ANTICIPATING INJURY & DEATH: CONTROLLING FOR NEW VARIABLES ON SOUTHERN CALIFORNIA HIGHWAYS

By

Jianming Ma
Graduate Student Researcher
The University of Texas at Austin
6.9 E. Cockrell Jr. Hall, Austin, TX 78712-1076
Email: mjming@mail.utexas.edu

Kara M. Kockelman
(Corresponding Author)
Clare Boothe Luce Assistant Professor of Civil Engineering
The University of Texas at Austin
6.9 E. Cockrell Jr. Hall, Austin, TX 78712-1076
Phone: 512-471-0210
FAX: 512-475-8744
Email: kkockelm@mail.utexas.edu

Presented at the 83rd Annual Meeting of the Transportation Research Board
January 2004, Washington D.C.

ABSTRACT

This study investigates the relationship between occupant injury and a host of other factors, including traffic and weather conditions present at the time of crash, road design, vehicle type, and occupant characteristics. Crash and traffic-detector data from six Orange County freeways were used in an ordered probit model. Crash outcomes were classified for each occupant, as no injury, non-visible injury, visible injury, severe injury, and fatal injury. Higher design speeds (holding speed limits fixed) and speeding contribute to injury severity, while lighting and pavement surface conditions appear to play no role. Consistent with the literature, sideswipe and rear-end collisions result in less severe injury, relative to other collision types (such as broadside, hit-object, and rollover), while females and older persons are at higher risk of injury. Information on current traffic conditions proved very valuable for injury severity prediction. And a variety of design and policy recommendations can be drawn, to enhance highway design and roadway safety.

Key words

Crash analysis, injury severity, traffic safety, ordered probit model, Southern California

Introduction

Highway traffic crashes result in more loss of human life (as measured in human-years) than any other mode of transport in the U.S. *(1)*. While vehicles and many roadways are being better designed and congestion is slowing crash speeds, traffic crashes are becoming more critical in many ways, particularly in societies that continue to motorize. The present cost of U.S. crashes per year is estimated to be $230.6 billion; per capita this figure is $820 in 2000 *(2)*. These costs do not include the cost of delays imposed on other travelers, which may be of the same order of magnitude. Schrank and

Lomax *(3)* estimate that roughly half (52-58%) of all traffic delays are due to non-recurring events, such as crashes, costing on the order of $1000 per capita per year, particularly in urban areas. So $500 of this may be attributable to crashes, in delays borne by other road users.

Most crashes are reported, and all can be classified by injury to those involved. This work characterizes occupant injuries by five classes *(4)*: no injury, non-visible injury, visible but not severe injury, severe injury, and fatal[1].

In the U.S., highways accommodate 87%[2] of the nation's vehicle miles traveled in 2001 *(5)*, thanks to their high speeds and capacity. Unfortunately, high-speed crashes tend to be much more severe than those on slower roads. The severe or fatal crash is not only of major concern to the public, but also is of concern to policymakers. Therefore, injury severity of occupants on high-speed roadways is the focus of this study.

Relatively complete crash related data sets were collected from six high-speed highways in Southern California via the Highway Safety Information System *(4)*. These data describe the nature of the crash, as well as features of the crash-involved vehicles, occupants, and roadway section where the crash occurred. An ordered probit model was used to investigate the effects of all available factors on the severity of injury (including no injury) to occupants. Based on the findings, certain policy and design enhancements can be suggested, in order to improve the safety of high-speed roadways.

The first section of this paper is a literature review of previous studies. The remaining sections describe the data sets and methods used here, as well as empirical findings. Conclusions, including policy implications and suggestions for further study, are then summarized.

Literature Review

Several efforts have investigated crash severity using multivariate regression modeling *(see, e.g., 6, 7, 8 ,9, 10, 11, 12, 13)*. For example, Duncan et al. *(6)* investigated the effects of some factors (e.g., lighting, road conditions, speed limits, road geometries) on injury severity of occupants involved in two-vehicle rear-end crashes on divided roads using the ordered probit modeling technique. There were five categories in terms of occupant injury severity in their data set from North Carolina: no injury; no visible injury; non-incapacitating; incapacitating; and fatal. Their results implied that darkness, high speed differentials, high speed limits, and grades have contributed greatly to more severe injury crashes. Golob et al. *(7)* examined the injury severity of truck-involved crashes using the data set collected from California, and they found that "hit object" collisions are the most severe crashes, rear-end and sideswipe collisions are the second and third most severe crash types, respectively.

There have been many efforts devoted to studying the effects of speed and variations in speeds on occupant injury severity. The importance of speeds, and variations in speeds across vehicles has gained much attention since the 1960s. Solomon *(14)* and Cirillo *(15)* concluded that vehicles have a higher probability of being involved in crashes if those vehicles travel at speeds above or below the average speed.

