IDENTIFICATION OF FACTORS RELATED TO URBAN AND RURAL HIGHWAY CRASHES

By

Indike Ratnayake
Graduate Research Assistant
Department of Civil Engineering
Kansas State University
Manhattan, KS 66506
Email: ratnayak@ksu.edu

(Word count = 5,637 + 250x4 = 6,637)

Submitted for Presentation at Fall Student Conference, Midwest Transportation Consortium,
November 19, 2004,

ABSTRACT

This paper presents the details of an investigation aimed at identifying critical factors contributing towards higher crash severity in rural and urban highway crashes. Ordered probit models were used to analyze the crash data from KARS (Kansas Accident Reporting System) database. Both urban and rural crashes were considered separately in this analysis and five categories of contributory factors, driver, vehicular, roadway, environmental and crash related, were included in the models. The results of this study showed that most of the contributory factors, which contribute towards crashes resulting in higher severities, are common for urban and rural crashes. Many driver related factors such as alcohol involvement, lack of seat belt usage, excessive speed, and driver ejection or being trapped due to the crash are more critical in resulting more severe crashes. It also showed that roadway geometry related parameters such as curved and graded roads are more contributory towards higher crash severity. In two vehicle crashes, when the collision type is head-on, angle or rear-ended, the propensity of resulting in a more severe crash is high in both rural and urban highways. In rural areas, single vehicle crashes are significant towards higher severities compared to two vehicle crashes and animal vehicle crashes, but in urban areas both single and two vehicle crashes are significant. In contrast, under wet road surface conditions the probability of having a more severe crash is low in rural areas. The reason might be the reduced speeds due to driver cautiousness under such conditions.

INTRODUCTION

Addressing of safety issues related to highways has become a serious concern as the social and economic impacts due to highway crashes are high. In year 2002, total of 42,815 people died due to highway crashes in the United States (*1*) and nearly 2 million people were injured. According to a study carried out by NHTSA *(2)*, the economic lost due to highway crashes in year 2000 was about 230 billion and it was approximately equal to 2.3% of the US Gross Domestic Product in that year. These facts indicate the social and economic impact of highway crashes and it emphasizes the need of improving the safety of overall highway system.

About 60% of fatal crashes in year 2002 was occurred on rural highways on which over 75% of the total highway mileage (3.9 million) accounts for rural highways in the USA. However, less than 30% of total travel occurs on rural highways *(3).* Thus, the nature of crash occurrence on rural and urban highways seems to be different and the prevailing conditions on these two types of highways may also be different in most of the cases. In other words, the nature of relevant contributing factors, which contribute towards higher severities in urban and rural highway crashes, may also be different.

Previous studies have identified that those contributing factors could mainly be categorized as driver, environmental, roadway, vehicular and crash related. Numerous attempts have been made to identify critical factors towards highway crashes through statistical methods. However, comparatively fewer studies have been carried out on rural highways. The main objective of this study was to identify contributing factors that would be capable of increasing the severity of both rural and urban highway crashes, which consequently would be useful in suggesting necessary countermeasures in addressing the highway safety issues.

LITERATURE REVIEW

In almost all the crash reporting data bases, the crash severity is reported in three or more categories, fatal, incapacitating, property damage only, etc. and thus makes it possible of ordering the severity level from most severe to less severe. In other words, the severity, the response variable in the model, could be considered as an ordinal variable. This phenomenon has been applied to model the crash severity using both ordered probit and ordered logit structure by O'Donnell et al (*4*). Khattak et al (*5*) have employed an ordered probit modeling approach in their study to investigate the relevant factors towards injury severities to older drivers. Khattak et al (*6*), Kockelman et al (*7*), Ma et al (*8*) also have applied the ordered probit structure in their studies.

Shankar et al (*9*) have applied nested logit structure to successfully develop a model to find out the relationship between crash severity and crash prediction factors. The factors comprised of driver related factors, roadway geometry and surface condition, time of the crash and light condition at crash occurrence, and collision type whether single vehicle, two vehicle or multi vehicle. The advantage of this method is that the effects of unobserved terms could be avoided as they are cancelled off in the estimation process. Abdel-Aty et al (*10*) have applied this nested logit structure to investigate the effect of lead vehicle's size on the rear-end crash configuration. They have calibrated different logit nests to estimate the probabilities of four rear-end crash configurations as a function of driver age and gender, vehicle type and maneuver, light condition, visibility of the driver and speed.

In another attempt by Ulfarsson et al (*11*) has applied the nested structure using multivariate multinomial logit models in modeling the effect of gender of the occupant on the severity of injuries they suffered in SUV, minivan, Pickup and passenger car crashes.

As many influential factors in highway crashes are categorical or dichotomous variables many researchers have employed categorical data analysis approaches in their studies. A logistic regression modeling approach has been applied by Dissanayake et al (*12*) to investigate influential factors towards older drivers in highway crashes. All the four types of influential factors, driver, environmental, vehicular and highway related have been used in their attempt to model the injury severity. This logistic regression method has been applied by many researchers, Farmer et al (*13*), Krull et al (*14*) in their attempts in identifying critical factors towards crash severity in different kinds of highway crashes.

