic profile). Other entries may be similarly interpreted.

The directions of the elasticity effects (for each altered state in Table 3) from both the IP and RCMVP models are consistent with the discussions in the previous section. However, there are clear differences in the elasticity estimates from the IP and RCMVP models. For instance, according to the IP model, the event of sustaining SIBD is 72.0% less likely when both drivers wear seat belts relative to when both drivers are unbelted, while, according to the RCMVP model, the same event is 96% less likely compared to when both drivers are unbelted. This underestimation by the IP model of the restraint effectiveness of seat belt use is, of course, because the model comingles restraint effectiveness with the increased aggressiveness caused by offsetting behavior related to seat belt use. For the same reason, the IP model underestimates the elasticity effects for the crash scenario where one driver is belted while the other is not. The IP model similarly also underestimates the effects of driving under the influence of alcohol and other altered states. This underestimation is particularly high for the altered states associated with roadway speed limits and crash characteristics. Overall, these elasticity effects demonstrate the pitfalls of assessing variable effects, and the potential danger in evaluating the effectiveness of crash injury severity reduction measures, when ignoring the crash-level nature of injury severities, the endogeneity of seat belt use, and unobserved heterogeneity in the effects of variables.

With regard to the relative effects of variables, crashes that (a) involve one or both drivers with a crime history (relative to crashes involving drivers without a crime history), (b)

involve one or both drivers being young (30 years or younger) (as opposed to other drivers), (c) involve frontal collisions (relative to other kinds of collisions), and (d) occur at high speed roadway locations (as opposed to crashes on low speed roadway locations), most increase the risk of both drivers sustaining severe injuries. On the other hand, crashes that (a) involve one or both drivers being belted (as opposed to both drivers being unbelted), (b) involve one or both drivers being married (relative to both drivers being unmarried), and (c) occur on four lane roads (as opposed to other roads), are the ones that are associated with the lowest risk of both drivers being severely injured.

5. CONCLUSIONS

The current study contributes to the existing injury severity modeling literature by developing a multivariate probit model of seat belt use and injury severity of both drivers involved in two-vehicle crashes. This resulting model system accounts for three key econometric issues: (1) correlation in the injury severity levels of the two drivers, (2) endogeneity of seat belt use in predicting injury severity levels, and (3) parameter heterogeneity in the effects of variables on both injury severity and seat belt use. To our knowledge, this is the first study in the injury severity literature that addresses all of these three issues at once. The proposed model is applied to analyze the injury severity of drivers involved in two-vehicle road crashes in Denmark. In addition to making a methodological contribution to the injury severity literature, the current paper also makes a substantive empirical contribution by controlling for the effects of cross-driver and cross-vehicle characteristics in jointly analyzing the injury severity of both drivers involved in the crash.

There are several important findings from the research. First, the results show the potentially misleading estimates of variable effects that can be obtained if one does not recognize important econometric considerations in the crash-level modeling of injury severity. Such misleading estimates can lead to misinformed policies as well as to inaccurate forecasts of injury severities in a crash. Second, there is a strong suggestion that people offset the restraint benefits of seat belt use by driving more aggressively. This implies a need to design information campaigns on seat belt use that not only underscore the substantial restraint benefits of seat belt use, but also highlight the danger to other road users of driving aggressively. A better understanding of the neuropsychological and cognitive mechanisms underlying aggressive

