Improving performance of public transit buses by minimizing driver distraction

K. A. D'Souza & S. K. Maheshwari
School of Business, Hampton University, USA

Abstract

This paper explores the problems of distracted driving for bus drivers at a local transit agency and uncovers factors that may cause the distractions. Data was collected on sources of driver distraction and perceived risks associated with a particular distracting activity along with potentially related independent variables like location, driving hours/week, age, gender, and driving experience. The seven highest distracting activities were categorized into three *risk zones* using a risk range system derived from the average distracting rating, average distracting duration, and driver's perception of risk. Multinomial logistic regression was utilized to model each *risk zone* distracting activity using levels of distraction as the dependent variable and correlating it with the factors as independent /predictor variables. A stepwise procedure included all the selected factors in the model initially; non-significant factors were eliminated until a good fit was achieved with significant factors. The model's goodness of fit was statistically tested and further verified graphically. The multinomial logistic regression outputs were analyzed for all seven *risk zone* distracting activities. Due to space limitation, an analysis of the highest risk distracting activity involving passenger using mobile phone is included in the paper.

The results revealed that the common sources of driver distractions were due to passenger-related activities. The male drivers are more likely than female drivers to get distracted by passengers, while female drivers are more likely to get distracted by the ticket machine than male drivers. Older drivers are less likely to get distracted by the ticket machine and passenger-related activities, although more driving experience increased the likelihood of distraction by passengers and ticket machines. The drivers with higher weekly driving hours are less likely to get distracted by ticket machines and climate controls. The recommendations made on the basis of the results could be used as a potential training tool to mitigate driver distraction and improve bus transit performance.
Keywords: sources of transit bus driver distraction, modelling bus driver distraction, multinomial logistic regression, stepwise procedure, predicting driver distraction risk, risk range system, risk zone.

1   Introduction

The US Department of Transportation [1] defines distracted driving as any non-driving activity a person engages in that has the potential to distract him or her from the primary task of driving and increase the risk of crashing. Distraction is categorized as: visual (taking your eyes off the road), manual (taking your hands off the wheel), and cognitive (taking your mind off what you are doing). In 2009, distracted driving claimed over 5,000 lives and nearly half a million injuries across the US [2]. Federal and state government have responded to this problem by introducing regulatory policies. Such policies may not be consistent due to insufficient information on transit bus driver distraction. Most studies on distracted driving have focused on personal and commercial vehicle drivers [3–5]. Research on transit bus driver distraction is limited although between 1999 and 2005, transit buses accounted for one third fatal crashes among all bus types [6]. In the case of passenger vehicles, most of the distraction is within the control of the driver. However, for transit vehicles, some distractions are caused by factors beyond the driver's control such as operating equipment or attending to passengers [7]. The accident reports filed by transit bus drivers rarely document distraction as the cause of accidents. For example, in the Commonwealth of Virginia, only 1.72% bus accidents were reported due to driver distraction and 59.28% due to no violations [8], possibly inclusive of distraction related accidents. Due to lack of reporting distractions by transit drivers, the associated risks and impact on performance is not well understood and consequently is difficult to study.

This paper presents results of an exploratory study on driver distraction conducted at the local bus transit agency by Hampton University's *Eastern Seaboard Intermodal Transportation Applications Center (ESITAC)*. The objective was to identify main sources of driver distraction, quantify risks associated with these distracting activities, and correlate potential independent factors with risk of distraction. It expands upon earlier studies on transit bus driver distraction [7–9] by identifying the high risk distracting activities and factors such as location, driving hours/week; and age, gender, and experience of the driver that are related to distractions. The seven distracting activities that could pose a safety hazard to transit operation were classified into *risk zones* and analyzed further by multinomial logistic regression models to estimate the magnitude and direction of factors on the level of distraction. Due to limited space, the statistical analysis for the highest *risk zone* distracting activity (passenger using mobile phone) is discussed in details and an overview of the results is presented for the remaining activities.