Kim et al (*15*) have applied log-linear models in their attempt to investigate the contribution of personal and behavioral factors towards injury severity in automobile crashes. Again, they have applied this method to study the effect of age, sex and vehicle type towards the driver being fault for the crash (*16*). Abdel-Aty et al (*17*) also have applied the log-linear method in their study to reveal the effect of driver age on crash involvement. However, this method is less applicable in a situation where there is large number of explanatory variables (influential factors) in consideration due to the sophistication of interpreting the outcomes.

A negative binomial modeling approach has been applied by Shankar et al (*18*) to study the effect of roadway geometrics (horizontal and vertical alignments) and environmental factors such as weather and other seasonal effects. Miaou (*19*) has considered three modeling structures to evaluate the performance of Poisson and negative binomial regression models in studying the relationship between truck accidents and roadway geometric design.

CRASH DATA AND VARIABLE SELECTION

The KARS (Kansas Accident Reporting System) database, from which the data used in this study was extracted, consisted of data for highway crashes occurred in all public roadways in Kansas and reported by the police officers. The data coverage was 10 years (1993 to 2002) and a preliminary analysis of those data was conducted to select the necessary data sample to be used in the modeling. After reviewing the preliminary analysis results of crash data, it was decided to use the crash data for last four years, 1999 to 2002 for both urban and rural crashes. The separation of the crashes in to urban and rural was based on the functional class of the roadway on which the crash had happened. The changes and improvements that have been made to the database and the highway system during this period were also considered in the selection of data sample. Each crash record contained driver, vehicular, roadway, environmental related details and some other crash related details like crash type, time of occurrence, emergency response details, etc.

Every crash has been categorized in to five injury severity levels, namely, fatal, disabling/incapacitating, non-incapacitating, possible, and property damage only (no injury). The severity of a crash was assigned according to the highest injury severity sustained by an involved occupant in the crash. For instance, if there is at least one fatality resulting from a crash, then it was defined as a fatal crash and when there is at least one incapacitating injury but no fatal injuries then it was classified as an incapacitating injury. Likewise this process continues and when there are no evident injuries reported it was fallen into no injury category.

In the data extraction process, the crash records related to more than two vehicles, pedestrians and trains were discarded from the selected data set. The reason is, nature of these crashes is different from others and their sample sizes are much smaller compared to other types. After discarding missing data fields, the final data set comprised altogether of 70,384 records for rural crashes and 121,261 records for urban crashes. A random sample was selected from each of these data sets to be used for model calibration and finally the sample sizes which were used in the modeling contained 70,384 records for rural crashes and 89,785 records for urban crashes. The intention of using large data samples was to minimize any biases result from inadequate data records in each severity category. On the other hand, large sample size would minimize errors caused by any assumptions made in the modeling process. For instance, the normality assumption for the error distribution assumed in this study could be considered as reliable since the sample size is large. Table 1 and Table 2 shows some important characteristics of rural and urban crash data utilized in the modeling process.

The factor selection process was based on both prior knowledge from previous studies and on the presumption that a particular factor would be significant towards the crash severity. Thus, the selected candidate vector of explanatory variables comprised of many factors some of which may or may not be critical in assessing the crash severity. The selected factors were categorized into driver-related, environmental-related, highway-related, vehicular-related and crash-related related factors such as emergency response time, time of the crash, crash type, etc. The selected factors and their representation in the model are shown in Table 3.

It should be noted that selection of some factors, which were believed to be important, was restricted by inadequate availability of data in the database. One such variable was the estimated speed of the vehicle at the time of crash. Many studies have revealed (*4,7,10,11*) this is as a very critical factor in highway crashes not only as a single contributing factor but also it makes substantial influence on other variables as well when considering the interaction effects. Thus, due to the lack of data on travel speed it was decided to consider the posted speed limit at the location of the crash instead of the travel speed of the vehicle.

METHODOLOGY

When a variable can be ranked or ordered but the difference between two levels are unknown that variable is called an ordinal variable. The response variable in this study, crash severity, can also be ordered as fatal, disabling/incapacitating, non-incapacitating, possible and no injury (PDO) and thus it can be considered as an ordinal response variable. Long (*20*) has discussed the applicability of ordered logit and probit models in analyzing this type of data. These ordered choice models are capable of capturing the qualitative difference between two ranked levels, in this case, between two injury categories (*12*).

The derivation of the ordered model is based on the measurement model (*20*),

$$y_i = m \quad \text{if } \tau_{m-1} \leq y^* < \tau_m \text{ for m = 1 to J} \tag{1}$$

where y^* is the injury risk, which is an unobserved continuous variable called latent variable ranging from $-\infty$ to ∞, which is mapped to an observed variable *y.* The τ values are called thresholds or cut off points and the extreme categories at m=1 and m= J are defined by openended intervals with $\tau_0 = -\infty$ and $\tau_J = \infty$. According to the measurement model the variable y is thought of as providing incomplete information about an underlying y*.