driving in general, and how individuals perceive their responsibility toward others vis-à-vis their own safety, should be helpful to design countermeasures to reduce aggressive driving (see also Miles and Johnson, 2003, Dahlen and White, 2006, and Paleti *et al*., 2010). Third, drivers hit by or colliding with vehicles driven by male drivers are more likely to sustain serious injuries compared to drivers hit by or colliding with vehicles driven by female drivers. This combined with the finding that male drivers, on average, are less likely to sustain serious injuries in the event of a crash point to the need to stress in information campaigns that men, if they drive aggressively, are not putting themselves as much at risk as they are putting others on the road at risk, especially women drivers. Fourth, crashes involving young and unmarried drivers, and drivers with a crime history, tend to lead to the most severe injuries for all drivers involved. These results need additional exploration to examine whether there is a need for specific driving education campaigns directed toward these driver groups. Fifth, drivers in lighter vehicles are more vulnerable to injuries than those in heavier vehicles. With regard to cross-vehicle effects, drivers hit by or colliding with lighter vehicles are less likely to sustain serious injuries than drivers hit by or colliding with heavier vehicles. These results suggest additional attention in driver education courses on driving in mixed traffic situations, and the continued need for vehicle manufacturers to invest in safety technology. Finally, lower speed limits on roadways and dividers (median barriers) to physically separate traffic flow in opposite directions appear to be good countermeasures to reduce injury severity conditional on a crash.

REFERENCES

Abay, K.A., 2012. Examining pedestrian-injury severity in a comprehensive and comparative model specification. Memo, Department of Economics, University of Copenhagen.

Adams, J.G.U., 1994. Seat belt legislation: the evidence revisited. *Safety Science* 18(2), 135-152.

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.

Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B* 35(7), 677-693.

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

Bhat, C.R., 2011. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B* 45(7), 923-939.

Bhat, C.R., Eluru, N., 2009. A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B* 43(7), 749-765.

Bhat, C.R., Srinivasan, S., 2005. A multidimensional mixed ordered-response model for analyzing weekend activity participation. *Transportation Research Part B* 39(3), 255-278.

Bhat, C.R., Varin, C., Ferdous, N., 2010. A comparison of the maximum simulated likelihood and composite marginal likelihood estimation approaches in the context of the multivariate ordered response model. In: Greene, W., Hill, R.C. (Eds.), *Advances in Econometrics: Maximum Simulated Likelihood Methods and Applications*, 65-106, Emerald Group Publishing Limited, Bingley, UK.

Castro, M., Paleti, R., Bhat, C.R., 2012. A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. *Transportation Research Part B* 46(1), 253-272.

Chen, F., Chen, S., 2011. Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways. *Accident Analysis and Prevention* 43(5), 1677-1688.

Christoforou, Z., Cohen, S., Karlaftis, G.M., 2010. Vehicle occupant injury severity on highways: An empirical investigation. *Accident Analysis and Prevention* 42(6), 1606-1620.

Cohen, A., Einav, L., 2003. The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *The Review of Economics and Statistics* 85(4), 828-843.

Cummings, P., Rivara, F.P., 2004. Car occupant death according to the restraint use of other occupants: A matched cohort study. *The Journal of the American Medical Association* 291(3), 343-349.

Cummings, P., Wells, J.D., Rivara, F.P., 2003. Estimating seat belt effectiveness using matched-pair cohort methods. *Accident Analysis and Prevention* 35(1), 143-149.

Dahlen, E.R., White, R.P., 2006. The big five factors, sensation seeking, and driving anger in the prediction of unsafe driving. *Personality and Individual Differences* 41(5), 903-915.

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

Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. *Accident Analysis and Prevention* 39(5), 1037-1049.

Eluru, N., Bhat, C.R., Hensher, D.A., 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 injury severity of multiple occupants of vehicles. *Transportation Research Record* 2165, 1-11.

EuroSafe, 2012. Child safety country profile, Denmark. European Child Safety Alliance. Available at: http://www.childsafetyeurope.org/reportcards/info/denmark-country-profile.pdf

Evans, L., 1986. The effectiveness of safety belts in preventing fatalities. *Accident Analysis and Prevention* 18(3), 229-241.

Evans, L., 1996. Safety-belt effectiveness: The influence of crash severity and selective recruitment. *Accident Analysis and Prevention* 28(4), 423-433.

