Targeting speeders on Perth roads

Zuhair Ebrahim1 and Hamid Nikraz2

1,2Civil Engineering Department - Curtin University, Western Australia
GPO Box U 1987, Perth Western Australia, 6845
(+61) 8 9266 6245 ; fax (+61) 8 9266 2681
e-mail: zuhairebrahim60@yahoo.com.au

ABSTRACT
Speeding on the road is not only a concern for developed countries i.e. USA, Europe and Australia, but continuing to be of alarming concern for many countries in Asia, Africa and the Middle East. This paper aim is to focus on highlighting speeders characteristics on 40 km/h on a busy urban road with high pedestrian movement. Thus, authorities decided to drop the limit on that road to 40 km/h from 60 km/h and introduced an electronic variable sign as an engineering measure. Speeding data was utilised from the WA (Western Australian) Police to examine the factors that may contribute to speeding. This data will be applied to compute the logit from the independent variables such as age, gender, time of day. Detection types are either ‘on the spot’ type where infringement notices are delivered on site to offenders, or ‘roadside’ type where notices are posted to offenders. Four speeding levels were recognised (low, medium, high and excessive) that was employed as a dependent variables. The low level speeding was considered as a reference category. All the above variables were modelled using multinomial logistic regression. A stepwise procedure included all the selected independent factors. Results showed that the likelihood of detected speeding does decrease with age and it also showed that this is a slightly stronger indicator for more dangerous speeding. Thus, younger drivers are slightly likely to drive with high speed. Age ranges between 17-83 years. Female drivers are less likely to be detected speeding than male drivers. As for the time of detection, it seems that drivers are prone to be detected during afternoon time more than morning time. This was consistent with road crashes between 2007-2011 road crash data on the same study road. Significantly, it was found that after the period of the installation of the electronic 40km/h electronic flashing signs, drivers are less likely to speed compared to before the installation adding to the model. The period of testing the signs benefits in reducing speeding may be short and may need more time to determine. Further to the above, the model needs to be simulated for further testing.

INTRODUCTION
Speed control is now a major road safety priority throughout the world. Excessive speeding is a leading cause of fatalities on WA (Western Australian) roads. According to the office of road safety in WA, 40 to 45% of drivers regularly drive above the posted speed limit. A report by the World Health Organisation, found that speed of motor vehicles is at the core of the road injury problem. It found that speed influences crash risk and their consequences, [1] Paden et al. (2004). A review of international speed limits (compared to other OECD countries) found that Australia’s limits tended to be higher than those found elsewhere, including in Europe and North America [2]. Australian speed limits have traditionally been set at the speed that 85% of drivers would choose when driving along a particular road section, but many now believe that this criterion is becoming a barrier to achieving the crash injury reduction targets posed by road safety strategies [2]. They recommended that to reduce road injury rates further, more must be done to reduce driving speeds in Australia, and lowering speed limits may be a critical component in achieving this outcome. The Austroads report revealed that EU countries have utilized harm minimisation principles as the basis for setting speed limits.

Literature search found that, on December 2001, the speed of 50 km/h was adopted as a default speed limit in the built-up area of Western Australia [3]. The primary motivation for introducing this speed reduction program in WA was the expectation that a 20 to 30 per cent reduction in serious crashes could be achieved, benefitting mostly the more vulnerable users of local area road networks such as pedestrians, children and bicyclists. This expectation was based on similar gains achieved in both local and overseas jurisdictions.

A study on a 50km/h speed limit by [4], found that the percentage of Perth drivers travelling 10 or more km/h above the speed limit was estimated to be 18.2 in 2008. In comparing it with Sweden, it was found that 11–46% of all drivers exceed the limit with 10 km/h or more of different speed limits, [5]. Most recent data shows that there is a far bigger problem in Perth, involving thousands of “good” drivers who took frequent smaller risks such as low level-speeding, Holman [6]. A recent pedestrian study in Israel concluded that vehicle–pedestrian conflicts need to be minimized by significantly reducing vehicle speeds in areas of pedestrian presence and activity [7]. Main Roads authority of WA found one stretch of road that had higher pedestrian accidents than any other road in the metropolitan area. This road had a posted speed of 60 km/h limit and they decided a trial to install 40 km/hr electronic speed signs along that road, see Figure 1 below.
The purpose was to introducing a variable speed zone by lowering travel speeds from 60 km/h to 40 km/h during peak periods of pedestrian activity. The initiative is a trial to improve safety for pedestrians and other road users. The signs stayed displaying 40 km/hr from 7.30 am till 10.00 pm (except Friday and Saturday they are extended till 1:00 am instead of 10:00 pm). The same signs will display 60 km/h outside the above mentioned time. Concerns were raised due to the high number of speeding fines [8]. The most severe increase came in the number of speeding motorists in 40 km/h (non-school zone such as this road understudy). There were 96 drivers were caught daily in the 40 km/h zone compared to the 83 the previous year. If drivers’ speeding behaviours continued to a similar level in that zone, it will have an increase by 4745 speeding fines annually in that stretch of roads if used the same enforcement intensity. According to [9] from South Australia, it was concluded that one - third of licensed drivers were caught speeding in 2007. Thus, for the importance of this issue, this paper will focus on testing the TIN’s (Traffic Infringement Notices) data, applying MLR (Multi Logistic Regression) to these speeding levels and four other explanatory variables, age, gender, time being detected speeding and installation phase before / after installation of variable signs.