Then the structural model can be considered as,

$$y^* = x_i\beta + \varepsilon_i \tag{2}$$

where x_i is a row of a vector of explanatory variables with a 1 in the first column for the intercept and the i th observation for x_k in the k+1 column. β is a vector of parameters to be estimated and ε_i is the error term which is assumed to be normally distributed. However, the KARS database does not comprise any information on injury risk, y^* as it is unobserved, but it includes details on the variable y which is observed at different levels of y^* at which, y =1 if there are no evident injuries, y =2 if the crash results only possible injuries, y=3 when the crash is non-incapacitating, y = 4 if it is a incapacitating crash and y=5 when crash is fatal. Thus, the measurement model (1) can be illustrated as,

$$y_i \begin{cases} 1 \text{ (No injury)} & \text{if } \tau_0 = -\infty \le y^* < \tau_1 \\ 2 \text{ (Possible)} & \text{if } \tau_1 \le y^* < \tau_2 \\ 3 \text{ (Non-incapacitating)} & \text{if } \tau_2 \le y^* < \tau_3 \\ 4 \text{ (Incapacitating)} & \text{if } \tau_3 \le y^* < \tau 4 \\ 5 \text{ (Fatal)} & \text{if } \tau_4 \le y^* < \tau_5 = \infty \end{cases} \tag{3}$$

where the threshold values τ_1, τ_2, τ_3 and τ_4 are parameters to be estimated. According to the measurement model the probability that the i^{th} victim of crash, suffer injury severity level of m (m = 1 to 5) is the probability that the injury propensity y^* takes a value between two cut off points. That is,

$$\Pr(y_i = m|x_i) = F(\tau_m - x_i\beta) - F(\tau_{m-1} - x_i\beta) \tag{4}$$

where F(x) is the cumulative distribution function of the unobserved error term ε_i evaluated at given x under the assumption that ε_i s are normally distributed with mean zero and constant variance as mentioned previously. For example, the probability that the victim i sustain a fatal injury due to the crash is,

$$\Pr(y_i = 1 | x_i) = 1 - F(\tau_4 - x_i\beta) \tag{5}$$

It should be noted that to these probabilities be positive the thresholds values should satisfy the order, $\tau_1 < \tau_2 < \tau_3 < \tau_4$ *(21)*.

The estimation of these model parameters can be carried out through the method of maximum likelihood. The log likelihood, which is the logarithm of the likelihood function, can be written as,

$$\ln L(\beta, \tau | y, X) = \sum_{j=1}^{J} \sum_{y_i = j} \ln[F(\tau_j - x_i\beta) - F(\tau_{j-1} - x_i\beta)] \tag{6}$$

where β is the vector of parameters from the structural model, first column consisting of the intercept and τ is the vector of threshold parameters. The procedure consists of maximizing this equation using numerical methods. To make the model estimable either one threshold value, possibly τ_1 or the intercept is constrained to be some arbitrary value usually zero. The software used in this analysis assume the intercept $\beta_0=0$ and estimate the other parameters. For more details on parameter estimation of ordered models using maximum likelihood procedure reader is directed to *Regression Models for Categorical and Limited Dependent Variables* (*20*).

The R^2 value which is called Generalized Coefficient of Determination is depicted as,

$$R^2 = 1 - \left\{ \frac{L(0)}{L(\hat{\beta})} \right\}^{\frac{2}{n}} \tag{8}$$

and $$R^2_{\max} = 1 - \{L(0)\}^{[2/n]}$$

where $L(0)$ is the likelihood of the model which includes only intercept terms, $L(\hat{\beta})$ is the likelihood of the specified model with all the significant factors, and n is the sample size (*22*). However, according to Nagelkerke (*22*) this R^2 value achieves its maximum when it is equal to 0.75 for models with dichotomous variables, which is the case in this study, which contradicts with the original definition of the coefficient of determination that it should be in the range of 1 and 0. Thus, he proposed an adjusted value for R^2, called $\overline{R}^2$, which is defined as,

$$\overline{R}^2 = \frac{R^2}{R^2_{\max}} \tag{9}$$

which has the maximum and minimum values of 0 and 1 respectively.

All the model estimations in this study were carried out by using SAS (Version 8) software.

MODEL ESTIMATION

Since the number of explanatory variables which were considered in this study was large, a parameter selection method had to be applied to minimize the higher demand of time and resources as, when the number of variables is high the time and resources (computer memory) needed for model estimation are also high. The stepwise selection method, a built-in facility in the software (SAS) for this purpose, was used in this analysis (*23*). In this method, the model starts with no explanatory variables but only the intercept terms in the model and adds one variable at a time if the residual chi-square value is significant at a given level of significance (95%). Once a variable is entered in to the model it is tested by backward selection method to make sure it is still significance over the variables already had in the model. In addition to the stepwise selection, the software also provides the capability to do backward selection. In this method, the model starts with all the variables and eliminates one variable at a time based on the significance of the residual chi-square value. Both these methods were applied in the model parameter estimation procedure and provided the same results.

