

# Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model

Xuedong Yan<sup>a,\*</sup>, Essam Radwan<sup>b</sup>, Mohamed Abdel-Aty<sup>c</sup>

<sup>a</sup> Department of Civil and Environmental Engineering, College of Engineering, University of Central Florida, 12172 Descartes Cr. Apt#1, Orlando, FL 32826, USA

<sup>b</sup> Center for Advanced Transportation Systems Simulation, University of Central Florida, Orlando, FL 32816-2450, USA

<sup>c</sup> Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816-2450, USA

Received 9 February 2005; received in revised form 4 May 2005; accepted 4 May 2005

## Abstract

Multi-vehicle rear-end accidents constitute a substantial portion of the accidents occurring at signalized intersections. To examine the accident characteristics, this study utilized the 2001 Florida traffic accident data to investigate the accident propensity for different vehicle roles (striking or struck) that are involved in the accidents and identify the significant risk factors related to the traffic environment, the driver characteristics, and the vehicle types. The Quasi-induced exposure concept and the multiple logistic regression technique are used to perform this analysis. The results showed that seven road environment factors (number of lanes, divided/undivided highway, accident time, road surface condition, highway character, urban/rural, and speed limit), five factors related to striking role (vehicle type, driver age, alcohol/drug use, driver residence, and gender), and four factors related to struck role (vehicle type, driver age, driver residence, and gender) are significantly associated with the risk of rear-end accidents. Furthermore, the logistic regression technique confirmed several significant interaction effects between those risk factors.

© 2005 Elsevier Ltd. All rights reserved.

**Keywords:** Rear-end accidents; Signalized intersections; Quasi-induced exposure; Multiple logistic regression; Striking role; Struck role

## 1. Introduction

Rear-end accidents are one of the frequently occurring types of accidents, accounting for almost one-third of all reported accidents in the US (1.848 million) and 11.8% of multi-vehicle fatal accidents (1923) (National Transportation Safety Board, 2001). Rear-end accidents are the most common accident type at signalized intersections since the diversity of actions taken increases due to signal change. Specific causes of rear-end accidents include the following drivers' inattentive driving and following too closely. A proper space cushion is needed to provide a driver enough reaction time to recognize a hazardous situation and make a stop decision.

Typically, driver, vehicle, and roadway/environment characteristics influence accident occurrence and injury severity. Moreover, since a rear-end accident is related to both driving behaviors and performances of the leading (struck) vehicle and the following (striking) vehicle, the accident risk is possibly associated with struck or striking role that a driver or vehicle would assume in a rear-end accident.

The driver age and gender were considered as main driver characteristics that might be associated with a rear-end accident. There is general consensus among researchers that older drivers tend to process information and take a corresponding action more slowly than younger driver. Slower reaction times for older versus younger drivers contribute to a disproportionately heightened degree of risk especially when older drivers are faced with two or more choices of action (Staplin et al., 2001). However, the younger drivers, especially under 25 years, are more likely to be involved in aggressive attitude and inattentive driving. A previous study on rear-end acci-

\* Corresponding author. Tel.: +1 407 823 5810; fax: +1 407 823 4676.

E-mail addresses: yxd22222@yahoo.com (X. Yan), aeradwan@mail.ucf.edu (E. Radwan), mabdel@mail.ucf.edu (M. Abdel-Aty).

dents indicated that drivers younger than 18 years were most vulnerable to roadway accidents followed by 18–24-year-old drivers; the propensity of drivers involved in accidents showed a decreasing trend with increasing age until the age of 69, after which the drivers again showed a higher accident involvement propensity as compared to the drivers who were 25–69 year old (Santokh, 2003). Additionally, among 18–24-year-old drivers, male drivers were found to be more prone to accident involvement as compared to their female counterparts. It was also found that in rear-end accidents, drivers up to the age of 25 years are more likely to be in the striking role than in the struck role. As drivers get older, they tend to be in the striking role less often than in the struck role.

For different type of vehicles, steering and braking performance are critical in the avoidance of accidents. Strandberg (1998) pointed out that except for the hazards due to unpredicted change in properties within one vehicle, differences between vehicles in braking performance are responsible for many rear-end accidents. Moreover, the size of the leading vehicle may influence the behavior of the following driver. Sayer et al. (2000) examined the effect that the leading vehicle size (specifically, height and width) has on a passenger car driver's gap maintenance under near optimal driving conditions (e.g. daytime, dry weather, free-flowing traffic). Results showed that passenger car drivers followed light trucks at shorter distances than they followed passenger cars, but at the same velocities. Specifically, it appears that when dimensions of lead vehicles permit following drivers to see through, over, or around them, drivers maintain significantly longer (i.e. safer) distances. Abdel-Aty and Abdelwahab (2004) examined the relationship between vehicle type (car or LTV including light truck, van, and SUV) and the role (striking or struck) played by each vehicle in the accident. Using a nested logit structure model, they analyzed the probabilities of the four rear-end accident configurations (car–car, car–LTV, LTV–car, and LTV–LTV) as a function of driver's age, gender, vehicle type, vehicle maneuver, light conditions, driver's visibility and speed. Results showed that driver's visibility and inattention in the following (striker) vehicle have the largest effect on being involved in a rear-end collision of configuration Car–truck (a regular passenger car striking an LTV).

The critical road environment conditions could play a significant role in rear-end accidents and they may contain all kinds of non-driver related factors such as lighting conditions, the roadway surface conditions, highway characteristics, traffic volume, the weather conditions, and so on. Braking performance of vehicle is substantially reduced in icy and snowy road surface condition and deceleration capacity may decrease by more than 90% compared to dry condition (Strandberg, 1998). The heavy traffic volume results in the smaller headway between gaps between leading and following vehicles, which definitely increases the possibility of rear-end accidents. Khattak (2001) reported that a majority of the accidents (54.9%) occurred during the peak times 7:00–9:00 a.m. and 3:00–6:00 p.m. A small portion (10.8%)

of the accidents occurred at night on unlit streets and a smaller portion (4.9%) occurred at night on lighted streets.

However, of those previous research findings, relatively few studies used the accident database and related statistical model to explore the propensity of rear-end accidents that occurred at signalized intersections. This paper presents the results of a thorough investigation into the relationship between the rear-end accidents and a series of potential risk factors classified by driver characteristics, road environments, and vehicle type. The quasi-induced exposure concept is used to compare the relative accident involvements between different risk conditions based on the 2001 Florida accident database. For striking role and struck role in the rear-end accidents respectively, multiple logistic regression models are used for hypothesis testing to identify the significant factors that contribute to the rear-end accidents.

## 2. Methodology

### 2.1. Accident database and rear-end accident identification

The 2001 accident database, obtained from the Florida Department of Highway Safety and Motor Vehicles (DHSMV), was used in this study. The DHSMV data constitute a relational database that includes seven files. Each file deals with a specific feature of a traffic accident. Files may be linked as needed to combine the information contained in each other. Three files used in the analysis presented here were the event (containing the characteristics and environment of the accident), drivers (containing the drivers' characteristics), and vehicles (information about the vehicles' characteristics and vehicles actions in the traffic accident) files.