The following presentation of the paper is organized as follows: a literature review is presented in Section 2; the methodology for the study is explained in Section 3; in Section 4, the statistical analysis of data is conducted and results are discussed; Section 5 contains conclusions drawn from the results and recommendations for mitigating driver distractions; finally, Section 6 discusses the limitations of the study.

2   Literature review

Driver distraction poses a significant safety problem in the personal and public transport sector drawing the interest of several researchers. A study funded by the AAA Foundation [3] identified the major sources of personal vehicle distraction contributing to crashes, developed taxonomy of driver distractions, and examined the potential consequences of these distractions on driving performance. The source of bus driver distractions at a major Australian public transport company was investigated using ergonomics methods through which, a taxonomy of the sources of bus driver distraction was developed along with countermeasures to remove/mitigate their effects on driver performance [7].

Demographic factors such as driver age, driving experience, and gender could impact distracting behaviour; however these studies have produced mixed results. According to National Highway Traffic Safety Administration (NHTSA), the less than 20 age group had the greatest proportion of distracted drivers [2], but Green *et al.* [10] found no consistent normal/distracted differences in specific driving parameters for each road type by driver age combination. Older female drivers were at increased risk of crash due to poor attention cognitive, executive and motor skills [11]. The age group > 75 years presented the highest risk due to age-related problems with physical and cognitive abilities [12]. A chi-square analyses to determine the relationship between exposures to distracted driving and driver demographics (age, race, and gender) showed significant associations in reporting distracted driving by the passenger in the vehicle [4]. Blower *et al.* [6] reported that age, sex, hours driving, trip type, method of compensation, and previous driving records are related to driver error. A multivariate model applied to study accidents in trucks, identified driver age and gender among the several other factors related to rear-end crashes [13].

The location of the driving route is likely to have an impact on the source of distraction. A driver who is less familiar with the driving route is more likely to be involved in rear-end accidents at signalized intersections [12]. A densely populated area would have greater number of passengers and higher external sources of distraction due to more frequent stops, other road users or pedestrians, work zone activities, and toll booths etc. [3].

Multinomial logistic regression models are widely used in transportation to study the relationship between dependent variable having more than two unordered categories and a set of continuous and categorical independent predictor variables [13–16]. A multinomial logistic regression model that included a categorical dependent variable having four categories (very satisfied, satisfied, somewhat dissatisfied, and very dissatisfied) with 12 independent variables was developed to identify the factors contributing to service quality and customer satisfaction in a public transport system [14]. Washington *et al.* [15] developed a multinomial logistic regression model consisting of 18 independent variables covering driver factors, traffic flow, distance, number of signals etc. to study factors that influence driver's selection of route on their morning commute to work. Yan *et al.* [13] utilized a multinomial logistic regression model to study the impact of potential factors such as driver factors, road layout, and

environmental conditions on rear-end truck to car, car to truck, and car to car crashes. A multinomial logit model was developed to investigate the socio-economic and demographic factors that significantly affect passenger satisfaction with airport security screening process [16], and to determine the choice of aircraft in the US, National Airspace System [17]. Odds ratios were applied to identify high risk tasks/variables [18] and used along with magnitude and direction of the coefficients of the independent variables [15, 19−21] to infer results from the multinomial logistic regression parameter estimates.

3   Methodology

A self-administered survey instrument developed by Salmon *et al.* [9] was redesigned to collect distraction data from a sample of drivers employed at a local transit agency. The survey was totally anonymous and completed voluntarily (18% response rate) by drivers wanting to participate in the study. The driver location was recorded according to the geographical districts covered while driving the bus: Northside and Southside. The survey collected driver demographics and driving pattern, source and extent of distraction, duration of distraction, and perceived effect of distraction. The ratings and durations for each activity were averaged and then each activity was ranked based on average rating as well as average duration. The activities involving perceived visual, manual, and cognitive effects risk to drivers were ranked based on the aggregate count. The seven activities belonging to the top five average distraction rating, average distraction duration, and driver's perception of risk were classified into *risk zones* using a risk range system.

