# A CHAIN MODEL FORMULATION AND ITS IMPLICATION FOR THE EFFECT OF SPEED LIMIT ON COLLISION SEVERITY TO AN INVOLVED DRIVER VIA ESTIMATED SPEED 

Chin-Zin YANG
Associate Professor
Department of Traffic and Transportation
Engineering and Management
Feng Chia University
100 Wenhwa Road, Taichung,
Taiwan 40724, R.O.C.
Fax: +886-4-24520678
E-mail: czyang@fcu.edu.tw

Chia-Ming AI
Instructor
Department of Traffic and Transportation
Engineering and Management
Feng Chia University
100 Wenhwa Road, Taichung, Taiwan 40724, R.O.C.

Fax: +886-4-24520678
E-mail: cmai@fcu.edu.tw


#### Abstract

Provided with in-field crash-investigation and reported information from 4476 real accident episodes, a log-linear model is used to capture the relationship between speed and its predictors such as speed limit. While by introducing a binary logit model, the estimated speed may contribute explanatory power on subject driver's injury severity. From the results of the first model, we can surmise that higher speed goes with higher speed limit and speeding violation happens frequently at lower or higher speed limit like an U shape. Most interestingly, the second model estimation shows that speed limit can influence collision severity via estimated speed, e.g., higher speed limit is associated with higher estimated speed, which in turn is related to more serious injury until the speed limit of expressway or freeway due to higher standard of engineering design.


Key Words: chain model, speed limit, collision severity

## 1. INTRODUCTION

A few former studies show higher speeds are inclined to cause more severe vehicle damages or human casualties in terms of momentum conservation theory when crashes happen and vehicles are suddenly stopped. Speed limit is one of the important factors that influences speed before a collision. From two examples, Evans (1991) demonstrates the dependence of speed on speed limit. Subjective estimates of actual speed are close to $62 \mathrm{~km} / \mathrm{h}$ while traveling along a highway with a $50 \mathrm{~km} / \mathrm{h}$ speed limit, and approximately equal to $108 \mathrm{~km} / \mathrm{h}$ on a freeway with a $100 \mathrm{~km} / \mathrm{h}$ speed limit. Evans depicts the influence of speed on accident risk in a way of higher speed leading to higher severity, i.e., speed, speed to the second power, and to the fourth power are positively related to crash rate, injury rate, and fatality rate respectively. Najjar et al. (2002) justifies the increase of unreasonable low speed limits by comparing before with after in real accident data of relevant highway locations. Where higher speed limits are posted, crash rates, fatal crash rates, and fatality rates for most of the new speed limit areas remain stable, but only $7 \%$ of suburban two-lane roadways do not. Some researches (Sunanda and Lu, 2002; Bedard et al., 2002; Al-Ghamdi, 2002; Scuffham and

Langley, 2002) have proved that speed is one of the most important factors that increases casualty probability but crash rate, that is only affected greatly by speed deviation of vehicles in traffic flows by viewpoints from other paper (Wilmot and Khanal, 1999).
This research first aims at studying the impacts of speed limit on speed. Later, drivers' injury levels are predicted by their speeds, or affected by speed limit via estimated speed. The major implication of study results may allow understanding the appropriateness of the speed limits for highways with several design standards, and whether or not it is possible to alleviate traffic incident severity by only changing these limits under different lighting conditions. The data regarding speed limits ranging from 20 to 70 along local highways in the urban or country areas, 80 on expressways, and 90 or higher on freeways, are easy to obtain from in-site incident reports in Taiwan. Speed before a collision is not available, except as estimated comprehensively by pavement cutting evidence made by braking maneuver or by drivers' self reports (the estimated results and detailed estimation procedure is given and designed in another report by Yang and Ai, 2003). Before this paper is intended to achieve its current goal, all necessary data regarding speeds were made ready and are called original speeds. This is distinguished from estimated speeds in the injury related model, which employed the output (i.e. estimated speeds) of the speed related model as an independent variable.
Other minor consideration goes to the choice of controlled covariates in the two relationships, speed limit to original speed and estimated speed to casualty. From the provided data items (to be described later), we subjectively select driver's blood alcohol concentration (BAC) level, driver's license holding status, environmental lighting, and weather conditions to be candidates for the former one. As for the latter one, there are some factors to be thought that are more related to injury than to speed. These include things such as driver's gender, driver's age level, driver's education level, both vehicle types (we are only interested in two vehicle cases), relative heading direction, driving movement, impact direction, and road type.