Using National Accident Sample System (NASS) data, Joksch *(16)* found that higher speed changes during a crash correspond to higher rates of severe and fatal crash involvement. Lave *(17)* predicted fatality rates for six broad classes of high-speed U.S. roads while controlling for average speeds and speed variation (approximated as the difference between 85th percentile and mean speeds observed during long sample periods). His resulting regression models suggested speed variation as a valuable, positive predictor of crash rates, with little effect due to average speed. Of course, crashes are extremely rare (occurring every 156,000 miles or so of driving *[9]*).

These and many other results are based on data that have been aggregated temporally and/or spatially, often over a year and hundreds, even thousands, of centerline miles. Such aggregation can obscure true relations. Davis *(18)* clearly demonstrated the ecological fallacies that can develop at the aggregate level, using examples of speed-crash rate relationships that hold at the individual level.

Kloeden et al. *(19)* examined the importance of speed for crash involvement risk on high-speed rural highways in Adelaide, Australia. After studying the characteristics of sampled crashes, the authors estimated a 24% reduction in injurious and fatal crashes if no vehicles traveled above the speed limit and a 32% reduction if all speed limits on undivided roads were lowered to 50 mph (the lowest speed limit in their sample).

There also have been several studies using ordered probit models to investigate the crash severity of individual crashes. Kockelman and Kweon *(8)* used such models to investigate driver injury using the U.S. General Estimates System (GES) data set and four severity categories. Their work emphasized the effects of vehicle type, while controlling for a host of other factors (such as crash type, weather, speed, and occupant characteristics). They found that pick-ups and SUVs tend to be less safe for their drivers than passenger cars in single-vehicle collisions; but in two-vehicle collisions, pick-up and SUV drivers are less likely to incur injury while those in other vehicle types are likely to suffer greater injury (by virtue of being crash-involved with an SUV or pick-up).

Quddus et al. *(10)* analyzed motorcycle-crash severity using ordered probit models and nine years of crash reports from Singapore. Motorcyclist injury was classified as slight, serious, or fatal. And motorcycle damage was classified as none, slight, extensive, and total. Those cycles with greater engine capacity and carrying passengers, and those cyclists driving during early morning hours, colliding with stationary objects, and/or not using their cycle's headlamps (during daytime) were likely to be involved in more severe crashes.

This study also examines crash severity, but like O'Donnell and Connor *(13)*, it considers all occupants. And, like most crash-severity work, it examines data from crashes that have already occurred. Certain populations may be at greater risk of crash involvement, even if those resulting crashes are less severe. Thus, for an overall perspective on injury risk, it is important to also understand crash exposure and crash involvement, as Kweon and Kockelman *(9)* have attempted. They observed that young drivers are much more likely to be involved in crashes, for every mile driven, particularly when driving pickups and SUVs (which are much more prone to rollovers). Crash rates of female drivers, while generally slightly lower, are higher than men's in pickups and SUVs. Of course, young drivers and female drivers also drive (and travel by motorized vehicle) less, on average, than their counterparts. So their *overall* crash risks can be quite a bit lower.

Every study suffers from some limitation. For example, previous research on crash severity has not controlled for roadway design and traffic conditions, examining only a subset of severe crash causes. Here, data from the HSIS and loop detectors just upstream of crash sites provide additional explanatory information, permitting new insights and recommendations.

Data Description

The data records are based on a total of 11,045 crashes, involving 34,416 occupants (including drivers) and 23,987 vehicles, recorded in 1998 on highways I-5, SR-22, SR-55, SR-57, SR-91, and I-405 in Southern California *(4)*. After merging this data set with road geometry, driver features, and vehicle information data sets, 15,201 complete crash-involved individual observations remained. This crash data set involved 12,201 vehicles and 5,843 crashes. However, given only (30-second) traffic detector data for the month of January 1998, other crash observations from 1998 had to be removed, leaving 2,348 crash-occupant observations. In merging the HSIS crash data with local loop detector data (for traffic conditions preceding and at time of crash) and with roadway design data, complete records for only 317 crashes remained. These crashes involved 838 occupants and 690 vehicles, and they spanned 190 distinct roadway segments.

The final data set consists of records for each of 838 crash-involved individuals, including both drivers and occupants. Each record details features of the crash, the occupant, the vehicle containing that occupant, traffic conditions (from the nearest upstream traffic detectors[3]) and the relevant roadway section. Table 1 describes each variable, and its mean and standard deviation across the 838-observation data set. The average roadway section is just 0.285 miles long; so each record's geometric details, a novelty in this type of model, are highly site specific. The average detector station lies within 1,668 feet of the crash-reported milepost, and traffic variables are computed based on 10-minute averages of conditions preceding the reported crash times, across all lanes (one way). As a result, the associated traffic details, another novelty of the data set, are highly site- *and* time-of-day specific.