MLR BACKGROUND

Logistic regression is valued as an important analysis tool to roads safety studies. It is a source of attraction for researchers in targeting the effect to dependent variables on a binary dichotomous outcome or multi-level outcome such as the case of MLR understudy. For example, in terms of speed and speeding concerns, [10] found that higher speed limits were contributing to higher risk levels in China. Whereas [11] in Saudi Arabia and using odds ratio concept as an estimate of the logit, found that speed as a cause of accident factors were significantly associated with accident severity. According to [12] in Australia, with using logistic regressions found there was no interaction between the effect of TINs and speed related publicity awareness on the frequency of casualty crashes. It further found that concluded that the increased level of speed camera ticketing was associated monotonically with road crash reduction. Demographics such as age and gender were also used in research studies. For instance, according to [13] and using a log-linear regression analysis, found that driver age was a contributing factor to crash severity. It was stated further by [14] that the striking drivers with relatively larger accident propensity tend to be male drivers, younger <26 years, or older drivers >75 years. For striking drivers, accident propensity appeared a decreasing trend with increasing age until the age of 56–65, and then increase to a higher accident involvement for the age group older than 75.

Gender, on the other hand was also used as a prominent explanatory variable in studies such as those found in [15; 16 &17] in older female drivers attributing the contributing causes to due to poor attention, cognition and other old age related reasons. In terms of time of day effect, [18] found that there was an impairments in driving performance in the early afternoon that are similar magnitude to those occurring in the late evening and early. On the contrary [19] in studying the industrial accidents in Queensland Australia, found that more injuries occur in the morning in Australia than in the afternoon on every working day of the week. In Spain, [20] concluded that most accidents in construction industry were more towards the afternoon periods attributing it to the “lunch effect” the study added that these accidents occurred in the hours around the lunch break (from 13:00 to 17:00 h) and are of greater severity and involve more fatalities than those that occur at other times of day. Other dimension was added by [21] in Australia in a survey of the construction workers reported that construction industry involves the consumption of alcohol which explain the afternoon accidents involvements. This current study has included morning and afternoon time to be included as explanatory variable in the model.
MODELLING APPROACH

It is known that driving speed is a continuous variable, which can be modelled with standard regression. This study is modifying the speeding variable and categories those speed data into four speeding levels (low, medium, high and extreme). This speeding data is considered as ordinal and according to [22] using unordered model such as multinomial logit Regression model to model ordered data. It added that more caution is needed if decided to use ordered model as it can lead to a serious biases in the estimation of the probabilities, unlike the use of unordered model such MLR where the model parameters estimates remain consistent, but there is a loss of efficiency. In addition to that [23] believe that if using the ordered probability models such as ordinal models, it will place restriction on how variables in X affect outcome probabilities therefore and due to the above reasons, unordered logistic model MRL will be adopted for this study despite the speeding levels of ordinal and ordered dependent variable. The study for the Multinomial Logistic regression will deal with three discrete outcomes as the low speeding levels will be used as a baseline (reference) level. The speeding outcome will be shown as described in Table 1.