The *risk zone's* distracting activities were further analyzed using multivariate regression to determine the impact of factors on the distraction levels. The categorical dependent variable, driver distraction had four levels: not distracted, slightly distracted, distracted, and very distracted. This is similar to the categorical dependent variable (very satisfied, satisfied, somewhat dissatisfied, and very dissatisfied) used for studying customer satisfaction in public transport systems [14]. The independent variables included categorical variables such as gender, location, shift etc., and continuous variables such as age, driving experience, and driving hours per week. The multinomial logistic regression model was applied to compute the logit from the independent variables.

The advantages of using multinomial logistic regression are the following: it is a multivariate model requiring fewer underlying assumptions such as normally distributed data or equality of variances necessary for multiple regression models. The survey-response data appear ordered (ranging from being Not Distracted to Very Distracted); hence one might consider using the ordinal logistic regression model. The ordinal logistic regression models place a restriction on how the variables affect outcome probabilities [15]. An ordinal logistic regression test model was developed for the passenger using mobile phone activity. However, the final model output exhibited a poor fit (p = 0.381) with no significant independent variables. Due to this limitation and the reported problems with the ordinal logistic regression model [15, 16], the unordered

logistic model (multinomial logistic regression) was utilized in this study although the dependent variable appears to be ordered.

4 Data analysis and results

An analysis of historical bus accident data for the past two years revealed different accident rates in the transit agency's Northside and Southside districts. The accidents/million miles were higher in the larger and more densely populated Southside (61.87) compared to the Northside (53.54). Out of the 520 Northside (Southside data was not available at the time of study) incidents and/or accidents that occurred over the past two fiscal years, 164 (32%) accidents were identified as preventable. A portion of these preventable accidents could possibly be due to driver distraction.

4.1 Identification of high risk activities

The survey instrument asked the transit agency's bus drivers to indicate how distracting they find the listed activities and approximately how long per shift (in a typical 8-hour shift), do they experience the distracting activities while driving the bus. The ratings and durations for each activity were averaged and ranked from highest to lowest [4, 9] average values. The *risk zone I, II, III, AND IV* ranges were established at 90%, 70%, 50%, and less than 50% of the highest value (Table 1). The survey also collected information on the driver's perception of visual, manual, and cognitive distraction. The activities were graded as a percentage of the highest visual (19 drivers), cognitive (33 drivers), and manual (11 drivers) effects of distraction and ranges were established in Table 1. Using the risk range system in Table 1 for the ratings and durations, and perceived effects of distraction, all the 18 distracting activities were classified into their respective *risk zone* (Table 2). The *risk zone* I has the highest distraction risk and *risk zone* IV has the lowest distraction risk. The seven high risk activities identified by the risk range system in Table 1 were classified into *risk zones* I, II, and III and the remaining 11 activities that caused minimal distraction risks were classified into *risk zone* IV (Table 2).

Table 1:　　Risk range system.

Average rating	Average duration	Number of driver's perception (out of total 48 drivers)	Risk zone
> 2.2	> 2.4 hrs	visual > 17; cognitive > 30; manual > 10	I
1.8 - 2.2	1.9 - 2.4 hrs	visual 13-17; cognitive 23-30; manual 8-10	II
1.2 - 1.8	1.3 - 1.9 hrs	visual 10-13; cognitive 17-23; manual 6-8	III
< 1.2	< 1.3 hrs	visual < 10; cognitive < 17 manual < 6	IV

Table 2:   Classification of distracting activities into risk zones.

Risk zone	Distracting activities
I	passenger using mobile phone, passenger talking to driver
II	passengers, fatigue/sickness
III	passenger not following etiquette, ticket machine, climate control
IV	remaining 11 distracting activities

4.2 Modelling high risk activities

The seven high risk distracting activities in Table 2 were statistically analyzed using multinomial logistic regression (MLR) models to compute risk of distraction due to the five factors. The standard multinomial logit formulation is defined by Washington *et al.* [15] as:

$$P_n(i) = \frac{e^{\beta_i X_{in}}}{\sum_{\forall I} e^{\beta_I X_{In}}} \tag{1}$$

where;

$P_n(i)$: Probability that driver n will get i^{th} distracting outcome ($i \in I$).

