

A MODEL OF INJURIES SUSTAINED IN TWO-VEHICLE
COLLISIONS ON URBAN ROADS

P Barnard
Director of Transport
National Farmers' Federation
Sydney
New South Wales

ABSTRACT:

The main objective of this paper is to develop an econometric model of road accident injury severity for two-vehicle collisions using individual accident data. The statistical technique utilised, ordered logit, provides a framework for the development of a casual model of road accident injury severity while recognising the ordinal measurement of injury data. Independent variables are chosen on the basis of an understanding of the physics of road crashes and results obtained from laboratory based biomechanical research. Estimates are provided of the role played by vehicle speeds, vehicle weights, restraint usage, occupant intoxication and socio-economic factors in determining injury severity given that an accident has occurred. The most important of these findings are (i) the probability of sustaining a severe injury is directly related to vehicle speeds and masses, (ii) seat belt usage reduces the chances of sustaining a severe injury by about one-third, and (iii) given an accident of a set severity level vehicle occupants under the influence of alcohol are about two and one-half times more likely to sustain a severe injury than sober occupants.

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

1. INTRODUCTION

Road accident injuries cost the Australian community an estimated \$1950 million per year (based on Atkins 1981). Any measure that can reduce road injuries, however marginally, therefore will yield considerable community savings. It is because of these large potential savings that much effort has been expended in attributing causes to road injuries.

Two distinct types of road injury studies may be identified. One type is based in the laboratory and focuses on the biomechanics of automobile collisions, using animals, anthropomorphic dummies or cadavers in place of live human beings. The laboratory setting enables the forces generated by automobile collisions to be studied under controlled conditions and very accurate measures to be taken. Results from these studies have traditionally formed the basis of road safety legislation. They are, however, expensive to run and are unable to exactly replicate the automobile collision environment with live human beings as subjects. There is a need to ensure that the fundamental relationships unearthed in laboratory based studies can be reproduced in the field.

Field studies collect data on actual road accidents and correlate observable road, vehicle and person characteristics with accident measures, including injury severity. The detail of data collected in these studies varies enormously. The principal data contained in 'mass accident' data records relate to the location of the crash, the registration of the vehicle, damages sustained in the accident, casualties sustained, plus basic person and vehicle data. Although useful in the compilation of general accident statistics these data bases often do not contain vital information necessary for an *understanding* of the forces at work in an accident such as vehicle speeds on impact, vehicle weights, vehicle alignment, road geometry, etc. The collection of detailed data of this type is left to specialised field studies.

Just as the level of detail of data gathered in field studies varies, so does the sophistication of statistical procedures applied to this data. It has not been uncommon to draw policy conclusions from results obtained applying univariate statistical procedures to the data. In the last 10 years, however, there has been a trend to the use of multivariate techniques, especially regression analysis (e.g. Campbell and Filmer 1980, Layton and Weigh 1983, Zlatoper 1984, Garcia-Ferrer and del Hoyo 1987) and log-linear models (as an Australian example, see Kerns and Goldsmith 1984). Recently a number of researchers have argued strongly for the application of latent variable models to road accident data (Barnard 1985, 1989, Chirachavala et al. 1984, Jovanis 1986).

Latent variable models are useful when the phenomena to be studied is in essence categorical or, at the least, is categorically measured. Examples of categorical variables abound at a disaggregate level in accident research; either an accident occurs (1) or it does not (0).

an injury is sustained (1) or is not sustained (0), given an injury is sustained it is of type i or type j , etc. These models predict a probability for each possible discrete outcome. The probabilistic base for these models is conceptually suited to the way we think about road accidents. They are by nature chancey events.

In this paper a latent variable model of injury severity is constructed using a field data set from Adelaide, Australia. The particular latent variable technique utilised is ordered logit. This technique recognizes that there is no means currently available to quantify injury severity on a continuous scale; rather it is only possible to grade injuries into a few ordered categories. The observed dependent variable is therefore discrete or categorical, not continuous.

The categorical nature of injury scales has been ignored by many authors. For instance, Carlson (1979) has justified the use of regression analysis by arguing that the use of continuous variable relationships provide a number of important advantages; in brief,

- (i) coefficients are readily estimable and easily understood,
- (ii) regression models can be estimated with few observations, and
- (iii) the continuous variable relationships established by regression techniques facilitates easy comparison with biomechanical relationships which also tend to be continuous in character.

It will be shown in this paper that all these advantages also apply to the ordered probability model. Other researchers, notably Krishnan and his colleagues (Krishnan 1983, Krishnan et al. 1983), while recognizing the discrete nature of injury severity data have chosen to simplify the problem by collapsing the injury categories into a dichotomy and then estimating the model on aggregate (proportions) data. Use of a disaggregate ordered latent variable model in preference to an aggregate binary latent variable model avoids this inefficient utilisation of data and allows a richer specification of variables.

The remainder of this paper is organized as follows. The next Section contains a mathematical specification for the ordered logit injury severity model. In Section 3 the likely determining factors of injury severity are explored. The data used to estimate the model are described in Section 4 and the model estimation results are discussed in Section 5.

2. MODEL SPECIFICATION

It is convenient to begin exposition of the injury severity model developed here by assuming that there exist only two injury classes; either an injury occurs at or above some specified severity level, denoted by '1', or it does not, denoted by '0'. The injury data available to the researcher are, therefore, observed as a series of 0s and 1s. In reality injury severity does not form a dichotomy but

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

rather varies continuously; it is only for ease of measurement that a limited number of injury classifications (in this case two) are created. Consequently it is natural to envisage a continuous variable underlying the observed dichotomy. It is also conceptually appealing to identify the change in the observed variable value (from 0 to 1) as occurring when the unobserved continuous variable crosses a threshold value.

The process described may be mathematically expressed as:

$$\begin{aligned} y_q &= 0 & \text{if } y_q^* \leq t_q^* \\ y_q &= 1 & \text{otherwise.} \end{aligned} \quad (1)$$

where y_q is the observed injury sustained by crash victim q , y_q^* is an underlying continuous variable measuring the severity of the crash and t_q^* is the threshold level above which an injury is observed. The threshold value is allowed to vary across individuals, reflecting different capacities to tolerate automobile collision forces without sustaining an (observed) injury.