<table>
<thead>
<tr>
<th>km above speed limit</th>
<th>Speeding level</th>
<th>Risky speeding outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;9km</td>
<td>Low</td>
<td>Lowdanger*</td>
</tr>
<tr>
<td>10km-19km</td>
<td>Intermediate</td>
<td>Moderatelydangerous</td>
</tr>
<tr>
<td>20km-29 km</td>
<td>High</td>
<td>Highlydangerous</td>
</tr>
<tr>
<td>&gt;30 km</td>
<td>Extreme</td>
<td>Excessively dangerous</td>
</tr>
</tbody>
</table>

* Reference category

The study is aiming to test the odds probability of the dependent variable is speeding level occurring as the values of the IVs change. In other word, the objective is to find out the magnitude of risk levels due to high speeding levels. This comes from the well-known concerns in the transportation literature about speeding above the limit which could result in serious and more severe crashes. Thus, the Y event is very unlikely to occur if f(Y) is close to 0 and that it will be unlikely that the driver will engage in risky behaviour. On the contrary, it is very likely to occur if f(Y) is close to 1. Low level speeding 0-9 km above posted speed limits is considered reference category compared to other three higher risky levels.

In order to build the model, the first step is to construct the model to get a logit score Y, then to the study will check further the predicted probabilities and scatter them against the Pearson residuals that the model produces.

The logit function is given by equation 1, that is derived from different predictors multiplied by their corresponding regression coefficients, where Y is calculated as follows:

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \]

Where \( \beta_0 \) is the intercept of the value of Y when all the predicting variables \( X_1, X_2, X_3 \ldots X_n \) are equal to zero. The variable Y (logit) is a measure of the sum of the input of all the independent predictor variables used in the model. The variable Y is defined as:

Where, Speeding = Four levels, and low levels is a reference category. Compared to other three speeding levels.
Age = Recorder age of detected speeding driver. Continuous variable.
Gender = Gender of detected speeding driver. Categorical variable, 0 = female, 1 = male.
Time = Time of day speed detected. Categorical variable, 0 = morning (6.00-11.59) am,
1= afternoon pm (12.00-19.00) pm.
Phase = Period of before and after electronic signs installed. Categorical variable, 0= before installation, 1= after installation.

Further to equation 2, it was found by [24] that SPSS would compare the reference category with other categories consecutively in pairs by This would results in three pairs, each comparison is using to the low level speeding category. Since this would calculate the log-odds, it is appropriate to calculate the probabilities of that logit, by using equation 3.

Where f(Y) is the probabilities of speeding occurring represented by Y and e is the base of natural logarithm. For several predictors, the equation would become

Substituting Y of equation (2) in (4), with all predictors, the final probability calculated will become:

(5)

Speeding data will therefore be calculated through MLR using speeding risk and their corresponding probabilities. Thus, SPSS output are evaluated and discussed below.

MODEL EVALUATION
Three evaluation steps are discussed below that involves different aspects of evaluations. Firstly the crucial fitting information that determines the model good fit, then the parameter estimates testing the coefficients and the odds ratios of the model and finally plotting of Pearson residuals against the predicted probabilities.

Model Fitting Information
Details in Table 2 below discuss the Log likelihood, AIC, Goodness of fit and other Pseudo $R^2$. It can be noted that all details are encroaching showing a good fitting model if these explanatory variables are used.

<table>
<thead>
<tr>
<th>Model fitting details</th>
<th>value</th>
<th>$p$</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2 LL ($x^2$)</td>
<td>236.48</td>
<td>0.001</td>
<td>The change is significant and it explains the decrease in unexplained variance and it is considered as a good improvement to the model.</td>
</tr>
<tr>
<td>AIC</td>
<td>1877-1677=200</td>
<td>-</td>
<td>Value of AIC has lowered when model introduced indicating a good fit.</td>
</tr>
<tr>
<td>Goodness of fit</td>
<td>Pearson =1268, Deviance=917</td>
<td>0.14</td>
<td>The predicted values are not significant and not different from the observed, thus the fit of the model is good.</td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
<td></td>
<td>1.00</td>
<td>Fairly similar values and fairly reasonable representing a good size effect according to [24].</td>
</tr>
</tbody>
</table>
Parameter estimates
In this section, three comparisons are executed revealing the parameters values of the mode. Most importantly, Table 3 below shows details of significant parameters, particular attention needs to be paid to the Odds ratios $\text{Exp}(B)$ for interpretation of the model values and the decrease and increase of the predictor's effect on the risky speed outcome.