## 2. DATA DESCRIPTION AND BASIC STATISTICS

The employed data, originally 11,282 observations, each consisting of 49 data fields, are from the provincial re-authentication organization in Taiwan. The primary task of this organization provides information of forwarding evidence to making a decision about liability split among involved parties in accident cases through group discussion. These data are collected in meetings held between March 2000 and August 2002, but for the first and second stage analyses only 4653 and 4476 observations, each containing 16 variables, are available for further analyses due to considering only cases with two road users involved and missing values of important data (e.g. speed limit, original speed, and casualty). Although traffic volume and road configuration seem to be more important to estimate the speed, these two data items are not available in the original data set.
The core of this study is the reconstructed speed, which connects models for two stages, one
as dependent variable and the other after estimation as independent variable. Therefore, it is necessary first to demonstrate how the speed just prior to a collision is determined. Due to the fact a driver does not always look into instrument panel where the value of speed odometer is available, and then does not estimate precisely, or even if he really knows the value fails to tell the truth, otherwise he does not remember or is not able to report because of serious injury and field investigation probably cannot find solid evidence at crash site, such as the hit other vehicle type, relative damage to both vehicles (if the other road user is not pedestrian), braking distance shown by cut pavement, all of which are ready to use in reconstructing speed before the crash.. The complicated procedure of reconstructing speed used in this research is available elsewhere (Yang and Ai, 2003) and is not the major interest of this paper. Here, we briefly describe the important results of that study. The reported speeds are compared to those derived from reconstruction based on braking distance and other information, which are used along with the impact of collision. Under specific conditions, i.e. given a non-rainy day, a truck or bus hit a motorcycle from right side when moving at a reported speed $40,50,60,70$, $80,90 \mathrm{~km} / \mathrm{h}$ respectively, the resultant shown braking distances (and in parenthesis, its relevant speed list on a conventional table by Cooper, 1990) are as following in sequence: $13.4 \mathrm{~m}(48.8 \mathrm{~km} / \mathrm{h}), 21.1 \mathrm{~m}(61.3 \mathrm{~km} / \mathrm{h}), 28.3 \mathrm{~m}(70.9 \mathrm{~km} / \mathrm{h}), 36.6 \mathrm{~m}(80.6 \mathrm{~km} / \mathrm{h}), 47.9 \mathrm{~m}$ $(92.3 \mathrm{~km} / \mathrm{h}), 62.8 \mathrm{~m}(106 \mathrm{~km} / \mathrm{h})$, therefore the underestimation factors are $0.820,0.816,0.846$, $0.868,0.867,0.852$. Using the original speed categorization in that research, the reported speeds can be reconstructed into approximately actual speeds by factors $1 / 0.820,1 / 0.831$, $1 / 0.867,1 / 0.852$ for speed ranges [-, 44.9], [45, 59.9], [60, 79.9], [80, -] accordingly.
In addition to the major hypotheses that higher speed limit introduces higher speed, and higher speed results in more serious injury, other factors to be considered as controlled covariates at two model stages have their assumed roles, which are subjectively predefined and depicted here: (1) higher BAC level promotes higher speed, (2) non-legal license holder has a tendency to drive faster, (3) darker environment increases the probability of speeding, (4) non-raining day makes faster driving easier; while after a crash, (5) female is inclined to more serious injury, (6) it is more difficult for a mature person to resist injury from a collision, (7) a more highly educated person (possibly because of lack of physical training) is more vulnerable to injury, (8) a lighter vehicle produces more casualties, (9) touching a heavier object generates larger scale consequence, (10) opposite traffic has a large influence, (11) the going ahead maneuver has a more direct effect, (12) head-on impact induces a lethal episode, and (13) signalized intersection decreases life-threatening situation.
For a better understanding of the background of selected cases, the descriptive statistics for each variable is summarized by the nature of data types such as being discrete or continuous. The abbreviated words in parenthesis will be used in latter tables. (1) Frequency of certain drinking status- $\mathrm{BAC}>0.11$ \%, 83; BAC falls in [0.05 \%, $0.11 \%$ ], 94; BAC falls in [0.00 \%, $0.049 \%]$, 69; non-drinker, 4407; (2) frequency of certain license holding status- legitimate (LGTM), 7137; illegitimate with age > 18 years (ILGTM, >18), 187; illegitimate with age <

