ANALYZING THE CONTINUUM OF FATAL CRASHES: A GENERALIZED ORDERED APPROACH

Shamsunnahar Yasmin
Department of Civil Engineering & Applied Mechanics
McGill University
Suite 483, 817 Sherbrooke St. W.
Montréal, Québec, H3A 2K6, Canada
Ph: 514 398 6823, Fax: 514 398 7361
Email: shamsunnahar.yasmin@mail.mcgill.ca

Naveen Eluru*
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Suite 301D, 4000 Central Florida Blvd.,
Orlando, Florida, 32816, USA
Phone: 407 823 2841 Fax: 407 823 3315
Email: naveen.eluru@ucf.edu

Abdul R. Pinjari
Department of Civil and Environmental Engineering
University of South Florida
Ph: 813-974- 9671, Fax: 813-974-2957
Email: apinjari@eng.usf.edu

*Corresponding author

ABSTRACT

In the United States, safety researchers have focused on examining fatal crashes (involving at least one fatally injured vehicle occupant) by using Fatality Analysis Reporting System (FARS) dataset. FARS database compiles crashes if at least one person involved in the crash dies within thirty consecutive days from the time of crash along with the exact timeline of the fatal occurrence. Previous studies using FARS dataset offer many useful insights on what factors affect crash related fatality, particularly in the context of fatal vs. non-fatal injury categorization. However, there is one aspect of fatal crashes that has received scarce attention in the traditional safety analysis. The studies that dichotomize crashes into fatal versus non-fatal groups assume that all fatal crashes in the FARS dataset are similar. Keeping all else same, a fatal crash that results in an immediate fatality is clearly much more severe than another crash that leads to fatality after several days. Our study contributes to continuing research on fatal crashes. Specifically, rather than homogenizing all fatal crashes as the same, our study analyzes the fatal injury from a new perspective by examining fatality as a continuous spectrum based on survival time ranging from dying within thirty days of crash to dying instantly (as reported in the FARS data). The fatality continuum is represented as a discrete ordered dependent variable and analyzed using the mixed generalized ordered logit (MGOL) model. By doing so, we expect to provide a more accurate estimation of critical crash attributes that contribute to death. In modeling the discretized fatality timeline, the Emergency Medical Service (EMS) response time variable is an important determinant. However, it is possible that the EMS response time and fatality timeline are influenced by the same set of observed and unobserved factors, generating endogeneity in the outcome variable of interest. Hence, we propose to estimate a two equation model that comprises of a regression equation for EMS response time and MGOL for fatality continuum with residuals from the EMS model to correct for endogeneity bias on the effect of exogenous factors on the timeline of death. Such research attempts are useful in determining what factors affect the time between crash occurrence and time of death so that safety measures can be implemented to prolong survival. The model estimates are augmented by conducting elasticity analysis to highlight the important factors affecting time-to-death process.

Keywords: Generalized Ordered Logit, Endogeneity, Two-stage residual inclusion, FARS, Elasticities

1. INTRODUCTION

Road traffic crashes and their consequences such as injuries and fatalities are acknowledged to be a serious global health concern. In the United States (US), motor vehicle crashes are responsible for more than 90 deaths per day (NHTSA, 2012). Moreover, these crashes cost the society $230.6 billion annually (GHSA, 2009). In an attempt to reduce the consequence of road traffic crashes and to devise countermeasures, transportation safety researchers study the influence of various exogenous variables on vehicle occupant injury severity. In identifying the critical factors contributing to crash injury severity, safety researchers have focused on either examining fatal crashes (involving at least one fatally injured vehicle occupant) or traffic crashes that compile injury severity spectrum at an individual level (such as no injury, possible injury, non-incapacitating injury, incapacitating injury and fatality). In the US, the former category of studies predominantly use the Fatality Analysis Reporting System (FARS) database (see Zador et al., 2000; Gates et al., 2013) while the latter group of studies typically employ the General Estimates System (GES) database (see Kockelman and Kweon, 2002; Eluru and Bhat, 2007; Yasmin and Eluru, 2013). FARS database compiles crashes if at least one person involved in the crash dies within thirty consecutive days from the time of crash. Further, FARS database reports the exact timeline of the fatal occurrence within thirty days from the time to crash.

A number of research efforts have examined the impact of exogenous characteristics (such as driver characteristics, vehicle characteristics, roadway design and operational attributes, environmental factors and crash characteristics) associated with fatal crashes employing crash data with at least one fatality. These studies employed two broad dependent variable categorizations – (1) fatal/non-fatal or (2) fatal/serious injury. The binary categorization was analyzed employing descriptive analysis or logistic regression methods for identifying the critical factors affecting fatal crashes (for example see Zhang et al., 2013; Al-Ghamdi, 2002; Huang et al., 2008; Travis et al., 2012). Several studies have also investigated the factors affecting the involvement in a fatal crash as a function of individual characteristics. The important individual behavioral determinants of fatal crashes include excessive speed, violation of traffic rules and lack of seat belt use (Siskind et al., 2011; Valent et Al., 2002; Sivak et al., 2010; Viano et al., 2010). Other driver attributes such as aggressive driving behavior, unlicensed driving and distraction during driving are identified to be the most significant contributors of fatal crashes for young drivers (Lambert-Bélanger et al., 2012; Hanna et al., 2012; Chen et al., 2000). Studies have also examined the effect of race/ethnicity in fatal crashes (Braver, 2003; Romano et al., 2006; Campos-Outcalt et al., 2003; Harper et al., 2000). On the other hand, most critical factors identified from earlier research for older drivers in fatal crashes are frailty and reduced driving ability (Baker et al., 2003; Lyman et al., 2002, Thompson et al., 2013). Gates et al. (2013) investigate the influence of stimulants (such as amphetamine, methamphetamine and cocaine) on unsafe driving actions in fatal crashes. Stübig et al. (2012) investigate the effect of alcohol consumption on preclinical mortality of traffic crash victims (see also Fabbri et al., 2002).

Many of the earlier studies also focused on the vehicular characteristics of fatal crashes (Fredette et al., 2008) and demonstrated that the relative risk of fatality is much higher for the driver of lighter vehicle (sedan, compact car) compared to those in the heavier vehicle (SUV, Vans, Pickups). Among the environmental factors, it was found that collision during night time (Arditi et al., 2007) has the most significant negative impact on fatality risk in a crash. In terms of crash characteristics, head-on crash and crashes on high speed limit road locations increased the probability of fatalities in a crash (Fredette et al., 2008; Bédard et al., 2002).

These studies offer many useful insights on what factors affect crash related fatality, particularly in the context of fatal vs. non-fatal injury categorization. However, there is one aspect of fatal crashes that has received scarce attention in the traditional safety analysis. The studies that dichotomize crashes into fatal versus non-fatal groups assume that all fatal crashes in the FARS dataset are similar. Keeping all else same, a fatal crash that results in an immediate fatality is clearly much more severe than another crash that leads to fatality after several days. In fact, there is evidence from epidemiological studies (Tohira et al., 2012) that the risk factors associated with early trauma deaths of crash victims are different from the risk factors associated with late trauma deaths. For instance, Tohira et al. (2012) reported that older drivers (aged 65 years or older) and/or crash victims with a depressed level of consciousness were at increased risk of late trauma death. Research attempts to discern such differences are useful in determining what factors affect the time between crash occurrence and time of death so that countermeasures can be implemented to improve safety situation and to reduce road crash related fatalities. Early EMS (Emergency Medical Service) response is also argued to potentially improve survival probability of motor vehicle crash victims (Clark and Cushing, 2002; Clark et al., 2013). In fact, Meng and Weng (2013) reported 4.08% decrease in the risk of death from one minute decrease in EMS response time, while Sánchez-Mangas et al. (2010) reported that a ten minutes EMS response time reduction could decrease the probability of death by one third. Given the import of this variable, it is also important to explore the effect of EMS response time in examining crash fatalities.