Methodology

The ordered probit model is used to track the order of response in qualitatively defined discrete variables, such as injury severity *(20)*. Such a specification was used here, with five possible injury outcomes per crash-involved individual. These are labeled 1 through 5 and represent no injury, no visible injury (but with occupant complaining of pain), visible but not severe injury, severe injury, and fatal injury (where the occupant dies within 30 days of the collision). Figure 1 offers a histogram of reported injury levels.

Formally, an ordered probit model's specification can be expressed as follows *(21)*:

$$Y_i^* = X_i'\beta + \varepsilon_i$$

where Y_i^* is a latent, continuous measure of injury severity for individual i, X_i' is a vector of measurable explanatory variables, β is a vector of parameters to be estimated, and ε_i is an unobservable error term. The error terms are assumed to be identically and independently distributed, according to a standard normal distribution (with mean zero and variance one).

The observed, discrete severity level variable, Y_i, can be computed using the following formula:

$$Y_i = \begin{cases} 1 & if \quad Y_i^* \le \mu_1 = 0 & \text{No injury} \\ 2 & if \quad \mu_1 < Y_i^* \le \mu_2 & \text{Complaint of pain} \\ 3 & if \quad \mu_2 < Y_i^* \le \mu_3 & \text{Visible injury} \\ 4 & if \quad \mu_3 < Y_i^* \le \mu_4 & \text{Severe injury} \\ 5 & if \quad Y_i^* > \mu_4 & \text{Fatal Injury} \end{cases}$$

where μ_i are three threshold values to be estimated, after normalizing the first threshold's value to zero[4]. μ_j is the threshold between severity level j and $j+1$, where j belongs to the set $\{1, 2, 3, 4\}$. Figure 2 illustrates the relationship between Y_i^* and Y_i.

The probabilities corresponding to each discrete injury severity can be obtained via the following equation:

$$\begin{aligned} P(Y_i = j) &= P(\mu_{j-1} < Y_i^* \le \mu_j) \\ &= P\left(\varepsilon_i \le \mu_j - X_i'\beta\right) - P\left(\varepsilon_i \le \mu_{j-1} - X_i'\beta\right) \\ &= \Phi\left(\mu_j - X_i'\beta\right) - \Phi\left(\mu_{j-1} - X_i'\beta\right) \end{aligned}$$

where $j = 1, 2, 3, 4$, or 5, and $\Phi(\ )$ represents the standard normal cumulative distribution function. For crash severity levels (Y_i) of 1 or 5, extreme thresholds μ_0 and μ_5 apply in this equation. These are negative and positive infinity, respectively, representing the two (limitless) tails of the normal distribution.

The model parameters ($\beta's$ and $\mu's$) were estimated using LimDep's *(22)* maximum likelihood estimation (MLE) routine. Model goodness of fit was assessed using the likelihood ratio index or LRI, also known as a rho-bar squared. *(23)*

$$\text{Adjusted LRI } = 1 - \frac{L(b) - K/2}{L(c)}$$

where $L(b)$ is the log-likelihood value of the model estimated, K is the number of estimated model parameters, and $L(c)$ is the log-likelihood value when only a constant term is used.

Empirical Findings

The results of model estimation are shown in Table 2. These include an initial model, involving all available control variables, and a final model, wherein control variables not exhibiting statistical significance at the 0.2 level have been removed, via a process of step-wise deletion *(21)*. As noted earlier, descriptions of all variables can be found in Table 1.

Variables of every type were found to be informative in the final models.

Females and older persons are estimated to be at greater risk of injury and death than others. There is no significant difference between drivers and other vehicle occupants in terms of the risk involvement in injury and fatal crashes. In practice, seated on the vehicle's left side, drivers are at greater risk in head-on collisions; they also may be the last to duck and cover in a crash situations, as they try to navigate to safety. Females and older persons may have less protective tissue (such as lower-density skeletons), on average, than males.

Rollover collisions result in far more severe injury to vehicle occupants than sideswipe or rear-end collisions. This is largely due to the relative weakness of vehicle roofs and the strong movements that occur during a rollover (particularly if occupants are without their seatbelts). Colliding with a stationary object (which is not another vehicle) also causes a more severe form of crash (than sideswipe or rear-end collisions), perhaps because the crash vehicle generally is off the road, traversing a sloped embankment and/or colliding with an unyielding tree, overpass structure or light pole. Broadside (side) and head-on collisions are also more severe than sideswipe or rear-end collisions, and are highly similar in severity. Though the kinetic energy that must be dissipated when opposing vehicles slam head-on into one another is quite high, cars and trucks are less strong from the side than from the front, with side doors relatively thin and flexible and stationary doorsills rather low (typically lying below an impacting vehicle's bumper).