18 years (ILGTM, <18), 100; unknown (LC-UN), 229; (3) frequency of certain lighting condition- day time, 3094; night with illumination (night w/ illum), 1059; night without illumination (night w/o illum), 280; unknown (LTCDT-UN), 220; (4) frequency of certain weather condition- non-raining, 4086; raining, 567; (5) frequency of certain speed limit (SPLM)- $25 \mathrm{~km} / \mathrm{h}, 19 ; 30 \mathrm{~km} / \mathrm{h}, 158 ; 40 \mathrm{~km} / \mathrm{h}, 2547 ; 50 \mathrm{~km} / \mathrm{h}, 622 ; 60 \mathrm{~km} / \mathrm{h}, 828 ; 70 \mathrm{~km} / \mathrm{h}, 221$; 80km/h, 31; 90km/h, 65; 100km/h, 162; (6) 7 values of 8 (a subjectively determined number) equal-frequency group boundaries for reconstructed speed-18, 24, 37, 43, 49, 60, 69; (7) frequency of certain gender- male, 3613; female, 863; (8) 3 values of 4 (a subjectively determined number) equal-frequency group boundaries for years of age (YOA)- 27, 35, 45; (9) frequency of certain educational level- undergraduate, 405; 5 year professional training after junior high school, 425; senior high school, 760; 3 year professional training after senior high school, 909; junior high school, 735; elementary school, 399; kindergarten, 2; illiterate 36; unknown. 805; (10) frequency of certain subjective(S) vehicle type- passenger car, 2288; pickup, 388; station wagon, 208; taxi, 128 (all of above, Small Cars); heavy motorcycle, 670; light motorcycle, 266; truck, 254; bus, 110; other vehicle types (OVT-S), 164; (11) frequency of certain objective(O) vehicle type or hit object- passenger car, 2065; pickup, 326; station wagon, 173; taxi, 111 (all of above 4, Small Cars); heavy motorcycle, 860; light motorcycle, 393; truck, 148; bus, 73; bicycle (BCC), 88; pedestrian (PDT), 123; others, 116; (12) frequency of certain relative heading direction- opposite, 1268; same, 1466; cross or perpendicular angle (left side), 892; cross or perpendicular angle (right side), 850; (13) frequency of certain driving maneuver- going straight ahead (SA), 3625; right turn (RT), 168; left turn (LT), 555; right U turn (RUT), 5; left U turn (LUT), 123; (14) frequency of certain collided impact- front, 1396; right front, 818; right, 438; right rear, 284; rear, 108; left rear, 194; left, 339; left front, 899; (15) frequency of certain road type- signalized intersection (SIG-INT), 806; flash light intersection, 627; non-signalized intersection, 1425; straight segment, 1422; curvature segment, 196 (all of above 4, non SIG-INT); (16) frequency of certain crash severity- casualty, 1356; non-casualty, 3120.

## 3. THE MODEL FOR THE EFFECT OF SPEED LIMIT ON SPEED

Since the independent factors influencing speed are all discrete, but the reconstructed speed is continuous originally, the log-linear model seems a suitable candidate to employ. This model uses the logarithm of speed as dependent variable, regressing on a linear combination of other relevant contributors including speed limit, therefore justifying the positive nature of speed value. Among the four available assumptions regarding the distribution of model residual and resultant dependent variable: exponential, Weibull, log-logistic, and log-normal, the last one has been chosen because the first is just a special case of the second, and the implication of the model is not easy to depict directly just from the calibration outputs for the second and third, where the physical meaning depends on the calculation of gamma function and the inverse of sin function. This log-linear model is well known in bio-medical literature (e.g.

Collett, 1994) as the accelerated failure time (AFT) model whose estimated target is time a positive value. According to the theory of AFT, the formulas consisting of the log-linear format and the expectation and 85th percentile of the targeted speed are listed below:
$\ln (\mathrm{S} / X)=\beta X+\sigma \mathrm{W}$
$\mathrm{E}(\mathrm{S} / X)=\exp \left(\beta X+0.5 \sigma^{2}\right)$
$\mathrm{E}(\mathrm{S} / X$ with all $X$ 's elements $=0)=\exp \left(\beta_{0}+0.5 \sigma^{2}\right)$
$\operatorname{ER}\left(\mathrm{S} / X\right.$ with $\mathrm{X}_{\mathrm{i}}=1$ over $X$ with $\mathrm{X}_{\mathrm{i}}=0$, all other elements being equal)
$=\mathrm{E}\left(\mathrm{S} / X\right.$ with $\left.\mathrm{X}_{\mathrm{i}}=1\right) / \mathrm{E}\left(\mathrm{S} / X\right.$ with $\left.\mathrm{X}_{\mathrm{i}}=0\right)=\exp \left(\beta_{\mathrm{i}}\right)$
$\operatorname{PER}_{85}(\mathrm{~S} / X)=\exp \left(\beta X+0.5 \sigma^{2}+1.04 \delta\right)$
$\delta=\left(\lambda^{2}+\xi^{2}+2 \rho \varepsilon \xi\right)^{0.5}$
Where $S$ denotes the speed to be predicted, $\ln (S)$ natural logarithm of $S, X$ a contributor (i.e., variable) vector to influencing velocity, $\beta$ a parameter (i.e., coefficient) vector to be estimated, $\sigma$ a scale factor also to be estimated, W the residual variable which is assumed from standard normal distribution, $\mathrm{E}(\mathrm{S})$ expectation of speed, $\mathrm{ER}(\mathrm{S})$ expected ratio of speed, $\mathrm{PER}_{85}(\mathrm{~S})$ 85th percentile of speed, $\delta$ the standard deviation for the sum variable regarding speed limit and a interesting variable, $\lambda$ the standard error for estimated speed limit's parameter, $\xi$ the standard error for estimated that variable's parameter, and $\rho$ the correlation between the above two estimated parameters.
After several trials of model calibrations, the selected result is the list in table 1. All but weather condition are more or less significant in terms of estimated parameter of certain dummy variable sufficiently departing from that of basic dummy variable (may be a new combined one after grouping for similar effects from different levels) for each categorical factor. In addition to constant and scale (the group of drivers with all $X$ 's element being equal to 0 , e.g., non-drinking legal day time driving has an estimated average speed $34.0 \mathrm{~km} / \mathrm{h}$ along areas with speed limit set as 25 or $30 \mathrm{~km} / \mathrm{h}$, please refer to equation (3)), there are totally 14 powerful dummy variables, 3 for drinking status (basic being non-drinker), 1 for license holding status (basic being legitimate, illegitimate with age > 18 years, or unknown) 3 for lighting condition (basic being day time), and 7 for speed limit (basic being 25 or $30 \mathrm{~km} / \mathrm{h}$ ) respectively. As expected, the indicated effects are in the right direction, however, these effects are not quite proportional in regard to drinking status. From information about the exp(coeff.) column in table 1, compared to the speed under non-drinking situation, the speed is increased by 21 to $25 \%$ under other drinking statuses, but there is no significant difference between speeds among those drinking statuses. An illegitimate driver age $<18$ years tends to speed 34 \% more than other license holding statuses. Driving at night allows facing less traffic and then eases 7 to 18 \% higher speed maneuver for the with illumination to without illumination condition respectively. Finally, the most important and significant factor is the speed limit. Appropriate speeds are generally chosen along with speed limits, there are 10, 26, $38,60,136,177$, and $196 \%$ speed surplus when the environments set limit as $40,50,60,70$, 80,90 , and $100 \mathrm{~km} / \mathrm{h}$, relative to the one setting as 25 or $30 \mathrm{~km} / \mathrm{h}$.