The inclusion of the response time in the model was a concern at the initial stage of the model estimation process. This is because, initially, emergency response time was introduced to the model as a continuous variable and parameter was estimated. However, the estimated parameter relevant to response time was found to be not explaining its effect correctly towards the crash severity in both rural and urban crashes. Preliminary analysis of crash data revealed that in 95% of all type of crashes, the emergency services had responded within one hour in rural areas while in almost all the crashes (99%) in urban areas have been covered within an hour. However, there were some cases that the response time was more than even 20 hrs, but all of them were property damage only crashes. This situation may lead to some unreliable predictions. Thus it was decided to apply this variable as a dichotomous variable into the model to obtain a better explanation of its effect towards the crash severity. Several modeling efforts were carried out using different categories of the response time and eventually came up with the category, as depicted in the Table 3, which was found to be the best way of explaining it.

MODEL FITTING INFORMATION

The estimated adjusted R^2 value for the model for rural crashes is found to be 0.38 and for urban crashes it is little bit lower, which is 0.16 from Table 3. Thus the rural crash model is more capable of predicting the crash severity with the selected explanatory variables. Even though there is no generally accepted rule for testing the accuracy of ordered multiple-choice models (*10*), it is very important to check the prediction accuracy of the developed model. The software used in this study (SAS) produces predicted probabilities for each observation using the fitted model (*23*). For instance, it gives the probability of an observation being a fatal, disabled, etc. and finally the predicted overall severity of the observation based on the largest individual probability of each severity group. These predicted probabilities were obtained for each data sample for urban and rural crashes, which had been separated from the original data set, using the fitted model. It was found that, in rural crashes, about 75% of observations were predicted correctly and about 73% of urban crashes were correctly predicted. It should be noted that theses are the overall prediction rates and the prediction accuracy of different severity levels may vary.

MODEL RESULTS

Estimated model coefficients using maximum likelihood method for the ordered probit model for urban and rural crashes are shown in Table 4. As the parameter estimation in ordered models assumes the injury risk and explanatory variables are linearly related (equation 2), the interpretation of the parameters should be accordingly. That is, a variable with positive parameter indicates that the effect of relevant variable has an increasing tendency towards the crash severity. A variable with a negative parameter has decreasing effects towards the crash severity. It can be seen from the results that most of the parameters have similar effects towards the severity of the crash for both urban and rural crashes except for few factors. For instance, almost all the driver attributes which were considered in this analysis have similar effects towards the crash severity except the variable DR_RESTRICT. In the case of drivers comply with restrictions, it causes to resulting in a less sever crashes in urban crashes (as the estimated coefficient is negative) while it is not significant in rural crashes. As far as crash type is concerned, single vehicle crashes are significant in increasing the crash severity in both urban and rural crashes but two vehicle crashes are only significant in urban crashes. However, if a two vehicle crash happens, then the type of collision, head-on, rear-ended, angle, is significant towards increasing the severity of the crash in both urban and rural crashes. The following sections consist of discussion on some important variables that were identified through model outputs as critical towards crash severity.

Driver Related Factors

As many previous studies have revealed (*10, 7, 11)*, the variable 'SPEED' has the tendency of increasing the severity of a crash as it has a positive estimated parameters irrespective of whether urban or rural crash. When, at least one of the drivers involved in the crash does not use seatbelts the risk of having a more severe crash is high. In addition, when the driver is ejected or trapped in the vehicle due to the crash there is a higher probability of resulting in a higher severe crash in both rural and urban highways. When the involved driver is a male the chance of having a high severity crash is less (15, 10, *7*). In other words, female drivers are more likely to be involved in high severity crashes in both rural and urban areas. This may be due to the fact that females are generally not much capable as males of bearing physical or mental trauma resulting in a crash (*10*).

When alcohol or drugs are involved in the crash it is more likely to be ended as a high severity crash in both types of highways as the relevant variable has positive parameters in both of the cases. In KARS database the alcohol involvement has been recorded as whether alcohol presented or alcohol contributed towards the crash based on the judgment made by the police officer. However, in some cases there might not be clear evidence available to make the decision whether alcohol contributed towards the crash or not. According to the Kansas Department of transportation (KDOT), they have revised the rule by introducing the new definition for the alcohol involved crash in 1990 by taking into account both facts, alcohol present or alcohol contributed (24), and thus this new definition was used in this study to define the alcohol involvement in crashes.

Crash Type

Single vehicle crashes are significant over two vehicle crashes and animal-vehicle crashes in increasing the severity of a crash in rural areas. This is proved by having positive parameters for rollover crashes and negative parameters for crashes that occur on the roadway. That is, when the crash occurs off the roadway there is a higher risk for having a severe crash. However, in urban areas, both the single vehicle and multi vehicle crashes are significant but crashes related to animals are non-significant towards the severity of the crash. According to the KARS database, significant amount (over 30%) of all the rural crashes accounts for animal-vehicle crashes but most of the time those result in minor injury crashes while only very few animal-vehicle crashes (less than 5%) occur in urban areas. Thus the model results seem to be in agreement with these facts as it gives negative coefficient estimate for that variable in rural crashes but non-significant in urban crashes.