I: set of all possible distraction outcomes (not distracted, slightly distracted, distracted, very distracted).

β_i: Estimated coefficient related with independent variable X_{in} for discrete outcome i, and,

X_{in}: Random variable that driver n will get i^{th} distracting outcome ($i \in I$).

The variable X_i represents a set of independent predictor variables (nominal or scalar) factors, such as location, driving hours/week; and driver age, sex, and driving experience. The variables X_i will also be employed in determining the extent of driver distraction.

The variable Y *(logit)* is a measure of the total contribution of all the independent predictor variables used in the model. The variable Y is usually defined as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots \ldots + \beta_k X_k \tag{2}$$

The intercept β_0 is the value of Y when all the independent variables are equal to zero. β_1, β_2, β_3, and so on, are the regression coefficients of X_1, X_2, and X_3 respectively. Each of the regression coefficients describes the size of the contribution of risk factor X_i relative to a reference category. A positive regression coefficient means that the explanatory variable increases the probability of the outcome, while a negative regression coefficient means that the variable decreases the probability of that outcome [15, 17]; a large regression coefficient means that the risk factor strongly influences the probability of that outcome, while a near-zero regression coefficient means that that risk factor has little influence on the probability of that outcome [19].

The numerator of eqn. (1) is the choice utility for discrete outcome I for driver n, and the denominator is the sum of choice utilities for all alternative outcomes for driver n. The function f (Y_i) represents the probability of a particular

outcome, such as the extent of distraction. Using eqn. (1), the probabilities of outcomes in each group are formulated as:

$$f(Y_1) = \frac{e^{Y_1}}{(e^{Y_1} + e^{Y_2} + e^{Y_3})} \tag{3}$$

$$f(Y_2) = \frac{e^{Y_2}}{(e^{Y_1} + e^{Y_2} + e^{Y_3})} \tag{4}$$

$$f(Y_3) = \frac{e^{Y_3}}{(e^{Y_1} + e^{Y_2} + e^{Y_3})} \tag{5}$$

Y_i is the multinomial logit defined by:

$$Y_i = \beta_0 + \beta_1 * LOCAT + \beta_2 * SEX + \beta_3 * AGE + \beta_4 * EXP + \beta_5 * DRIVING/WK \tag{6}$$

where,
LOCAT: Location of driver, a categorical variable, 1 = Northside, 2 = Southside.
SEX: Gender of driver, a categorical variable, 1 = Male, 2 = Female.
AGE: Reported age of driver, a continuous variable.
EXP: Number of years of experience driving a bus, a continuous variable.
DRIVING/WK: Weekly driving hours, a continuous variable.

Following the approach of Field [20], the dependent variable is broken down into a set of comparisons between two categories. The variable Not Distracted was chosen as a reference category. The four (k = 4) categorical dependent variables, Not Distracting, Slightly Distracting, Distracting, and Very Distracting) were converted into three (k − 1) output groups Y_1, Y_2, and Y_3. The first output (Y_1) tests whether or not the driver is more likely to get Slightly Distracted versus Not Distracted. Similarly, the second output (Y_2) tests whether or not the driver is more likely to get Distracted versus Not Distracted. And the third output (Y_3) tests whether or not the driver is more likely to get Very Distracted versus Not Distracted. The coefficients computed by the MLR models are relative to the reference category and are utilized to predict the probability of the extent that a driver finds an activity distracting versus not distracting from the binary logistic function [20]:

$$f(Y) = \left(\frac{1}{1 + e^{-Y}} \right) \tag{7}$$

where f (Y) is the probability of event Y occurring, given known values of Xs [20]. In this study, f (Y) is the probability of a driver getting slightly distracted, distracted, and very distracted. The probability values from the function f (Y) will vary between 0 and 1. The event Y is very unlikely to occur if f (Y) is close to 0 and very likely to occur if it is close to 1.