Geweke, J., 1991. Efficient simulation from the multivariate normal and student-*t* distributions subject to linear constraints and the evaluation of constraint probabilities. In: Keramidas, E.M. (Ed.), *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, 571-578. Interface Foundation, Fairfax Station, VA.

Gras, M.E., Cunill, M., Sullman, M.J.M., Planes, M., Font-Mayolas, S., 2007. Predictors of seat belt use amongst Spanish drivers. *Transportation Research Part F* 10(3), 263-269.

Hajivassiliou, V.A., McFadden, D.L., 1998. The method of simulated scores for the estimation of LDV models. *Econometrica* 66(4), 863-896.

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

Heckman, J. J., Vytlacil, E. J., 2005. Structural equations, treatment effects and econometric policy evaluation. *Econometrica* 73(3), 669-738.

Huang, H., Chin, H.C., Haque, M.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.

Hutchinson, T.P., 1986. Statistical modeling of injury severity, with special reference to driver and front seat passenger in single-vehicle crashes. *Accident Analysis and Prevention* 18(2), 157-167.

Jung, S., Qin, X., Noyce, D., 2010. Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident Analysis and Prevention* 42(1), 213-224.

Keane, M.P., 1994. A computationally practical simulation estimator for panel data. *Econometrica* 62(1), 95-116.

Kim, J.K., Ulfarsson, G.F., Shankar, V.N., Mannering, F.L., 2010. A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis and Prevention* 42(6), 1751-1758.

Levitt, S.D., Porter, J., 2001. Sample selection in the estimation of air bag and seat belt effectiveness. *The Review of Economics and Statistics* 83(4), 603-615.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A* 44(5), 291-305.

Maddala, G.S., 1983. *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

Malyshkina, N., Mannering, F., 2009. Markov switching multinomial logit model: An application to accident-injury severities. *Accident Analysis and Prevention* 41(4), 829-838.

McCarthy, P., Talley, W.K., 1999. Evidence on risk compensation and safety behaviour. *Economics Letters* 62(1), 91-96.

McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. *Applied Econometrics* 15(5), 447-470.

Miles, D.E., Johnson, G.L., 2003. Aggressive driving behaviors: Are there psychological and attitudinal predictors? *Transportation Research Part F* 6(2), 147-161.

Milton, J.C., Shankar, V.N., Mannering, F.L., 2008. Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis and Prevention* 40(1), 260-266.

Moore, D.N., Schneider IV, W.H., Savolainen, P.T., Farzaneh, M., 2011. Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. *Accident Analysis and Prevention* 43(3), 621-630.

Murphy, S.L., Xu, J.Q., Kochanek, K.D., 2012. Deaths: Preliminary Data for 2010. National Vital Statistics Reports 60(4), National Center for Health Statistics, Hyattsville, MD. Available at: http://www.cdc.gov/nchs/data/nvsr/nvsr60/nvsr60_04.pdf

National Highway Traffic Safety Administration (NHTSA), 2009. Lives saved calculations for seat belts and frontal air bags. DOT HS 811206, U.S. Department of Transportation, Washington, D.C. Available at: http://www-nrd.nhtsa.dot.gov/Pubs/811206.pdf

National Highway Traffic Safety Administration (NHTSA), 2012. 2010 Motor vehicle crashes: overview. DOT HS 811552, National Center for Statistics and Analysis, U.S. Department of Transportation. Available at: http://www-nrd.nhtsa.dot.gov/pubs/811552.pdf

Paleti, R., Eluru, N., Bhat, C.R., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. *Accident Analysis and Prevention* 42(6), 1839-1854.

Quddus, M.A., Wang, C., Ison, S.G., 2010. Road traffic congestion and crash severity: Econometric analysis using ordered response models. *Journal of Transportation Engineering* 136(5), 424-435.

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

Reinfurt, D., Williams, A., Wells, J., Rodgman, E., 1996. Characteristics of drivers not using seat belts in a high belt use state. *Journal of Safety Research* 27(4), 209-215.