Speeding has been commonly considered an important factor contributing to injury severity of occupants. The results in the model suggest that occupants involved in speeding related crashes tend to be severely injured. One possible reason is that drivers have less time to react to the sudden changes in road environment due to speeding; the other reason might be that occupants involved in speeding related crashes are greatly impacted by kinetic energy which is dissipated when collisions happen.

The design speed on a road also has an important effect on the injury severity of crashes. As can be seen from Table 2, occupants involved in crashes occurring on roadways with higher design speeds are more likely to be severely injured. People, of course, tend to drive faster on higher-design-speed roadways; so, when they do lose control and/or cannot avoid a crash, the impact is more severe. However, the coefficient associate with *overall traffic* speed (SPEED) preceding the crash suggests that people crashing during periods of higher speeds (holding design speed, density, and all other explanatory variables constant, of course) tend to be involved in slightly less severe injury crashes. This is an interesting result. Perhaps those who travel below the speed of traffic are driving relatively cautiously, knowing/sensing that they are more easily injured.

Inclusion of the traffic data actually is very important to this model. Without actual speed and density variables, the data set can be expanded to all sorts of sites, without loop detector data. However, when this was done here, coefficients on SPEEDING and DESIGNSPD took on counter-intuitive, negative signs. The resulting coefficient estimates were unreasonable, though highly statistically and highly practically significant. It is rather well agreed that higher speeds are associated with more severe crashes, once a crash occurs *(see, e.g., 16, 8)*. There is a slight positive correlation ($\rho = +0.18$) between the DENSITY and (binary) SPEEDING variables, suggesting that under more congested conditions crash-involved drivers tend to be speeding; of course, congested conditions tend to involve slower speeds overall (and thus less injurious outcomes). It appears that this work's control for actual, observed speeds counteracts the bias that such correlation with unobserved information would introduce.

Conclusions

Vehicle crashes result in more death and loss than a great many other factors, combined. Findings from this study can assist transportation policy makers and highway engineers in formulating safety policy and roadway design. Given an estimate of crash counts, crash severity prediction becomes paramount. And these results suggest that severity is linked to not only traveler and crash characteristics, but roadway design, environmental features and traffic conditions.

By combining Highway Safety Information System *(4)* data on crashes, involved vehicles and occupants, weather, and roadway design with traffic detector data, and applying ordered probit models for severity, this research illuminates the influence of more variables than ever before.

Consistent with findings in prior work, by researchers like Kockelman and Kweon *(24)* and O'Donnell and Connor *(13)*, females, older persons, and those in passenger cars are more at risk of reported injury. Also, rear-end collisions are the least severe form of crash, followed by sideswipe collisions.

For the first time, these models quantify the role of various design features. Lighting, pavement surface and weather conditions appear to play no role in injury levels, though they most certainly can affect crash rates. Dense traffic flow helps to reduce the occupants' risk of involvement in severe crashes. These relationships should be kept in mind when evaluating the benefits of more generous roadway design, since severity is a key component of overall traveler risk.

Many variables of interest are not available in the HSIS or detector data sets for these high-speed Southern California roadways. However, they may be very useful for consideration. These include information on horizontal curvature, sight distance, vehicle weight (versus type), occupant health at time of crash, distance to nearest hospital, and other features of crashes and their victims. A longer time series of crashes for a greater array of highway types also would be very useful. More details on crash exposure and frequency would also be invaluable, for holistic assessments of risk, as a function of driver, vehicle, roadway, and environmental features.

Current traffic information turns out to be highly valuable is predicting crash outcomes, in terms of injury severity for occupants. These key variables have not been controlled for in previous studies of injury severity. Further work should be undertaken to assess traffic-oriented models of injury prediction.

Finally, it may be fruitful to consider other model specifications, such as a two-period joint probit model and the Beta-logistic model *(25)*. Crash risk and, more specifically, crash severity are key transportation topics that will remain with us for some time.

Endnotes

[1] In cases where the injury is not visible, occupants must complain of pain arising from the crash. Severe injury describes an "(i)njury which prevents the injured party from walking, driving, or performing activities he/she normally was capable of before the collision." *(26, p. 156)* In a fatal injury, the occupant dies within 30 days of the collision, due to injuries sustained during the collision. The California Highway Patrol (CHP) is responsible for reporting highway crashes. In the corridors studied here, the thresholds for reportable crashes are $500 in property damage or loss of life. Such thresholds can and do differ. For example, some California counties report only tow-away crashes, others report only crashes with property damage over $1,000, and some do not require reports on any non-injury crashes *(4, p. 3)*.

[2] "Highways" refers to interstates, major collectors, minor arterials, minor collectors, other principal arterials in urban and rural areas, and other urban freeways and expressways. Rural local roads and urban local roads are not considered highways in the definition used for VMT calculations.

[3] Most detectors were within 5,000 ft of the crash site's reported milepost. The maximum distance was 10,348 feet.