Roadway Factors

Irrespective of the crash occurrence area, the variable related with the roadway geometry (RDCUR_GRAD) results in a positive parameter. This implies the fact that when the roadway is not leveled and straight it is more likely to be resulting in a high severity crash. When a crash occurs on an urban or rural interstate or local road the probability of having a more severe injury is less, compared to arterials and collectors. This may be due to the fact that, when people drive in local roads they might be more careful and also there might be lesser vehicular interactions due to the low traffic volumes on those highways. On interstates, the decreasing trend in having more severe injuries may be due to high safety attributes available on those highways almost uniform travel speed conditions.

Environmental Factors

When the crash occurs on a wet road surface, which indeed has less skid resistance, it seems to be ended with a lesser severe crash in both urban and rural roadways as the variable related to the road surface condition gives a negative parameter. This may be due to the fact that drivers are more cautious under severe weather conditions and try to maintain lower driving speeds under these conditions. On the other hand, when the crash occurs under dark or unlit conditions in urban areas, the severity of the crash is going to be higher. However, this variable, LIGHT_CON, is non-significant in rural areas.

Vehicular Factors

When the vehicle maneuver before the crash is straight and following the roadway, the propensity of hgaving a more severe crash is high in both urban and rural crashes. The consideration of this factor was in comparison with other types of vehicle maneuvers such as, right turning, left turning, lane changing, etc. This is in consistency with finding of this study as when two vehicles collide head-on the severity of the crash is increased. In addition, the vehicular faults have the tendency in resulting higher severe crashes in urban roadways. The vehicular faults include faults in tires, wheels, brakes, and windshield.

According to parameter estimations, when the vehicle is registered in the state of Kansas, chance of having more severe crashes is less in rural areas but in urban areas it is going to increase the severity of the crash. The selection of this variable was based on the intention of assessing the effect of the driver familiarity with the surrounding.

Emergency Response Time

When the emergency response time is less than 5 minutes, the tendency of having a more severe crash is less compared to longer response times in both urban and rural areas. However, it should be noted that, there is no hard-and-fast rule in defining this threshold value of 5 minutes. In fact, even though this cut-off value was come up with the data used in this study, there is a possibility of having another threshold value under different circumstances. Therefore a more general interpretation, saying that - longer the emergency response time is higher the probability of having a more severe crash - would be more appropriate.

CONCLUSIONS

Having the intention of identifying contributing factors, which contributes towards higher severe crashes, in urban and rural highway crashes, an ordered probit model was developed. Different types of contributing factors, driver-related, environmental-related, roadway-related, vehicular-related, and crash-related factors, were considered in the study. The results of the study show that the effects of most of the contributory factors towards the injury severity are similar in both rural and urban crashes. Factors such as alcohol or drug involvement, posted speed limit, driver being in fault for the crash, seat belt violation, driver being ejected or trapped, roadway geometry with curve and grades seem to increase the severity of both rural and urban highway crashes. In addition, single vehicle crashes are more critical towards increasing the severity over two vehicle and animal vehicle crashes in rural areas while both two vehicle and single vehicle crashes are significant in urban areas. If the crash is a collision of two vehicles, then the type of collision, head-on, angle or rear ended, are having tendency of increasing the severity irrespective of the area of occurrence.

One of the important findings in this study is that higher injury risk for drivers who do not use safety equipments at the time of crash irrespective of whether the crash happens on a rural or urban highway. Since Kansas has secondary seat belt law, this finding might highlight the need of considering legislation of a primary seatbelt law. On the other hand, when the driver is ejected or trapped due to the crash, it is more likely to be ended up as a high severity crash. It is important to note that, when the driver does not wear any seat belt the probability of ejecting due to the crash might be high and thus the driver is in a more vulnerable situation. The data used in this analysis were based on police reports and thus the accuracy of the findings is subjected to the accuracy of the data used. Especially in the case of seat belt usage, the accuracy of data is questionable since not everybody may accept the truth and in many situations driver might be already out of the vehicle when police officers arrive the scene.

REFERENCES

1. Traffic Safety Facts 2002, National Highway Traffic Safety Administration, US Department of Transportation, 2002.
2. Blincoe, L., A. Seay, E. Zaloshnja, T. Miller, E. Romano, S. Luchter, and R.Spicer. The Economic Impact of Motor Vehicle Crashes, 2000, Publication DOT HS 809 446, National Highway Traffic Safety Administration, U.S. Department of Transportation, 2002
3. Office of Highway Policy Information, Federal Highway Administration, US Department of Transportation ,http://www.fhwa.dot.gov/policy/ohim/hs02/re.htm, Accessed July 2003.
4. O'Donnell, C.J., and D.H. Conner. Predicting the Severity of Motor Vehicle Accident Injuries Using Models of Ordered Multiple Choice, Accident *Analysis and Prevention*, Vol.28, No. 6, 1996, pp 739-753.
5. Khattak, A.J., M. D. Pawlovich, R.R. Souleyrette, and S.L. Hallmark. Factors Related to More Severe Older Driver Traffic Crash Injuries, *Journal of Transportation Engineering*, Vol. 128, No. 3, 2002, pp 243-249.
6. Khattak, A. J., J.R. Schneider, and F. Targa. Risk factors in Large Truck Rollovers and Injury Severity: Analysis of Single-Vehicle Collisions, *In Transportation Research Record*, CD-ROM-Transportation Research Board, 82nd Annual Meeting, Washington, D.C., 2003.
7. Kockelman, K.M., and Y. Kweon. Driver Injury Severity: An Application of Ordered Probit Models, *Accident Analysis and Prevention*, Vol.34, No.3, 2002 pp 313-321.
8. Ma, J., and K.M. Kockelman. Anticipating Injury & Death: Controlling for new Variables on Southern California Highways, 83rd Annual Meeting of the Transportation Research Board, January 2004, Washington, D.C.