The objective of our study is to identify the associated risk factors of driver fatalities while recognizing that fatality is not a single state but rather is made up of a timeline between dying instantly to dying within thirty days of crash (as reported in the FARS data). The detailed information available in FARS provides us a continuous timeline of the fatal occurrences from the time of crash to death. This allows for an analysis of the survival time of victims before their death. To be sure, earlier research efforts also focused on examining the factors influencing the time period between road crash and death (Golias and Tzivelou, 1992; Marson and Thomson, 2001; Feero et al., 1995; Al-Ghamdi, 1999; Gonzalez et al., 2006; Gonzalez et al., 2009; Brown et al, 2000). These studies demonstrated that nature of injury, EMS response time and pre-hospital trauma care were the main factors affecting the time till death and concluded that timely EMS response with proper pre-hospital trauma care may improve the survival outcome. For analysis of the time to death data, these studies employed univariate statistical analysis (such as descriptive analysis or Fisher's exact test, Student t test). Most recently, Ju and Sohn (2014) analyzed the factors that are potentially associated with variation in the expected survival time by using Weibull regression approach and identified that survival probabilities and expected survival times are related to changes in delta V, alcohol involvement, and restraint systems. But, none of these studies investigate the timeline of death at the disaggregate level as a function of exogenous characteristics for a crash victim. Our study builds on existing fatality analysis research by developing a disaggregate level model for the discrete representation of the continuous fatality timeline using the FARS dataset. The fatality timeline information obtained through FARS is categorized as an ordered variable ranging from death in thirty days to instantaneous death in seven categories as follows: died between 6th-30 days of crash, died between 2nd-5 days of crash, died between 7th-24 hours of crash, died between 1st-6 hours of crash, died between 31st-60 minutes of crash, died between 1st-30 minutes of crash and died instantly.

Due to the inherent ordered nature of the fatality variable created, an ordered discrete outcome modeling approach is an appropriate framework for examining the influence of exogenous factors on the timeline of death. However, the traditional ordered outcome models restrict the impact of exogenous variables on the outcome process to be same across all alternatives (Eluru et al, 2008). The recent revival in the ordered regime has addressed this limitation by allowing the analyst to estimate individual level thresholds as function of exogenous variables as opposed to retaining the same thresholds across the population (as is the case in the standard ordered logit (OL)). The approach is referred to as the Generalized Ordered Logit (GOL) (or partial proportional odds logit) (Yasmin and Eluru, 2013; Eluru, 2013; Mooradian et al, 2013) model. At the same time, the conventional police/hospital reported crash databases may not include individual specific behavioural or physiological characteristics and vehicle safety equipment specifications for crashes. Due to the possibility of such critical missing information, it is important to incorporate the effect of unobserved attributes within the modeling approach (see for example Srinivasan, 2002; Eluru et al., 2008; Kim et al., 2013). In non-linear models, neglecting the effect of such unobserved heterogeneity can result in inconsistent estimates (Chamberlain, 1980; Bhat, 2001). Hence, we employ the mixed generalized ordered logit (MGOL) framework to examine driver fatalities characterized as an ordinal discrete variable of an underlying severity continuum of fatal injuries.

In modeling the discretized fatality timeline, the EMS response time variable is an important determinant. However, it is possible that the EMS response time and fatality timeline are influenced by the same set of observed and unobserved factors, generating endogeneity in the outcome model of interest. In fact, it was identified that EMS response time are affected by several external environmental and regional factors (Brodsky, 1992; Meng and Weng, 2013). Such correlations impose challenges in using the EMS response variable as an explanatory variable in examining fatality outcome of crashes. For example, consider two potential crash scenarios. In scenario 1 a relatively major crash occurs and in scenario 2 a minor crash occurs. When the information of a crash is provided the urgency with which the EMS teams are deployed for the first scenario is likely to be higher than the urgency for the second scenario. So, we potentially have a case where EMS time for arrival is lower for scenario 1 but potentially the consequences of the crash for scenario 1 are much severe i.e. survival time is much smaller. So, in a traditional modeling approach one would conclude that lower EMS arrival times are associated with smaller survival times. This is a classic case of data endogeneity affecting the modeling results. Hence, it is necessary to account for this endogeneity in the modeling process. In our study, we propose to apply an econometric approach to accommodate for this. Specifically, we propose to estimate a driver-level fatal injury severity model while also accounting for endogeneity bias of EMS arrival time using ordered outcome modeling framework with endogeneity treatment. In doing so, the correction for endogeneity bias is pinned down in the ordered outcome models by employing a two-stage residual inclusion (2SRI) approach.

In summary, the current research makes a three-fold contribution to the literature on vehicle occupant injury severity analysis. First, our study is the first attempt to analyze the fatal injury from a new perspective and examine fatality as a continuous spectrum based on survival time ranging from dying within thirty days of crash to dying instantly. Second, we propose and estimate a two equation model that comprises of regression for EMS response time and MGOL with residuals from the EMS model to correct for endogeneity bias on the effect of exogenous factors on the timeline of death. Finally, we compute elasticity measures to identify important factors affecting survival time after motor vehicle crash.

The rest of the paper is organized as follows. Section 2 provides details of the econometric model framework used in the analysis. In Section 3, the data source and sample formation procedures are described. The model estimation results and elasticity effects are presented in

Section 4 and 5, respectively. Section 6 concludes the paper and presents directions for future research.

2. MODEL FRAMEWORK

The focus of our study is to examine driver-level fatal injury at a disaggregate level while also accounting for endogeneity bias of EMS arrival time by using a MGOL model framework with endogeneity treatment. In doing so, the correction for endogeneity bias is pinned down in MGOL model by employing a 2SRI approach[1] (as opposed to the two-stage predictor substitution approach). The framework used for MGOL model with endogenous treatment consists of a two-stage procedure. In the first stage, the residuals are computed from the linear regression estimates of the endogenous variable (EMS arrival time). In the second stage, MGOL model is estimated by including the first-stage residuals as additional regressor along with the endogenous variable in examining the outcome of interest. In this section, econometric formulation for MGOL model with the 2SRI treatment is presented.

2.1 First Stage

Let i $(i = 1,2, \dots \dots, I)$ and j $(j = 1,2, \dots \dots, J)$ be the indices to represent driver and the time between crash occurrence and time of death for each fatally injured driver i. In this paper, index j takes the values of: died between 6th to 30 days of crash $(j = 1)$, died between 2nd to 5 days of crash $(j = 2)$, died between 7th to 24 hours of crash $(j = 3)$, died between 2nd to 6 hours of crash $(j = 4)$, died between 31st to 60 minutes of crash $(j = 5)$, died between 1st to 30 minutes of crash $(j = 6)$ and died instantly $(j = 7)$ for all fatally injured drivers. Let us also assume that y_i represents the discrete levels of time to death, $\boldsymbol{x}_{\mathrm{i}}$ is a column vector of observable exogenous variables, $\boldsymbol{u}_i$ is a set of e $(e = 1,2, \dots \dots, E)$ endogenous variables and $\boldsymbol{q}_i$ is a $1 \times E$ set of unobservable endogenous variables possibly correlated with both the outcome and the endogenous variables, generating endogeneity bias in the outcome model. In our analysis, we hypothesize that EMS arrival time may be correlated with the unobservable determinants of fatal injury severity of drivers, thus we have $e = 1$ in the current study context. Following Terza et al. (2008), we present the endogeneity of $\boldsymbol{u}_i$ by assuming an idiosyncratic influence of the same latent variables $\boldsymbol{q}_i$ on both the outcome and endogenous variables as a linear regression model as:

$$L_i = \rho \boldsymbol{w}_i + \boldsymbol{q}_i \tag{1}$$

where,

$\boldsymbol{w}_i = [\boldsymbol{x}_i \; \boldsymbol{v}_i]$ and $\boldsymbol{v}_i$ is a set of at least E instrumental variables

ρ is a corresponding row vector of parameter estimates

The residuals of endogenous variables can be computed as:

$$q_i^R = \boldsymbol{u}_i - Pr(\boldsymbol{u}_i | \boldsymbol{w}_i) \tag{2}$$

[1] The reader is referred to Terza et al. (2008) for a detailed discussion of why the two stage residual inclusion method provides consistent estimates in non-linear models, while the two stage predictor substitution method does not.

where, $Pr(\boldsymbol{u}_i|\boldsymbol{w}_i)$ is the predictor of $\boldsymbol{u}_i$.

2.2 Second Stage

In the proposed two-stage model, the modeling of discrete levels of fatal crashes is undertaken using MGOL specification. The MGOL accommodates unobserved heterogeneity in the effect of exogenous variable on injury severity levels in both the latent injury risk propensity function and the threshold functions (Srinivasan, 2002; Eluru et al., 2008). In the MGOL model, the discrete levels of time to death (y_i) are assumed to be a mapping (or partitioning) of an underlying continuous latent variable (y_i^*) as follows:

$$y_i^* = (\beta + \alpha_i)\boldsymbol{x}_i + \sigma\boldsymbol{u}_i + \lambda\boldsymbol{q}_i + \varepsilon_i\ ,\ \ y_i = j, if\ \tau_{i,j-1} < \ y_i^* < \tau_{i,j} \qquad (3)$$

where,

β, σ and λ are corresponding row vectors of associated parameters for $\boldsymbol{x}_i$, $\boldsymbol{u}_i$ and $\boldsymbol{q}_i$, respectively.