[4] If the variance of the random error component were not specified, a second threshold's value would require specification. If the model's constant term were to be set to zero, no threshold terms would be specified. Such specifications permit statistical identification of model parameters.

Acknowledgements

The authors thank Thomas F. Golob and Wilfred Recker for generously sharing the crash and loop detector data and the National Cooperative Highway Research Program for funding this research under contract number 17-23. The authors are grateful to Yusuf Mohamedshah of the Federal Highway Administration (FHWA) who provided data sets used in this research, Young-Jun Kweon and Karl Petty for offering useful discussions related to the data and methods of analysis, and Annette Perrone for editorial assistance.

Disclaimer

While the National Cooperative Highway Research Program (NCHRP) sponsored this research project, the opinions expressed here do not necessarily reflect the policy of NCHRP. The results of this work are preliminary and have not yet been approved by the NCHRP 17-23 project panel for publication as an NCHRP report.

References

1. Bureau of Transportation Statistic (BTS). Pocket Guide to Transportation. BTS03-01. 2003. http://www.bts.gov/publications/pocket_guide_to_transportation/2003/pdf/entire.pdf. Accessed July 29, 2003.

2. U.S. DOT-NHTSA *The Economic Impact of Motor Vehicle Crashes 2000*. http://www.nhtsa.dot.gov/people/economic/EconImpact2000/introduction.htm. Accessed July 19, 2003.

3. Schrank, D. and Lomax, T. *The 2002 Urban Mobility Report*. Texas Transportation Institute, The Texas A&M University System. 2002.

4. HSIS. *Guidebook for the California State Data Files.* Volume I: SAS file formats. #FHWA-RD-00-137. 2000.

5. Bureau of Transportation Statistics (BTS). The Intermodal Transportation Database – Annual Vehicle Miles Traveled. 2001. http://transtats.bts.gov/. Accessed July 31, 2003.

6. Duncan, C.S., Khattak, A.J., and Council, F.M. Applying the ordered probit model to injury severity in truck-passenger car rear-end collisions. In *Transportation Research Record* No. 1635, pp. 63-71. 1998.

7. Golob, T.F., Recker, W.W. and Leonard, J.D. An analysis of severity and incident duration of truck-involved freeway accidents. *Accident Analysis and Prevention* 19 (5), pp. 375-395. 1987.

8. Kockelman, K. and Kweon, Y.J. Driver injury severity: An application of ordered probit models. *Accident Analysis and Prevention* 34 (3), pp. 313-321. 2002.

9. Kweon, Y.J. and Kockelman, K.M. Overall injury risk to different drivers: combining exposure, frequency, and severity models. *Accident Analysis and Prevention* 35 (4), pp. 441-450. 2003.

10. Quddus, M.A., Noland, R.B. and Chin, H.C. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. *Journal of Safety Research* 33 (4), pp. 445-462. 2002.

11. Abdel-Aty, M.A., Chen, C.L., and Schott, J.R. An assessment of the effect of driver age on traffic accident involvement using log-linear models. *Accident Analysis and Prevention* 30 (6), pp. 851-861. 1998.

12. Farmer, C.M., Braver, E.R. and Mitter, E.L. Two-vehicle side impact crashes: the relationship of vehicle and crash characteristics to injury severity. *Accident Analysis and Prevention* 29 (3), pp. 399-406. 1997.

13. O'Donnell, C.J. and Connor, D.H. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. *Accident Analysis and Prevention* 28 (6), pp. 739-753. 1996.

14. Solomon, D. *Accidents on Main Rural Highways Related to Speed, Driver, and Vehicle*. Federal Highway Administration, Washington, DC, (Reprinted 1974). 1964.

15. Cirillo, J.A. Interstate system accident research study II, Interim Report II. In *Public Roads* 35 (3), pp. 71-75. 1968.

16. Joksch, H.C. Velocity change and fatality risk in a crash: a rule of thumb. *Accident Analysis and Prevention* 25 (1), pp. 103-104. 1993.

17. Lave, C. Speeding, coordination, and the 55-MPH limit. *The American Economic Review* 75(5), pp. 1160-1164. 1985.

18. Davis, G.A. Is the claim that 'variance kills' an ecological fallacy? *Accident Analysis and Prevention* 34 (3), pp. 343-346. 2002.

19. Kloeden C.N., Ponte G. and McLean A.J. *Traveling Speed and the Risk of Crash Involvement on Rural Roads.* Road Accident Research Unit. Adelaide University. Report # CR 204. 2001.

20. McKelvey, R.D. and Zavoina, W. A statistical model for the analysis of ordinal level dependent variables. *Journal Mathematical Sociology* 4, pp. 103-120. 1975.