The output is split into three tables since the dependent categorical variables are compared in pairs. Table 3 presents statistical test ratios and the parameter

estimates for final model of passenger using mobile phone. Due to space restrictions, the MLR model output for only one distracting activity is presented.

Table 3:   MLR model outputs for passenger using mobile phone.

Model Chi-Square (χ^2) = 71.56 (24)**** Pearson Stat (NS) Deviance Stat(NS)	$R^2 = 0.833$ (Cox & Snell); 0.897 (Nagelkerke); 0.678(McFadden)	AIC final model (87.93) **B**IC final model (133.53)		
Independent Predictor Variables and Interactions	Coeff β (SE)	Wald Statistic	Odds Ratio Exp (B)	95% CI includes 1
Slightly distracting vs. Not distracting				
Intercept	-105.49 (47.85)**	4.86		
LOCAT = 1	-9.48 (3.05)***	9.64	< 1	No
LOCAT = 2	0.00			
SEX =1	82.41 (21.78)****	14.31	> 1	No
SEX = 2	0.00			
AGE	1.65 (0.82)**	4.01	> 1	No
EXP	2.57 (1.23)**	4.34	> 1	No
DRIVING/WK	1.89 (1.01)*	3.49	> 1	Yes
AGE*DRIVING/WK	-0.03 (0.02)*	2.90	< 1	Yes
SEX=1*DRIVING/WK	-1.70 (0.45)****	14.45	< 1	No
AGE*EXP	-0.04 (0.02)*	3.46	< 1	Yes
Distracting vs. Not distracting				
Intercept	+156.58 (51.15)***	9.37		
LOCAT = 1	-5.82 (2.02)***	8.29	< 1	No
LOCAT = 2	0.00			
SEX =1	20.06 (9.68)**	4.29	> 1	No
SEX = 2	0.00			
AGE	-2.72 (0.88)***	9.51	< 1	No
EXP	3.69 (1.38)***	7.09	> 1	No
DRIVING/WK	-3.79 (1.23)***	9.54	< 1	No
AGE*DRIVING/WK	0.070 (0.02)***	9.80	> 1	No
SEX=1*DRIVING/WK	0.36 (0.20)*	3.16	< 1	Yes
AGE*EXP	-0.06 (0.02)**	6.55	< 1	Yes
Very distracting vs. Not distracting N/S				

*p < 0.10; **p < 0.05; ***p < 0.01; ****p < 0.001. N/S = Not Significant.

4.2.1 Assessment of the model

The likelihood ratio test using model fitting information is presented in Table 3. It shows that the difference in the -2Log Likelihood between the intercept only (without any independent variables) and the final model (with all the independent variables) provides the chi-square (χ^2) = 71.56 (24). It signifies a good improvement in the model fit. It follows that the independent variables contribute significantly to the outcome of the distraction level. Lower values of

the AIC (87.93) and the BIC (133.53) in the final model also indicate a good fit. The BIC appears higher but according to Field [20], models with lower values of AIC or BIC produces a better fit. Table 3 shows the model's Goodness of Fit as indicated by the multiple statistics. The p values for Pearson and Deviance (both test the same results) chi-square (χ^2) = 1.00 (p = 1) proving no significance. Hence, the predicted values of the model are not significantly different from the observed values at all outcome levels – the model fits the data well. The measures of Pseudo R^2 (0,833, 0.897, and 0.678) are reasonably similar and high values resulting in a good fit.

