Accident Analysis and Prevention 40 (2008) 1703–1712

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Presence of passengers: Does it increase or reduce driver's crash potential?

Chris Lee [a,*], Mohamed Abdel-Aty [b]

[a] Department of Civil and Environmental Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada
[b] Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816, USA

ARTICLE INFO

Article history:
Received 21 February 2007
Received in revised form 2 January 2008
Accepted 5 June 2008

Keywords:
Passenger
Driver
Freeway crash
Bivariate probit
Binary logit

ABSTRACT

This study examines the impact of passengers on the driver's crash potential on freeways. To estimate the impact, a set of bivariate probit models were developed using the 5-year (1999–2003) crash records on a 36.3-mile stretch of Interstate-4 freeway (I-4) in Orlando, Florida. Bivariate probit models identify the correlation between potentially inter-related choices of three passenger characteristics and three crash characteristics. The analysis using bivariate probit models showed that there exist strong correlations between passenger and crash characteristics. It was found that drivers generally display safer driving behavior when they are accompanied by passengers, and more passengers reduce driver's crash potential. It was also found that younger driver's crash potential increases with the presence of a younger passenger only. In addition, the analysis of crash type using traffic flow parameters at the time of crashes showed that young drivers with only younger passengers are more likely to be involved in single-vehicle crashes in high-speed and low-volume conditions. The findings in this study provide insight into how the presence of passengers has an impact on driver behavior and traffic safety in various conditions.

1. Introduction

There have been debates whether the presence of passengers positively or negatively affects driver behavior leading to crash occurrence. While some studies suggested that passengers prevent drivers' risky driving behavior, some studies suggested that passengers distract drivers and consequently increased crash potential. These contradictory effects take place depending on the characteristics of drivers and passengers such as age and sex. For instance, crash potential is reduced when middle-age passengers accompany teenage drivers as guardians whereas crash potential increases when teenage passengers accompany teenage drivers. Thus, it is important to identify the correlation between passenger characteristics and crash characteristics to evaluate the effect of passengers on crash potential.

In this regard, some studies have shown that passengers may reduce driver's risky behavior (Isaac et al., 1995; Vollrath et al., 2002; Hing et al., 2003; Rueda-Domingoa et al., 2004) while others have shown that they distract and increase driver's crash potential (Doherty et al., 1998; Cooper et al., 2005; Simons-Morton et al., 2005). The latter findings have been focused on teenagers and younger drivers. In fact, these contradictory effects depend on driver's and passenger's demographic factors (e.g. age, gender). For instance, Rueda-Domingoa et al. (2004) found that the presence of passengers had more protective effect for older drivers than younger drivers in Spain. However, they found that the protective effect of passengers differ by the age of passengers. Some studies focused more on the specific high-risk driver age group (e.g. teenage drivers) and their crash risk associated with the presence of passengers. Doherty et al. (1998) found passengers have a negative effect on crash rates for particularly teenage drivers. Similarly, Cooper et al. (2005) observed that younger drivers are distracted, rather than being protected, when they are accompanied by younger passengers. They observed that restricting teenage passengers who accompany new teenage drivers significantly reduced the crash involvements of teenagers in California. Similar findings were also reported in Lam et al. (2003). Williams and Shabanova (2002) also found that teenage drivers are less likely to wear seatbelts when they are carrying more teenage passengers. Some researchers suggested that passenger's gender also has influence on risky driving. For example, Simons-Morton et al. (2005) observed that teenage drivers displayed more aggressive driving behavior (indicated by mean speed higher than posted speed limit and shorter headways) when they were accompanied by male teenage passengers than female teenage passengers.

In some cases, injury severity of crashes was also related to the presence of passengers. Preusser et al. (1998) reported that the young drivers (younger than 24) are more likely to be involved in

* Corresponding author. Tel.: +1 519 253 3000; fax: +1 519 971 3686.
E-mail addresses: cclee@uwindsor.ca (C. Lee), mabdel@mail.ucf.edu (M. Abdel-Aty).

0001-4575/$ – see front matter
doi:10.1016/j.aap.2008.06.006

fatal crashes than older drivers when they are accompanied by passengers. Similarly, Lin and Fearne (2003) found that the presence of passengers increases the likelihood of a fatal injury of young drivers. On the other hand, Isaac et al. (1995) reported that alcohol-impaired drivers are less likely to be fatally injured when they are accompanied by passengers. They suggested that this is because unimpaired passengers can reduce the risk of alcohol-involved crashes. Some studies focused on the injury severity of passengers involved in crashes. Williams and Wells (1995) observed that the death rate of teenage passengers relative to older passengers is higher than the death rate of teenage drivers relative to older drivers. They also reported that two-thirds of total teenage passengers killed in crashes were driven by teenage drivers. In particular, the passenger's seat (e.g. front or back/driver-side or passenger-side seat) in vehicles is also closely related to the injury severity of child passengers (Glass and Graham, 1999).