Table 1. A Log-normal AFT Model for Speed Limit to Original Speed Relationship

| LOG LIKELIHOOD $=-4105.9$ |  | sample sizes $=4653$ |  |  |
| :---: | :---: | :---: | :---: | :---: |
| GLOBAL CHI-SQUARE $=756.9 \quad$ D.F. $=14 \quad \mathrm{P}=0.0000$ |  |  |  |  |
| VARIABLE | COEFFICIENT | Standard Error | COEFF./S.E. | EXP(COEFF.) |
| CONSTANT | 3.3563 | 0.0441 | 76.0341 | 28.6832 |
| BAC>0.11 \% | 0.1929 | 0.0653 | 2.9548 | 1.2128 |
| $0.05 \%<$ BAC $<0.11 \%$ | 0.1918 | 0.0613 | 3.1289 | 1.2115 |
| $0.00 \%<$ BAC $<0.049 \%$ | 0.2238 | 0.0711 | 3.1450 | 1.2508 |
| Non-Drinker | 0 | - | - | 1 |
| LGTM | 0 | - | - | 1 |
| ILGTM, >18 | 0 | - | - | 1 |
| ILGTM, <18 | 0.2951 | 0.0592 | 4.9827 | 1.3432 |
| LC-UN | 0 | - | - | 1 |
| Day Time | 0 | - | - | 1 |
| Night w/ Illum | 0.0705 | 0.0211 | 3.3425 | 1.0730 |
| Night w/o Illum | 0.1658 | 0.0370 | 4.4774 | 1.1803 |
| LTCDT-UN | 0.0829 | 0.0410 | 2.0234 | 1.0864 |
| SPLM, 25 | 0 | - | - | 1 |
| SPLM, 30 | 0 | - | - | 1 |
| SPLM, 40 | 0.0982 | 0.0455 | 2.1579 | 1.1032 |
| SPLM, 50 | 0.2346 | 0.0500 | 4.6970 | 1.2644 |
| SPLM, 60 | 0.3188 | 0.0486 | 6.5586 | 1.3755 |
| SPLM, 70 | 0.4693 | 0.0591 | 7.9358 | 1.5989 |
| SPLM, 80 | 0.8580 | 0.1143 | 7.5042 | 2.3584 |
| SPLM, 90 | 1.0195 | 0.0849 | 12.0047 | 2.7717 |
| SPLM, 100 | 1.0865 | 0.0639 | 16.9934 | 2.9638 |
| SCALE | 0.5848 | 0.0061 | - | 1.7946 |

Under a specific situation with a non-drinking legitimate driver, the velocity values of expectation and 85th percentile for a combination (8*3 $=24$ episodes) of speed limit by lighting condition are estimated in table 2. From the table, we can surmise that higher speed goes with higher speed limit and speeding violation happens frequently at lower or higher speed limit like an U shape, and darker environment increases the violating events significantly, while driving in the areas with 60 or $70 \mathrm{~km} / \mathrm{h}$ as speed limits has least possibility of speeding. According to the standard of traffic engineering, the posted value can prevent 85 \% of drivers from speeding, but there seems some speed limit change buffer for the areas where speed limits are equal to or under 30 and are equal to or above 80 . The limit might even be set differently between day and night taking speed limits 40 or $50 \mathrm{~km} / \mathrm{h}$ as examples.