21. Greene, W.H. *Econometric Analysis*, *Fourth Edition.* Prentice Hall, New Jersey. 2000.

22. Greene, W.H. *LIMDEP VERSION 7.0 User's Manual, Revised Edition*. Econometric Software, Inc. 1998.

23. Horowitz, J. An evaluation of the usefulness of two standard goodness-of-fit indicators for comparing non-nested random utility models. In *Transportation Research Record* No. 874, pp. 19-25. 1982.

24. Kockelman, K. To LDT or Not to LDT: An assessment of the principal impacts of light-duty trucks. In *Transportation Research Record* No. 1738, pp. 3-10. 2000.

25. Khanal, M. and Recker, W.W. *Models of Dynamic Commuter Behavior Using Longitudinal Data*. Report #: UCI-ITS-WP-95-3. 1995.

26. CHP 1998 Annual Report of Fatal and Injury Motor Vehicle Traffic Collisions. 2002. http://www.chp.ca.gov/html/switrs1998.html. Accessed July 3, 2003.

List of Figures

List of Tables

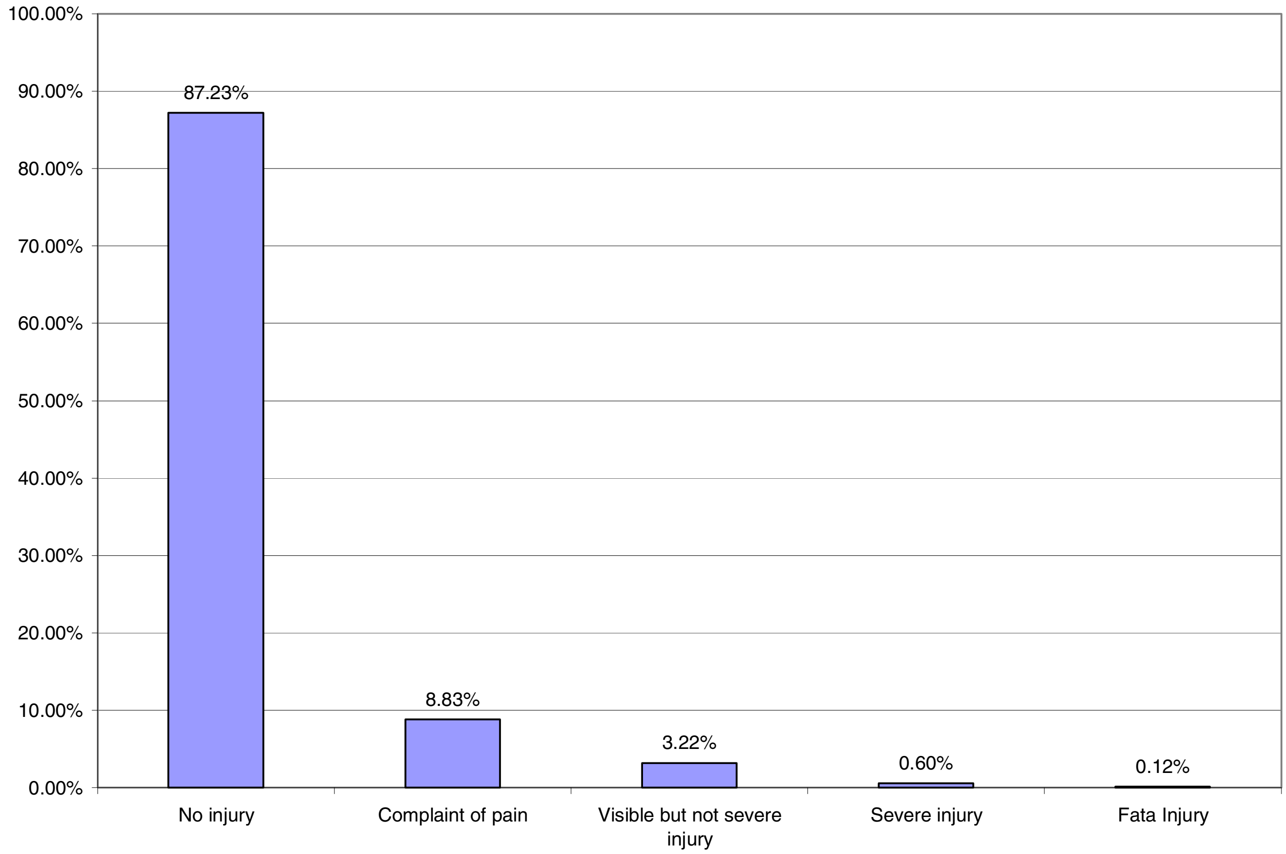

Figure 1. Histogram of Occupant Injury Severity

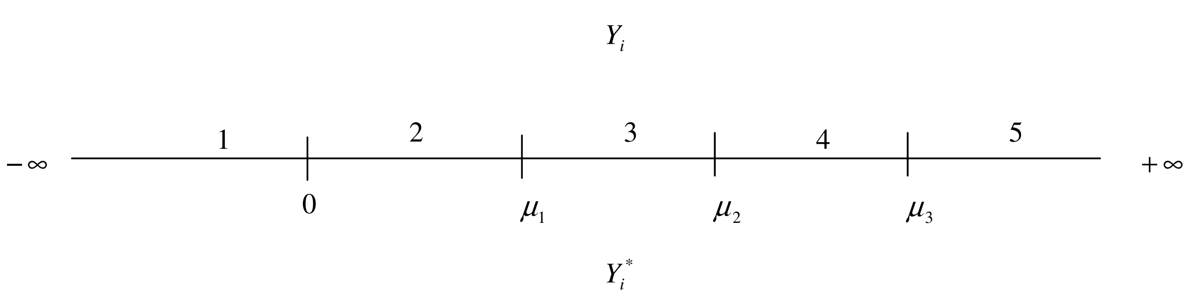

Figure 2. Relationship between Latent and Discrete Injury Severity Variables

Table 1. Description of Variables

Variables	Variable description	N_{obs}	Min	Max	Mean	Std. Dev.
INJSVRTY	Occupant Injury severity: 1 = No injury (87.2% of cases); 2 = Complaint of pain (8.8%); 3 = Visible but not severe injury (3.2%); 4 = Severe injury (0.6%); 5 = Fatal injury (death within 30 days as a result of the crash; 0.1%)	838	1	5	1.1754	0.50943
AGE	Occupant's age (years)	832	0	82	32.9267	16.22008
NUMVEHS	Number of vehicles being involved in crashes	838	1	7	2.5561	1.1033
MEDWID	Median width (ft)	838	6	99	20.9821	18.79769
NO_LANES	Total number of lanes	838	6	15	9.6074	2.62327
PAV_WDL	Left paved shoulder width (ft)	838	0	15	4.9702	3.36548
RSHLDWID	Right shoulder width (ft)	838	0	13	9.5406	1.25682
FEMALE	If the occupant sex is female=1, otherwise=0	838	0	1	0.3604	0.4804
PASSCAR	If the vehicle is a passenger car, passenger car with a trailer, or motorcycle=1, otherwise=0	838	0	1	0.6814	0.46622
PICKUP	If the vehicle is a pickup/panel truck, or pickup/panel truck with a trailer=1, otherwise=0	838	0	1	0.2589	0.43832
TRUCK	If the vehicle is a truck/truck tractor, or truck/truck tractor with trailer=1, otherwise=0	838	0	1	0.0298	0.17023
DISMPEDE	If the vehicle type is coded as dismounted pedestrian=1, otherwise=0	838	0	1	0.0024	0.04882
OTHTYPVH	If the vehicle is a school bus, emergency vehicle, other motor vehicle, other non-motor vehicle, spilled loads, or disengaged tow=1, otherwise=0	838	0	1	0.0274	0.16348