α_i is a row vector representing the unobserved factors specific to driver i and his/her trip environments

ε_i is a random disturbance term assumed to be standard logistic

$\tau_{i,j}$ represents the thresholds

Once the linear regression for the endogenous variable is estimated, we can insert the computed residuals of equation 2 as additional regressors in equation 3 for the outcome of interest. Thus, substituting the residuals for the unobservable latent factors, we can re-write equation 3 as:

$$y_i^* = (\beta + \alpha_i)\boldsymbol{x}_i + \sigma\boldsymbol{u}_i + \lambda q_i^R + \varepsilon_i\ ,\ \ y_i = j, if\ \tau_{i,j-1} < \ y_i^* < \tau_{i,j} \qquad (4)$$

In the above setting, the endogeneity of $\boldsymbol{u}_i$ will be absent if λ turns out to be zero. Moreover, in equation 4, $\tau_{i,j}$ $(\tau_{i,0} = -\infty\ , \tau_{i,J} = \infty)$ represents the upper threshold associated with driver *i* and time scale *j*, with the following ordering conditions: $\left(-\infty < \tau_{i,1} < \tau_{i,2} < \ \dots\dots\dots < \tau_{i,J-1} < +\infty\right)$. To maintain the ordering conditions and allow the thresholds to vary across drivers, Eluru et al. (2008) propose the following non-linear parameterization of the thresholds as a function of exogenous variables:

$$\tau_{i,j} = \tau_{i,j-1} + exp[\left(\delta_j + \gamma_{i,j}\right)\boldsymbol{z}_{i,j}] \qquad (5)$$

where, $\boldsymbol{z}_{ij}$ is a set of exogenous variable associated with j th threshold; δ_j is a time to death-specific row vector of parameters to be estimated (we need to restrict δ_1 to be a row vector of zero values for identification reason) and γ_{ij} is another row vector representing the unobserved factors specific to driver i and his/her trip environments. The traditional OL model assumes that the thresholds $\tau_{i,j}$ remain fixed across drivers ($\tau_{i,j} = \tau_j\ \ \forall\ \ i$); that is, it assumes that δ_j has all zero elements for all j values (except for the constant). Thus, the model will collapse to a simple OL model if α_i turns out to be zero in equation 4 and $\tau_{i,j}$ remain fixed across driver in equation 5. On the other hand, if α_i and $\gamma_{i,j}$ terms of equation 4 and 5 are found to be zero in model estimation, then the model will collapse to simple GOL model.

In equations 4 and 5, we assume that α_i and γ_{ij} are independent realizations from normal distribution for this study. Thus, conditional on α_i and γ_{ij}, the probability expression for individual i and alternative j in MGOL model with the 2SRI treatment take the following form:

$$\begin{aligned}\pi_{ij} &= Pr\left(y_i = j \middle| \alpha_i, \gamma_{ij}\right) \\ &= \Lambda[(\delta_j + \gamma_{i,j})\, \boldsymbol{z}_{i,j} - \{(\beta + \alpha_i)\boldsymbol{x}_i + \sigma \boldsymbol{u}_i + \lambda q_i^R\}] - \Lambda[(\delta_{j-1} + \gamma_{i,j-1})\, \boldsymbol{z}_{i,j} \\ &\qquad - \{(\beta + \alpha_i)\boldsymbol{x}_i + \sigma \boldsymbol{u}_i + \lambda q_i^R\}]\end{aligned} \tag{6}$$

The unconditional probability can subsequently be obtained as:

$$P_{ij} = \int_{\alpha_i, \gamma_{ij}} [Pr\left(y_i = j \middle| \alpha_i, \gamma_{ij}\right)] * dF\left(\alpha_i, \gamma_{ij}\right) d(\alpha_i, \gamma_{ij}) \tag{7}$$

The parameters to be estimated in the MGOL model with the 2SRI treatment are: the parameters corresponding to the linear regression (ρ), the parameters corresponding to the propensity ($\beta, \sigma, \lambda\ and\ \alpha_i$) and the parameters corresponding to thresholds ($\delta_j\ and\ \gamma_{i,j}$). In this study, we use a quasi-Monte Carlo (QMC) method proposed by Bhat (2001) for discrete outcome model to draw realization from its population multivariate distribution. Within the broad framework of QMC sequences, we specifically use the Halton sequence (4,000 Halton draws) in the current analysis (see Eluru et al., 2008 for a similar estimation process).

3. DATA

3.1 Data Source

The data for the current study is sourced from the FARS database for the year 2010. FARS data is a census of all fatal crashes in the US and compiles crashes if at least one person involved in the crash dies within thirty consecutive days from the time of crash. The FARS database has a record of 30,196 fatal crashes with 32,885 numbers of fatalities for the year 2010. This data base is obtained from the US Department of Transportation, National Highway Traffic Safety Administration's National Center for Statistics and Analysis (ftp://ftp.nhtsa.dot.gov). The FARS dataset provides a continuous timeline of the fatal occurrences from the time of crash until thirty days. It also provides information on a multitude of factors (driver characteristics, vehicle characteristics, roadway design and operational attributes, environmental factors, crash characteristics and situational variables) representing the crash situation and events.

3.2 Sample Formation and Description

This study is focused on fatality outcome of passenger vehicles' drivers who were involved in either a single or two vehicle crashes. The crashes that involve more than two vehicles are excluded from the dataset. Commercial vehicles involved collisions are also excluded to avoid the potential systematic differences between commercial and non-commercial driver groups. From the dataset, only the drivers who were fatally injured are considered for the current study. The final FARS dataset, after removing records with missing information for essential attributes consisted of about 5,102 driver records. The continuous timeline (computed as the difference between declared death time and crash time) provided in FARS was then discretized as a seven point discrete ordinal variable to represent the scale of fatal injury severity of drivers involved in these crashes - from least severe to most severe fatal crashes as follows: 1) Died between 6th to 30 days of crash, 2) Died between 2nd to 5 days of crash, 3) Died between 7th to 24 hours of crash, 4) Died between 2nd to 6 hours of crash, 5) Died between 31st to 60 minutes of crash, 6) Died between 1st to 30 minutes of crash and 7) Died instantly. The distributions of driver fatalities over the fatality scale in our final estimation sample are presented in Table 1. We adopted a seven alternative discrete spectrum for our analysis based on observed frequencies and time to death groupings of policy interest. It is important to note that, within an ordered outcome structure, it would be relatively easy to incorporate a larger number of alternative categories, if needed, while still retaining a parsimonious specification. From Table 1 we can see that more that 60% drivers died within one hour of crash and almost one third of these crash victims are reported to die instantly. Also, only 5.9% of the drivers can evade mortality more than five days of crashes.

Table 2 offers a summary of the sample characteristics of the exogenous factors in the estimation dataset. From the descriptive analysis, we observe that a large portion of crashes occur on high speed limit road (54.5%), on rural road (62.8 %), during dry weather condition (70.6%) and at non-intersection location (75.9%). The majority of drivers are aged between 25 and 64 (57.1%). In addition to the variables describing the crash situation and events presented in Table 2, FARS database also provides information on crash notification time, EMS response time and time of EMS arrival at hospital. From this information, it is possible to compute EMS response time (as the difference between EMS arrival time at the crash scene and crash time) and hospital arrival time (as the difference between EMS arrival at hospital and EMS arrival at crash scene). However, EMS arrival time at hospital is available only for the crash victim who arrived first at hospital among all other crash victims (if present) for that specific crash. Therefore, hospital arrival time is not available for all fatal records of driver, and, hence is not considered in our final estimation sample. On the other hand, the sample we use in the current study provides information about the EMS response time. From the descriptive statistics of this variable we observe that EMS response time exceeds one hour – most popularly referred to as the "golden hour" – only for 3.1% of records. The median EMS response time is about 11 minutes, with a range of 0 minute to approximately 9.5 hours.

4. EMPIRICAL ANALYSIS

4.1 Variables Considered

In our analysis, we selected a host of variables from six broad categories: <u>driver characteristics</u> (including driver age, alcohol consumption and previous driving conviction records), <u>vehicle characteristics</u> (including vehicle age), <u>roadway design and operational attributes</u> (including speed limit, traffic control device, roadway functional class and land use), <u>environmental factors</u> (including time of day, lighting condition and weather condition), <u>crash characteristics</u> (including manner of collision and collision location) and <u>situational variable</u> (including driver ejection, number of passengers and EMS response time). The final specification of the model development was based on combining the variables when their effects were not statistically different and by removing the statistically insignificant variables in a systematic process based on statistical significance (95% confidence level). For continuous variables, linear, polynomial and spline forms were tested.