Table 2. The Estimated Values of $\mathrm{E}(\mathrm{V})$ and $\mathrm{PER}_{85}(\mathrm{~V})$ by Speed Limit and Lighting Condition

|  | Day Time |  | Night w/ illumination |  | Night w/o illumination |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| SPLM | $\mathrm{E}(\mathrm{V})$ | $\operatorname{PER}_{85}(\mathrm{~V})$ | $\mathrm{E}(\mathrm{V})$ | $\operatorname{PER}_{85}(\mathrm{~V})$ | $\mathrm{E}(\mathrm{V})$ | PER $_{85}(\mathrm{~V})$ |
| 25 | 34.16 | 35.77 | 36.66 | 38.51 | 40.33 | 42.78 |
| 30 | 34.16 | 35.77 | 36.66 | 38.51 | 40.33 | 42.78 |
| 40 | 37.69 | 38.20 | 40.44 | 41.32 | 43.28 | 46.23 |
| 50 | 43.21 | 44.33 | 46.37 | 47.74 | 51.01 | 53.32 |
| 60 | 47.00 | 48.07 | 50.43 | 51.84 | 55.48 | 57.79 |
| 70 | 54.60 | 56.92 | 58.59 | 61.26 | 64.46 | 68.00 |
| 80 | 80.57 | 89.91 | 86.45 | 96.56 | 95.12 | 106.49 |
| 90 | 94.74 | 102.19 | 101.66 | 109.87 | 111.84 | 121.56 |
| 100 | 101.61 | 106.67 | 109.03 | 114.79 | 119.95 | 126.84 |

## 4. THE MODEL FOR THE EFFECT OF SPEED LIMIT ON CASUALTY VIA SPEED

The estimated speed for each driver can be obtained individually by substituting his data (including drinking status, license holding status, environmental lighting condition, and speed limit) into formula (2) with outputs from table 1 . Then the factors influencing speed can be implied in the estimated speed through which driver's casualty is further predicted. The speed limit to speed and then speed to casualty chain relationship can thus be explored. In other words, studying the resultant contribution of speed limit on casualty is possible from now on. In order to depict the precise effect of speed on casualty, we arbitrarily make the continuous value of estimated reconstructed speed into 8 equal frequency speed groups, and assume that there are different effects on casualty between groups but within groups. Allowing groups with similar effects can be united into a new group. The estimated descriptive results report 7 values of 8 equal-frequency group boundaries for estimated reconstructed speed (ERSP)- 38, $41,47,55,65,75,88$. These values are a little higher when compared to the original ones since the reconstructed speed with value 0 is omitted from further consideration at the last step (stage one).
The reason for choosing logistic regression as the candidate model at stage two is the binary outcome nature of two casualty levels (with and without). Also and its model formulation starts at assuming the logistic function of probability of having casualty, i.e., natural logarithm of odds of having casualty is a linear combination of related covariates including estimated reconstructed speed. This function measuring the tendency to being involved an accident with casualty is primarily called utility function in economics or transportation demand literature (e.g. Ben-Akiva and Lerman, 1997). Then the question becomes how to make a probability estimate with this tendency value (the tendency value of having no casualty is defaulted as zero for comparison and easy computation purpose), i.e., this estimate is equal to the logit function of that tendency value. The formulas indicating the randomness of tendency function