Suppose that y_q^* and t_q^* could be observed; then a regression relationship for y_q^* could be specified as:

$$y_q^* = Z_q \beta + \epsilon_q \quad (2)$$

where Z_q is a row vector of independent variables, β is a parameter vector and ϵ_q is an error term. Likewise a regression relationship could be established for the maximum tolerance capacities:

$$t_q^* = X_q \delta + u_q \quad (3)$$

where X_q , δ and u_q are analogous to Z_q , β and ϵ_q . Z_q should contain determinants of crash severity, measured in terms of the forces acting on the bodies of vehicle occupants. Included in this vector would be variables such as change in vehicle velocity on impact, vehicle size, and occupant seating position and restraint usage. Similarly X_q should contain determinants of the ability of an occupant's body to withstand collision forces without injury, an example being the age of the occupant (Viano et al. 1978). By comparing the expected values of y_q^* and t_q^* obtained from regression analyses applied to equations (2) and (3), the expected number of injured (1s) and non-injured (0s) car accident victims could be calculated.

However, neither y_q^* nor t_q^* can be observed and consequently equations (2) and (3) cannot be estimated directly using standard regression techniques. Nevertheless by combining (1), (2) and (3) it is possible to make a statement of the probability of an injury being observed. This probability can be expressed as:

$$\text{Prob}(y_q = 1) = \text{Prob}(\epsilon_q - u_q > X_q \delta - Z_q \beta) \quad (4)$$

Where the $\epsilon_q - u_q$ are independently and identically distributed (iid) logistically, by integration of equation (4), the injury probabilities are defined by a logit model:

$$\text{Prob}(y_q = 1) = \frac{1}{1 + \exp(X_q \delta - Z_q \beta)} \quad (5)$$

If the vectors X_q and Z_q contain only categorical variables (or continuous variables artificially categorised) then the data can be arranged in the form of a contingency table and the model (5) estimated by weighted least squares using the observed proportions in each cell as:

$$\log \left[\frac{\tilde{P}_q}{1 - \tilde{P}_q} \right] = Z_q \beta - X_q \delta, \quad (6)$$

where \tilde{P}_q is the observed proportion of injured individuals in category q . This was the approach adopted by Krishnan and his co-authors. However, when Z_q or X_q contain continuous variables, such as velocity change on impact, or when the data size is small, it is better to estimate (5) directly using maximum likelihood methods (McFadden 1974).

So far a binary injury classification scheme has been assumed. Typically injury classification schemes are more complex than this. A commonly used scheme is the abbreviated injury score (AIS). The scale is:

- 0 = no injury.
- 1 = minor injury.
- 2 = moderate injury.
- 3 = severe, but not life threatening, injury.
- 4 = severe, life threatening injury, but recovery is probable.
- 5 = critical injury, involving non-immediate death or a permanent impairment of bodily function such as paralysis, and.
- 6 = death within 24 hours.

This is an example of ordinal scale measurement. The categories bear a ranked relationship to one another (e.g. a moderate injury is 'worse' than a minor injury which in turn is 'worse' than no injury), but the numbers assigned do not indicate distance between categories, as would be the case for interval or ratio scale measurement. To emphasize this last point, the distance of 3 scale points between no injury and a serious injury (e.g. a broken thumb) should not be taken as equivalent to the distance of 3 scale points between a serious injury and death within 24 hours. Clearly, when measuring the 'actual' severity of injuries, the distance between the latter two points in the scale is greater than the former; but this is not indicated from the assigned numerical values. Regression analysis requires the dependent variable to be measured on an interval or ratio scale. Consequently, regression analysis is an unsuitable technique for analysing variations in injury severity as measured by the AIS.

Our previous exposition may be readily extended to embrace the analysis of ordered categorical variables. It is an easy matter to envisage a continuous variable (y_q^*) underlying the discrete responses of the AIS scale. The discrete points of the AIS scale can then be viewed as arising from this unseen variable crossing, not one, but a number of threshold values. The first value (say t_0^*) will be the crossing point between no observed injury and a minor injury, the second (t_1^*) between a minor injury and a moderate injury, and so on. There will be six such threshold values for the AIS scale. It is also clear that $t_{0q}^* < t_{1q}^* < t_{2q}^* < t_{3q}^* < t_{4q}^* < t_{5q}^*$

As before these threshold values may be endogenised by setting,

$$t_q^* = X_q \delta + u_q, \tag{7}$$

and relating the individual threshold values to t_q^* by,

$$\begin{aligned} t_{0q}^* &= t_q^* \\ t_{1q}^* &= t_q^* + k_1 \\ t_{2q}^* &= t_q^* + k_2 \\ &\dots \\ t_{5q}^* &= t_q^* + k_5 \end{aligned} \tag{8}$$

with $k_1 < k_2 < \dots < k_5$

3. EXPLANATORY VARIABLES OF INJURY SEVERITY

Biomechanical studies of automobile collisions have shown that the single most important variable contributing to crash severity is the change in vehicle velocity on impact. As it relates directly to the change in momentum, it best reflects the forces acting on vehicle occupants' bodies during a collision. During the collision period the occupants' initial velocities must be changed to the new velocity of the vehicle compartment, resulting in occupant contact with the vehicle interior and/or with a restraint system. The occurrence and nature of the injuries sustained will largely depend upon the de-acceleration time history of the collision (Krishnan et al. 1983).