Rifaat, S.M., Tay, R., de Barros A., 2011. Effect of street pattern on the severity of crashes involving vulnerable road users. *Accident Analysis and Prevention* 43(1), 276-283.

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

Shinar, D., Compton, R., 2004. Aggressive driving: an observational study of driver, vehicle, and situational variables. *Accident Analysis and Prevention* 36(3), 429-437.

Ulfarsson, G.F., Mannering, F.L., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis and Prevention* 36(2), 135-147.

Winston, C., Maheshri, V., Mannering, F., 2006. An exploration of the offset hypothesis using disaggregates data: The case of airbags and antilock brakes. *Journal of Risk and Uncertainty* 32(2), 83-99.

World Health Organization, 2009. Global status report on road safety: Time for action. Geneva, Switzerland.

Xie, Y., Zhang, Y., Liang, F., 2009. Crash injury severity analysis using Bayesian ordered probit models. *Journal of Transportation Engineering* 135(1), 18-25.

Yamamoto, T., Shankar, V.N., 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.

Yan, X., Ma, M., Huang, H., Abdel-Aty, M., Wu, C., 2011. Motor vehicle–bicycle crashes in Beijing: Irregular maneuvers, crash patterns, and injury severity. *Accident Analysis and Prevention* 43(5), 1751-1758.

Ye, F., Lord, D., 2011. Investigation of effects of underreporting crash data on three commonly used traffic crash severity models. *Transportation Research Record* 2241, 51-58.

Zhu, X., Srinivasan, S., 2011. A comprehensive analysis of factors influencing the injury-severity of large-truck crashes. *Accident Analysis and Prevention* 43(1), 49-57.

Table 1a: Descriptive Statistics of Seat Belt Use of a Driver and the Injury Severity Level Sustained by that Driver

Injury severity	Seat belt[a]		Overall injury severity level sample shares
	Used	Not used	
No injury	50.6	23.5	47.7
Minor injury	32.0	42.2	33.1
Serious injury	17.4	34.3	19.2
Overall seat belt use and non-seat belt use shares	89.4	10.6	100.0

[a] The cross-tabulation cell values are percentages taken across rows for each column (so that the sum of the figures in each column, not including the last row, is 100%).

Table 1b: Descriptive Statistics of Seat Belt Use of a Driver and the Injury Severity Level Sustained by Other Driver

Injury severity	Seat belt[b]		Overall injury severity level sample shares
	Used	Not used	
No injury	46.4	57.7	47.7
Minor injury	34.1	25.0	33.1
Serious injury	19.5	17.3	19.2
Overall seat belt use and non-seat belt use shares	89.4	10.6	100.0

[b] The cross-tabulation cell values are percentages taken across rows for each column (so that the sum of the figures in each column, not including the last row, is 100%).

Table 2: Estimated Coefficients for the Preferred Model (RCMVP)