U , the logistic function, the probability estimate of having casualty (the logit function at the same time). The odds (or odds ratio) estimate of being involved in a casual event are list as following:
$\mathrm{U}=\mathrm{V}+\varepsilon$
$\mathrm{V}=\operatorname{logistic}(\mathrm{P}($ casualty $/ Y))=\ln (\mathrm{P}($ casualty $/ Y) /(1-\mathrm{P}($ casualty $/ Y))=\alpha Y$
AVGP(casualty $/ Y$ ) $=\operatorname{logit}(\mathrm{V})=1 /(1+\exp (-\mathrm{V}))$
LBP(casualty/Y) $=1 /(1+\exp (-(\alpha Y-1.96 \eta))$
$\operatorname{UBP}($ casualty $/ Y)=1 /(1+\exp (-(\alpha Y+1.96 \eta))$
$\mathrm{O}($ casualty $/ Y)=\exp (\mathrm{V})=\exp (\alpha Y)$
O (casualty $/ Y$ with all $Y$ 's elements $=0)=\exp \left(\alpha_{0}\right)$
OR(casualty $/ Y$ with $\mathrm{Y}_{\mathrm{i}}=1$ over $Y$ with $\mathrm{Y}_{\mathrm{i}}=0$, all other elements being equal)
$=\mathrm{O}\left(\right.$ casualty $/ \mathrm{Y}$ with $\left.\mathrm{Y}_{\mathrm{i}}=1\right) / \mathrm{O}\left(\right.$ casualty $/ \mathrm{Y}$ with $\left.\mathrm{Y}_{\mathrm{i}}=0\right)=\exp \left(\alpha_{\mathrm{i}}\right)$
$\eta=\left(\phi^{2}+\psi^{2}+2 \zeta \phi \psi\right)^{0.5}$
Where U represents random tendency function, V the deterministic part of $\mathrm{U}, \varepsilon$ the residual assumed as standard Weibull distributed, AVGP(casualty) $=\mathrm{P}$ (casualty) a average probability estimate for a driver being involved in an accident and injured, $\alpha$ an estimated vector of parameters (i.e., coefficients), $Y$ an vector of necessary independent variables (i.e., terms) which affect casualty possibility, LBP(casualty) a lower bound probability estimate, $\eta$ the standard deviation for the sum variable regarding estimated speed and a interesting variable, $\phi$ the standard error for estimated speed's parameter, $\psi$ the standard error for estimated that variable's parameter, $\zeta$ the correlation between the mentioned two estimated parameters, UBP(casualty) a upper bound probability estimate, O an odds estimate related to P , $\alpha_{0}$ the constant estimator, OR an odds ratio estimate of two interested groups one with Y's specific element $\mathrm{Y}_{\mathrm{i}}$ being equal to 1 , the other 0 , both assumed with all other elements being equal. When all significant variables are retained after removal of unsuitable ones step by step, table 3 shows a chosen estimation for interested parameters and other important information. Generally speaking, the model is moderately good in terms of near $70 \%$ of correct prediction by definition of geometric mean, which is equal to $\exp (-1682.4 / 4476)=0.687$. The educational level is the only one factor with categories of which none has effect significantly different form others. As derived from formula (13), when a crash happens, the group of drivers with basic conditions (a male 45 years of age or younger driver uses passenger car going straight ahead and hitting opposite direction the same vehicle type at neither right nor front impact near non- signalized intersection where speed limit is smaller than 41 or greater than $75 \mathrm{~km} / \mathrm{h}$ ) has odds of casualty being equal to 0.255 . The results of estimated parameters seem more or less in the right direction in term of the effects of certain categories being compared to those of the relevant defaulted categories, however unjustified relationship is found regarding the effects of driving or hitting other vehicle compared to those of two big vehicle types (truck or bus) and the effect of right impact compared to that of front impact, it is hard to image why both driving and hitting other vehicle types are more dangerous than
driving and hitting two big vehicle types, and why right impact is almost equal fatal to front one. Besides constant, there are 21 dummy variables being determined qualified for further analysis, one for gender (basic being male), one for ages (basic being equal to or under 45 years), three for subjective vehicle type (basic being passenger car), one for road type (basic being other types but signalized intersection), two for driving movement (basic being straight ahead), three for relative heading direction (basic being opposite), two for collision impact (basic being other sides but right or front), four for estimated speed (basic being under 41 or over $75 \mathrm{~km} / \mathrm{h}$ ), and four for hit object (basic being passenger car). When an accident occurs, the casualty odds ratios (shown in parenthesis after comparison description) for a certain category relative to its basic one among each variable can be revealed from the $\exp (c o e f)$ column in table 3. A female is more likely to be an injury victim at a ratio of 1.29, older driver (45 or more years of ages) is vulnerable to injury (1.29), motorcycle protects less from injury (47.1) while driver of a big car (consisting of truck and bus) or other type vehicle is not so influenced by outer impact ( 0.42 or 0.60 respectively). A crash at signalized intersection generates injury relatively rarely ( 0.61 ). Right (or right $U$ ) turn or left (or left $U$ ) turn does not seem a movement with an injury prone ( 0.49 or 0.75 ). When facing another vehicle from same or left or right direction one is more free from injury ( 0.45 or 0.63 or 0.57 in sequence). Right or front impact contributes more events of serious collision (1.44 or 1.50 respectively). Driving at four estimated speed ranges (41-47 or $47-55$ or $55-65$ or $65-75$ ) make a crash more severe ( 1.59 or 1.75 or 3.10 or 5.06 in sequence). Touching a large vehicle (i.e., truck or bus) or other vehicle type one suffers from greater injury impact ( 2.32 or 2.82 respectively). Meeting smaller objects (e.g. motorcycle or bicycle/pedestrian) is safer in terms of severity ( 0.284 or 0.145 respectively).
Further analyses with the effects of two or more variables are available based on the estimated results of the correlation matrix of parameters, hence, one possible example pertaining to the combined influence of gender and estimate speed is provided in table 4 if other variables are all predefined as being the defaulted categories. There are 6(estimated speed range groups)*2(gender groups) $=12$ possible values for each estimated average casual probability (AVGP), each estimated lower bound and each upper bound of 95 \% confidence interval (CI) for casual probability (LBP and UBP respectively). From the calculated outputs, there is no significant difference between the probability values regarding speed range groups 41-47 and 47-55 no matter in what gender groups, and the 95 \% CI of estimated probability becomes larger when estimated speed gets bigger except the last estimated speed group, which means the point estimator of probability turns not reliable gradually until the one before the last one, meanwhile, female has the tendency to being an injury victim more probably than male, however, clear discrimination is only found among the first two groups.