HEADON	If the crash is a head-on collision=1, otherwise=0	838	0	1	0.0155	0.12366
SIDESWIP	If the crash is a sideswipe collision =1, otherwise=0	838	0	1	0.1897	0.39233
REAREND	If the crash is a rear-end collision =1, otherwise=0	838	0	1	0.6134	0.48727
BROADSID	If the crash is a broadside collision=1, otherwise=0	838	0	1	0.0227	0.14895
HITOBJCT	If the crash is a hit-object collision=1, otherwise=0	838	0	1	0.142	0.34926
OVERTURN	If the crash is an overturned collision=1, otherwise=0	838	0	1	0.006	0.07706
OTHCRTYP	If the crash is other type of collision =1, otherwise=0	838	0	1	0.0107	0.10314
ALCOHOL	If the crash is due to influence of a alcohol =1, otherwise=0	838	0	1	0.0453	0.20819
FOLW2CLS	If the crash is because of following too closely=1, otherwise=0	838	0	1	0.006	0.07706
IMPRTURN	If the crash is resulting from improper turn=1, otherwise=0	838	0	1	0.08	0.27138
SPEEDING	If the crash is due to speeding=1, otherwise=0	838	0	1	0.6217	0.48525
OTHVILXN	If the crash is because of other violations (hazardous)=1, otherwise=0	838	0	1	0.247	0.43153
DAYLIGHT	If the crash occurred in daylight=1, otherwise=0	838	0	1	0.6301	0.48307
DUSKDAWN	If the crash occurred in dusk-dawn periods=1, otherwise=0	838	0	1	0.0203	0.14106
DRKSTLGT	If the crash occurred in dark with street lights=1, otherwise=0	838	0	1	0.1623	0.36894
DRKNOLGT	If the crash occurred in dark without street lights=1, otherwise=0	838	0	1	0.1874	0.39043
DRY	If the crash occurred on a dry road=1, otherwise=0	838	0	1	0.7673	0.4228
WET	If the crash occurred on a wet road=1, otherwise=0	838	0	1	0.2327	0.4228
CLEAR	If the crash occurred on a clear day=1, otherwise=0	838	0	1	0.5955	0.49109
CLOUDY	If the crash occurred while cloudy=1, otherwise=0	838	0	1	0.2936	0.45566
RAINING	If the crash occurred while raining=1, otherwise=0	838	0	1	0.1086	0.31131
FOG	If the crash occurred during foggy conditions=1, otherwise=0	838	0	1	0.0024	0.04882
DRIVER	If the occupant is driver=1, otherwise=0	838	0	1	0.8234	0.38157
DESGNSPD	Design speed (mph)	838	60	70	69.5465	2.08186
SPEED	Average speed 10 min before crash, across all lanes (mph)	838	3.22	107.8	54.7215	22.60668
DENSITY	Average density 10 min before crashes (#vehicles per lane per mile = 5280*OCC/24.5, where 24.5 is the average assumed effective length of vehicles)	838	0.01	127.3	29.9369	25.11451
V_CRATIO	Average V/C ratio 10 min before crashes (Defined as VOL/16.67, where 16.67 veh/30 sec. = 2000 vph)	838	0.09	1.49	0.5543	0.20516