Table 3: Model parameter estimates

<table>
<thead>
<tr>
<th>Predictors including interactions</th>
<th>B (SE)</th>
<th>Wald $x^2$</th>
<th>95 CI limits</th>
<th>Lower</th>
<th>Upper</th>
<th>Odds Ratio</th>
<th>$Exp(B)$</th>
<th>P -value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Moderately dangerous vs. Low danger</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.41(0.27)</td>
<td>80.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>Gender = 0</td>
<td>0.68 (0.30)</td>
<td>5.05</td>
<td>1.09</td>
<td>3.54</td>
<td>1.96</td>
<td>0.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender = 1</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time = 0</td>
<td>-1.11(0.21)</td>
<td>27.46</td>
<td>0.22</td>
<td>0.50</td>
<td>0.21</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time = 1</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase = 0</td>
<td>0.23(0.28)</td>
<td>0.73</td>
<td>0.39</td>
<td>1.26</td>
<td>1.26</td>
<td>N/S*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase = 1</td>
<td>0.74 (0.30)</td>
<td>6.11</td>
<td>1.17</td>
<td>3.92</td>
<td>2.10</td>
<td>0.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.01 (0.01)</td>
<td>3.50</td>
<td>0.98</td>
<td>1.00</td>
<td>0.01</td>
<td>N/S*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time*Period</td>
<td>2.29(0.58)</td>
<td>15.55</td>
<td>3.17</td>
<td>31.11</td>
<td>9.93</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Highly dangerous vs. Low danger</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.23</td>
<td>0.32</td>
<td>14.49</td>
<td></td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender = 0</td>
<td>0.84 (0.33)</td>
<td>6.66</td>
<td>1.22</td>
<td>4.37</td>
<td>2.31</td>
<td>0.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender = 1</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time = 0</td>
<td>-1.26 (0.26)</td>
<td>23.00</td>
<td>0.17</td>
<td>0.48</td>
<td>0.29</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time = 1</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase = 0</td>
<td>0.74 (0.30)</td>
<td>6.11</td>
<td>1.17</td>
<td>3.92</td>
<td>2.10</td>
<td>0.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase = 1</td>
<td>-0.02 (0.01)</td>
<td>8.81</td>
<td>0.97</td>
<td>0.99</td>
<td>0.98</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The value of Y calculated from equation 2, is equal to the value of β when all other independent variables are equal zero. The coefficient is either a positive or a negative one, indicating an increase or a decrease of the probability respectively. Similarly, if the coefficient is high or low means the risk due to speeding is strongly or weakly affecting the outcome respectively [25]. Therefore the parameter estimates are as follows:

**Gender**: the coefficient of a person gender (0=female, 1=male) is significantly predicted whether they drove with moderately or highly dangerous speeding relative to low danger speeding. The coefficient of extreme speeders relative to low danger speedings is found to be not significant. For these two significant coefficients and as female relative to males are found to be 0.68 and 0.84 with odds ratio > 1. Thus female drivers are less likely to be detected speeding than male drivers. The odds of a female driver moderately dangerous speeding compared to low danger speeding (1/odds ratio) = (1/1.96) = 0.51 times less than for a male driver. Similarly the odds of a female detected with highly dangerous speeding compared to detected with low level danger speeding is (1/2.31) = 0.43 times less likely than male driver detected.

**Time**: the coefficients for time of detection, where 0 = morning and 1 = afternoon shows to be significant predictors in all three binary comparisons. The negative coefficients are -1.11, -1.26 & -1.67 are decreasing the probability of the outcome which is moderately dangerous, highly dangerous and extremely dangerous speeding respectively. Since odd ratios for the three comparison are <1, this indicates that afternoon time is more prone for driver to be detected than morning by (1/0.21), (1/0.29) & (1/0.19) respectively. This means that driver could be prone to be detected during afternoon time by 4.76, 3.45 & 5.26 times than morning time respectively.

**Phase**: this coefficient is of high importance to this study since it determines the usefulness of using the electronic flashing 40km/h sign in reducing speeding. Period variable is used to determine the speeding levels difference between the two periods before installation and after the installation of the electronic 40km/h electronic flashing 40km/sign. It was found that period speeding is significantly predicted whether they drove with moderately or highly dangerous speeding relative to low danger speeding. Whereas the coefficient of extreme speeders relative to low danger speeding is found to be not significant. Use comparing highly dangerous driving and extremely dangerous driving with low level speeding with from the two coefficients. The coefficients for periods are before = 0 and 1 = after. They are 0.74 and 1.36 are with both odds ratios > 1. Thus before are before are less likely to speed compared to The odds for a driver to highly dangerous speeding compared to low speed before the installation of the signs is driver danger speeding (1/odds ratio) = (1/2.10) = 0.48 times less than for after driver. Similarly the odds of a driver detected with highly dangerous speeding compared to detected with low level danger speeding before the installation is (1/3.91) = 0.26 times more than for after installation.
Age: the coefficient for high speeders compared to low danger is found to be -0.02 and the odds ratio (Exp B) is 0.98. Thus, if the speeders age is increased by one unit (year), the likelihood of detected with high speeding versus low speeding is expected to decrease by 0.02 units. Younger drivers are slightly likely to drive with high speed. In the case of extreme speeding, the coefficient is found to be -.05 with an odds ratio of 0.96. This means that, if the speeder age is increased by one year, the likelihood of detected extremely dangerous speeding versus low speeding is decreased by 0.05 units. Hence as age increases the likelihood of dangerous speeding is reduced slightly. In summary, it can be noticed that both age coefficients are negative indicating a decrease of the probability. Both of the coefficients are of low values suggesting little effect on the probability of the outcomes.