In Florida, when an accident occurs and the local police department is notified, the responding officer will determine whether to fill out a long- or short-form accident report. If an accident involves an injury or a felony (e.g. hit and run), the accident must be filed on a long-form. If an accident involved only property damage (a minor accident), usually it is up to the officer to report it on a long or a short-form. Accident forms are then forwarded to the respective counties, which choose whether or not to file accidents reported on short-forms. From here, only the accidents reported on long-forms are forwarded onto the DHSMV, which maintain records based on only accidents reported on long-forms. Since most accidents that occur involve only property damage and not a serious injury or a felony, it can be argued that the FDOT and DHSMV accident databases under-represent the property-damage-only accidents. Abdel-Aty et al. (2005) investigated the significant differences in the important accident-related factors between models based solely on accidents reported on long-forms and models based on accidents reported on both long- and short-forms (i.e. models based on restricted and complete datasets). They found that for rear-end, right-turn

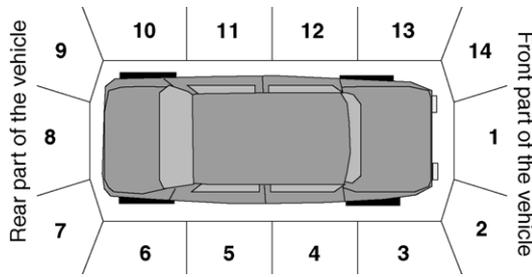


Fig. 1. Code for collision positions of vehicles in DHSMV.

and sideswipe accidents, the important factors are fairly consistent between the models created by complete and restricted datasets. The results show that factors in accidents causing mostly injuries for rear-end, right-turn and sideswipe accidents are generally the same as factors causing non-injury or minor accidents. This indicates that models based on complete and restricted datasets for these types of accidents are roughly equivalent. Therefore, this rear-end accident analyses based on the 2001 accident database from the DHSMV long-form data should present similar results if based on a complete accident database.

Rear-end accident scenarios may involve three or more vehicles. To simplify the assignment of driver culpability and easily identify accident roles of vehicle (leading or following) in the accident, the analysis was restricted to two-vehicle collisions, in which both vehicles proceeded straight at signalized intersections. The collision positions of vehicles are coded and recorded in the DHSMV, as shown in Fig. 1. In a two-vehicle rear-end accident, the code 7, 8, or 9 can identify the struck vehicle and the vehicle with the code 1, 2, or 14 should be the striking one. In most rear-end cases, the leading vehicles did not take a major responsibility for the accidents but they could have performed an unexpected stop; on the other hand, the main contributing causes to the accident for the striking vehicle are careless driving and following too closely.

2.2. Quasi-induced exposure technique

To test accident propensity and explore the traffic accident database, quasi-induced exposure technique (Carr, 1970; Haight, 1973; Stamatiadis and Deacon, 1995, 1997) is becoming widely used in traffic safety research. Stamatiadis and Deacon (1995) developed the term relative accident involvement ratio (RAIR), as the measure of accident causing propensity used in the quasi-induced exposure analysis. It is equal to the ratio of the percentage of a specific subgroup in at-fault drivers to the percentage of the same subgroup in not-at-fault drivers. The at-fault drivers are those who were mostly responsible for the accident occurrence and the not-at-fault drivers are those less responsible for the accidents. The key assumption is that the distribution of not-at-fault drivers closely represents the distribution of all drivers exposed to accident hazards. Previous studies had successfully applied

the quasi-induced exposure method to analyze traffic accident risks of drivers and vehicles under a given set of road environment conditions (Stamatiadis and Deacon, 1995; Aldridge et al., 1999; Hing et al., 2003). However, few of them focused on the investigation of non-driver-related (road environment) factors as exclusive main effects on the traffic safety.

To introduce the road environment factors into statistical model and test their exclusive main effects on accidents, this research extended the application of the quasi-induced exposure. In this study, first, two-vehicle accidents occurring at signalized intersections are identified, which are categorized into two groups: rear-end accidents and non-rear-end accidents (exposure). Those rear-end accidents include information of struck drivers/vehicles and striking drivers/vehicles, as well as the corresponding road environment conditions. Those non-rear-end accidents include information of drivers/vehicles that had no improper driving action but were involved in the accidents, as well as their corresponding road environment conditions. Thus, the comparisons between drivers/vehicles/environments distribution in the rear-end group and the non-rear-end group can be used to investigate the accident propensities. To increase the comparability between rear-end drivers and not-at-fault drivers, the driving behaviors for both groups are limited in ‘going-through’ the signalized intersections when the accidents were happening. For better understanding the principle of the data classification, Fig. 2 illustrated the relationship between rear-end and non-rear-end groups. According to the accident definition, the dataset identified 7666 two-vehicle rear-end accidents that happened at signalized intersections and 15,734 non-rear-end accidents involved by not-at-fault drivers as exposure.

Based on the above categorization, three types of relative accident involvement ratios are calculated to test the main effects of driver, vehicle, and environment factors related to rear-end accidents. The extended assumption here is that the distributions of road environment factors in non-rear-end accidents may represent the distributions of road environment factors confronted by all vehicles/drivers. Comparing the variable distributions in non-rear-end accidents to those in all two-vehicle accidents at intersections showed similar trend. Using the RAIR formula developed by Stamatiadis and Deacon (1995), the relative accident involvement ratio is defined as Eq. (1):

$$RAIR_i = \frac{\sum D1_i}{\sum D2_i} \quad \text{or} \quad RAIR_i = \frac{\sum V1_i}{\sum V2_i} \quad \text{or} \quad RAIR_i = \frac{\sum E1_i}{\sum E2_i} \quad (1)$$

where  $RAIR_i$  is the relative accident involvement ratio for type  $i$  drivers/vehicles/environments,  $D1_i$  the number of striking drivers of type  $i$  in rear-end accidents,  $D2_i$  the number of not-at-fault drivers of type  $i$  in non-rear-end accidents,  $V1_i$  the

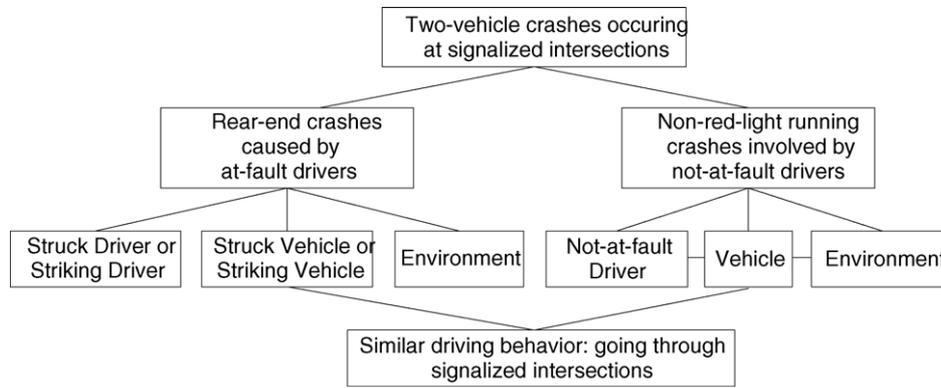


Fig. 2. Rear-end and not-rear-end accident classification used for quasi-induced exposure analysis.

number of striking vehicles of type  $i$  in rear-end accidents,  $V2_i$  the number of not-at-fault vehicles of type  $i$  in non-rear-end accidents,  $E1_i$  the number of rear-end accidents involving environment type  $i$ , and  $E2_i$  is the number of non-rear-end accidents involving environment type  $i$ .

Furthermore, to test the interaction between type  $i$  drivers/vehicles/environments and type  $j$  drivers/vehicles/environments, the RAIR can be defined as Eq. (2):

$$RAIR_{i,j} = \frac{\sum \sum N1_{i,j}}{\sum \sum N2_{i,j}} \quad (2)$$

where  $RAIR_{i,j}$  is the relative accident involvement ratio types  $i$  and  $j$  drivers/vehicles/environments,  $N1_{i,j}$  the number of rear-end drivers, vehicles, or the related environments of types  $i$  and  $j$  in rear-end accidents, and  $N2_{i,j}$  is the number of not-at-fault drivers, vehicles, or the related environments of types  $i$  and  $j$  in non-rear-end accidents.