The graphical methods were found suitable to assess the best fit of the binary logistic regression model [23] since model inadequacies are generally reflected in the pattern of residuals generated by the SPSS 17,0 [22] software package. The model's output data was analyzed by plotting the Pearson residuals versus the Predicted Probability. The residuals are scattered randomly above and below the line without any pattern over the entire range of the predicted values (Figure 1).
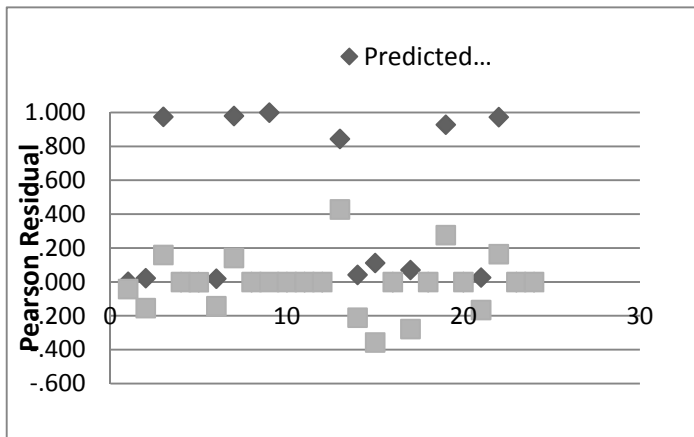

Figure 1:   Scatter plot of Pearson residuals.

4.3 Results of the model for high risk activities

The MLR models established a relationship between the level of bus driver distraction (dependent variable) and the distracting factors (independent variables) for the *risk zone* distracting activities. From Table 3, the magnitude and direction of each constant term (β_0) and significant independent variable coefficient (β_i) along with odds ratios (Exp B) for passenger using mobile phone were used to illustrate the impact of changes in each on the outcome of the categorical dependent variable. Keeping the other independent variables constant, a higher or positive amount of a coefficient of the independent variable [15−17,19] along with an odds ratio > 1 [18] would result in a higher dependent variable outcome and vice versa. The impact of each variable on the distracting levels for passenger using mobile phone is discussed as follows:

LOCATION: The coefficients for Northside relative to Southside are -9.48 and – 5.82 and odds ratio < 1. Hence, drivers covering Southside driving routes are more likely than drivers covering Northside driving routes to get slightly distracted followed by distracted.

SEX: The coefficients for males relative to females are 82.41 and 20.06 and odds ratio > 1. Hence, males are more likely than females to be slightly distracted followed by distracted.

AGE: The coefficients for slightly distracting is 1.65 and odds ratio > 1. Hence, if the driver age is increased by one year, the likelihood of getting slightly distracted versus not distracted would be expected to increase by 1.65 units. Older drivers are more likely to get slightly distracted by passengers using mobile phones. The coefficients for distracted is -2.72 and adds ratio < 1. Hence, if the driver age is increased by one year, the likelihood of getting distracted versus not distracted would be expected to decrease by 2.72 units. Older drivers are less likely to get distracted.

EXP: The coefficients are 2.57 and 3.67 and odds ratio > 1. Hence, if the driver experience increases by one year, the likelihood of getting slightly distracted followed by distracted versus not distracted would increase by 2.57 units and 3.67 units. Experienced drivers are more likely to get slightly distracted or distracted.

DRIVING/WK: The level of significance is > 0.05 and the 95% CI of odds ratio includes 1 but because of its importance to traffic safety, it is discussed further. The coefficient for slightly distracted is 1.89 and odds ratio > 1. Hence, if the driving hours/week increases by one hour, the likelihood of getting slightly distracted versus not distracted would increases by 1.89 units. Drivers who drive greater number of hours/week are more likely to get slightly distracted. The coefficient for distracted is – 3.79 and odds ratio < 1. Hence, if the driving hours/week is increases by one hour, the likelihood of getting distracted versus not distracted would decrease by 3.79 units. Drivers who drive greater number of hours/week are less likely to get distracted.