Table 3. A Logistic Regression for Estimated Speed to Casualty Probability Relationship

| LOG LIKELIHOOD $=-1682.4$ sample sizes |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| GOODNESS OF FIT CHI-SQ (2*O*LN(O/E) $)=1383.8$ D.F. $=1524 \quad \mathrm{P}=0.995$ |  |  |  |  |
| TERM | COEFFICIENT | Standard Error | COEF/SE | EXP(COEF) |
| Female | 0.2518 | 0.116 | 2.17 | 1.29 |
| Male | 0 | - | - | 1 |
| YOG, >45 | 0.2525 | 0.105 | 2.41 | 1.29 |
| YOG, 45 or $<45$ | 0 | - | - | 1 |
| Motorcycles (S) | 3.853 | 0.129 | 30.0 | 47.1 |
| Truck or Bus (S) | -0.8687 | 0.208 | -4.18 | 0.419 |
| OVT-S | -0.5072 | 0.254 | -2.00 | 0.602 |
| Small Cars (S) | 0 | - | - | 1 |
| SIG-INT | -0.4913 | 0.127 | -3.86 | 0.612 |
| Non SIG-INT | 0 | - | - | 1 |
| RT or RUT | -0.7138 | 0.313 | -2.28 | 0.490 |
| LT or LUT | -0.2922 | 0.134 | -2.19 | 0.747 |
| SA | 0 | - | - | 1 |
| Direction, Same | -0.7914 | 0.118 | -6.68 | 0.453 |
| Direction, Left | -0.4657 | 0.131 | -3.55 | 0.628 |
| Direction, Right | -0.5575 | 0.133 | -4.20 | 0.573 |
| Opposite | 0 | - | - | 1 |
| Impact, Right | 0.3653 | 0.153 | 2.39 | 1.44 |
| Impact, Front | 0.4074 | 0.976E-01 | 4.17 | 1.50 |
| Impact, Others | 0 | - | - | 1 |
| ERSP, 41 to 47 | 0.4651 | 0.105 | 4.44 | 1.59 |
| ERSP, 47 to 55 | 0.5624 | 0.148 | 3.80 | 1.75 |
| ERSP, 55 to 65 | 1.130 | 0.178 | 6.36 | 3.10 |
| ERSP, 65 to 75 | 1.621 | 0.494 | 3.28 | 5.06 |
| ERSP, Others | 0 | - | - | 1 |
| Motorcycles (O) | -1.260 | 0.122 | -10.3 | 0.284 |
| Truck or Bus (O) | 0.8430 | 0.180 | 4.69 | 2.32 |
| BCC or PDT (O) | -1.931 | 0.239 | -8.09 | 0.145 |
| OVT-O | 1.036 | 0.226 | 4.58 | 2.82 |
| Small Cars (O) | 0 | - | - | 1 |
| CONSTANT | -1.365 | 0.121 | -11.3 | 0.255 |

Table 4. The Estimated Values of AVGP, LBP and UBP by Estimated Speed and Gender

| Estimated <br> Speed | Male |  |  | Female |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | AVGP | LBP | UBP | AVGP | LBP | UBP |
| 41 or $<41$ | 0.2034 | 0.1677 | 0.2446 | 0.2473 | 0.1963 | 0.3065 |
| $(41,47]$ | 0.2891 | 0.2404 | 0.3432 | 0.3434 | 0.2768 | 0.4169 |
| $(47,55]$ | 0.3094 | 0.2456 | 0.3814 | 0.3653 | 0.2833 | 0.4567 |
| $(55,65]$ | 0.4415 | 0.3530 | 0.5339 | 0.5043 | 0.3992 | 0.6089 |
| $(65,75]$ | 0.5637 | 0.3279 | 0.7739 | 0.6243 | 0.3818 | 0.8173 |
| 75 or $>75$ | 0.2034 | 0.1677 | 0.2446 | 0.2473 | 0.1963 | 0.3065 |

The information from both table 2 and table 4 indicates that the effect of speed limit on casualty probability can be projected via the chain of the effect of speed limit on estimated speed and the effect of estimated speed on casualty probability. A general trend gives a comprehensive remark that the higher the speed limit, the faster the estimated speed, finally, the greater likelihood the crash generates casualty until the estimated reach $75 \mathrm{~km} / \mathrm{h}$, or the speed limit is equal to $70 \mathrm{~km} / \mathrm{h}$. Therefore, driving along higher speed limit areas promotes higher speed, in turn, higher estimated speed increases probability of severe crash except along the location where the design standard of roadway is highly elevated, e.g., expressway or freeway has entrance/exit control. This merit discourages the happening of mortal collision types, such as frontal impact from opposite direction and side impact from right or left direction.