### 2.3. Statistical modeling

Previous studies had appropriately applied logistic regression analysis to test the significance of traffic accident risk factors based on the technique of induced exposure (Stamatiadis and Deacon, 1995; Hing et al., 2003). Logistic regression belongs to the group of regression methods for describing the relationship between explanatory variables and a discrete response variable. A binary logistic regression is proper to use when the dependent is a dichotomy (an event happened or not) and can be applied to test association between a dependent variable and the related potential factors, to rank the relative importance of independents, and to assess interaction effects. Binary logistic regression is used in this study since the dependent variable  $Y$  (accident classification) can only take on two values:  $Y=1$  for rear-end accidents, and  $Y=0$  for non-rear-end accidents. The probability that a rear-end accident will occur or not is modeled as logistic distribution in Eq. (3):

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad (3)$$

The logit of the multiple logistic regression model (Link Function) is given by Eq. (4):

$$g(x) = \ln \left[ \frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n \quad (4)$$

where  $\pi(x)$  is conditional probability of a rear-end accident, which is equal to the number of rear-end accidents divided by the total number of accidents.  $x_n$  are independent variables (driver/vehicle/environment factors). The independent variables can be either categorical or continuous, or a mixture of both. Both main effects and interactions can generally be accommodated.  $\beta_n$  is model coefficient, which directly determines odds ratio involved in the rear-end accident. The odds of an event are defined as the probability of the outcome event occurring divided by the probability of the event not occurring. The odds ratio that is equal to  $\exp(\beta_n)$  tells the relative amount by which the odds of the outcome increase (OR greater than 1.0) or decrease (OR less than 1.0) when the value of the predictor value is increased by 1.0 units (David and Lemeshow, 1989). Especially for categorical independent variables, the odds ratios represent the accident risk comparison among different levels of drivers/vehicles/environments.

The previous studies (Stamatiadis and Deacon, 1995; Hing et al., 2003) clearly expressed the relationship between logistic regression and RAIR in the quasi-induced exposure analysis. In fact, for a specific type of drivers/vehicles/environments, the odds generated from the logistic regression model are analogous to the corresponding RAIRs, and the odds ratios from the model are equivalent to the comparisons among those RAIRs. Furthermore, a significant  $P$ -value (e.g.  $P \leq 0.01$ ) for the Wald  $\chi^2$  statistic is evidence that a regression coefficient in the model is nonzero, which also indicates the statistical importance of those RAIR comparisons between different types of drivers/vehicles/environments.

In addition, since either striking or struck drivers/vehicles might be associated with the rear-end accidents, two separate models were developed for different roles in the accidents (striking or struck), to investigate the relationship between rear-end accidents and a series of potential risk factors. For

simplicity and ease of interpretation of the results, all those factors are chosen as categorical variables to fit models. The SAS program procedure, LOGISTIC, was used for model development and the hypothesis testing was based on 0.01 significance level because of the large sample size.

### 3. Modeling results for environments and striking drivers/vehicles

#### 3.1. Main effect estimation

Based on the main effect model, the multivariate logistic regression analysis identified significant factors directly associated with rear-end accidents happening at signalized intersections. Table 1 lists the model estimation and odds ratios properly adjusting other factors for significant independent variables. The following sections document the interpretation of the regression results for those variables classified by road environment factors and striking driver/vehicle characteristics.

##### 3.1.1. Road environment factors

Seven road environment factors including number of lanes, divided/undivided highway, accident time, road surface condition, highway character, urban/rural, and speed limit shows significant association with the risk of rear-end accidents. Fig. 3 illustrates univariate comparisons of relative accident involvement ratios between difference conditions for each road environment factor.

The number of lanes at the accident site as originally recorded by the reporting officer includes both sides of the median where applicable. This analysis focuses on intersections located on the most typical highways with 2, 4, and 6 lanes. As shown in Fig. 3, relative accident involvement ratio (RAIR) for 6-lane highway is larger than 2-lane highway and (RAIR) for 4-lane highway is smaller than 2-lane highway. However, adjusting other variables in the main effect model, Table 1 shows that there is no significant difference between 6- and 2-lane highways ( $P=0.216$ ), but the risk of rear-end accidents happening at 4-lane highways could be 27% lower than the 2-lane ones with a 0.01 level of significance. Comparing RAIRs, rear-end accidents are more likely to happen at divided highways than undivided highways. Adjusting for other variables, the risk for divided highways could be 16% higher than the risk for undivided highways with a 0.01 level of significance. However, the crash trend for number of lanes and divided/undivided highway may be confounded by the traffic volume effect because the highways with high heavy traffic tend to have more lanes and be separated by medians.

The results show that most rear-end accidents occurred at daytime (71.7%), versus 28.3% ones at night. An interesting finding is that the RAIR (1.06) for daytime is also higher than that (0.88) for night, as shown in Fig. 3. Based on the model, the accident risk for the night condition could be 50% lower than that for the daytime condition at a 0.0001 level of sig-

nificance. The presumable reasons are that the traffic volume at daytime is higher than that at night; the morning peak and afternoon peak may affect driving attitude and contribute to the accidents.

Fig. 3 shows that there is a clear trend between the RAIR and the road surface condition. The plot shows that as RAIR increases the road surface conditions become more and more slippery. Compared to a dry road surface, the accident risk on the wet surface could be 1.788 times higher and that on the slippery surface could be 3.302 times higher. The results testified that reduction of braking capacity will definitely contribute to rear-end accidents.

The highway characters are classified as four levels in the database, including straight-level (S-L), straight-up/downgrade (S-U/D), curve-level (C-L), and curve-upgrade/downgrade (C-U/D). Fig. 3 illustrates that as the highway character becomes more complex, the rear-end accident is more likely to occur. At signalized intersections, the risks of curve and grade approaches could be 79% and 28% higher than straight approaches. If both curve and grade are present at the same time, the rear-end risk could be twice as that for normal straight highways. Generally, the grade of an intersection approach may significantly affect the time and the distance needed for a motorist to stop a vehicle at an intersection. If approaching the intersection on a downhill grade, motorists may not account for vehicle mass and momentum which will require a longer stopping time. Moreover, if the intersection is located on a horizontal curve, the potential sight obstruction on the inside of curves, such as cut slopes, walls, buildings, bridge piers, and longitudinal barriers, could restrict stop sight distance and result in rear-end accidents.

The analysis also shows that rear-end is particularly relevant to urban accidents (64%) and the accident risk in urban area could be 20% higher than rural area, presumably because urban areas are more dominated by signalized intersections and have higher traffic volume than rural area. As shown in Table 1 and Fig. 3, there is a clear trend between RAIR and speed limit. As the speed limit increases, the risk of the rear-end accidents increases. For 55 mph speed limit, the odds of being involved in an accident could be 4.6 times higher than 25 mph speed limit. At signalized intersections, with the higher speed limits, generally, drivers are more likely to fall into the dilemma zone where they possibly can neither execute the intersection crossing nor execute the stopping maneuver safely and comfortably at the onset of yellow. Such a negative situation may result in rear-end accidents due to relatively higher operation speeds. On the other hand, at the lower operation speed, the following driver can more easily change lane or brake to avoid striking the lead vehicle.