4.2 Model Specification and Overall Measures of Fit

In the research effort, initially we estimated three different models: 1) OL, 2) GOL and 3) MGOL, by considering EMS response time as an explanatory variable in our empirical analysis. In our initial specifications of all the three aforementioned models we obtained a counterintuitive result for EMS response time indicating that the likelihood of early death decreases with an increase in EMS response time. Therefore, to further explore the effect of this indicator variable, several specifications (log transformation, dummy categories) of EMS response time have been explored in OL, GOL and MGOL frameworks. However, for all the aforesaid specifications, we observe that a longer EMS response time has negative impact on the survival probability of drivers in the current study context. The result could be a manifestation of endogeneity between crash seriousness and EMS response time i.e. severe crashes are likely to have shorter EMS times while less severe crashes are likely to have longer EMS times. So, in such scenarios the early arrival of EMS coincides with early death causing a non-intuitive parameter estimate. Thus, to control for the endogeneity of EMS response time with fatal crash outcomes, we include a residual variable through 2SRI method in examining the fatality spectrum. To that extent, we have further estimated the following three ordered outcome models with endogenous treatment: 1) OL with 2SRI treatment, 2) GOL with 2SRI treatment and 3) MGOL model with 2SRI treatment. After controlling for the endogeneity, the coefficient on the logarithm of EMS response time is found out to be positive in all three model specifications indicating that the likelihood of early death increases with an increase in EMS response time.

Prior to discussing the estimation results, we compare the performance of these models in this section. At first, the exogeneity of regressors $\boldsymbol{u}_i$ in equation 4 is tested for $\lambda = 0$ by using likelihood ratio (LR) test within each set of models. The LR test statistic is computed as $2[LL_U - LL_R]$, where LL_U and LL_R are the log-likelihood of the unrestricted and the restricted models, respectively. The computed value of the LR test is compared with the χ^2 value for the corresponding degrees of freedom. These estimates are presented in Table 3. From the first three rows of LR test values in table 3 we can see that all three models with 2SRI treatment outperform the corresponding models without 2SRI treatments at any significance level. The LR test comparisons confirm the importance of accommodating endogoneity between EMS response time and fatal injury outcome in the analysis of driver fatalities. Further, we also compare the estimated ordered models with 2SRI treatments by using LR test for selecting the preferred model among those. The results are presented in last three rows of Table 3. The LR test values indicate that MGOL model with 2SRI treatment outperforms the OL model with 2SRI treatment at any level of statistical significance. The MGOL model with 2SRI treatment outperforms the GOL model with 2SRI treatment at the 0.05 significance level. The comparison exercise clearly highlights the superiority of the MGOL model with 2SRI treatment in terms of data fit compared to all the other ordered models.

4.3 Estimation Results

In presenting the effects of exogenous variables in the model specification, we will restrict ourselves to the discussion of the MGOL model with 2SRI treatment. Table 4 presents the estimation results. To reiterate, the dependent variable under consideration is the 7 point ordinal variable defined as: died between 6th-30 days of crash, died between 2nd-5 days of crash, died between 7th-24 hours of crash, died between 1st-6 hours of crash, died between 31st-60 minutes of crash, died between 1st-30 minutes of crash and died instantly. Estimation results of Table 4 has six different columns. The first column corresponds to the propensity and represents the estimates of the parameters of equation 4. From second to sixth columns of estimation results in Table 4 corresponds to the thresholds and represent parameters of equation 5. In MGOL model, when the threshold parameter is positive (negative) the result implies that the threshold is bound to increase (decrease); the actual effect on the probability is quite non-linear and can only be judged in conjunction with the influence of the variable on propensity and other thresholds. In the following sections, the estimation results are discussed by variable groups.

Driver Characteristics: The effect of driver age is found to have significant impact on the length of hospital stay before death. The parameter characterizing the effect of young driver (age 24 & less) suggests that the likelihood of dying earlier is lower for young driver compared to middle-aged (age 25-64) driver. The negative sign of latent propensity associated with old driver (age 65 & above) suggests that the likelihood of dying earlier is lower for older driver compared to middle-aged driver. On the other hand, the impacts of old driver on both of the fourth and fifth thresholds are negative. The results suggest an increased probability of dying within 6th-30 days of crash and, also in general, a decreased possibility of instant death, presumably due to the declined wound healing and immune competence of drivers with advancing age after surviving the early phase of trauma (Tohira et al., 2012).

As expected, MGOL model estimates related to alcohol impairment indicate a higher likelihood of early mortality risk of alcohol impaired drivers compared to the sober drivers. At the same time, the positive values of the second threshold of alcohol impaired driver reflects an increase in the probability of dying within 2nd-5 days of crash. Intoxicated drivers are identified to be less immune to post traumatic response and suffer from more severe abdominal injuries (Zeckey et al., 2011; Stübig et al., 2012). Furthermore, higher impact speed differential due to the risk taking disposition of alcohol intoxicated driver presumably reduces the time to death of this group of drivers (Soderstrom et al., 2001).

Previous driving records also have significant influence on time to death after crash. The results associated with previous recorded suspension and revocation of driving licence indicates that an increase in number of previous recorded suspension and revocation deceases the likelihood of early mortality. The result is perhaps indicating more cautious driving of this group of driver to avoid any further conviction while driving. Also, the result indicates that drivers are less likely to evade early mortality with an increasing record of other previous record of harmful motor vehicle convictions (other than previous recorded suspension and revocation of driving licence, previous recorded crashes, previous drinking convictions and previous speeding convictions). However, the effect of other previous record of harmful motor vehicle convictions variable results in an estimate that is normally distributed with mean 0.104 and standard deviation of 0.208 implying that almost 71% of the drivers with higher records of earlier harmful motor vehicle convictions involved in the collision sustain early death.

Vehicle Characteristics: Among different vehicle characteristics explored in this study, only vehicle age is significant in the final model specification. Vehicle age result does not have any effect on the propensity of time to death after crash, but demonstrates a higher likelihood of death within 1st-30 minutes of crash for the driver of old vehicles (vehicle age≥11 years) and in general, a higher probability of instant death in a crash. The result highlights the advantages of newer vehicle fleet – presence of advanced safety technologies (electronic stability control, improvement in air bag design, crash cage, energy-absorbing steering columns, crash-resistant door locks and high-penetration-resistant windshields) and designs of newer vehicle with improved crash worthiness (O'Neill, 2009; Ryb et al., 2011).

Roadway Design and Operational Attributes: The results for speed limit indicate that the propensities to die earlier are higher for crashes occurring on roads with medium or higher speed limit roads relative to crashes on lower speed limit roads. As is expected, within the two speed categories considered, the higher speed category has a larger impact relative to the medium speed category, which underscores the fact that the probability of early mortality risk increases with the increasing speed limits of roadways. MGOL model estimates for higher speed limit results in a parameter that is normally distributed with a mean 0.359 and standard deviation 0.447, which indicates that almost 78% of the drivers cannot evade early death for the crashes occurring on higher speed limit roads. Higher speed, representing average driving speed, significantly increases the kinetic energy of crashes (Elvik, 2004; Sobhani et al., 2011) resulting in medical complications with multiple injuries and traumatic brain injury to the victims (Weninger and Hertz, 2007). Further, the cabin intrusion caused by high mechanical force of such crash might also increase the extrication time of victims from the damaged vehicle (Weninger and Hertz, 2007). Crashes at stop-sign controlled or other traffic controlled (such as warning sign, regulatory sign, railway crossing sign) intersections seem to increase the likelihood of early death relative to crashes at other locations, possibly suggesting non-compliance with these traffic control devices and judgment problems (Chipman, 2004; Retting et al., 2003).

Environmental Factors: With respect to time of day, the latent propensities for off peak and evening peak periods (related to morning peak and nigh-time) are found negative, indicating lower likelihood of early mortality, may be a result of traffic congestion and slow driving speeds during these periods. At the same time, the effect of off peak period on the threshold indicates a lower probability of dying within 1st-30 minutes after crash. The weather condition effects simplified to a simple binary representation of cloudy condition. The result indicates that if collisions occur during cloudy weather (relative to those during other weather conditions) the drivers are less likely to evade early death, perhaps because of the reduced visibility, which presumably results in reduced perception-reaction and reduced ability to take evasive actions at the crash incident (Tay et al., 2011). The effect of cloudy weather condition on the threshold also indicates increased likelihood of death within 2nd-5 days of crash.