Table 2. The Models of Injury Severity Using the Ordered Probit Method

Variables	The Initial Model				The Final Model			
	Coeff.	Std.Err.	Stdd coef	P-value	Coeff.	Std.Err.	Stdd coef	P-value
CONSTANT	-3.82947	3.30812		0.247029	-3.91047	2.79185		0.161313
AGE	0.008414	0.004361	0.267908	0.053697	0.008777	0.003825	0.279467	0.021748
NUMVEHS	-0.04387	0.069185	-0.09501	0.526036				
MEDWID	-0.00438	0.003531	-0.16173	0.214458				
NO_LANES	-0.01128	0.027177	-0.05811	0.677971				
PAV_WDL	0.00458	0.022088	0.03026	0.835717				
RSHLDWID	0.006379	0.055872	0.015738	0.909098				
FEMALE	0.20987	0.141707	0.197911	0.138603	0.201275	0.119751	0.189805	0.092807
PICKUP	-0.154	0.163478	-0.1325	0.346195				
TRUCK	-0.2526	0.5269	-0.08441	0.631646				
DISMPEDE	3.13202	9.36477	0.30015	0.738042				
OTHTYPVH	-0.52502	0.551553	-0.16848	0.341152				
SIDESWIP	-0.80882	0.511925	-0.6229	0.114117	-0.65662	0.189465	-0.50569	0.000529
REAREND	-0.73819	0.508203	-0.70608	0.146348	-0.66537	0.171378	-0.63642	0.000103
BROADSID	-0.06075	0.62674	-0.01776	0.922784				
HITOBJCT	-0.08439	0.574609	-0.05786	0.883235				
OVERTURN	-0.32163	0.83916	-0.04865	0.701512				
OTHCRTYP	-7.08975	1.23E+11	-1.4354	1				
FOLW2CLS	0.547144	0.906867	0.082765	0.546286				
IMPRTURN	0.196131	0.395059	0.104482	0.61957				
SPEEDING	0.446274	0.364208	0.425092	0.220452	0.217268	0.165902	0.206955	0.190326
OTHVILXN	0.37658	0.377009	0.318995	0.317861				
DUSKDAWN	-8.50077	2.17E+13	-2.35384	1				
DRKSTLGT	-0.01381	0.206465	-0.01	0.946679				
DRKNOLGT	0.105961	0.167429	0.081209	0.526816				
WET	0.155529	0.237327	0.129081	0.512251				
CLOUDY	-0.0155	0.174756	-0.01386	0.929334				
RAINING	-0.03192	0.299083	-0.01951	0.91501				
FOG	-10.546	6.32E+13	-1.01065	1				
DRIVER	0.158821	0.166561	0.118959	0.340322				
DESGNSPD	0.053943	0.044656	0.220447	0.22706	0.054812	0.039493	0.223998	0.165162
SPEED	-0.01087	0.006438	-0.48242	0.091297	-0.0106	0.005346	-0.4705	0.047342
DENSITY	-0.01574	0.00697	-0.77597	0.023923	-0.01691	0.005063	-0.83357	0.000839
V_CRATIO	-0.16075	0.444775	-0.06474	0.717787				
Mu(1)	0.710403	0.100593		1.64E-12	0.677443	0.079963		2.89E-15
Mu(2)	1.61245	0.250285		1.18E-10	1.43833	0.15963		2.89E-15
Mu(3)	2.38106	0.378688		3.22E-10	2.03842	0.332623		8.88E-10
LogLik value at constant				-404.5413				-404.5413
LogLik value (full model)				-362.2781				-378.198
Adjusted LRI				0.058741				0.050287
$N_{obs.}$				838				838