The number of passengers is also a significant factor affecting crash risk. Keall et al. (2004) found that the crash risk of driving with more than one passenger is higher than the risk of driving alone or driving with a single passenger. Doherty et al. (1998) suggested that the number of passengers should be restricted for novice/young drivers to prevent the distraction from passengers. Hing et al. (2003) found that older drivers (75 or older) were more involved in crashes when they carried two or more passengers. They suggested that this may be because drivers are more likely to be distracted by more passengers.

In fact, the presence of passengers is more effective in reducing crash risk under certain traffic, environmental and road geometric conditions. For instance, Vollrath et al. (2002) observed that the presence of passengers can help reduce crash risk particularly when traffic is moving slowly and lighting is dark. This finding seems to suggest that passengers help relieve drivers' impatience during traffic congestion and improve drivers' visibility in dark condition (or potentially relieve driver's fatigue during nighttime driving). The positive impact of passengers on reducing crash risk in adverse weather condition, particularly for older drivers, was also found in Hing et al. (2003). This study also found that crash risk is higher with the presence of passengers on curved road sections with higher grade.

From the review of the past studies, driver and passenger characteristics are important factors affecting crash risk and particularly the certain combinations of driver–passenger age groups (e.g. young drivers accompanied by young passengers) have higher crash risk. Also, the effects of traffic, environmental and road geometric factors on crash risk vary with the presence or absence of passengers. The presence of passengers has generally positive effects on reducing crash risk under congested traffic and adverse weather conditions.

However, there exist some limitations in these past studies. First, most statistical models used in the analyses do not properly take into account the potential correlations of inter-related variables such as the presence of passengers and crash potential. In other words, most models only describe the relationship between one dependent variable and a set of explanatory variables, but do not consider the correlation of multiple dependent variables in different equations that are simultaneously estimated. Second, most studies overlooked the combined effect of traffic flow parameters with the presence of passengers in the analysis. Although Vollrath et al. (2002) considered this effect, the parameters used in their study are subjective and qualitative measures of traffic density (e.g. dense, slow traffic) and it is difficult to determine the general quantitative effect of traffic flow on crash potential. Thus, more objective and quantitative measures of traffic flow should be used to identify its impact on crash potential associated with the presence of passengers.

Thus, the objectives of this study are (1) to develop the models that consider the correlation of inter-related choices associated with the presence of passengers and crash potential, and (2) to identify the traffic conditions leading to crashes with the presence of passengers using the short-term traffic flow data at the crash time. From this analysis, we can better understand the complex interactions among driver/passenger characteristics, traffic/environmental conditions, road geometry and crash potential.

2. Data description and preliminary analysis

This study used 5-year crash data that have occurred on a 36.3-mile stretch of Interstate-4 freeway (I-4) in Orlando, Florida. This stretch was chosen since inductive loop detectors were installed in every 0.5 mile on this section of the freeway. These data were extracted from the Florida Traffic Crash Records Database. A total of 2817 crashes occurred on the mainline freeway from 1999 to 2003. Since more than one driver was involved in some of these crashes, one driver was randomly selected from each multi-vehicle crash (regardless of driver citation) to avoid over-representation of crashes where many drivers are involved. For each crash, the information on the selected driver and his/her passenger(s) (if the driver was accompanied by at least one passenger) including driver's age and gender, and passenger's age was collected from the database. Due to the absence of the information on passenger's gender from 1999 to 2001, passenger's gender was not considered in this analysis. In addition, the information on the time of crash occurrence and environmental conditions at the crash time was also collected. It should be noted that since the presence of passengers was only known for the drivers who were involved in crashes, not for driving population, exposure could not be measured. Instead, the sample of non-cited drivers was considered as a comparison group relative to cited drivers in order to estimate the likelihood of individual driver's citation when passengers were present (i.e. the effect of the presence of passengers on driver citation). Thus, the term "crash potential" in this study essentially implies the relative likelihood of crash causation by specific driver group with or without passengers compared to other driver groups.

Given that traffic and road geometric conditions are also important factors affecting crash potential associated with the presence of passengers, traffic flow characteristics prior to the crash time were obtained from loop detectors close to the crash sites, and the curvature of road sections was also obtained. Traffic flow parameters include 5-min average speed, volume and occupancy 5–10 min prior to the time of crashes. The road section is classified into the curved section ($0 <$ radius of curvature ≤ 3000 feet) and the straight road section (otherwise).