Crash Characteristics: With respect to manner of collision, the time to death propensity is observed to be lower for front-to-rear collision relative to other manners of collision. The results associated with a head-on collision reflect a higher probability of death within 1st-6 hours of crash and in general indicate the anticipated increased likelihood of early death. Head-on collisions are often caused by drivers violating traffic rules, crossing the centerline by mistake and losing control of their vehicles (Zhang and Ivan, 2005). The pre-impact speed vectors of motor vehicles are directed in opposing directions during a head-on collision, resulting in greater dissipation of kinetic energy and heavier deformation of motor vehicle bodies (Prentkovskis et al., 2010), resulting in higher risk of injury.

As observed in several previous studies (Al-Ghamdi, 2002), the results related to crash location of our study reflect an increased injury risk propensity for collision at non-intersection location (related to crashes at intersection and other locations). However, the effects of "non-intersection location" indicator in threshold parameterization are relatively complex. It has a positive impact on the threshold between 1st-6 hours and 31st-60 minutes crash outcome categories; while it has a negative impact on the threshold between 31st-60 minutes and 1st-30 minutes categories. In general, the net implication is that collision at non-intersection location has a higher probability of sustaining early death (the specific impact of other fatal crash categories on driver fatalities are context-specific).

Situational Variables: As identified in several previous studies (Palanca et al., 2003), the result related to driver ejection indicate an increased early death propensity. Number of passenger in vehicle at the time of collision is also found to have significant impact on the time to death of driver. The results related to presence of more passengers reflect an increased early death propensity, perhaps indicating inattentiveness to the driving task due to distraction caused by in vehicle interactions among occupants.

The last two rows of estimation results in Table 4 represent the associated results of: (1) the logarithm of EMS response time and (2) the residual obtained from regressing the logarithm of EMS response time variable on morning peak, late-night, dark-not lighted, rain, snowy, rural, principle arterial and minor arterial indicator variables[2]. The role of the residual variable is to control for the endogeneity of the EMS response time variable in examining the time to death. From Table 4, we can see that after controlling for endogeneity, the coefficient on the logarithm of EMS response time is positive and statistically significant indicating that EMS response time has the expected impact on severity once we control for the endogeneity bias. Specifically, as can be observed from the coefficient of the residual term, the non-intuitive impact of EMS time was a result of the correlation between EMS time and unobserved determinants[3]. Through our approach, by accounting for the endogeneity we were able to differentiate between the observed impact of EMS time and the spurious effect due to the unobserved factors.

5. ELASTICITY EFFECTS

The parameter effects of the exogenous variables in Table 4 do not provide the magnitude of the effects on time to death of drivers. For this purpose, we compute the aggregate level "elasticity effects" for all categories of independent variable (see Eluru and Bhat, (2007) for a discussion on the methodology for computing elasticities) and present the computed elasticities in Table 5. The effects are computed for all categories of fatal crashes. The results in the table can be interpreted as the percentage change (increase for positive sign and decrease for negative sign) in the probability of the fatal severity categories due to the change in that specific exogenous variable.

[2] The estimation results for the linear regression model are presented in Appendix A.
[3] Several specifications were tested to examine the coefficient on the residual from regression of Logarithm of EMS arrival time. The coefficient was relatively stable across these specifications and we chose the model that offered the best fit.

The following observations can be made based on the elasticity effects of the variables presented in Table 5. First, the results in Table 5 indicate that there are considerable differences in the elasticity effects across different fatal crash categories, suggesting that fatality is not a single state but rather is made up of multiple discrete states from dying instantly to dying within thirty days of crash. Second, the most significant variables in terms of lower survival probability for drivers are crashes on high speed limit roads, crashes on medium speed limit roads and head-on crashes. A forgiving road environment should be designed for a high and medium speed limit road locations to allow the drivers more space to recover from a driving error. Moreover, policies concerning enforcement for reducing traffic violations have the potential to reduce head-on crashes. Third, in terms of longer survival probability, the important factors are old driver, front-to-rear crash and crashes during off peak period. Fourth, elasticity estimates of EMS response time in Table 5 emphasize the importance of early EMS response. Finally, the elasticity analysis assists in providing a clear picture of attribute impact on driver time-to-death variables. The elasticity analysis conducted provides an illustration of how the proposed model can be applied to determine the critical factors contributing to reducing the survival time.

6. CONCLUSIONS

In the United States, safety researchers have focused on examining fatal crashes (involving at least one fatally injured vehicle occupant) by using Fatality Analysis Reporting System (FARS) dataset. FARS database compiles crashes if at least one person involved in the crash dies within thirty consecutive days from the time of crash along with the exact timeline of the fatal occurrence. Previous studies using FARS dataset offer many useful insights on what factors affect crash related fatality, particularly in the context of fatal vs. non-fatal injury categorization. However, there is one aspect of fatal crashes that has received scarce attention in the traditional safety analysis. The studies that dichotomize crashes into fatal versus non-fatal groups assume that all fatal crashes in the FARS dataset are similar. Keeping all else same, a fatal crash that results in an immediate fatality is clearly much more severe than another crash that leads to fatality after several days. Research attempts to discern such differences are useful in determining what factors affect the time between crash occurrence and time of death so that countermeasures can be implemented to improve safety situation and to reduce crash related fatalities.

To that extent, the current research makes a three-fold contribution to the literature on vehicle occupant injury severity analysis. First, our study is the first attempt to analyze the fatal injury from a new perspective and examine fatality as a continuous spectrum based on survival time ranging from dying within thirty days of crash to dying instantly. For the empirical analysis, the fatality timeline information obtained through FARS was categorized as an ordered variable ranging from death in thirty days to instantaneous death in seven categories as follows: died between 6th-30 days of crash, died between 2nd-5 days of crash, died between 7th-24 hours of crash, died between 1st-6 hours of crash, died between 31st-60 minutes of crash, died between 1st-30 minutes of crash and died instantly. Second, we estimated two-equation model that comprises of regression for EMS response time and ordered outcome model with residuals from the EMS model to correct for endogeneity bias on the effect of exogenous factors on the timeline of death. In doing so, the correction for endogeneity bias was pinned down in the ordered outcome models by employing a two-stage residual inclusion (2SRI) approach. In the research effort, we estimated the following three ordered outcome models with endogenous treatment: 1) OL with 2SRI treatment, 2) GOL with 2SRI treatment and 3) MGOL model with 2SRI treatment while employing a comprehensive set of exogenous variables (driver characteristics, vehicle characteristics, roadway design and operational attributes, environmental factors, crash characteristics and situational variables). The comparison exercise highlighted the superiority of the MGOL model with 2SRI treatment on the sample in terms of data fit compared to the other ordered outcome models in the current study context.

From the empirical analysis we found that, the factors that contributed to an increase in the likelihood of early death include: alcohol impairment, previous record of other harmful motor vehicle convictions, medium and higher speed limit, presence of stop sign, presence of other traffic control device, cloudy weather, head-on crashes, collision at non-intersection locations, driver ejection, presence of more passengers and longer EMS response time. The factors that contributed to a decrease in the likelihood of early death include: young driver, previous record of license suspension and revocation, crashes during off-peak and evening peak periods and front-to-rear crashes. In our research, to further understand the impact of various exogenous factors, elasticity effects were estimated. Moreover, we found that after controlling for endogeneity, the coefficient on the logarithm of EMS response time was intuitive and statistically significant indicating that EMS response time is correlated with unobserved determinants generating endogeneity in the outcome model of the time to death of drivers.

In our research, to further understand the impact of various exogenous factors, elasticity effects were estimated. The elasticity effects indicated that there were considerable differences in the elasticity effects across different fatal crash categories, suggesting that fatality is not a single state but rather is made up of multiple discrete states from dying instantly to dying within thirty days of crash. The most significant variables in terms of lower survival probability for drivers were crashes on high speed limit roads, crashes on medium speed limit roads and head-on crashes. In terms of longer survival probability, the important factors were old driver, front-to-rear crash and crashes during off-peak period. Moreover, the elasticity analysis assisted in providing a clear picture of attribute impact on driver time-to-death variables.

The study is not without limitations. In our research effort, we categorized the spectrum of fatal crashes in seven refined categories of fatalities ranging from fatality after thirty days to instant death. However, some of the earlier studies (Trunkey, 1983) argued that the distribution of survival times after traffic crash is “trimodal”. There are also studies (Clark et al., 2012) that contradict the trimodal distribution of survival time after crash. Thus, it will be an interesting exercise to explore the impact of the fatality spectrum discretization in examining the impact of exogenous variable within the MGOL model structure. In our analysis, we adopted an ordinary least squares based instrumentation approach for EMS time. However, it might be useful to consider alternative instrumentation approaches such as a duration model based instrumentation in future efforts. Finally, we do recognize that many relevant variables on medical treatment offered to injured drivers is unavailable in FARS data. Efforts to augment FARS data with such detail will substantially enhance empirical findings from the model estimated in our research.