Variables considered	Injury Severity		Seat Belt Use	
	Coefficient	*t*-stat	Coefficient	*t*-stat
Driver characteristics				
Seat belt use	-1.379	-8.84		
Under the influence of alcohol	0.102	1.73		
Male	-0.516	-18.21	-0.272	-6.57
*S.d**	0.397	8.97		
Age *(base $31 \leq age \leq 60$ years)*				
Young (age < 25 years)			-0.320	-7.61
Adult (age 25-30 years)			-0.209	-4.20
Old (age > 60 years)	0.142	3.71	0.114	1.85
S.d	0.249	1.54		
Married			0.214	5.99
Crime history indicator			-0.485	-12.50
Other driver's characteristics				
Male	0.202	7.57		
Seat belt use	0.318	4.63		
Own vehicle's characteristics (base: van $\leq 2000kg$)				
Private car	0.252	4.30		
Van (weight > 2000 kg)	-0.252	-3.38		
Other driver's vehicle characteristics				
Private car	-0.305	-5.51		
Van (weight > 2000 kg)	0.227	3.33		
Roadway characteristics (base: speed limit ≤ 50 kmph)				
Speed limit = 60 *kmph*	0.127	3.25		
Speed limit = 70 *kmph*	0.177	4.19		
Speed limit = 80 *kmph*	0.254	9.31		
Speed limit > 80 *kmph*	0.256	4.60		
Road type (base: undivided two lane road)				
Two lane with central divider	-0.104	-2.72	-0.233	-5.44
Four lane roads			0.144	2.55
Crash characteristics (base: no frontal impact)				
Struck by other vehicle moving frontally	0.264	10.06		
Strike other vehicle frontally	0.089	2.93		
Both collided frontally	0.351	9.31		
S.d	0.473	8.93		
Accident happened after 2005			0.277	7.15
Threshold parameters				
ψ_1	-0.924	-5.97		
ψ_2	0.146	0.91		
μ_1			-1.510	-30.40
Correlation parameters				
ρ_{self}	0.346	3.94		
ρ_{IS}	-0.323	-11.96		

* S.d stands for estimated standard deviation of the parameters.

Table 3: Elasticity Effects for the "Serious Injury for Both Drivers" Injury Level from the IP and RCMVP models

Crash Scenario	IP Model Estimate	RCMVP Model Estimate
Driver characteristics		
Seat belt use (*base: both unbelted*)		
One driver belted and the other not	-55.16	-84.51
Both drivers belted	-72.03	-96.00
Driving under the influence (DUI) of alcohol (*base: both drivers sober*)		
One driver DUI and the other sober	16.83	24.49
Both drivers DUI	37.02	56.82
Gender of both drivers (*base: both drivers women*)		
One of the driver is a man and the other woman	8.20	-8.61
Both drivers are men	34.71	-4.29
Age of drivers (*base: both drivers are middle age: 31 ≤ age ≤ 60 years*)		
One driver is age < 25, other middle age	79.49	122.06
One driver is age 25-30, other middle age	49.10	73.62
One driver is age > 60, other middle	-7.83	-0.58
Both drivers age < 25 years	223.96	402.63
Both drivers age 25-30 years	123.86	207.68
Both drivers age > 60 years	-13.45	2.45
One driver is age < 25, other driver age is 25-30 years	167.63	288.93
One driver is age < 25, other driver age > 60 years	65.37	120.58
One driver is age is 25-30 years, other driver age > 60 years	37.66	73.14
Marital status *(base: both drivers are not married)*		
One driver married, other single	-32.20	-44.00
Both are married	-53.77	-68.28
Crime history *(base: both drivers do not have crime history)*		
One driver has crime history, other not	136.91	221.10
Both drivers have crime history	463.24	952.82
Vehicle's characteristics (base: both vehicles are vans < 2000 kg)		
One of the vehicles is private car and other vehicle is Van < 2000 kg	-8.91	-15.44
One of the vehicles is Van > 2000 kg and other vehicle is Van < 2000 kg	-6.90	-8.16
Both vehicles are private car	-3.73	-17.40
Both vehicles are Vans > 2000 kg	2.01	-4.04
One vehicle is private car and other vehicle is Van > 2000 kg	-3.82	-29.57
Roadway characteristics (base: speed limit = 50 kmh)		
Speed limit = 60 *kmh*	35.32	74.35
Speed limit = 70 *kmh*	52.10	111.51
Speed limit = 80 *kmh*	78.56	189.82
Speed limit > 80 *kmh*	78.81	194.50
Road type (base: undivided two lane road)		
Two lane with central divider	87.89	111.05
Four lane roads	-39.35	-53.28
Crash characteristics (base: no frontal impact)		
One driver strikes other vehicle frontally	28.55	49.26
One driver struck by other vehicle frontally	80.06	193.94
Both drivers collided frontally	120.70	310.54
Accident happened after 2005	-69.52	-77.87