##### 3.1.2. Striking driver/vehicle characteristics

Vehicle type and four factors related to driver characteristics including driver age, alcohol/drug use, driver residence, and gender shows significant association with the risk of rear-end accidents. Fig. 4 illustrates comparisons of relative

Table 1  
Model estimation and odds ratios for significant independent variables for striking role

Parameter	Coefficient estimate	Odds ratio	95% Wald confidence limits		Wald $\chi^2$	P-value
Intercept	-1.221				82.429	<0.0001
Number of lanes					66.122	<0.0001
Others vs. 2-lane	-0.104	0.901	0.803	1.012	3.072	0.080
6-lane vs. 2-lane	-0.071	0.931	0.832	1.042	1.532	0.216
4-lane vs. 2-lane	-0.319	0.727	0.657	0.804	38.556	<0.0001
Divided/undivided highway					15.948	<0.0001
Undivided vs. divided	-0.152	0.859	0.798	0.926	15.948	<0.0001
Accident time					341.878	<0.0001
Night vs. daytime	-0.701	0.496	0.461	0.534	341.878	<0.0001
Road surface condition					209.789	<0.0001
Wet vs. dry	1.195	3.302	2.255	4.835	37.713	<0.0001
Slippery vs. dry	0.581	1.788	1.642	1.947	177.863	<0.0001
Highway character					38.376	<0.0001
C-U/D vs. S-L	0.723	2.06	1.34	3.169	10.833	0.001
C-L vs. S-L	0.581	1.788	1.353	2.364	16.665	<0.0001
S-U/D vs. S-L	0.243	1.275	1.113	1.46	12.285	0.001
Urban/rural					26.718	<0.0001
Urban vs. rural	0.178	1.195	1.117	1.278	26.718	<0.0001
Posted speed					532.690	<0.0001
55 mph vs. 25 mph	1.528	4.607	3.433	6.183	103.551	<0.0001
50 mph vs. 25 mph	1.362	3.903	2.937	5.186	88.195	<0.0001
45 mph vs. 25 mph	1.022	2.779	2.177	3.548	67.249	<0.0001
40 mph vs. 25 mph	0.578	1.783	1.39	2.286	20.744	<0.0001
35 mph vs. 25 mph	0.483	1.62	1.267	2.073	14.760	0.000
30 mph vs. 25 mph	-0.029	0.972	0.753	1.253	0.049	0.825
Driver age					152.224	<0.0001
>75 vs. <26	0.116	1.123	0.946	1.334	1.764	0.184
66–75 vs. <26	-0.392	0.676	0.578	0.789	24.373	<0.0001
56–65 vs. <26	-0.618	0.539	0.472	0.616	83.104	<0.0001
46–55 vs. <26	-0.420	0.657	0.592	0.729	62.062	<0.0001
36–45 vs. <26	-0.343	0.71	0.649	0.777	55.663	<0.0001
26–35 vs. <26	-0.290	0.748	0.686	0.817	42.105	<0.0001
Alcohol/drug use					938.944	<0.0001
6 vs. 1	2.408	11.116	6.422	19.242	74.012	<0.0001
5 vs. 1	2.260	9.584	7.513	12.226	331.021	<0.0001
4 vs. 1	3.811	45.206	16.181	126.293	52.864	<0.0001
3 vs. 1	4.604	99.83	24.337	409.499	40.863	<0.0001
2 vs. 1	5.008	149.563	95.419	234.429	476.914	<0.0001
Residence code					51.372	<0.0001
4 vs. 1	0.582	1.789	1.114	2.875	5.781	0.016
3 vs. 1	0.369	1.446	1.196	1.748	14.533	0.000
2 vs. 1	0.288	1.334	1.212	1.468	34.879	<0.0001
Gender					9.921	0.002
Female vs. male	-0.105	0.9	0.843	0.961	9.921	0.002
Type of vehicle					25.429	<0.0001
Large vehicle vs. car	0.224	1.251	1.068	1.464	7.734	0.005
Light truck vs. car	0.189	1.208	1.108	1.317	18.501	<0.0001
Van vs. car	0.138	1.148	1.024	1.287	5.573	0.018

accident involvement ratios between different levels of those factors.

The 10-year interval was chosen to group drivers by age. The seven driver age groups included younger than 26, 26–35, 36–45, 46–55, 56–65, 66–75, and older than 75 years. Comparing relative accident involvement ratios, the graph for

driver age illustrated in Fig. 4 doesn't show a smooth U-shape pattern. The result is consistent with the previous analysis (Santokh, 2003): the propensity of drivers involved in accidents showed a decreasing trend with increasing age until the age of 56–65, after which the drivers increase to a higher accident involvement propensity for the age group older than

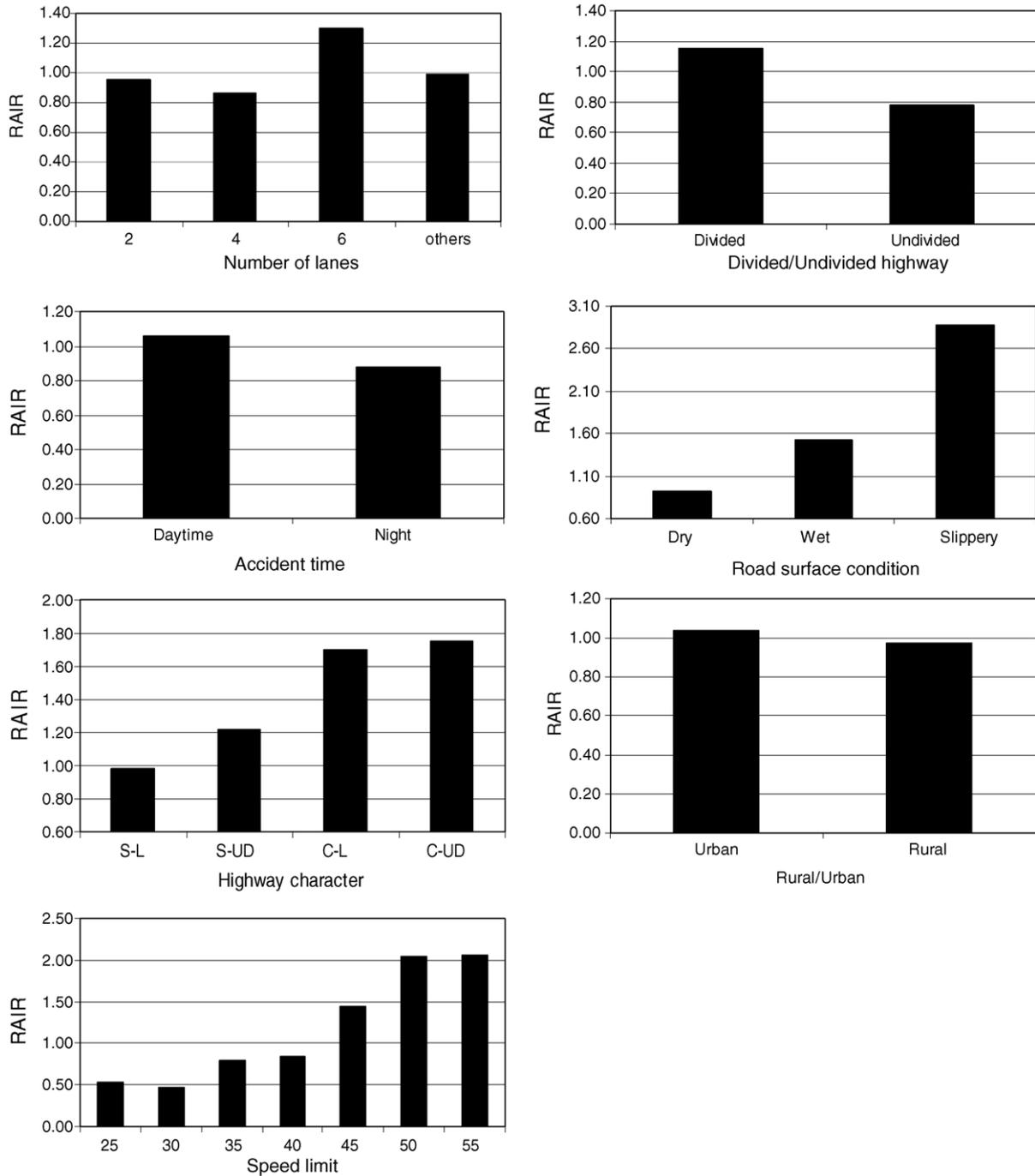


Fig. 3. Relative accident involvement ratios by road environment factors.

75, as compared to the drivers who were younger than 26 year old. However, adjusting for other factors in the model, there is no significant difference between the oldest group and youngest group ( $P$ -value = 0.184). The 56–65 years drivers who have lowest risk on rear-end accidents may benefit from lower following speed and conservative driving attitude. The oldest group (>75 years) presented the highest risk, presumably because of age-related deterioration of their physical and cognitive abilities. The younger group has a larger accident

propensity, presumably because of risk-taking and attitudinal factors.