Following Marquardt (1974), Carlson (1979), Hutchinson (1983) and Krishnan et al. (1983), from momentum considerations the change in vehicle velocity on impact can be calculated as:

$$\Delta v_c = \frac{(v_c^2 + v_o^2 + 2v_c v_o \cos(\theta))^{0.5}}{1 + m_c/m_o} \quad (12)$$

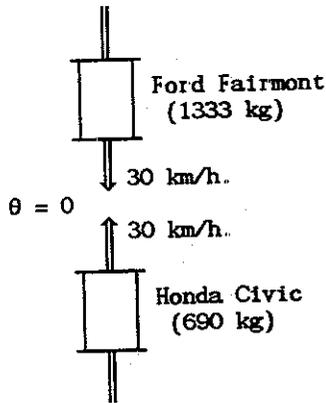
where v_c is the initial velocity of the case vehicle, v_o is the velocity of the other vehicle, m_c and m_o are the masses of the vehicles and θ is the angle of alignment of the vehicles at the point of impact.¹ The change in velocity on impact increases as the initial velocities of the vehicles increase and as the ratio m_c/m_o increases.

To aid appreciation of the index hypothetical crash configurations are depicted in Figures 1 and 2. Both figures involve a two-vehicle collision between a Honda Civic of mass 690 kg and a Ford Fairmont of mass 1333 kg. The vehicles are assumed to be both travelling at 30 km/h immediately prior to impact. In the head-on crash $\theta = 0$, so that the relative collision velocity is 60 km/h and the change in velocity for the Honda Civic is -40 km/h. If both cars were travelling at 40 km/h the change in velocity of the Honda on impact would be -53 km/h. The corresponding changes in velocity for the Ford Fairmont are -20 km/h and -27 km/h. For the accident depicted in the second diagram the change in velocity on impact for the vehicles is exactly half that of the Figure 1 accident. The reduction in impact velocity change is entirely due to the different alignment of the vehicles, measured by the $\cos(\theta)$ term in equation (12).

A feature of equation (12) is that the change in vehicle velocity on impact is contingent upon the ratio of vehicle masses. This means,

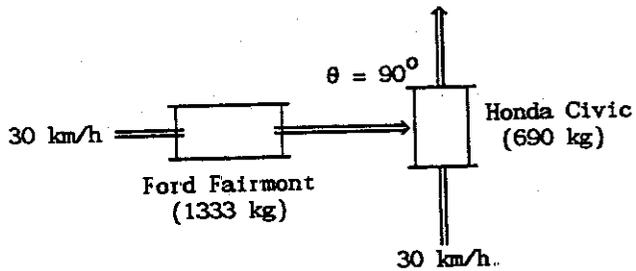
¹The formula of equation (12) takes into consideration the change in velocity along both the x and y axes. Some of the studies mentioned only included the velocity change in one direction.

FIGURE 1
EXAMPLE OF IMPACT VELOCITY CHANGE CALCULATION:
HEAD-ON COLLISION



$$\Delta v_{\text{Honda}} = \frac{[30^2 + 30^2 + 2(30)(30)]^{0.5}}{1 + \frac{690}{1333}} = 40 \text{ km/h.}$$

FIGURE 2
EXAMPLE OF IMPACT VELOCITY CHANGE CALCULATION:
SIDE-ON COLLISION



$$\Delta v_{\text{Honda}} = \frac{[30^2 + 30^2 + 2(30)(30)\cos(90)]^{0.5}}{1 + \frac{690}{1333}} = 28 \text{ km/h}$$

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

given initial velocities and vehicle alignment, the velocity change for a collision involving two equivalently weighted light vehicles will be the same as for two equivalently weighted heavy vehicles. An understanding of collision physics, however, leads to the conclusion that the severity of the crash will be greater for light vehicle occupants than for massive vehicle occupants. Larger vehicles provide a protective effect to occupants. The larger the vehicle the greater is the proportion of initial vehicle energies that can be absorbed by metal deformation without intrusion into the occupant compartment. There is also more room in the vehicle cabin for the occupant to travel without striking an object.

In addition to the main effects of impact velocity change and vehicle mass the severity of the crash to the occupant will be affected by seating position and restraint usage. The driver, in particular, is likely to be more vulnerable than other vehicle occupants because of the close proximity to the body of the steering apparatus. The role of seat belts in reducing injuries has been well documented (e.g. Lave and Webb 1970, Trinca 1980, Layton and Weigh 1983).

There is less evidence on the factors contributing to the different capacities of individuals to tolerate collision forces. Age and sex are the two variables used in the current analysis.

4. THE ADELAIDE IN-DEPTH ACCIDENT STUDY

The data source for the current research is the Adelaide in-depth accident study. This study, sponsored by the Office of Road Safety and the Australian Road Research Board, obtained an 8% sample of accidents in the Adelaide Metropolitan Area to which an ambulance was called during the period March 1976 - March 1977. The inclusion criterion means that the sample consists of a non-random subset of the total accident population. Accident researchers invariably work with non-random samples. Typically the sample inclusion criterion is that damages should exceed a specified monetary amount (currently \$1000 in South Australia). In practice many accidents with monetary losses well in excess of the specified monetary amount go unreported. It is likely that the Adelaide in-depth sample inclusion criterion is not very different to normal reporting criterion. Ambulances tend to be called routinely to accidents of even moderate severity. In the sample used for modelling an ambulance was called, but was not required, in over 50% of accidents. The Adelaide in-depth sample was found to be representative of the population of accidents, characterised by mass accident data records, for a number of key variables (Road Accident Research Unit 1979). In total the study collected information on 494 vehicle accident involvements.

The available sample size from this source was small compared to that available from mass accident data tapes. It is, however, of a higher quality with the data on each accident reflecting the combined on-site talents of an engineer, psychologist and a medical officer. In selecting this data source we were particularly mindful of having

accurate information on the velocities of vehicles just prior to impact and vehicle alignment when calculating the velocity change index of crash severity. This information is not included on Australian mass accident data records. We were also aware that the statistical techniques adopted in this study may be successfully applied, and stable parameter estimates obtained, with as few as 70 observations (Tye et al. 1982, p. 27).

No automobile mass information was collected in the Adelaide data; however, information on vehicle make and model was obtained. The information needed to calculate impact velocity change was derived by linking the Adelaide data with a data set on automobile characteristics, collected under the Dimensions of Automobile Demand Project sponsored by the National Energy Research Development and Demonstration Program (NERRDP) (Hensher 1986).