ACKNOWLEDGEMENTS

The corresponding author would like to acknowledge financial support from Natural Sciences and Engineering Research Council (NSERC) of Canada under the Discovery Grants program for undertaking the research. The authors would also like to acknowledge the contribution of Tonjanika Robinson who contributed through a preliminary literature review effort for the research.

REFERENCES

Al-Ghamdi, A., 1999. Injury Severity and Duration of Hospital Stay for Urban Road Accidents in Riyadh. Transportation Research Record 1665, 125-132.

Al-Ghamdi, A.S., 2002. Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity. Accident Analysis and Prevention 34 (6), 729-741.

Arditi, D., Lee, D.E., Polat, G., 2007. Fatal Accidents in Nighttime vs. Daytime Highway Construction Work Zones. Journal of Safety Research 38 (4), 399-405.

Baker, T.K., Falb, T., Voas, R., Lacey, J., 2003. Older Women Drivers: Fatal Crashes in Good Conditions. Journal of Safety Research 34 (4), 399-405.

Bédard, M., Guyatt, G.H., Stones, M.J., Hirdes, J.P., 2002. The Independent Contribution of Driver, Crash, and Vehicle Characteristics to Driver Fatalities. Accident Analysis and Prevention 34 (6), 717-727.

Bhat, C.R., 2001. Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. Transportation Research Part B 35 (7), 677-693.

Braver, E.R., 2003. Race, Hispanic Origin, and Socioeconomic Status in Relation to Motor Vehicle Occupant Death Rates and Risk Factors among Adults. Accident Analysis and Prevention 35 (3), 295-309.

Brodsky, H., 1992. Delay in Ambulance Dispatch to Road Accidents. American Journal of Public Health 82 (6), 873-875.

Brown, L.H., Khanna, A., Hunt, R.C., 2000. Rural vs Urban Motor Vehicle Crash Death Rates: 20 Years of FARS Data. Prehospital Emergency Care 4 (1), 7-13.

Campos-Outcalt, D., Bay, C., Dellapena, A., Cota, M.K., 2003. Motor Vehicle Crash Fatalities by Race/Ethnicity in Arizona, 1990-96. Injury Prevention 9 (3), 251-256.

Chamberlain, G., 1980. Analysis of Covariance with Qualitative Data. The Review of Economic Studies 47 (1), 225-238.

Chen, L., Baker, S.P., Braver, E.R., Li, G., 2000. Carrying Passengers as a Risk Factor for Crashes Fatal to 16- and 17-Year-Old Drivers. Journal of American Medical Association 283 (12), 1578-1582.

Chipman, M.L., 2004. Side Impact Crashes - Factors Affecting Incidence and Severity: Review of the Literature. Traffic Injury Prevention 5 (1), 67-75.

Clark, D.E., Cushing, B.M., 2002. Predicted Effect of Automatic Crash Notification on Traffic Mortality. Accident Analysis and Prevention 34 (4), 507-513.

Clark, D.E., Qian, J., Sihler, K.C., Hallagan, L.D., Betensky, R.A., 2012. The Distribution of Survival Times after Injury. World Journal of Surgery 36 (7), 1562-1570.

Clark, D.E., Winchell, R.J., Betensky, R.A., 2013. Estimating the Effect of Emergency Care on Early Survival after Traffic Crashes. Accident Analysis and Prevention 60, 141-147.

Eluru, N., 2013. Evaluating Alternate Discrete Choice Frameworks for Modeling Ordinal Discrete Variables. Accident Analysis and Prevention 55, 1-11.

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.

Elvik, R., 2004. To What Extent can Theory Account for the Findings of Road Safety Evaluation Studies?. Accident Analysis and Prevention 36 (5), 841-849.

Fabbri, A., Marchesini, G., Morselli-Labate, A.M., Rossi, F., Cicognani, A., Dente, M., Iervese, T., Ruggeri, S., Mengozzi, U., Vandelli, A., 2002. Positive Blood Alcohol Concentration and Road Accidents. A Prospective Study in an Italian Emergency Department. Emergency Medicine Journal 19 (3), 210-214.

Feero, S., Hedges, J.R., Simmons, E., Irwin, L., 1995. Does Out-of-Hospital EMS Time Affect Trauma Survival? The American Journal of Emergency Medicine 13 (2), 133-135.

Fredette, M., Mambu, L.S., Chouinard, A., Bellavance, F., 2008. Safety Impacts due to the Incompatibility of SUVs, Minivans, and Pickup Trucks in Two-Vehicle Collisions. Accident Analysis and Prevention 40 (6), 1987-1995.

Gates, J., Dubois, S., Mullen, N., Weaver, B., Bédard, M., 2013. The Influence of Stimulants on Truck Driver Crash Responsibility in Fatal Crashes. Forensic Science International 228 (1-3), 15-20.

Golias, J.C., Tzivelou, H.S., 1992. Aspects of Road-Accident Death Analyses. Journal of Transportation Engineering 118 (2), 299-311.

Gonzalez, R.P., Cummings, G., Mulekar, M., Rodning, C.B., 2006. Increased Mortality in Rural Vehicular Trauma: Identifying Contributing Factors through Data Linkage. Journal of Trauma - Injury, Infection and Critical Care 61 (2), 404-409.

Gonzalez, R.P., Cummings, G.R., Phelan, H.A., Mulekar, M.S., Rodning, C.B., 2009. Does Increased Emergency Medical Services Prehospital Time Affect Patient Mortality in Rural Motor Vehicle Crashes? A Statewide Analysis. The American Journal of Surgery 197 (1), 30-34.

Governors Highway Safety Association (GHSA), 2009. Toward Zero Deaths: Every Life Counts, GHSA, Washington, D.C.

Hanna, C.L., Laflamme, L., Bingham, C.R., 2012. Fatal Crash Involvement of unlicensed Young Drivers: County Level Differences According to Material Deprivation and Urbanicity in the United States. Accident Analysis and Prevention 45, 291-295.

Harper, J.S., Marine, W.M., Garrett, C.J., Lezotte, D., Lowenstein, S.R., 2000. Motor Vehicle Crash Fatalities: A Comparison of Hispanic and Non-Hispanic Motorists in Colorado. Annals of Emergency Medicine 36 (6), 589-596.

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.

Ju, Y.H., Sohn, S.Y., 2014. Time to Death Analysis of Road Traffic Accidents in Relation to Delta V, Drunk Driving, and Restraint Systems. Traffic Injury Prevention 15 (8), 771-777.

Kim, J.K., Ulfarsson, G.F., Kim, S., Shankar, V.N., 2013. Driver-Injury Severity in Single-Vehicle Crashes in California: A Mixed Logit Analysis of Heterogeneity due to Age and Gender. Accident Analysis and Prevention 50, 1073-1081.

Kockelman, K.M., Kweon, Y.J., 2002. Driver Injury Severity: An Application of Ordered Probit Models. Accident Analysis and Prevention 34 (3), 313-321.

Lambert-Bélanger, A., Dubois, S., Weaver, B., Mullen, N., Bédard, M., 2012. Aggressive Driving Behaviour in Young Drivers (Aged 16 through 25) Involved in Fatal Crashes. Journal of Safety Research 43 (5-6), 333-338.

Lyman, S., Ferguson, S.A., Braver, E.R., Williams, A.F., 2002. Older Driver Involvements in Police Reported Crashes and Fatal Crashes: Trends and Projections. Injury Prevention 8 (2), 116-120.

Marson, A.C., Thomson, J.C., 2001. The Influence of Prehospital Trauma Care on Motor Vehicle Crash Mortality. Journal of Trauma Injury - Infection and Critical Care 50 (5), 917-921.

Meng, Q., Weng, J., 2013. Uncertainty Analysis of Accident Notification Time and Emergency Medical Service Response Time in Work Zone Traffic Accidents. Traffic Injury Prevention 14 (2), 150-158.

Mooradian, J., Ivan, J.N., Ravishanker, N., Hu, S., 2013. Analysis of Driver and Passenger Crash Injury Severity Using Partial Proportional Odds Models. Accident Analysis and Prevention 58, 53-58.