Predicted probabilities

Graphically, it can be seen in Figure 2 below that the random scattered of the Pearson residual data against the predicted probabilities showing no trend or pattern formed [26; 27].

CONCLUSIONS

Four predictors were modelled, three of them namely age, gender of the driver and time of detected speeding are explanatory predictors whereas the forth one is testing the use of electronic sings and it is complimenting the other three predictors. Data showed that the likelihood of detected speeding does decrease with age and it also showed that this is slightly stronger indicator for more dangerous speeding. Thus, the younger drivers are slightly likely to drive with higher speed that the limit as age increases. Female drivers are less likely to be detected speeding over the limit than male drivers. The odds of a female driver moderately dangerous speeding compared to low danger speeding is 0.51 times less than for a male driver and is 0.43 times less likely than male driver to be detected of more highly dangerous speed. As for the time of detection, it seems that drivers could be prone to be detected during afternoon time than morning time. If coupled with risk factors, road crashes may occur in the afternoon noon more than the morning. It would be of benefit to read further into the time period and its risk factors to control speed management to protect road users particularly the safety of vulnerable ones such as the pedestrians as mentioned in [28] also along such busy shopping strips such the strip under study emphases in [29] under the safe system aspirations in such areas such high pedestrian activity.

It is paramount to mention that road crashes between the years 2007-2011 were consistent with findings the analysis showed that between mornings period of (6:00 - 11:59) am pedestrian crashes consist of nearly 33% if compared with the afternoon from (12:01 - 5:59) pm. whereas all other crash types consist of around 25% in the morning compared with the afternoon period.

It was found that after the period of the installation of the electronic 40km/h electronic flashing signs, drivers are less likely to speed over the limit compared to before the installation. The odds for a driver to highly dangerous speeding compared to low speed before the installation of the signs is driver danger speeding is 0.48 times less than for after driver. Similarly the
odds of a driver detected with highly dangerous speeding compared to detected with low level danger speeding after the installation is 0.26 times more than for after installation. The period of testing the signs benefits in reducing speeding may be short and may need more time to determine. In addition to the above conclusions it was stated by [6] that drivers are generally slowing down which may indicates a safer speed culture may be spreading. To conclude, the results of the study may be beneficial to determine drivers’ age that can contribute to risky behaviours and factors of vulnerable road users including their gender such male drivers and particularly young age drivers. Time of the day may shade more light on how drivers speed and how it can be monitored to detect more speeders in endeavour to deter them including the way to provide engineering remedy to those crucial hours such the introduction with the lower speed limit such as 30km/h limits to ensure risk elimination to pedestrians. Further, the model needs to be simulated to further tests its workability and which need to be discussed in the next stage of this research study. In any events, the study was constrained with certain limitations that are discussed below.

**MODEL LIMITATIONS**

Certain limitations are worth mentioning such as:

- The period of the study need to be expanded to get a reliable data particularly the before / after data need to be unified on same number of hours bases.
- The experience of the driver licence need to be included along with other explanatory variables, this will contribute to target licencing and procedures involves novice drivers licencing, training, etc.
- The study looked purely into speed as a risk factor and ignored other safety issues e.g. seatbelt or unsafe vehicle etc.
- Extra information on driver history and repeat offending need to be included that could have added to the variables.

**ACKNOWLEDGEMENT**

We would like to thank Ms Chris Canny the Assistant Director of the Academic Development of the WA Police, for her valuable support and acknowledgement of this paper. Thanks to the Academic Research Administration Unit of the WA police for supplying the data and clarifying labelling and other data inquiries. We are also indebted to Thandar Lim of Main Roads WA for making available the roads crash data.
REFERENCES