The current research is confined to an examination of two-vehicle collisions. In the Adelaide data 303 vehicles were involved in collisions of this type. The total number of occupants in these vehicles was 561 persons. Automobile-only two-vehicle collisions involved 196 automobiles which were occupied by 394 persons. The distribution of AIS injuries sustained by vehicle occupants in two-vehicle collisions and in automobile-only two-vehicle collisions are displayed in Table 1. Unfortunately for this analysis no injuries were observed in AIS classes 5 or 6 (notably there were no fatalities). Other relevant summary statistics for the data are shown in Table 2.

5. INJURY SEVERITY MODEL ESTIMATION RESULTS

Four models were developed using the Adelaide/NERRDP data. These related to (i) two-vehicle collisions, all occupants, (ii) two-vehicle collisions, drivers only, (iii) automobile-only two-vehicle collisions, all occupants, and (iv) automobile-only two-vehicle collisions, drivers only. The form for $y_q^* - t_q^*$ used in these models was:

$$y_q^* - t_q^* = \delta_0 + \beta_0 + \delta_1 \text{VELCHNG} + \delta_2 \log(\text{MASS}) + \delta_3 \text{SBELT} + \delta_4 \text{DRIVER} + \delta_5 \text{INTOX} + \beta_1 \text{AGEGO} + \beta_2 \text{FEMALE}, \quad (13)$$

where VELCHNG is the velocity change on impact ($= \Delta v$), MASS is the mass of the case vehicle ($= m_c$), 'log' denotes the natural logarithm, SBELT is a binary variable taking value 1 if in the investigator's judgement a seat belt was certainly or probably worn and 0 otherwise, DRIVER is a binary variable if the occupant was seated in the driving position and 0 otherwise, INTOX is a binary variable if the occupant was slightly, moderately or severely intoxicated and 0 otherwise, AGEGO is a binary variable taking the value 1 if the occupant was aged more than 60 years and 0 otherwise, and FEMALE is a binary variable

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

TABLE 1

DISTRIBUTION OF AIS CLASSIFIED INJURIES

(a) Two-vehicle collisions.

| AIS Class | Proportion of Sample |
|-----------|----------------------|
| 0 | 0.46 |
| 1 | 0.34 |
| 2 | 0.14 |
| 3 | 0.05 |
| 4 | 0.01 |

(b) Automobile-only two-vehicle collisions.

| AIS Class | Proportion of Sample |
|-----------|----------------------|
| 0 | 0.44 |
| 1 | 0.39 |
| 2 | 0.14 |
| 3 | 0.03 |
| 4 | 0.01 |

TABLE 2

SUMMARY SAMPLE STATISTICS

| Variable Description | Sample Statistic |
|---|------------------|
| (a) Two-vehicle collisions | |
| average number of vehicle occupants | 1.85 persons |
| average speed on impact | 36 km/h |
| average collision impact velocity | 53 km/h |
| average velocity change on impact | 27 km/h |
| average vehicle mass | 1160 kg |
| percent of head-on collisions | 6% |
| percent of side-on collisions | 81% |
| percent of rear-end collisions | 13% |
| percent of occupants wearing a seat belt | 64% |
| percent of occupants aged more than 60 years | 7% |
| percent of female occupants | 42% |
| percent of drivers intoxicated | 16% |
| (b) Automobile-only two-vehicle collisions | |
| average number of vehicle occupants | 2.01 persons |
| average speed on impact | 36 km/h |
| average collision impact velocity | 55 km/h |
| average velocity change on impact | 28 km/h |
| average vehicle mass | 1134 kg |
| percent of head-on collisions | 5% |
| percent of side-on collisions | 88% |
| percent of rear-end collisions | 7% |
| percent of occupants wearing a seat belt | 66% |
| percent of occupants aged more than 60 years | 8% |
| percent of female occupants | 48% |
| percent of drivers intoxicated | 17% |

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

taking the value 1 if the occupant was a female and 0 if the occupant was a male.

A number of features of the model implied by equations (11) and (13) should be noted. First, the parameters δ_0 and β_0 cannot separately be estimated. The model estimates a single constant term equal to $\delta_0 + \beta_0$. A consequence of this is that the separate indices for y_q^* and t_q^* cannot be recovered, only the combined index, $y_q^* - t_q^*$. This inability is not a cause for concern, however, because we are only interested in knowing how the independent variables VELCHNG, MASS, SBELT, DRIVER, INTOX, AGE60 and FEMALE affect the probability of sustaining an injury at AIS level i , $i = 0, 1, 2, 3, 4$. Second, for the drivers-only models, DRIVER is a constant so that δ_4 cannot be estimated. The parameter δ_4 is absorbed into the constant term and other parameter estimates remain unbiased. Third, intoxication data were only collected for drivers. This term was therefore omitted in the 'all occupants' models.

Estimation results for the four models are shown in Tables 3 - 6. All included variables, except the absolute mass and seat belt variables in the 'automobile-only driver-injuries' model, are statistically significant at the 2.5% level using a one-tailed T-test, and even these variables are significant at the 5% level using this test. Parameter estimates attached to the two dominant indices of crash severity, VELCHNG and MASS, are of the anticipated sign. The sign of the parameter estimate attached to VELCHNG is positive signifying that, ceteris paribus, an increase in the initial speeds of the vehicles or a decrease in the mass of the case vehicle relative to the other vehicle involved in the collision will increase the likelihood that occupants of the case vehicle will be injured. Vehicle mass exerts a further influence on the probability of being injured through the MASS variable. This variable is measuring the degree of protection offered by travelling in a vehicle of larger absolute mass. The negative sign on this variable indicates that the probability of injury decreases as the absolute mass of the case vehicle increases.