9. Shanker, V., F. Mnnering, and W. Barfield. Effects of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies, *Accident Analysis and Prevention*, Vol.27, No. 3, 1995, pp 371-389.
10. Abdel-Aty, M., and H. Abdelwahab. Modeling Rear-end Collisions Including the Role of Driver's Visibility and Light Truck Vehicles Using Nested Logit Structure, *Accident Analysis and Prevention*, Vol.36, No. 3, 2004, pp 447-456.
11. Ulfarsson, G. F., and F.L. Mannering. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents, *Accident Analysis and Prevention*, No.36, No. 2, 2004, pp 135-147.
12. Dissanayake, S., and J.J. Lu. Factors Influential in Making an Injury Severity Difference to Older Drivers Involved in Highway Crashes, *Accident Analysis and Prevention*, Vol. 34, No. 5, 2002, pp 609-618.
13. Farmer, C. M., and A. K. Lund. Rollover Risk of Cars and Light Trucks after Accounting for Driver and Environmental Factors, *Accident Analysis and Prevention*, Vol.34, No.2, 2002, pp 163-173.
14. Krull K. A., A. J. Khattak, and F. M. Council. Injury Effects of Rollovers and Events Sequence in Single-Vehicle Crashes, *In Transportation Research Record*, CD-ROM-Transportation Research Board, 79th Annual Meeting, Washington, D.C., 2003.
15. Kim, K., L. Nitz, R. James, and L. Li. Personal and Behavioral Predictors of Automobile Crash and Injury Severity, *Accident Analysis and Prevention*, Vol.34, No. 5, 1995, pp 469-481.
16. Kim, K., L. Li, R. James, and N. Lawrence. Drivers at Fault; Influence of Age, Sex, and Vehicle Type, *Journal of Safety Research*, Vol.29, No. 3, 1998, pp 171-179.
17. Abdel-Aty, M, C.L. Chen, and J.R. Schott. An Assessment of the Effect of Driver Age on Traffic Accident Involvement Using Log-linear Models, *Accident Analysis and Prevention*,Vol.30, No. 6, 1998, pp 851-861.
18. Shanker, V., F. Mnnering, and W. Barfield. Statistical Analysis of Accident Severity on Rural Freeways, *Accident Analysis and Prevention*, Vol.28, No. 3, 1996, pp 391-401.
19. Miaou, S. The Relationship between Truck Accidents and Geometric Design of the Road Sections: Poisson Versus Negative Binomial Regression, *Accident Analysis and Prevention*, Vol.26, No. 4, pp 471-482, 1994.
20. Long, J.S., *Regression Models for Categorical and Limited Dependent Variables,* SAGE Publications, Inc., California, 1997.
21. Greene, W.H. *Econometric Analysis*, 3rd Edition, Prentice Hall, Upper Saddle River, N.J., 1997.
22. Nagelkerke, N.J.D. A Note on a General Definition of the Coefficient of Determination, Biometrika, Vol.78, pp 691-692, 1991.
23. SAS Online Doc: Version 8, SAS Institute Inc. Cary, NC, 1998.
24. *Alcohol Involvement in Motor Vehicle Accidents 1993-2002*, Kansas Department of Transportation, 2003.