Alcohol reduces alertness, interferes with judgment and impairs vision. Most drugs that affect the central nervous system may have the potential to impair driving ability. There are six levels of potential alcohol/drug use recorded in the Florida accident database: (1) No, (2) Alc-under influence, (3) Drug-under influence, (4) Alc&Drug-under influence, (5) Had been drinking, and (6) Pending BAC test results. As shown in

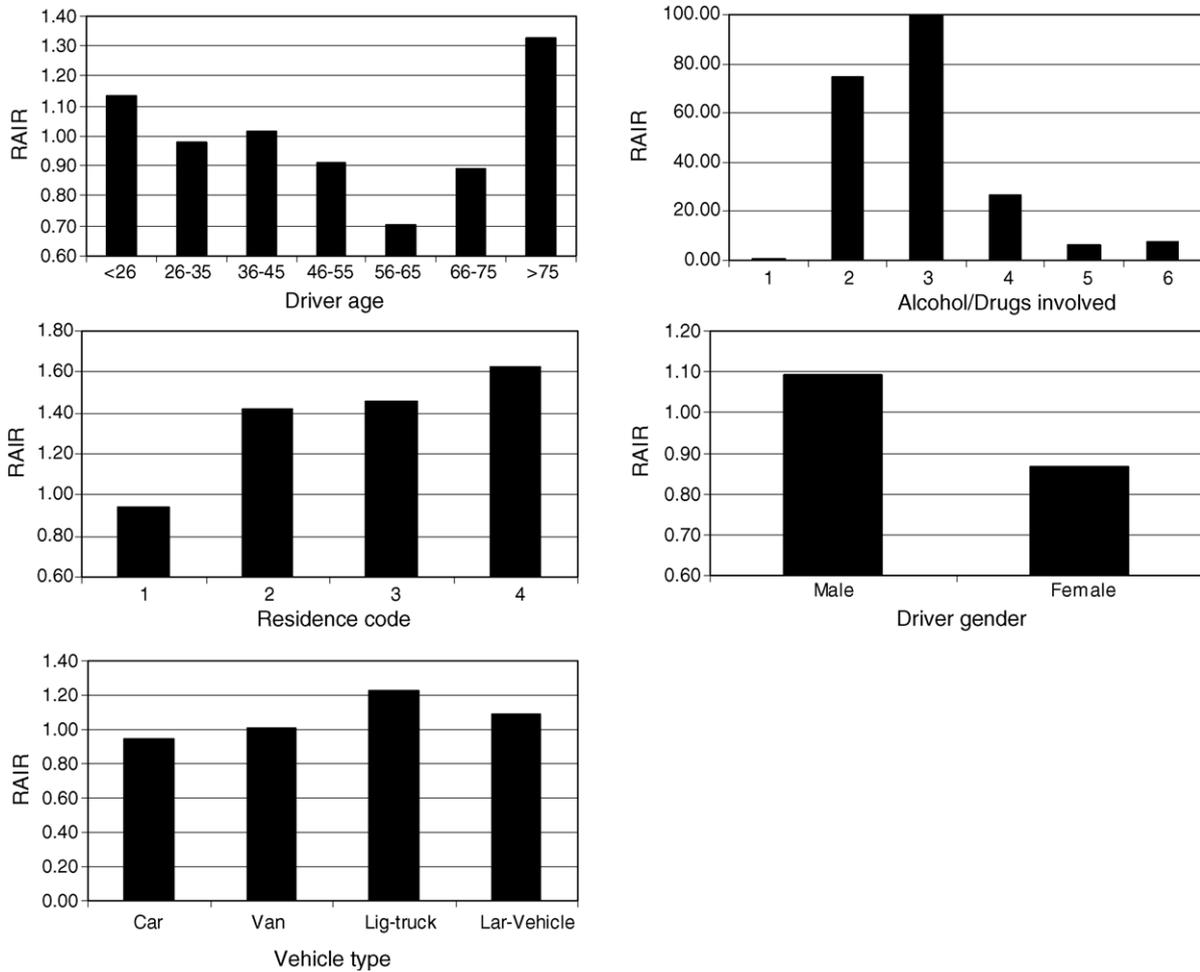


Fig. 4. Relative accident involvement ratios by striking driver characteristics and vehicle types.

Table 1 and Fig. 4, the drivers under influence of alcohol, drug, or alcohol and drug are substantially over-involved in accidents and their accidents risk can be 149.6, 99.8, and 45.2 times higher than those of normal drivers, adjusting for other factors. The results showed that both alcohol and drug have extremely negative effect on drivers involving rear-end accidents. Even drivers who had been drinking also show much higher involvement risk than non-drinking drivers (OR = 9.58). However, as not-at-fault drivers, those with alcohol/drug use performed worse to avoid accidents than normal ones. There is a higher chance to be underestimating their risk involving accidents based on the induced exposure technique. In addition, the data sample sizes in some levels of alcohol/drug use are very small, so that the point estimates of their odds ratios are less significant due to the large standard errors.

The accident database also provided driver classification by residence. The level 1, 2 3, or 4 stand for the residence of drivers living in the local county, elsewhere in the state of Florida, other state, or other country. Normally, local drivers can benefit from their driving experiences to the familiar traffic environments so as to avoid the adverse traffic conditions.

As shown in Fig. 4, it appears there is a clear trend which indicates that as the degree of drivers' familiarity with the driving environment decreases, they are more likely to be involved in rear-end accidents. Especially for the foreign drivers, their risk rate is around 79% higher than the county residents.

The results also showed significant gender difference in rear-end accidents ( $P$ -value = 0.002). According to the model, the male risk of involvement in accidents could be 19% higher than female drivers. The analysis is consistent with previous studies: gender is significant in predicting involvement in accidents (Ulleberg, 2001). Storie (1977) found that whereas men were involved more often than women in accidents caused by speeding and driving under the influence of alcohol, women were more frequently involved in accidents caused by judgment errors.

In addition to driver characteristics, vehicle type was also found to significantly affect accident propensity at a 0.0001 level of significance. There are total 13 types of vehicles classified by accident vehicles in the database. Four types of vehicles are focused on in the study including (1) passenger car, (2) passenger van, (3) pickup/light truck, and (4) large size vehicle. Large size vehicle is combined with medium

truck, heavy truck, truck-tractor, motor home, and bus, since the sample size of these types of vehicles is relatively small. The other vehicles such as motorcycle, moped, terrain vehicle, and train were excluded from further analysis. Results show that there are 69.38% automobiles, 8.06% passenger vans, 18.5% pickup/light trucks, and 4.06% large size vehicles involved in rear-end accidents as striking vehicle. As shown in Fig. 4, the RAIRs for passenger van, pickup/light truck, and large size vehicle are higher than passenger car. In the logistic regression model, adjusting other factors, it shows a significant trend that the accident risk is increasing with increment of vehicle size (see Table 1). The odds of large size vehicles could be 25% higher than automobiles. This result is consistent with the conclusion drawn by the Federal Motor Carrier Safety Administration (FMCSA, 2004) that trucks strike other vehicles in the rear much more often than they are struck by other vehicles. Large size vehicles are heavier than automobiles in the traffic stream. Relatively, they are less maneuverable and take longer to stop. Bus or truck drivers especially sit up much higher than passenger vehicle drivers and can see much further down the road, but they may have difficulty responding to brake light of the leading car with a small headway.

3.2. Interaction effects

After confirming the main effect model, the next regression analysis is to explore the possible significant interactions between these factors. It is found that there are three inter-

action factors associated with rear-end accidents including: driver age and number of lanes ( $P$ -value = 0.008), driver age and accident time ( $P$ -value = 0.002), and speed limit and number of lanes ( $P$ -value = 0.000). Based on 0.01 significance level, driver age and gender is not statistically significant, but its  $P$ -value is very small (0.012). Since they are always important to the traffic safety research, they are also included in the discussion of this study.