To appreciate the effect of vehicle size on injuries consider a head-on two-vehicle collision between a small (650 kg) and large (1300 kg) automobile both travelling at 30 km/h. The model predicts that for the injurious impact of the collision on small vehicle occupants to be the same as the injurious impact on large vehicle occupants one of the vehicles would have had to be travelling more than four times as fast; that is, at 120 km/h instead of 30 km/h. A diagrammatic depiction of the effect of vehicle mass on the probability of sustaining a severe injury in collisions involving various impact

TABLE 3

ORDERED LOGIT MODEL OF OCCUPANT INJURY SEVERITY
IN TWO-VEHICLE CRASHES

| Variable Mnemonic | Parameter Estimate | Standard Error | T-Statistic |
|--------------------------------------|-----------------------|-------------------|-------------|
| CONSTANT | 3.27464 | 1.2670 | 2.59 |
| VELCHNG | 0.06411 | 0.0071 | 8.97 |
| log(MASS) | -0.75606 | 0.1652 | -4.58 |
| DRIVER | 0.46227 | 0.2134 | 2.17 |
| SBELI | -0.30994 | 0.0956 | -3.24 |
| AGE60 | 0.95468 | 0.2893 | 3.30 |
| FEMALE | 0.78249 | 0.2041 | 3.83 |
| k_1 | 2.05903 | 0.1346 | 15.30 |
| k_2 | 3.74015 | 0.2411 | 15.51 |
| k_3 | 6.40987 | 0.6249 | 10.26 |
| Number of observations | | | 561 |
| Log Likelihood at $\bar{\delta} = 0$ | | | -660.98 |
| Log Likelihood at convergence | | | -557.48 |
| R^2 | | | 0.36 |

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

TABLE 4

ORDERED LOGIT MODEL OF OCCUPANT INJURY SEVERITY
IN AUTOMOBILE-ONLY TWO-VEHICLE CRASHES

| Variable Mnemonic | Parameter Estimate | Standard Error | T-Statistic |
|--------------------------------------|-----------------------|-------------------|-------------|
| CONSTANT | 4.04485 | 3.2690 | 1.24 |
| VELCHNG | 0.05813 | 0.0103 | 5.65 |
| log(MASS) | -0.86774 | 0.4418 | -1.96 |
| DRIVER | 0.55297 | 0.2437 | 2.27 |
| SBELT | -0.28259 | 0.1075 | -2.63 |
| AGE60 | 1.16618 | 0.3363 | 3.47 |
| FEMALE | 1.15146 | 0.2423 | 4.75 |
| k_1 | 2.10255 | 0.1543 | 13.63 |
| k_2 | 3.92270 | 0.3037 | 12.92 |
| k_3 | 5.65502 | 0.6583 | 8.59 |
| Number of observations | | 394 | |
| Log Likelihood at $\bar{\delta} = 0$ | | -651.01 | |
| Log Likelihood at convergence | | -408.84 | |
| R^2 | | 0.24 | |

TABLE 5

ORDERED LOGIT MODEL OF DRIVER INJURY SEVERITY
IN TWO-VEHICLE CRASHES

| Variable Mnemonic | Parameter Estimate | Standard Error | I-Statistic |
|--------------------------------------|-----------------------|-------------------|-------------|
| CONSTANT | 3.46723 | 1.5760 | 2.20 |
| VELCHNG | 0.06221 | 0.0098 | 6.34 |
| log(MASS) | -0.72935 | 0.2127 | -3.43 |
| SBELT | -0.31620 | 0.1252 | -2.53 |
| INTOX | 1.10009 | 0.3335 | 3.30 |
| AGE60 | 0.90783 | 0.3694 | 2.46 |
| FEMALE | 0.61840 | 0.2883 | 2.15 |
| k_1 | 2.13719 | 0.1927 | 11.09 |
| k_2 | 3.75491 | 0.3104 | 12.10 |
| k_3 | 6.21678 | 0.6771 | 9.18 |
| Number of observations | | | 303 |
| Log Likelihood at $\bar{\delta} = 0$ | | | -373.96 |
| Log Likelihood at convergence | | | -302.71 |
| R^2 | | | 0.42 |

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

TABLE 6

ORDERED LOGIT MODEL OF DRIVER INJURY SEVERITY
IN AUTOMOBILE-ONLY TWO-VEHICLE CRASHES

| Variable Mnemonic | Parameter Estimate | Standard Error | I-Statistic |
|--------------------------------------|-----------------------|-------------------|-------------|
| CONSTANT | 6.07375 | 4.7140 | 1.29 |
| VELCHNG | 0.06644 | 0.0152 | 4.37 |
| log(MASS) | -1.14107 | 0.6420 | -1.78 |
| SBELT | -0.27665 | 0.1519 | -1.82 |
| INTOX | 1.12041 | 0.4069 | 2.75 |
| AGE60 | 0.96767 | 0.4426 | 2.19 |
| FEMALE | 0.97779 | 0.3525 | 2.77 |
| k_1 | 2.10935 | 0.2255 | 9.35 |
| k_2 | 3.61398 | 0.3714 | 9.73 |
| k_3 | 5.30452 | 0.7399 | 7.17 |
| Number of observations | | 196 | |
| Log Likelihood at $\bar{\delta} = 0$ | | -234.31 | |
| Log Likelihood at convergence | | -203.79 | |
| R^2 | | 0.31 | |

velocities with a second vehicle of mass 1000 kg is shown in Figure 3.¹

Models were also estimated with impact velocity, the numerator of equation (12), included separately from the ratio of vehicle masses, the denominator of equation (12), and the logarithm of absolute vehicle mass. The separate terms were all statistically significant at the 2.5% level. However, the overall goodness of fit measures for these models and the models of Tables 3 - 6 were virtually identical, lending support to the use of the VELCHNG variable, derived from a theoretical consideration of the physics of two-vehicle collisions, as an index of crash severity. Models which omitted either of the mass effects or the impact velocity term were significantly inferior to the model in which all these terms were included. The correlation between log(MASS) and VELCHNG was always less than 0.35.

Parameter estimates attached to the remaining set of (binary) variables also took the anticipated signs. The positive sign attached to the AGE60 variable signifies that the elderly have an increased probability of being injured in a collision, of a given severity level, between two vehicles. The relationship between age and the probability of sustaining a severe injury is quantified diagrammatically in Figure 4. It is interesting to note that, ceteris paribus, females have a higher probability of being injured in road accidents than do males.