## 5. CONCLUSION AND RECOMMENDATION

This research demonstrates how to estimate the effect of speed limit on casualty probability from crash through a mediator, estimated speed, which is the dependent variable at the first stage modeling speed limit to original speed relationship, while a independent variable at the second stage modeling estimated speed to casualty probability relationship. In order to keep the effect of influence clear without disturbance from other important factors, these factors are subjectively chosen at two different stages. The results show most of the factors have been quite significant except weather condition at the first stage and educational level at the second stage. According to the continuous nature of speed and discrete nature of severity outcome, ln-linear regression (or positive value regression since it preserves the positive attribute for the dependent value) and logistic regression (or ligit model, since the logistic function of probability of having casualty is a linear combination of independent variables, while the logit function of tendency to having casualty is the probability of having that casualty. These two functions are mutually inverse and the outputs can be used to predict if an observation is involved in a crash with casualty) are used and justified separately at the two stages.
The treatment that no zero value of speed is considered in the ln-linear model might lead to overestimating the absolute value of speed and jeopardizing the appropriateness of relevant
inference about estimated speed and its effect on casualty probability. This treatment needs further investigation in order to make a precise conclusion.

Due to the data selection procedure in this study, i.e., only the cases, which receive authentication twice, are included, the included episodes are so serious and thus so controversial that this deserves the attention of related committees again. One in local county, the other in the whole island level of Taiwan, therefore, the collision injury tends to be severe, and the implication of the analysis from the fitted model is limited.
The sign of the effect of chosen variable are mostly reasonable and consistent with the preliminary assumption, although intra-group effect difference within each variable is not so significant, and the effect magnitude is not so proportional to the degree of change in a certain variable and as expected, part of the resultant outcomes still provide valuable insight into how speed limit influences speed, and how estimated speed affects probability of having casualty. From the comparison of the table 2 and table 4, there seems some buffer of changing speed limits applicable to areas with values of 80,90 , or $100 \mathrm{~km} / \mathrm{h}$, where the casualty likelihood is not elevated due to higher design standard on expressway or freeway. Most interestingly, the speed limits can be designed differently for three lighting conditions one with daytime light, the other with illumination at night, and another without illumination at night. What the proper speed limits will be along these high standard roadways needs more objective evidence.
The application of the chain model allows finding the indirect effect of speed limit on crash severity, and this effect has physical implication in practice. If at a particular location a threshold level of probability of generating crash injury is to be maintained under an intervention from a countermeasure, the average traffic speed near this location should be controlled under a certain range using the outputs from the second stage model. Then specifying and rearranging the speed limit according to the relationship of first model would be a possible intervention.
Further analysis of this research reveals that on the local highways, the ratio of injury likelihood is nearly proportional to the ratio of square speed. This relationship is close to that of Evans (1991), and therefore, there is a dilemma between mobility and safety, driving faster means possibly a more serious crash, while slower means less serious.

## REFERENCES

Evans, L. (1991) Traffic Safety and the Driver. Van Nostrand Reinhold, New York.
Najjar, Y.M., Russell, E.R., Stokes, R.W. and Abu-Lebdeh, G.. (2002) New speed limits on Kansas highways: impact on crashes and fatalities, Transportation Quarterly, Vol. 56, No. 4, 119-147.
Sunanda, D. and Lu, J.J. (2002) Factors influential in making an injury severity difference to older drivers involved in fixed object-passenger car crashes, Accident Analysis and Prevention, Vol. 34, 609-618.

Bedard, M., Guyatt, G.H., Stone, M.J. and Hirdes, J.P. (2002) The independent contribution of driver, crash, and vehicle characteristics to driver fatalities, Accident Analysis and Prevention, Vol. 34, 717-727.
Al-Ghamdi, A.S. (2002) Using logistic regression to estimate the influence of accident factors on accident severity, Accident Analysis and Prevention, Vol. 34, 729-741.
Scuffham, P.A. and Langley, J.D. (2002) A model of traffic crashes in New Zealand, Accident Analysis and Prevention, Vol. 34, 673-687.
Wilmot, C.G. and Khanal, M. (1999) Effect of speed limits on speed and safety: a review, Transport Reviews, Vol. 19, No. 4, 315-329.
Yang, C.Z. and Ai, C.M. (2003) A comprehensive study about the influence of speed limits on speed and safety. Research Report No. NSC-91-2211-E-035-023, National Science Council, Taipei, Taiwan, R.O.C.
Cooper, G.W. (1990) Work, energy, and speed from damage in traffic accidents. Topic 870 of the Traffic Accident Investigation Manual, Northwestern University Traffic Institute.
Collett, D. (1994) Modeling Survival Data in Medical Research. Chapman \& Hall, London. Ben-Akiva, J.P. and Lerman, S.R. (1997) Discrete Choice Analysis: Theory and Application to Travel Demand, Seventh Printing. Cambridge, Massachusetts.