Fig. 5 illustrates the effects of these interaction factors by relative accident involvement ratios. The results show that the driver gender difference is mainly present in the younger and middle groups, and after 55 years of age, the gender difference is not apparent. Besides the oldest group, male drivers younger than 55 years have the larger rear-end propensity compared to female drivers. The driver age effect is also influenced by driving environments. The figure shows that groups older than 55 years are more sensitive to the larger number of lanes than younger groups and they show higher accident involvement risk on 4 or 6-lane highways than 2-lane ones. Moreover, the results confirm that accident time effects are difference in age groups. For youngest and oldest groups, rear-end accidents more likely to happen during daytime than night; for middle age groups, the difference is not very apparent. Generally, older road users react more slowly to events that are not expected and take significantly longer to make decisions than younger road users. This difference becomes more exaggerated in complex situations at intersections with larger number of lanes and higher traffic volume.

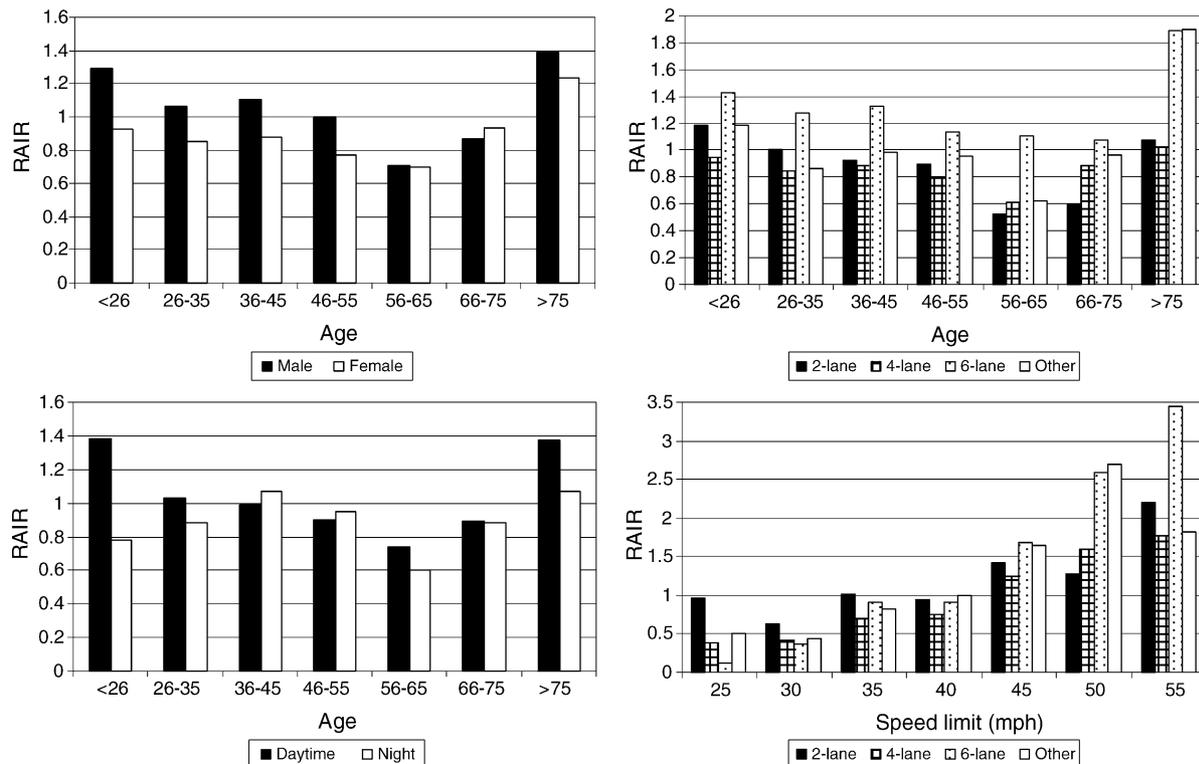


Fig. 5. Effects of interaction factors by relative accident involvement ratios.

Additionally, the results show interaction effect between speed limit and number of lanes. For lower speed limits (25–35 mph), the rear-end risk of 2-lane highway is larger than those with 6 lanes; however, for higher speed limits (45–55 mph), the rear-end risk of 2-lane highway is lower than those with 6 lanes, presumably because the intersections located on 2-lane highways with lower speed limit are more likely to be located in urban areas with higher traffic volume and those with higher speed limit are more likely to be in rural areas with lower traffic volume compared to 6-lane highways.

**4. Modeling results for struck drivers/vehicles**

In most two-vehicle rear-end cases, the leading vehicles were stopping or staying at signalized intersections and had no improper action. However, their sudden stops due to signal change or emergent situation with unexpectedly higher deceleration rates may contribute to rear-end accidents. Therefore, it is possible to find some significant patterns for the struck vehicle/driver associated with the accident risk. The term relative accident involvement ratio and logistic regression method can still be applied to perform this analysis. Then in Eq. (1), the  $D1_i$  is the number of struck drivers of type  $i$  in rear-end accidents;  $V1_i$  is the number of struck vehicles of type  $i$  in rear-end accidents; but the definitions for  $D2_i$  and  $V2_i$  keep unchanged.

Based on the results of logistic regression, there are four factors found at a 0.0001 level of significance, including vehicle type, driver age, driver residence, and gender. Table 2 lists the model estimation and odds ratios properly adjusting other factors and Fig. 6 shows the relative accident involvement ratios by struck driver characteristics and vehicle types.

For the leading driver age, the figure shows an opposite-U-shape trend: the older group and younger group have similar and smaller accident propensity, but middle age group has larger propensity. Drivers of 46–55 years of age are most likely to be involved in rear-end accidents as struck drivers, whose accident risk could be 74% higher than the younger drivers less than 26 years. The presumable reason is that for middle age drivers, better driving experience and physical performance may be helpful to detect a potential conflict earlier than younger and older drivers, but their emergent stop could contribute to their struck role in a rear-end accident. Furthermore, it is necessary to point out that for the quasi-induced exposure, the native weakness of the technique to obtain exposure from the database is that it does not account for the fact that skilled driving can be used to reduce not-at-fault accident involvements (Preusser et al., 1998). For the middle age drivers, they may be under-represented in the exposure group (victims in non-rear-end accidents) due to their better driving performance, but as struck drivers in rear-end accidents, they have no advantage to avoid the accidents compared to other groups. Therefore, their rear-end involvement ratio as struck drivers is possibly overestimated. The results also show that the female divers are more likely to be in the struck role as compared to the male drivers ( $P$ -value < 0.0001), presumably because the female drivers are more likely to brake when faced with critical traffic events (for example, yellow signal) than the male drivers.

As similar as striking drivers, if drivers are unfamiliar with the driving environment, they are more likely to be involved in the accidents as struck drivers. As shown in Table 2, the struck risk of the non-local drivers who lived in Florida area, other state, or other country could be 52%, 99%, and 126% higher than the local drivers. The non-local drivers have more difficulties to find their destination and possibly have abnor-

Table 2  
Model estimation and odds ratios for significant independent variables for struck role

Parameter	Coefficient estimate	Odds ratio	95% Wald confidence limits		Wald $\chi^2$	P-value
Intercept	-1.254				1250.918	<0.0001
Driver age					184.1119	<0.0001
>75 vs. <26	0.148	1.159	0.963	1.395	2.425	0.119
66–75 vs. <26	0.416	1.516	1.323	1.737	35.851	<0.0001
56–65 vs. <26	0.497	1.643	1.475	1.832	80.674	<0.0001
46–55 vs. <26	0.556	1.744	1.592	1.912	142.112	<0.0001
36–45 vs. <26	0.422	1.525	1.403	1.659	97.216	<0.0001
26–35 vs. <26	0.312	1.365	1.255	1.485	52.811	<0.0001
Residence code					166.8871	<0.0001
4 vs. 1	0.813	2.256	1.522	3.343	16.410	<0.0001
3 vs. 1	0.687	1.988	1.698	2.327	73.091	<0.0001
2 vs. 1	0.418	1.518	1.394	1.654	91.102	<0.0001
Gender					91.6433	<0.0001
Female/male	0.280	1.323	1.249	1.401	91.643	<0.0001
Type of vehicle					36.6844	<0.0001
Large vehicle vs. car	-0.323	0.724	0.615	0.853	14.963	0.000
Light truck vs. car	0.163	1.177	1.09	1.272	17.052	<0.0001
Van vs. car	-0.006	0.994	0.897	1.101	0.015	0.902