Among 2817 crashes, approximately 62% of the selected drivers drove alone and 38% of them were accompanied by at least one passenger. The list of variables related to crashes is summarized in Table 1. Using the data set including all the selected drivers, the composition and behavior of drivers with and without passengers were compared in terms of the following factors (categories in parentheses): driver's age (younger, middle or older), driver's gender (male or female), residence (local or non-local), alcohol/drug use (yes or no), seatbelt use (yes or no), citation for crash occurrence (at-fault or not-at-fault) and day of week (weekdays or weekends). A Chi-square test was performed to compare the proportions expressed as percentages of drivers with and without passengers for the corresponding category. The p-value less than 0.05 indicates that two proportions are significantly different at a 95% confidence level.

As shown in Table 2(a), As shown in Table 2(a), the results of the Chi-square test (i.e. low p-values) indicate that among the drivers involved in crashes, the proportion of the drivers with passengers

Table 1
List of variables related to crashes

Type	Variable	Categories
Driver characteristics	Driver's age	Younger (16–24) Middle (25–59) Older (60 and over)
	Driver's gender	Male Female
	Driver's residence	Local area Non-local area
	Driver's alcohol/drug use	No alcohol/drug use Alcohol/drug use
	Driver's seatbelt use	No seatbelt use Seatbelt use
	Driver citation	Not cited for crash occurrence Cited for crash occurrence
Passenger characteristics	Presence of passengers	Drive alone Drive with at least one passenger
	Number of passengers	Drive with one passenger Drive with more than one passengers
	Dummy for younger driver accompanied by only younger passenger(s)	All passengers are 16–24 At least one passenger is younger than 16 or older than 24
Crash characteristics	Time of crash occurrence	
	Time of day	Night (7 pm–3 am) Otherwise
	Day of week	Weekdays Weekends
	Type of crash	Single-vehicle crash Multi-vehicle crash
	Driver's injury severity	Fatal/severe injury Otherwise
Environmental condition at the crash time	Weather	Normal Adverse
	Road surface	Dry Non-dry
Traffic condition prior to the crash time	5-min average speed (miles/h) 5-min average volume (vehicles) 5-min average occupancy (%)	
Road geometry	Curvature	Curved (0 < radius of curvature ≤ 3000 feet) Straight (otherwise)

is significantly different from the proportion of the drivers without passengers in terms of driver's residence, alcohol/drug use, seatbelt use, day of week and driver citation. This implies that the likelihood of involvement in crashes is significantly correlated with the presence of passengers. It was observed that the proportions of the drivers who were from local area, were alcohol-impaired, drove during weekdays, did not wear seatbelt and were cited for crash occurrence were significantly lower when the drivers were accompanied by passengers. On the other hand, some factors were not found to be correlated with the likelihood of involvement in crashes associated with the presence of passengers. As shown in Table 2(a), the proportions of driver's age and gender do not differ between drivers with and without passengers.

It is interesting to note that the likelihood of involvement in crashes associated with the presence of passengers varies across three driver age groups. As shown in Table 2(b), the proportions of at-fault and middle-age drivers with passengers are significantly lower than the proportions of at-fault drivers without passengers (p-values less than 0.05) whereas the proportions of at-fault older drivers are not significantly different (p-value greater than 0.05). This suggests that passengers have more positive effect in preventing driver citation for younger and middle-age drivers than older drivers.

In terms of road geometry, the curvature of road sections was not found to be correlated with the presence of passengers and crash potential as the likelihoods of driver's being cited on curved sections were almost the same as the likelihood on straight sections regardless of the presence of passengers. In addition, there were no significant effects of weather and road surface conditions on crash potential associated with the presence of passengers. In terms of traffic characteristics, traffic conditions prior to crash occurrence were not dramatically different between crashes of "at-fault" drivers with and without passengers. However, assuming that traffic is defined to be "congested" when a 5-min average occupancy exceeds critical occupancy of 0.13, crashes are slightly less likely to occur in congested traffic condition when at-fault drivers carry passengers (35% of total drive-with-passenger crashes) than when at-fault drivers drive alone (37% of total drive-alone crashes). In spite of this small difference, this suggests that passengers have more positive effect on drivers in reducing crash potential in congested condition than uncongested condition.

The number of passengers who accompany drivers and the combination of driver's and passenger's age groups were also found to be correlated with the likelihood of involvement in crashes. The results of the Chi-square test in Table 2(c) show that younger drivers with fewer younger passengers are more involved in crashes and younger drivers with only younger passengers are more likely to be cited for crash occurrence compared to older drivers with passengers.