National Highway Traffic Safety Administration (NHTSA), 2012. Traffic Safety Facts: 2010 Motor Vehicle Crashes Overview, NHTSA, Washington, D.C.

O'Neill, B., 2009. Preventing Passenger Vehicle Occupant Injuries by Vehicle Design- A Historical Perspective from IIHS. Traffic Injury Prevention 10 (2), 113-126.

Palanca, S., Taylor, D.M., Bailey, M., Cameron, P.A., 2003. Mechanisms of Motor Vehicle Accidents that Predict Major Injury. Emergency Medicine 15 (5-6), 423-428.

Prentkovskis, O., Sokolovskij, E., Bartulis, V., 2010. Investigating Traffic Accidents: A Collision of Two Motor Vehicles. Transport 25 (2), 105-115.

Retting, R.A., Weinstein, H.B., Solomon, M.G., 2003. Analysis of Motor-Vehicle Crashes at Stop Signs in Four U.S. Cities. Journal of Safety Research 34 (5), 485-489.

Romano, E., Voas, R., Tippetts, S., 2006. Stop Sign Violations: The Role of Race and Ethnicity on Fatal Crashes. Journal of Safety Research 37 (1), 1-7.

Ryb, G.E., Dischinger, P.C., McGwin, G., Griffin, R.L., 2011. Crash-Related Mortality and Model Year: Are Newer Vehicles Safer? Paper presented at the Annals of Advances in Automotive Medicine, 55, 113-121.

Sánchez-Mangas, R., García-Ferrrer, A., De Juan, A., Arroyo, A.M., 2010. The Probability of Death in Road Traffic Accidents. How Important is a Quick Medical Response? Accident Analysis and Prevention 42 (4), 1048-1056.

Siskind, V., Steinhardt, D., Sheehan, M., O'Connor, T., Hanks, H., 2011. Risk Factors for Fatal Crashes in Rural Australia. Accident Analysis and Prevention 43 (3), 1082-1088.

Sivak, M., Schoettle, B., Rupp, J., 2010. Survival in Fatal Road Crashes: Body Mass Index, Gender, and Safety Belt Use. Traffic Injury Prevention 11 (1), 66-68.

Sobhani, A., Young, W., Logan, D., Bahrololoom, S., 2011. A Kinetic Energy Model of Two-Vehicle Crash Injury Severity. Accident Analysis and Prevention 43 (3), 741-754.

Soderstrom, C.A., Ballesteros, M.F., Dischinger, P.C., Kerns, T.J., Flint, R.D., Smith, G.S., 2001. Alcohol/Drug Abuse, Driving Convictions, and Risk-Taking Dispositions among Trauma Center Patients. Accident Analysis and Prevention 33 (6), 771-782.

Srinivasan, K.K., 2002. Injury Severity Analysis with Variable and Correlated Thresholds: Ordered Mixed Logit Formulation, Transportation Research Record 1784, 132-142.

Stübig, T., Petri, M., Zeckey, C., Brand, S., Müller, C., Otte, D., Krettek, C., Haasper, C., 2012. Alcohol Intoxication in Road Traffic Accidents Leads to Higher Impact Speed Difference, Higher ISS and MAIS, and Higher Preclinical Mortality. Alcohol 46 (7), 681-686.

Tay, R., Choi, J., Kattan, L., Khan, A., 2011. A Multinomial Logit Model of Pedestrian–Vehicle Crash Severity. International Journal of Sustainable Transportation 5 (4), 233-249.

Terza, J.V., Basu, A., Rathouz, P.J., 2008. Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. Journal of Health Economics 27 (3), 531-543.

Thompson, J.P., Baldock, M.R.J., Mathias, J.L., Wundersitz, L.N., 2013. An Examination of the Environmental, Driver and Vehicle Factors Associated with the Serious and Fatal Crashes of Older Rural Drivers. Accident Analysis and Prevention 50, 768-775.

Tohira, H., Jacobs, I., Mountain, D., Gibson, N., Yeo, A., 2012. Differences in Risk Factors between Early and Late Trauma Death after Road Traffic Accidents, 2012 IRCOBI Conference Proceedings - International Research Council on the Biomechanics of Injury, 1-9.

Travis, L.L., Clark, D.E., Haskins, A.E., Kilch, J.A., 2012. Mortality in Rural Locations after Severe Injuries from Motor Vehicle Crashes. Journal of Safety Research 43 (5-6), 375-380.

Trunkey DD., 1983. Trauma. Scientific American 249 (2), 28–35.

Valent, F., Schiava, F., Savonitto, C., Gallo, T., Brusaferro, S., Barbone, F., 2002. Risk Factors for Fatal Road Traffic Accidents in Udine, Italy. Accident Analysis and Prevention 34 (1), 71-84.

Viano, D.C., Parenteau, C.S., 2010. Severe-to-Fatal Injury Risks in Crashes with Two Front-Seat Occupants by Seat Belt Use. Traffic Injury Prevention 11 (3), 294-299.

Weninger, P., Hertz, H., 2007. Factors Influencing the Injury Pattern and Injury Severity after High Speed Motor Vehicle Accident-A Retrospective Study. Resuscitation 75 (1), 35-41.

Yasmin, S., Eluru, N., 2013. Evaluating Alternate Discrete Outcome Frameworks for Modeling Crash Injury Severity. Accident Analysis and Prevention 59, 506-521.

Zador, P.L., Krawchuk, S.A., Voas, R.B., 2000. Alcohol-Related Relative Risk of Driver Fatalities and Driver Involvement in Fatal Crashes in Relation to Driver Age and Gender: An Update Using 1996 Data. Journal of Studies on Alcohol 61 (3), 387-395.

Zeckey, C., Dannecker, S., Hildebrand, F., Mommsen, P., Scherer, R., Probst, C., Krettek, C., Frink, M., 2011. Alcohol and Multiple Trauma-Is There an Influence on the Outcome?. Alcohol 45 (3), 245-251.

Zhang, C., Ivan, J.N., 2005. Effects of Geometric Characteristics on Head-On Crash Incidence on Two-Lane Roads in Connecticut, Transportation Research Record 1908, 159-164.

Zhang, X., Yao, H., Hu, G., Cui, M., Gu, Y., Xiang, H., 2013. Basic Characteristics of Road Traffic Deaths in China. Iranian Journal of Public Health 42 (1), 7-15.

LIST OF TABLES

TABLE 1 Distribution of Fatal Injury Severity Categories

TABLE 2 MGOL Estimates

TABLE 3 Measures of Fit in Estimation Sample

TABLE 4 MGOL Estimates

TABLE 5 Elasticity Effects

TABLE 1 Distribution of Fatal Injury Severity Categories

Fatal Crash Categories	Frequency	Percentage
Died between 6th to 30 days of crash	302	5.9%
Died between 2nd to 5 days of crash	270	5.3%
Died between 7th to 24 hours of crash	233	4.6%
Died between 2nd to 6 hours of crash	1175	23.0%
Died between 31st to 60 minutes of crash	824	16.1%
Died between 1st to 30 minutes of crash	1086	21.3%
Died instantly	1212	23.8%
Total	5102	100.0%

TABLE 2 Crash Database Sample Statistics

Categorical Explanatory Variables	Sample Share	
	Frequency	Percentage
Driver Characteristics		
Driver age		
Age 24 & less	1144	22.423
Age 25-64	2915	57.134
Age 65 & above	1043	20.443
Under the influence of alcohol	1778	34.849
Vehicle Characteristics		
Vehicle age		
Vehicle age<11 years	2822	55.312
Vehicle age≥11 years	2280	44.688
Roadway Design and Operational Attributes		
Speed limit		
Speed limit less than 26 mph	261	5.116
Speed limit 26 to 50 mph	2059	40.357
Speed limit above 50mph	2782	54.528
Traffic control device		
No traffic control, traffic signal and yield sign	4271	83.712
Stop sign	401	7.860
Other traffic control device	430	8.428
Roadway functional class		
Principal Arterial	1680	32.928
Minor Arterial	997	19.541
Collector	1208	23.677
Local Road	1217	23.853
Land use		
Rural	3206	62.838
Urban	1896	37.162
Environmental Factors		
Time of day		
Morning Peak	548	10.741
Off-peak	1266	24.814
Evening peak	828	16.229
Late evening	1311	25.696
Late night	1149	22.521
Lighting condition		
Daylight and other lighting condition	2910	57.036
Dark-not lighted	1430	28.028