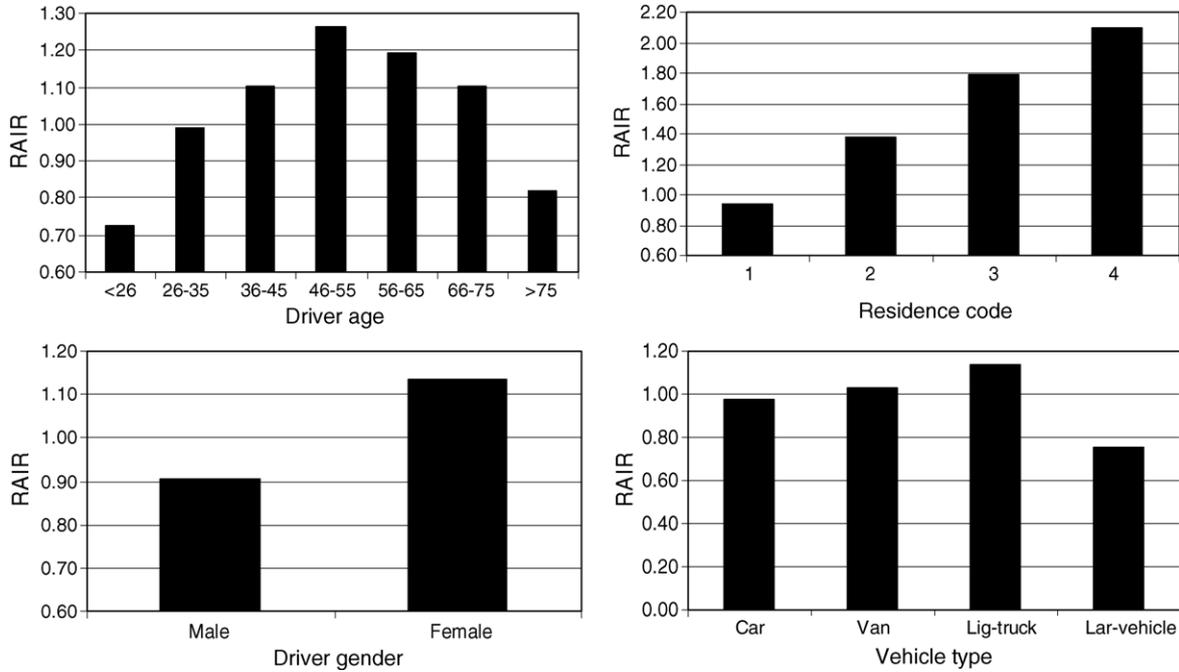


Fig. 6. Relative accident involvement ratios by struck driver characteristics and vehicle types.

mal behaviors at intersections, such as improper changing lane or sudden stop for right/left-turn, which may contribute to rear-end accidents.

The effect of struck vehicle type supports earlier findings. The light trucks are 18% more likely to be in the struck role as compared to the passenger car ( $P$ -value  $< 0.0001$ ). Abdel-Aty and Abdelwahab (2004) explained that light truck vehicles obscure drivers' visibility of other passenger cars. They may prevent drivers in cars behind them from being aware of traffic situation ahead, therefore more susceptible to collide with the light trucks in case of sudden application of breaks. Moreover, the results indicate that the risk of large size vehicle could be 28% lower than passenger cars. Normally, large size vehicles are slowly and noticeable in the traffic stream, their deceleration rate is lower than automobiles, and those aggressive drivers dislike following a truck or bus. So, the large size vehicle (trucks or buses) are less likely to be in the struck role.

### 5. Discussions on the effect of traffic volume on rear-end accidents

Since the DHSMV database only recorded the average daily traffic data for those state highway intersections, the full traffic volume data is not available in the database. To check the effect of traffic volume on the model results, the model was recalibrated based on the available data from Dade County (in Florida) which include 1248 observations that recorded the average daily traffic (ADT) on the major road at the intersections where the accidents occurred.

Comparing the rebuilt model results of multiple logistic regression without ADT to that with ADT, the coefficients of most independent variables did not show significant changes. However, it was found that the number of lanes is apparently confounded with the ADT. Normally, only the roadways with larger daily traffic have more lanes to handle heavy traffic, so the number of lanes and ADT are positively correlated. This analysis indicated that either of them can be used to fit the model but they may not appear in the same model together.

According to univariate logistic regression analyses, the ADT, treated as continuous variable, is significantly associated with the rear-end risk ( $P = 0.0018$ ). Generally, higher traffic volumes result in smaller gaps between vehicles and consequently higher rear-end probability. The odds ratio estimator for ADT is 1.12 and its confidence interval at the 0.95 significance level is [1.043, 1.203]; without considering other factors, drivers on the highway with the higher average

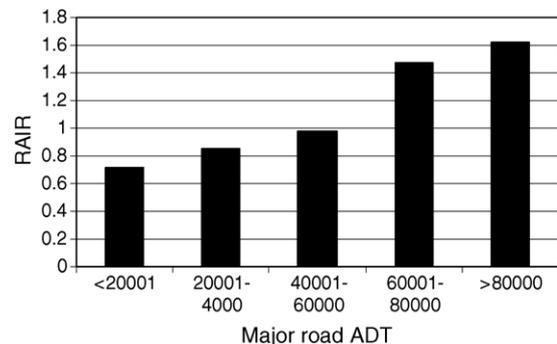


Fig. 7. Relative accident involvement ratios by major road ADT.

daily traffic flow might be 12% more likely to be involved in rear-end accidents compared to those with the average daily traffic flow that is 10000 vehicle/day lower. Segmenting ADT by 20000-vehicle/day interval, as a categorical variable, five ADT levels include ADT less than 20 001, 20 001–40 000, 40 001–60 000, 60 001–80 000, and more than 80 001 vehicles. Comparing relative accident involvement ratios, the graph for ADT illustrated in Fig. 7 shows a clear trend that the rear-end accident risk increases as the traffic volume increases.

## 6. Conclusions and suggestions

Because signalized intersections are accident-prone areas for rear-end accidents, in order to develop effective rear-end accident countermeasures, it is important to understand the driver characteristics, vehicle types, and road environment features of rear-end accidents. Using the 2001 Florida traffic accident database, this study investigated the overall characteristics of rear-end accidents at signalized intersections based on the quasi-induced exposure concept and logistic regression method.

The analysis showed that the risk of rear-end accidents for 6-lane highways is higher than 2-lane and 4-lane highways; the rear-end accidents are more likely to happen at divided highways than undivided highways. However, the crash trend for higher number of lanes and divided highways may be confounded by the traffic volume effect because the highways with high heavy traffic tend to have more lanes and be separated by medians; the results confirmed that the higher traffic volume contributes to the higher rear-end risk. It was found that the relative accident involvement ratio for night is apparently lower than daytime; compared to a dry road surface, the wet and slippery road surface could greatly contribute to rear-end accidents; furthermore, as the highway character is more complex, the rear-end accident is more likely to occur. When horizontal curve and grade are present at the same time, the rear-end risk could be twice as that for normal straight highways. Moreover, the analysis showed a clear trend that as the speed limit increases, the risk of the rear-end accidents increases, especially when the speed limit is higher than 40 mph.