Conversely, as expected, the wearing of a seat belt reduces the probability of an injury; see Figure 5. In the Adelaide sample it is estimated that in comparing a situation where all car occupants wore a seat belt with a situation where no car occupants wore a seat belt, the number of severe injuries (AIS classes 3 and 4) is reduced by 33% and the number of minor injuries (AIS class 1) by 12%. These estimates are considerably lower than the estimates of 60% and 30%, respectively, quoted by the United States National Highway Traffic Safety Administration (cited in Arnould and Grabowski 1983). The latter of my estimates, however, is in reasonable accord with an estimated 44% reduction in the number of fatalities from seat belt wearing derived from the work of Layton and Weigh (1983), based on Queensland data, and an estimated reduction of 30% in severe and fatal accidents by Krishnan (1983) based on two U.S. data sets.² In making

¹The variable levels assumed in Figures 3 - 6, excepting those explicitly set in each Figure, are $m_o = 1000$ kg, $m_c = 1000$ kg, SBELT = 1 (seat belt worn), DRIVER = 0 (occupant not seated in driver's position), INTOX = 0 (occupant sober), AGE60 = 0 (occupant younger than 60 years of age) and FEMALE = 0 (occupant is a male).

²Layton and Webb calculate that fatalities were reduced by 37% as a result of legislation making seat belt wearing compulsory, if fitted, after allowing for the proportion of vehicles in the fleet without seat