Dark-lighted	762	14.935
Weather condition		
Dry	3601	70.580
Rain	422	8.271
Snowy	210	4.116
Cloudy	850	16.660
Other weather condition	19	0.372
Crash Characteristics		
Manner of collision		
Front to rear	124	2.430
Head-on	897	17.581
Other type of collision	4081	79.988
Collision location		
Non-Intersection	75.931	75.931
Intersection	15.759	15.759
Other Location	8.310	8.310
Situational Variables		
Driver ejection		
Ejected	1197	23.461
Not ejected	3905	76.539
Ordinal/Continuous Explanatory Variables	**Mean**	
Previous Recorded suspensions and revocations	0.444	
Previous record of other harmful motor vehicle convictions	0.323	
Number of passengers	0.400	
Logarithm of EMS response time (in minute)	2.473	

TABLE 3 Measures of Fit in Estimation Sample

Summary Statistic	OL	GOL	MGOL
Log-likelihood at zero	-9928.0	-9928.0	-9928.0
Log-likelihood at sample shares	-9016.3	-9016.3	-9016.3
Number of observations	5102	5102	5102
Summary Statistic	**Without 2SRI Treatment**		
Log-likelihood at convergence	-8844.8	-8794.9	-8793.7
Number of parameters	18	28	30
Summary Statistic	**With 2SRI Treatment**		
Log-likelihood at convergence	-8839.8	-8790.8	-8787.4
Number of parameters	19	29	31
Log-likelihood (LR) test	**LR Test Values**		
OL without 2SRI/OL with 2SRI	9.9 (1 degree of freedom)		
GOL without 2SRI/GOL with 2SRI	8.2 (1 degree of freedom)		
MGOL without 2SRI/MGOL with 2SRI	12.6 (1 degree of freedom)		
OL with 2SRI/GOL with 2SRI	98.1 (10 degrees of freedom)		
OL with 2SRI/MGOL with 2SRI	104.8 (12 degrees of freedom)		
GOL with 2SRI/MGOL with 2SRI	6.8 (2 degrees of freedom)		

TABLE 4 MGOL Estimates

Variables	**Latent Propensity**	τ_2	τ_3	τ_4	τ_5	τ_6
Constant	-1.712(-5.229)	-0.441(-6.101)	-0.854(-13.236)	0.141(2.296)	-0.104(-1.386)	0.069(1.915)
Driver Characteristics						
Driver age (Base: Age 25-64)						
Age 24 & less	-0.147(-2.207)*	–	–	–	–	–
Age 65 & above	-1.015(-10.966)	–	–	-0.281(-4.334)	-0.182(-2.071)	–
Under the influence of alcohol	0.488(3.488)	0.434(3.261)	–	–	–	–
Previous Recorded suspensions and revocations	-0.068(-3.264)	–	–	–	–	–
Previous record of other harmful motor vehicle convictions	0.104(2.598)	–	–	–	–	–
SD	0.208(3.596)	–	–	–	–	–
Vehicle Characteristics						
Vehicle age (Base: Vehicle age<11 years)						
Vehicle age≥11 years	–	–	–	–	-0.157(-2.689)	–
Roadway Design and Operational Attributes						
Speed limit (Base: Speed limit<26 mph)						
Speed limit 26 to 50 mph	0.251(2.117)	–	–	–	–	–
Speed limit above 50mph	0.359(2.981)	–	–	–	–	–
SD	0.447(2.707)	–	–	–	–	–
Traffic control device (Base: No traffic control, traffic signal and yield sign)						
Stop sign	0.223(1.975)	–	–	–	–	–
Other traffic control device	0.171(2.148)	–	–	–	–	–
Environmental Factors						
Time of day (Base: Morning Peak, Late evening and Late Night)						
Off peak	-0.218(-3.157)	–	–	–	–	-0.161(-2.323)
Evening peak	-0.151(-2.012)	–	–	–	–	–
Weather condition (Base: Dry, Rain, Snowy and Other weather condition)						
Cloudy	0.467(2.987)	0.276(1.872)	–	–	–	–
Crash Characteristics						
Manner of collision (Base: Other type of collision)						
Front to rear	-0.317(-1.765)	–	–	–	–	–

Head-on	0.661(5.312)	–	–	0.261(3.781)	–	–
Collision location (Base: Intersection and Other location)						
Non-intersection	0.362(3.741)	–	–	0.217(3.346)	-0.234(-3.168)	–
Situational Variables						
Driver ejection (Base: Not ejected)						
Ejected	0.267(3.651)	–	–	–	0.145(1.963)	–
Number of passengers	0.159(4.874)	–	–	–	–	–
EMS response time						
Logarithm of EMS response time *(in minutes)*	0.247(1.993)					
Residual from regression of Logarithm of EMS arrival time (in minutes) on morning peak, late night, dark-not lighted, rain, snowy, rural, principle arterial and minor arterial	-0.363(-2.929)	–	–	–	–	–
τ_2= Threshold between 1st-5 days/ 7th-24 hours; τ_3= Threshold between 7th-24 hours/ 1st-6 hours; τ_4 = Threshold between 1st-6 hours/ 31st-60 minutes; τ_5 = Threshold between 31st-60 minutes/ 1st-30 minutes; τ_6 = Threshold between 1st-30 minutes/ Died Instantly						

*t-stats are presented in parenthesis

TABLE 5 Elasticity Effects

Variables	Died between 6-30 days	Died between 2-5 days	Died between 7-24 hours	Died between 2-6 hours	Died between 31-60 minutes	Died between 1-30 minutes	Died instantly
Driver Characteristics							
Driver age (Base: Age 25-64)							
Age 24 & less	13.694	11.328	9.375	6.150	0.648	-4.202	-10.384
Age 65 & above	108.781	93.747	74.488	2.099	-18.729	-17.964	-35.615
Under the influence of alcohol	-39.798	22.427	-7.732	-5.153	-0.658	3.421	8.611
Previous Recorded suspensions and revocations	6.270	5.246	4.367	2.871	0.312	-1.950	-4.823
Previous record of other harmful motor vehicle convictions	-6.784	-6.263	-5.725	-4.674	-2.074	1.511	8.782
Vehicle Characteristics							
Vehicle age (Base: Vehicle age<11 years)							
Vehicle age≥11 years	0.000	0.000	0.000	0.000	-15.654	3.016	7.977
Roadway Design and Operational Attributes							
Speed limit (Base: Speed limit<26 mph)							
Speed limit 26 to 50 mph	-22.181	-18.747	-15.721	-10.576	-1.590	6.632	18.132
Speed limit above 50mph	-25.631	-23.452	-21.341	-16.467	-5.703	7.403	28.948
Traffic control device (Base: No traffic control, traffic signal and yield sign)							
Stop sign	-18.668	-15.959	-13.637	-9.645	-2.112	5.457	16.739
Other traffic control device	-14.440	-12.444	-10.636	-7.420	-1.487	4.369	12.715
Environmental Factor							
Time of day (Base: Morning Peak, Late evening and Late Night)							
Off peak	20.164	17.115	14.241	9.234	0.793	-18.852	-4.193
Evening peak	14.059	11.726	9.714	6.322	0.590	-4.386	-10.591
Weather condition (Base: Dry, Rain, Snowy and Other weather condition)							
Cloudy	-36.654	1.822	-14.118	-9.567	-1.671	5.906	16.468
Crash Characteristics							
Manner of collision (Base: Other type of collision)							
Front to rear	31.998	25.907	20.893	12.738	-0.027	-9.992	-21.155
Head-on	-49.427	-44.114	-39.224	4.763	-1.926	7.301	19.919
Collision location (Base: Intersection and Other location)							
Non-intersection	-34.151	-29.006	-24.066	12.236	-24.944	7.878	17.758

Situational Variables							
Driver ejection (Base: Not ejected)							
Ejected	-22.368	-19.547	-16.760	-11.686	12.997	4.252	11.816
Number of passenger	-13.339	-11.499	-9.842	-6.893	-1.428	4.007	11.851
EMS response time	-2.410	-2.001	-1.706	-1.157	-0.227	0.690	2.033

APPENDIX A Linear Regression Estimates

Variables	**Coefficient**	**t-stat**
Constant	2.190	80.670
Roadway functional class (Base: Collector and Local road		
Principal Arterial	-0.074	-2.709
Minor Arterial	-0.118	-3.699
Land use (Base: Urban)		
Rural	0.363	14.183
Time of day (Base: Off-peak, Evening peak and Late evening)		
Morning Peak	0.070	1.794
Late night	0.213	6.877
Lighting condition (Base: Daylight and other lighting condition and Dark-lighted)		
Dark-not lighted	0.120	4.161
Weather condition (Base: Dry, Cloudy and Other weather condition)		
Rain	0.088	2.052
Snowy	0.145	2.419