5   Conclusions and recommendations

Out of the 18 named distraction activities identified for bus drivers at the local transit agency, the seven high risk distraction activities were classified into *risk zones* I, II, and III using a risk range system. Passenger-related activities (which are beyond the control of bus driver) were reported as most distracting and passengers using mobile phones caused the highest risk of distraction. A challenge for the transit agency is to develop effective policies for handling passenger behaviour so passengers will be less likely to engage in using mobile phones, standing or moving around the bus, noisy behaviour etc that may cause bus driver distraction. Personal electronic devices could be allowed to be used in the latter half of the bus (*Electronic-Free Zone*). Radios, pagers, electronic games and/or other devices must also be maintained at minimum noise levels. Drivers must avoid unnecessary communications with passengers. If

conversation cannot be avoided, it must be done with caution while concentrating on the road ahead, or when the bus is stopped.

The results from this study could be applied to reduce driver distraction and improve overall transit performance. The MLR model could be used to predict the probability (eqns. (3), (4), (5), and (7)) of an existing or new driver getting distracted by a risk zone activity. Relevant training programs can then be developed to mitigate risk of distraction. Training needs to be focused towards male and younger drivers who are more likely to get distracted by passengers-related activities. Educational training program should be designed on the proper use of technological devices mounted in the cab or issued to the driver, and hazards associated with utilizing these devices while driving. The design of ticket machine, control panel, and other devices must be user-friendly, and not require long glances away from the forward roadway.

6   Limitations of the study

This was an exploratory study conducted on a sample of localized transit bus drivers. Hence, results may not be generalized for other transit agencies in the Commonwealth of Virginia or the United States. The limited sample size resulted in a higher standard error for the coefficients of the variables and multicollinearity detected from the Pearson Correlation analysis making it complicated to determine the relative importance of each independent variable on the regression model and the effects on the dependent variable. Some of the odds ratios were extremely high or low indicating errors associated with these odds ratio estimates; thus, according to Hickman *et al.* [18], it is difficult to report the odds ratio in any meaningful sense other than to report there was a strong positive or negative relationship between the level of distraction and the factors causing them.

The self-administered survey was unable to capture the actual measures of visual, cognitive, and manual distraction, which has been linked in the literature to poorer driving performance and increased likelihood of crashing. Such studies have typically been carried out in more controlled settings, using driving simulators or instrumented vehicles (or drivers) on test tracks [7, 10, 11]. The MLR data was not tested to show it meets the Independence of Irrelevant Alternative (IIA) specifications which require the ratio of probabilities of selecting any two alternatives to be independent of the third choice [24].

Only five factors: location, driving hours/week, and driver gender, age, and driving experience were included as the independent variables. Other variables such as environmental, vehicle type, roadway designs etc could also have an impact on driver distraction. An important point to bear in mind while increasing the number of independent variables, the preferred sample size must be around 20 times the number of variables [19]. In order to overcome the limitations of this study and make the results applicable state-wide or nationally, a larger study encompassing a wider range of transit agencies is necessary. The study at the local transit agency was a step in that direction.

References

[1]   U.S. Department of Transportation. Statistics and Facts about Distracted Driving. National Highway Traffic Safety Administration, Washington, DC 20590, www.distraction.gov/stats-and-facts.

[2]   National Highway Traffic Safety Administration (NHTSA). *An examination of driver distraction as recorded in NHTSA databases*. DOT HS 811 216, 2009. NHTSA's National Center for Statistics and Analysis, Washington, DC.

[3]   AAA Foundation for Traffic Safety. 607 14th Street, NW, Suite 201, Washington D.C., www.aaafoundation.org.

[4]   Cross, G.W., Hanna, H., Garrison, T. & McKee, C. D*istracted driving in Mississippi: a statewide survey and summary of related research and policies*. Final Report. The Center for Mississippi Health Policy, Mississippi State University, Family and Children Research Unit, 2010.

[5]   Olson, R., Hanowski, R.J., Hickman, J.S. & Bocanegra, J. D*river distraction in commercial vehicle operations*. Report No. FMCSA-RRR-09-042. U.S. Department of Transportation, Washington D.C, 2009.

[6]   Blower, D., Green, P.E. & Matteson, A. *Bus operator types and driver factors in fatal bus crashes: results from the buses involved in fatal accidents survey*. Report No. FMCSA-RRA-09-041. U.S. Department of Transportation, Washington D.C, 2008.