Statistic

1.29

4.37

-1.78

-1.82

2.75

2.19

2.77

9.35

9.73

7.17

FIGURE 3
INJURY SEVERITY BY VEHICLE WEIGHT

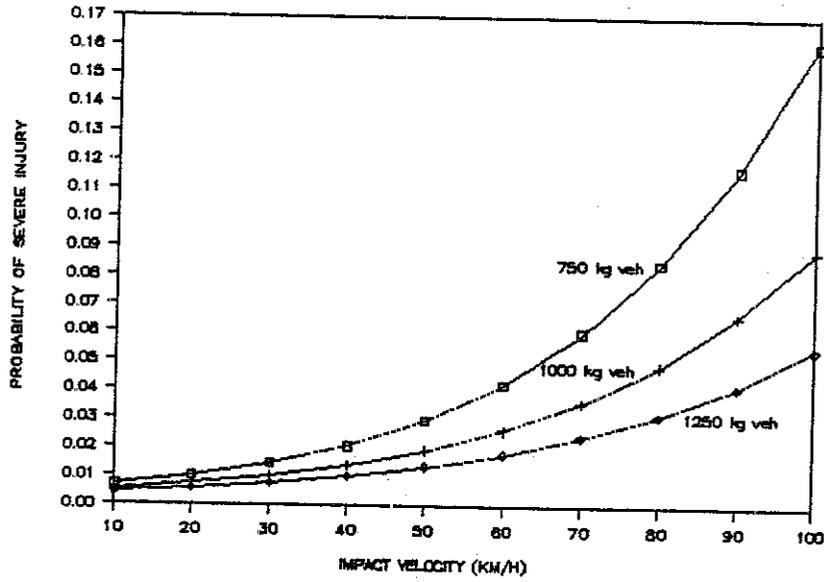


FIGURE 4
INJURY SEVERITY BY OCCUPANT AGE

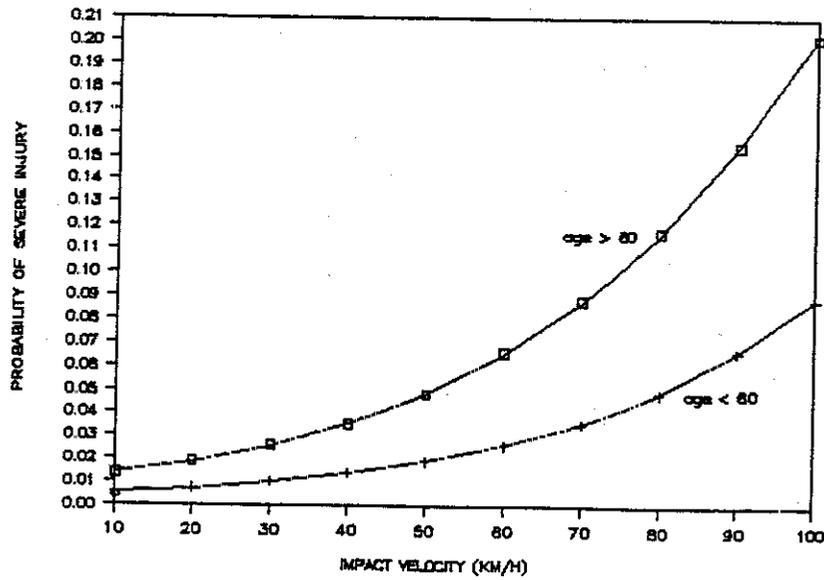


FIGURE 5

INJURY SEVERITY BY SEAT BELT USAGE

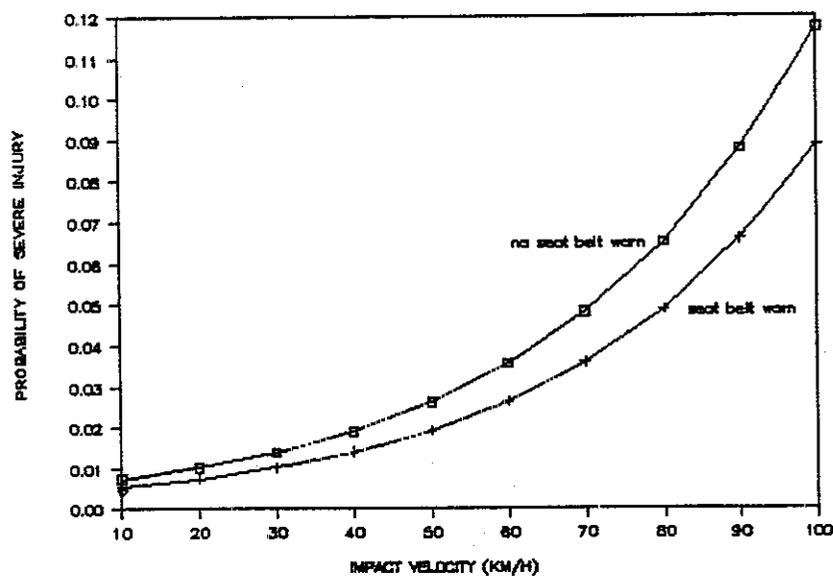
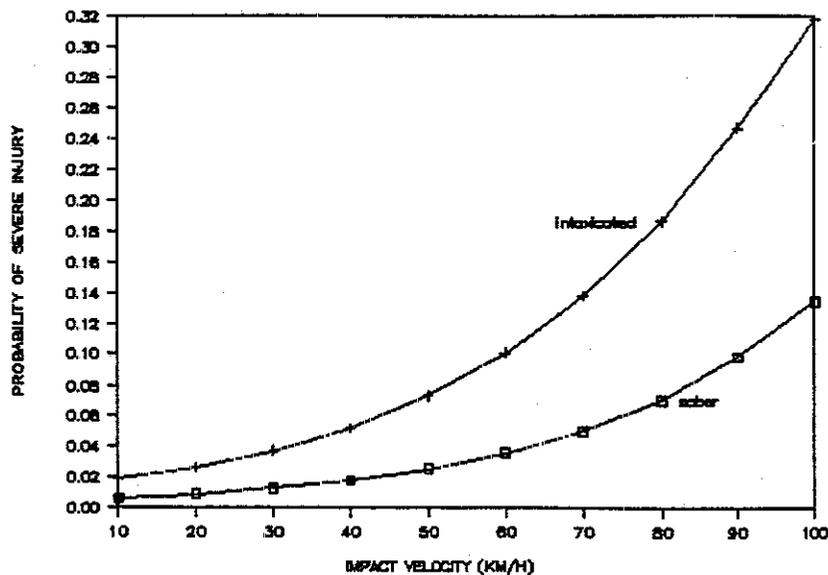


FIGURE 6

INJ. SEVERITY BY STATE OF INTOXICATION



A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

these comparisons it is important to note that the current research was restricted to an analysis of urban road accidents.

Finally, as shown in Figure 6, inebriated drivers were more likely to be injured than non-inebriated drivers. Other studies (e.g. Raymond 1974, Johnson 1978) have concluded that those affected by alcohol are more likely to be involved in accidents in general, particularly, severe accidents. The current study adds to the store of knowledge on the relationship between alcohol and road safety by concluding that given an accident of a set severity level those under the influence of alcohol are about two and one-half times more likely to sustain a severe (AIS classes 3 and 4) injury than a sober occupant.³

The two main measures used to assess the overall goodness of fit of the ordered logit model are the value of the log-likelihood function at convergence and R^2 . The latter of these is analogous to the R^2 of regression analysis and is given by:

$$R^2 = \frac{\sum_q (\hat{\tau}_q - \bar{\tau})}{\left[\sum_q (\hat{\tau}_q - \bar{\tau}) \right] + \frac{\pi^2}{3} Q} \quad (14)$$

where $\hat{\tau}_q = \hat{y}_q^* - \hat{t}_q^*$, $\bar{\tau} = \sum_q \hat{\tau}_q / Q$, and Q is the sample size.⁴ In

belts fitted. We have adjusted this estimate upwards to also allow for the proportion of the travelling population not using a seat belt when available (see Lay 1984).

³Dr. Max Lay has drawn to my attention research by Waller et al. (1986) which reached a similar conclusion about the effect of alcohol on injury severity. Waller and his colleagues concluded that alcohol involved drivers were 3.85 times more likely to be killed than sober drivers, once differences in vehicle deformation and accident type had been taken into account. Their estimate of alcohol induced injury severity is slightly higher than mine. Their study, however, did not control for seat belt usage differences between sober and inebriated drivers, sex differences and age differences. These additional factors are taken into account in the current study.

⁴It will be noted that the R^2 of equation (14) and the R^2 of regression analysis are equivalent except the sum of squared residuals and the total sum of squares are both estimates rather than actual values. The R^2 of equation (14) is therefore an estimate of the true R^2 . To fully utilise this estimate knowledge of the sample distribution of the true

assessing the fit of these models it must be emphasized that indeterminacy pervades the road accident injury process. Even in a low severity accident a fragment of flying glass can cause severe injury, possibly death, if it becomes lodged in a vulnerable point of the human body. No model can hope to capture this level of detail. In explaining between 24% - 42% of the variation in road accident injuries, using only a limited set of variables, the models do surprisingly well.

6. CONCLUSION

As emphasised in the introduction, many items of interest in the road accident field are measurable only on a nominal or ordinal scale. The responses of road accident researchers in analysing such data have fallen into three main classes. Like their sociological counterparts (see Winship and Mare 1984) those involved in the analysis of categorical road accident data have tended either (i) to ignore the nominal or ordinal scale of measurement, treating the item as if it were continuous, or (ii) have utilized special techniques that are not integrated into established frameworks for multivariate structural analysis, or (iii) have adopted frequency or contingency table approaches when a structural regression model would have been more appropriate. Latent variable models allow the retention of a regression structure while recognizing the measurement characteristics of categorical variables. They also provide an appealing probabilistic framework for the analysis of accident data.

In this paper a particular latent variable model, the ordered logit model, was applied to an analysis of injuries sustained in two-vehicle collisions on urban roads. A theoretical model was constructed of the injury process and an ordered logit model derived directly from this theory. At the heart of this theory were two determining factors for the severity of injuries sustained; the severity of the collision and the ability of vehicle occupants to withstand collision forces without injury. These primary factors were endogenised using characteristics of the vehicles involved in the collision and the occupants of these vehicles. Vehicle speeds, vehicle weights, seat belt usage, seating position, state of intoxication, age and sex were all found to be important determinants of the severity of injuries sustained in an accident.