LIST OF TABLES AND FIGURES

TABLE 1 Characteristics of Urban Crash Data Used for Modeling

TABLE 2 Characteristics of Rural Crash Data Used for Modeling

TABLE 3 Explanatory Variables Considered in the Study

TABLE 4 Maximum Likelihood Estimations of Parameters

TABLE 1 Characteristics of Urban Crash Data Used for Modeling

Factor		Crash Severity					Total	% *
		Fatal	Incapaci-tating	Non-Incapaci-tating	Possible	No Injury		
Alcohol or drug Involvement	Involved	18	78	471	249	993	1809	2.01
	Not Involved	137	1223	8813	11987	65816	87976	97.99
Light Condition	Day Light	75	888	6576	9453	51000	67992	75.73
	Dark/ Unlit	80	413	2708	2783	15809	21793	24.27
Road Surface Condition	Dry Road	133	1108	7684	9922	53364	72211	80.43
	Wet road	22	193	1600	2314	13445	17574	19.57
Roadway Geometry	Level and straight	87	859	6800	9356	50188	67290	74.95
	curve/grade	68	442	2484	2880	16621	22495	25.05
Road surface type	Black top	96	837	6164	8220	47036	62353	69.45
	Other	59	464	3120	4016	19773	27432	30.55
Utility Zones	Construction /maintenance	4	32	166	230	1345	1777	1.98
	Not a utility zone	151	1269	9118	12006	65464	88008	98.02
Functional Class	Interstate	47	300	1391	1614	9282	12634	14.07
	Arterial	76	724	5619	8015	37703	52137	58.07
	Collector	13	103	727	919	5561	7323	8.16
	Local	19	174	1547	1688	14263	17691	19.70
Crash Type	Single Vehicle	69	358	1930	1260	7414	11031	12.29
	Two Vehicle	85	937	7306	10927	57302	76557	85.27
	Animal-Vehicle	1	6	48	49	2093	2197	2.45
	Rollover crash	11	75	296	112	250	744	0.83
	Head-on	18	73	287	230	706	1314	1.46
	Rear-end	8	196	1962	4683	19334	26183	29.16
	Angle	52	603	4511	5351	25356	35873	39.95
Response Time (min)	<5	79	619	4384	5277	28378	38737	43.14
	5_15	55	578	4093	5769	31755	42250	47.06
	15-60	20	97	740	1098	5991	7946	8.85
	>60	1	7	67	92	685	852	0.95

** Percentage based on total number of 89,785 crashes*

TABLE 2 Characteristics of Rural Crash Data Used for Modeling

Factor		Crash Severity					Total	%*
		Fatal	Incapacitating	Non-Incapacitating	Possible	No Injury		
Alcohol or drug Involvement	Involved	170	317	688	360	889	2424	95.68
	Not Involved	458	1358	5512	4593	41719	53640	4.32
Light Condition	Day Light	326	1043	3685	3018	17456	25528	45.53
	Dark/ Unlit	302	632	2515	1935	25152	30536	54.47
Road Surafce Condition	Dry Road	557	1422	5035	3947	34932	45893	81.86
	Wet road	71	253	1165	1006	7676	10171	18.14
Roadway Geometry	Level/straight	348	957	3719	3119	28817	36960	65.92
	curve/grade	280	718	2481	1834	13791	19104	34.08
Road surface type	Black top	457	1206	4096	3249	31003	40011	71.37
	Other	171	469	2104	1704	11605	16053	28.63
Utility Zones	Construction/ maintenance	16	44	166	113	810	1149	2.05
	Not a utility zone	612	1631	6034	4840	41798	54915	97.95
Functional Class	Interstate	44	253	697	477	4180	5651	10.08
	Arterial	318	720	2141	1777	16485	21441	38.24
	Collector	173	421	1892	1441	12704	16631	29.66
	Local	93	281	1470	1258	9239	12341	22.01
Crash Type	Single Vehicle	319	1003	3798	2663	10829	18612	33.20
	Two Vehicle	304	627	2041	1821	10649	15442	27.54
	Animal-Vehicle	5	45	361	468	21128	22007	39.25
	Rollover crash	150	345	1111	752	1645	4003	7.14
	Head-on	91	107	150	82	200	630	1.12
	Rear-end	29	122	540	657	2787	4135	7.38
	Angle	160	346	1049	794	3716	6065	10.82
Response Time (min)	<5	82	337	1329	1203	10511	13462	24.01
	5_15	282	842	2987	2141	14133	20385	36.36
	15-60	246	471	1752	1515	15756	19740	35.21
	>60	18	25	132	94	2208	2477	4.42