Corresponding to the adverse road environment factors, appropriate engineering countermeasures need to be considered to reduce the rear-end crash rate. From the perspective of the intersection design and operation, improvement of configuration conditions (geometrics) may contribute to reducing reaction and stopping times, eliminating motorist confusion, and improving visibility of traffic control devices. It is suggested to avoid designing intersections located on a horizontal curve and vertical curve if possible. For an existing intersection with a curve or up/down grade, sufficient sight distance not only to the signal head but also to the other approaches should be satisfied, in order that the drivers going-through the intersection can detect potential conflicting vehicles in

time. In addition, motorist information countermeasures may provide advance information to the driver about the signal ahead, such as advanced warning signs installed upstream at the intersection. That may reduce sudden-stop behaviors because of insufficient reaction time due to a signal change (dilemma zone). In the area where drivers frequently drive in the rainy weather condition, the drainage design should ensure that the rainwater is removed from the intersection surface in time. If intersections with the 50 or 55 mph speed limit have been identified to have higher rear-end crash rates, reducing the speed limit to 40 or 45 mph may efficiently contribute to the lower accident rate.

The modeling results indicated that the driver factors as the striking role have different effects on the rear-end risk compared to the struck role. The struck drivers are more likely to be middle-age and female drivers, while the striking drivers with relatively larger accident propensity tend to be male, younger (<26), or older drivers (>75). For striking drivers, the 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. In general, the middle age drivers are usually under-involved in accidents due to better physical conditions and more driver experiences compared to the younger and older drivers. However, as the struck role in rear-end accidents, the middle age drivers have no advantage to void the accidents and even their quick reaction to the hazard in front of them might contribute to the struck role in a rear-end accident. Based on vehicle type, the rear-end accident risk for the striking role is increasing with increment of vehicle size, but compared to the passenger car, the large trucks or buses are less likely to be in the struck role and light trucks are more likely to be struck. For the driver residence, the non-local drivers always have higher accident risk compared to the local drivers, whatever as the striking role or struck role in a rear-end accident.

Corresponding to the higher risk driver populations, generally, younger drivers have less driving experience and tend to drive in situations conditions that increase their risk. An education program to emphasize the rear-end risk at signalized intersections is strongly suggested for the younger group. According to a precious study (Eby and Molnar, 1998), some of these young drivers may engage in risky driving behaviors because they are risk taking (i.e. they perceive the risk and do it anyway) and others engage in this behavior because they are risk ignorant (i.e. they do not perceive the risk in their behavior). Those who are risk ignorant may benefit from the education countermeasure and become safer drivers by having a better understanding of the risk inherent in their driving behaviors. For the older drivers, their higher rear-end risk may result from deteriorating physical conditions, decreasing judgment ability, and vision problem. It is necessary to make a further analysis of the criteria of driving license issuance for the older drivers related to driver age and health condition. In the other hand, properly designed and implemented, the avoidance warning systems (CAWS) would notify drivers

about potential dangers from roadway departures and other automobiles, particularly in rear-end collisions (James and Sarah, 2003). As a sort of in-vehicle technologies, CAWS may help reduce the crash risk for those drivers with weak driving abilities; CAWS may also effectively help bus and truck drivers to perceive the brake behavior of the leading vehicle since they may have difficulty responding to brake light of the leading car with a small headway because of a higher eyesight position. Moreover, the analysis confirmed the substantial effect of alcohol/drug use on driver's safety. Even drivers who had been drinking under legal alcohol use level could be 9.58 times more likely involved in a rear-end accident than non-drinking drivers, which strongly supports that the threshold of legal alcohol use level on the road should be reduced. If the data of illegal blood alcohol concentration (BAC) are available in traffic accident database, the further quantitative analysis of relationship between BAC and accident risk (not limited in rear-end accident) is strongly suggested.

Lastly, it is worth mentioning that the analyses in this paper pertain to identifying the significant factors and comparing relative accident involvements between different traffic conditions but not to predicting rear-end occurrence rate. The accident propensity analyses in this paper may provide a better understanding of the rear-end risks and more information to seek effective accident countermeasures.

## References

- Abdel-Aty, M., Abdelwahab, H., 2004. Modeling rear-end collisions including the role of driver's visibility and light truck vehicles using a nested logit structure. *Accid. Anal. Prev.* 36, 447–456.
- Abdel-Aty, M., Keller, J., Brady, P.A., 2005. Analysis of the types of crashes at signalized intersections using complete crash data and tree-based regression. In: *Proceedings of the TRB 84th Annual Meeting*, Washington, DC, January 9–13, 2005.
- Aldridge, B., Himmler, M., Aultman-Hall, L., Stamatiadis, N., 1999. Impact of passengers on young driver safety. *Transport. Res. Rec.* 1693, 25–30.
- Carr, B.R., 1970. A statistical analysis of rural Ontario traffic accidents using induced exposure data. In: *Proceedings of the Symposium on the Use of Statistical Methods in the Analysis of Road Accidents*, OECD, Paris, pp. 86–72.
- David, W.H., Lemeshow, S., 1989. *Applied Logistic Regression*. John Wiley, New York.
- Eby, D.W., Molnar, L.J., 1998. Matching safety strategies to youth characteristics: a literature review of cognitive development. Technical Report. US Department of Transportation, Washington, DC.
- FMCSA, Federal Motor Carrier Safety administration. <http://www.fmcsa.dot.gov/safetyprogs/research/briefs/Rear.html>. Access date: October 2004.
- Haight, F.A., 1973. Induced exposure. *Accid. Anal. Prev.* 5, 111–126.
- Hing, J.Y.C., Stamatiadis, N., Aultman-Hall, L., 2003. Evaluating the impact of passengers on the safety of older drivers. *J. Saf. Res.* 34, 343–351.
- James, P.B., Sarah, A.A., 2003. Alarm mistrust in automobiles: how collision alarm reliability affects driving. *Appl. Ergon.* 34, 499–509.
- Khattak, A.J., 2001. Injury severity in multivehicle rear-end crashes. *Transportation Research Record 1746*. TRB, National Research Council, Washington DC, 2001, pp. 59–68.
- National Transportation Safety Board, 2001. Special Investigation Report—Vehicle- and Infrastructure-Based Technology for the Prevention of Rear-End Collisions, NTSB/SIR-01/01, PB2001-917003. <http://www.nts.gov/Publictn/2001/SIR0101.pdf>.
- Preusser, D.F., Williams, A.F., Ferguson, S.A., Ulmer, R.G., Weinstein, H.B., 1998. Fatal crash risk for older drivers at intersections. *Accid. Anal. Prev.* 30, 151–159.
- Santokh, S., 2003. Driver Attributes and Rear-end Crash Involvement Propensity, National Highway Traffic Safety Administration, Report No. DOT HS 809 540.
- Sayer, J., Mefford, M., Huang, R., 2000. The effect of lead-vehicle size on driver following behavior. Technical Report No. UMTRI-2000-15, UMTRI-93805. Human Factors Division, Transportation Research Institute, Michigan University, Ann Arbor.
- Stamatiadis, N., Deacon, J.A., 1995. Trends in highway safety: effects of an aging population on accident propensity. *Accid. Anal. Prev.* 27, 443–459.
- Stamatiadis, N., Deacon, J.A., 1997. Quasi-induced exposure: methodology and insight. *Accid. Anal. Prev.* 29, 37–52.
- Staplin, L., Lococo, K., Byington, S., Harkey, D., 2001. *Highway Design Handbook for Older Drivers and Pedestrians*, U.S.DOT/FHWA Publication No. FHWA-RD-01-103. Washington, DC.
- Storie, V., 1977. Male and female car drivers: drivers observed in accidents, Report no. 761. Transport and Road Research Laboratory, Crowthorne, UK.
- Strandberg, L., 1998. Winter braking tests with 66 drivers, different tires and disconnectable ABS. Paper presented at International Workshop on Traffic Accident Reconstruction, Tokyo.
- Ulleberg, P., 2001. Personality subtypes of young drivers. Relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. *Transport. Res. Part F: Traffic Psychol. Behav.* 4, 279–297.