R^2 is required, but the sample distribution of the true R^2 is presently unknown (McKelvey and Zavoina 1975).

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

ACKNOWLEDGEMENTS

The interest of the author in safety research began while working at the Australian Road Research Board and has been in part funded by the Federal Office of Road Safety. This paper was written while the author was on leave as a Visiting Fellow at the Transport Research Group, Macquarie University.

REFERENCES

- ARNOULD, R.J. and H. GRABOWSKI (1983). Auto safety regulation: an analysis of market failure. *The Bell Journal of Economics*, Vol. 14, pp. 27 - 48.
- ATKINS, A.S. (1981). The economic and social costs of road accidents in Australia. Centre for Environmental Studies, CR 21, University of Melbourne, 127pp.
- BARNARD, P.O. (1985). Disaggregate Econometric Analysis of Road Accident Data. Consultants Proposal to the Australian Federal Office of Road Safety, 17pp.
- BARNARD, P.O. (1989). How some recent advances in econometrics might be used in road accident analysis. Draft consultants report to the Office of Road Safety, 91pp.
- CAMPBELL, R. and R.J. FILMER (1980). Deaths from road accidents in Australia. Unpublished mimeo.
- CARLSON, W.L. (1979). Crash injury prediction model. *Accident Analysis and Prevention*, Vol. 11, pp. 137 - 53.
- CHIRACHAVALA, T., D.E. CLEVELAND and L.P. KOSTYNIUK (1984). Severity of large-truck and combination-vehicle accidents in over-the-road service: a discrete multivariate analysis. *Transportation Research Record*, No. 975, pp. 23 - 36.
- GARCIA-FERRER, A. and J. del HOYO (1987). Analysis of the car accident indexes in Spain: a multiple time series approach. *Journal of Business and Economic Statistics*, Vol 5, No. 1, pp. 27 - 38.
- HENSHER, D.A. (1986). Dimensions of automobile demand: an overview of an Australian research project. *Environment and Planning A*, Vol. 18, pp. 1339 - 74.
- HUTCHINSON, I.P. (1983). A bivariate normal model for intra-accident correlations of driver injury with application to the effect of mass ratio. *Accident Analysis and Prevention*, Vol. 15, No. 3, pp. 215 - 24.

JOHNSON, I.R. (1978). The implications of alcohol impairment for research into road and traffic system design and management. *Australian Road Research*, Vol. 8, No. 4, pp. 57 - 62.

JOVANIS, P. and H. CHANG (1986). Disaggregate model of highway accident occurrence. Paper presented at the 65th Annual Meeting of the Transportation Research Board, January 13 - 16, Washington.

KERNS, I.B. and H.J. GOLDSMITH (1984). The impact on traffic crashes of the introduction of random breath testing in New South Wales. Proceedings of the Australian Road Research Board Conference, Vol. 12, Part 7, pp. 81 - 95.

KRISHNAN, K.S. (1983). Empirical estimates of seat belt effectiveness in two-car collisions. *Accident Analysis and Prevention*, Vol. 15, No. 3, pp. 227 - 36.

KRISHNAN, K.S., J.V. CARNAHAH and M.J. BECKMANN (1983). An injury threshold model for two-car collisions. *Management Science*, Vol. 29, pp. 909 - 26.

LAVE, L. and W. WEBBER (1970). A benefit-cost analysis of auto safety features. *Applied Economics*, Vol. 2, pp. 265 - 75.

LAY, M.G. (1984). *Source Book for Australian Roads*. Australian Road Research Board, Vermont South, Victoria, 2nd Edition, 551pp.

LAYTON, A.P. (1983). The impact of increased penalties on Australian drink / driving behaviour. *Logistics and Transportation Review*, Vol. 19, No. 3, pp. 261 - 66.

LAYTON, A.P. and J.C. WEIGH (1983). The efficacy of some recent Australian road safety policy initiatives. *Logistics and Transportation Review*, Vol. 19, No. 3, pp. 267 - 78.

MARQUART, J.F. (1974). Vehicle and occupant factors that determine occupant injury. Paper presented at SAE Automotive Engineering Congress, Detroit, Michigan.

McKELVEY, R.D. and W. ZAVOINA (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, Vol. 4, pp. 103 - 20.

McFADDEN, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, New York.

ROAD ACCIDENT RESEARCH UNIT (1979): Adelaide In-Depth Accident Study, 1975 - 79, Parts 1 - 10. University of Adelaide, Adelaide, Australia.

A MODEL OF INJURIES SUSTAINED IN TWO VEHICLE COLLISIONS

IRINCA, G. (1980). Medical aspects of seat belt usage. *Journal of Traffic Medicine*, Vol. 8, No. 3, pp 36 - 38.

IYE, W.B., L. SHERMAN, M. KINNUCAN, D. NELSON and I IARDIFF (1982). Application of disaggregate travel demand models. National Cooperative Highway Research Program, Report 253, Transportation Research Board, Washington D.C., 207pp.

VIANO, D.C., R.C. HAUT, M. COLOCOUSKY and K. ABSOLOM (1978). Factors influencing biomechanical response and closed chest trauma in experimental thoracic impacts. Paper presented at the 22nd Conference of the American Automobile Association for Automotive Medicine, Ann Arbor, Michigan.

WALLER, J.A., J.R. STEWART, A.R. HANSEN, J.C. STUTTS, C.I. POPKIN and E.A. RODGMAN (1986). The potentiating effects of alcohol on driver injury. *Journal of the American Medical Association*, Vol. 256, pp. 1461 - 66.

WINSHIP, C. and R.D. MARE (1984). Regression models with ordinal variables. *American Sociological Review*, Vol. 49, pp. 512 - 25.

ZLATOPER, I. J. (1984). Regression analysis of time series data on motor vehicle deaths in the United States. *Journal of Transport Economics and Policy*, Vol. XVIII, No. 3, pp. 263 - 74.