** Percentage based on total number of 56,064 crashes*

TABLE 3 Explanatory Variables Considered in the Study

Variable	Mean		Description
	Urban	Rural	
ACC_TIME	-	-	Time of the crash in 24 hr clock
ALCOHOL	0.02	0.04	=1 if alcohol or drug involved, =0 otherwise
ANGLE_CR	0.40	0.11	=1 if two vehicles collide angle, =0 otherwise
ANM_VEH_CR	0.02	0.39	=1 if an animal-vehicle crash, =0 otherwise
ARTERIAL	0.58	0.38	=1 if occur on an arterial, =0 otherwise
BLACK_RD_TOP	0.69	0.71	=1 if occur on a black road surface, =0 otherwise
COLLECTOR	0.08	0.30	=1 if occur on a collector, =0 otherwise
DR_AT_FLT	0.87	0.43	=1 if the driver is at fault for the crash, =0 otherwise
DR_EJECT	0.01	0.03	=1 if at least one driver ejected due to the crash, =0 otherwise
DR_LICENSED	0.94	0.97	=1 if driver has a valid license, =0 otherwise
DR_MALE	0.57	0.62	=1 if the driver (both drivers in two-vehicle crashes) is male, =0 otherwise
DR_NO_STBLT	0.13	0.16	=1 if at least one driver not used safety equipments, =0 otherwise
DR_OLD	0.21	0.33	=1 if driver age (both drivers in two-vehicle crashes) is >55, =0 otherwise
DR_RESTRICT	0.48	0.45	=1 if at least one driver complied with restrictions =0, otherwise
DR_YOUNG	0.39	0.44	=1 if driver age (both drivers in two-vehicle crashes) is <25, =0 otherwise
HDON_CR	0.01	0.01	=1 if a head-on crash, =0 otherwise
INTERSTATE	0.14	0.10	=1 if occur on an interstate, =0 otherwise
INTR_SECN	0.49	0.17	=1 if occur at an intersection, =0 otherwise
LIGHT_CON	0.24	0.54	=1 if crash happens in dark or unlit conditions, =0 otherwise
LOCAL	0.20	0.22	=1 if occur on a local road, =0 otherwise
ON_RDWAY	0.64	0.21	=1 if occur on the roadway, =0 otherwise
PKTIME	0.11	0.11	=1 if occur during 6:45 to 9:00 am, =0 otherwise
RD_CNT_MNT	0.02	0.02	=1 if occur at a construction or maintenance zone, =0 otherwise
RD_CUR_GRAD	0.25	0.34	=1 if roadway is not straight and level, =0 otherwise
REAR_END_CR	0.29	0.07	=1 if a rear-end crash =0 otherwise
RES_TIME	11.90	27.20	Emergency response time in minutes
RES_TIME_BINARY	0.54	0.29	=1 if response time <= 5 minutes =0 otherwise
ROLLOVER_CRH	0.01	0.07	=1 if a rollover crash =0 otherwise
SNG_VEH_CR	0.12	0.33	=1 if a single vehicle crash, =0 otherwise
SPEED	37.35	55.13	Speed limit in mph*
TWO_VEH_CR	0.85	0.28	=1 if a two-vehicle crash, =0 otherwise
VEH_ATFLT	0.02	0.02	=1 if at least one vehicle is at fault for the crash, =0 otherwise
VEH_AUTMBLE	0.96	0.94	=1 if at least one vehicle is an automobile, =0 otherwise
VEH_KS	0.89	0.87	=1 if vehicle (both vehicles in two-vehicle crashes) is registered in Kansas, =0 otherwise
VEH_MNR_STGT	0.60	0.72	=1 if vehicle (both vehicles in two-vehicle crashes) maneuver is straight before crash, =0 otherwise
WEEK_DAY	0.79	0.72	=1 if occur on a weekday, =0 otherwise
WET_RD_SURF	0.20	0.18	=1 if the road surface wet, =0 otherwise

* *1 mph = 1.6 kmph*

TABLE 4 Maximum Likelihood Estimation of Parameters

Factor	Rural		Urban	
	Estimated Parameters	Chi-Square statistic	Estimated Parameters	Chi-Square statistic
ACC_TIME	*	*	*	*
ALCOHOL	0.1712	184.5475	0.1533	110.0138
ANGLE_CR	0.3425	526.7614	0.3038	1161.874
ANM_VEH_CR	-0.3705	514.7419	*	*
ARTERIAL	*	*	0.0347	16.3952
BLACK_RDTOP	*	*	-0.037	53.4571
COLLECTOR	*	*	*	*
DR_AT_FLT	0.1527	396.5721	0.1145	210.5741
DR_EJECT	0.8090	2937.3087	0.862	2079.323
DR_LICENSED	-0.0733	23.5667	-0.0959	125.1182
DR_MALE	-0.0658	104.4052	-0.0533	134.1104
DR_OLD	0.0222	4.5329	0.0288	6.0369
DR_RESTRICT	*	*	-0.0145	9.9305
DR_STBLT	0.2909	1406.4535	0.3197	2702.537
DR_YOUNG	*	*	*	*
HDON_CR	0.6524	668.3326	0.5413	923.6405
INTERSTATE	-0.0785	48.4772	-0.0291	4.9137
INTR_SECN	0.0566	12.9985	0.0459	79.6121
LIGHT_CON	*	*	0.0347	38.4961
LOCAL	-0.0462	30.6673	-0.0366	14.1216
ON_RDWAY	-0.0706	20.7591	*	*
PKTIME	-0.0227	5.2308	*	*
RD_CNT_MNT	*	*	*	*
RD_CUR_GRAD	0.0244	14.1424	0.0232	19.0099
REAR_END_CR	0.2263	213.4816	0.2499	749.0772
RES_TIME_BINARY	-0.0185	6.8321	-0.0211	21.6144
ROLLOVER_CR	0.1595	233.6846	0.2551	136.5125
SNG_VEH_CR	0.2559	352.2112	0.5293	474.0109
SPEED	0.0171	708.8580	0.00951	229.2427
TWO_VEH_CR	*	*	0.2014	64.3719
VEH_ATFLT	*	*	0.085	6.5408
VEH_AUTMBLE	*	*	-0.0478	15.5639
VEH_KS	-0.0436	23.4420	0.0231	9.5006
VEH_MNR_STGT	0.0640	66.1673	0.1271	710.2423
WEEK_DAY	*	*	-0.0232	17.6201
WET_RD_SURF	-0.1284	247.5906	-0.0541	85.8404
τ_1	-2.0359	875.0145	-1.5053	662.597
τ_2	-1.1226	278.3613	-0.4302	68.712
τ_3	-0.0698	1.0802	0.7100	188.6006
τ_4	0.3849	32.8609	1.2967	627.4301
R^2	0.3095		0.1292	
Adjusted R^2	0.3837		0.1620	

* *Variables are not significant*