[7]   Salmon, P. M., Young K. L. & Regan, M. A. Distraction 'on the buses': A novel framework of ergonomics methods for identifying sources and effects of bus driver distraction. *Applied Ergonomics*, **42**, pp. 602-610, 2011.

[8]   Commonwealth of Virginia, 2008 Virginia Traffic Crash Facts. *Virginia crashes involving buses*. Virginia Highway Safety Office, Department of Motor Vehicle, PP. 58, 2008.

[9]   Salmon, P.M., Young, K.L. & Regan, M. A. *Bus driver distraction stage 1: A*nalysis of risk for State Transit Authority New South Wales bus drivers. *Final report*. Monash University Accident Research Centre, Victoria, Australia, 2006.

[10] Green, P.E., Wada, T., Oberholtzer, J., Green, P.A., Schweitzer, J. & Hong E. *How do distracted and normal driving differ: an analysis of the ACAS naturalistic driving data*. Report No: UMTRI-2006-35. The University of Michigan Transportation Research Institute, 2007.

[11] Oxley, J., Charlton, J., Fildes, B., Koppel, S., Scully, J., Congiu, M. & Moore, K. C*rash risk of older female drivers*. Report No. 245. Monash University Accident Research Centre, Victoria, Australia, 2005.

[12] Yan, X., Radwan, E. & Abdel–Aty, M. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. *Accident Analysis and Prevention*, **37**, pp. 983-995, 2005.

[13] Yan, X., Radwan, E. & Mannila, K.K. Analysis of truck-involved rear-end crashes using multinomial logistic regression. *Advances in Transportation Studies: an International Journal*, Section A **17**, pp. 39-52, 2009.

[14] Morfoulaki, M., Tyrinopoulos, Y. & Aifadopoulou, G. Estimation of satisfied customers in public transport systems: a new methodological approach. *Journal of the Transportation Research Forum*, **46 (1),** pp. 63-72, 2007.

[15] Washington, S.P., Karlaftis, M.G. & Mannering, F.L. *Statistical and Econometric Methods for Transportation Data Analysis*. Chapman & Hall/CRC, A CRC Press Company, Boca Raton, FL, pp. 257-295, 2003.

[16] Gkritza, K., Niemeier, D. & Mannering, F. Airport security screening and changing passenger satisfaction: an exploratory assessment. *Journal of Air Transport Management*, **12**, pp. 213-219, 2006.

[17] Bhadra, D. Choice of aircraft fleets in the U.S. domestic scheduled air transportation system: findings from a multinomial logit analysis. *Journal of the Transportation Research Forum*, **44 (3),** pp. 143-162, 2005.

[18] Hickman, J. S., Hanowski, R. J. & Bocanegra, J. Distraction in commercial trucks and buses: assessing prevalence and risk in conjunction with crashes and near-crashes. Report No. FMCSA-RRR-10-049. U.S. Department of Transportation, Washington, D.C, 2010.

[19] Petrucci, C. J. A primer for social worker researchers on how to conduct a multinomial logistic regression. *Journal of Social Service Research*, **35 (2),** pp. 193-205, 2009.

[20] Field, A. *Discovering Statistics Using SPSS*. SAGE Publications Ltd., 3rd Edition, pp. 264-315, 2009.

[21] UCLA. Annotated SPSS Output. Multinomial Logistic Regression Academic Technology Services, Statistical Consulting Group, http://www.ats.ucla.edu/stat/sas/notes2.

[22] SPSS 17.0. SPSS Inc., Chicago, IL 60606-6412, 2008.

[23] Landwehr, J. M., Pregibon, D., & Shoemaker, A. C. Graphical methods for assessing logistic regression models. *Journal of the American Statistical Association*, **79 (385)**, pp. 61-71, 1984.

[24] Small, K. A., & Hsiao, C. Multinomial logit specification tests. *International Economic Review*, **26 (3)**, pp. 619-627, 1985.