Fig. 1 illustrates the correlation of driver's age and passenger's age. This figure shows which drivers' age groups are more likely to be accompanied by which passenger's age groups. It was observed from the figure that drivers are generally accompanied by the passengers whose ages are similar to drivers' ages (e.g. friend, spouse)

Table 2
Descriptive statistics of crashes by presence of passengers

(a) Composition and behavior of drivers with and without passengers

Category	Variable	Drive alone (no. of observation = 1746)	Drive with passengers (no. of observation = 1071)	Chi-square (p-value)*
Driver composition	Driver's age			
	Younger (16–24)	444 (25.4%)	273 (25.5%)	
	Middle (25–59)	1181 (67.7%)	704 (65.7%)	
	Older (60 and over)	121 (6.9%)	94 (8.8%)	
	Driver's gender			
	Male	1135 (65.0%)	717 (67.0%)	
	Female	611 (35.0%)	354 (33.0%)	
	Driver's residence			
	Local area	902 (51.7%)	419 (39.1%)	41.8
	Non-local area	844 (48.3%)	652 (60.9%)	(<0.0001)
Driver behavior	Driver's alcohol/drug use			
	Yes	106 (6.1%)	39 (3.6%)	8.0
	No	1640 (93.9%)	1032 (96.4%)	(0.0047)
	Driver's seatbelt use			
	No	157 (9.0%)	52 (4.9%)	15.6
	Yes	1589 (91.0%)	1019 (95.1%)	(0.0001)
	Day of week of driving			
	Weekdays	1415 (81.0%)	773 (68.4%)	57.4
	Weekends	331 (19.0%)	338 (31.6%)	(<0.0001)
	Driver citation			
	At-fault	727 (41.6%)	251 (23.4%)	96.3
	Not-at-fault	1019 (58.4%)	820 (76.6%)	(<0.0001)

(b) Comparison of driver citation across driver age groups with and without passengers

Driver citation	Driver's age					
	Younger (16–24)		Middle (25–59)		Older (60 and over)	
	Drive alone	Drive with passengers	Drive alone	Drive with passengers	Drive alone	Drive with passengers
Not-at-fault	251 (56.5%)	189 (69.2%)	690 (58.4%)	559 (79.4%)	78 (64.5%)	72 (76.6%)
At-fault	193 (43.5%)	84 (30.8%)	491 (41.6%)	145 (20.6%)	43 (35.5%)	22 (23.4%)
Total	444 (100%)	273 (100%)	1181 (100%)	704 (100%)	121 (100%)	94 (100%)
Chi-square (p-value)*		11.0 (0.0009)		86.0 (< 0.0001)		3.1 (0.0772)

(c) Comparison of driver behavior across different combinations of driver–passenger ages

	Younger drivers with only younger passengers	Older drivers with passenger	Chi-square (p-value)*
Carry one passenger	129 (58.4%)	380 (47.6%)	7.6
Carry more than one passengers	92 (41.6%)	418 (52.4%)	(0.0057)
Total	221 (100%)	798 (100%)	
Driver at fault	71 (32.1%)	167 (20.9%)	11.5
Driver not at fault	150 (67.9%)	631 (79.1%)	(0.0007)
Total	221 (100%)	798 (100%)	

* The p-value less than 0.05 indicates that two proportions are significantly different at a 95% confidence level.

as indicated by the cluster of data points along a diagonal line with the slope of 1. Another cluster of data points below the diagonal line indicates that a substantial portion of middle-age drivers carries younger passengers (e.g. children).

From this preliminary analysis, it was found that the presence of passengers generally encourages driver's safer driving behavior; however, this positive impact of passengers varies across driver's age—the presence of younger passengers is less helpful for younger drivers than older passengers. In spite of some significant findings, this descriptive statistical analysis cannot clearly identify the association of multiple factors in complex relationships. Thus, this study proposes to use a discrete choice model which accounts for the correlations among choices that are made simultaneously.

3. Analysis using bivariate probit models

When we hypothesize that the passenger-related factors would impact on the crash-related factors, we need to predict the probability of crashes given that passengers are present or not (conditional probability). For this purpose, we have to instrument the passenger-related variable (i.e. estimate it using a model then use the estimated value rather than the observed value as an explanatory variable). Thus, the presence of passengers should be defined as a dependent variable, not an explanatory variable. In this case, one can suggest using two logistic regression models to identify the association of passengers with crashes. However, a dependent variable in one regression model cannot be used as an explanatory variable in the other regression model because of heterogeneity.

In this regard, a bivariate probit model was developed since it can account for a potential correlation of "unobserved" effects between inter-related choices (or responses), such as the presence of passengers and crash potential. For instance, two choices can be defined as follows: the first choice is whether a driver is driving alone ($y_1 = 0$) or accompanied by one or more passengers ($y_1 = 1$), and the second choice is whether a driver is not cited for crash occurrence ($y_2 = 0$) or cited for crash occurrence ($y_2 = 1$). The bivari-

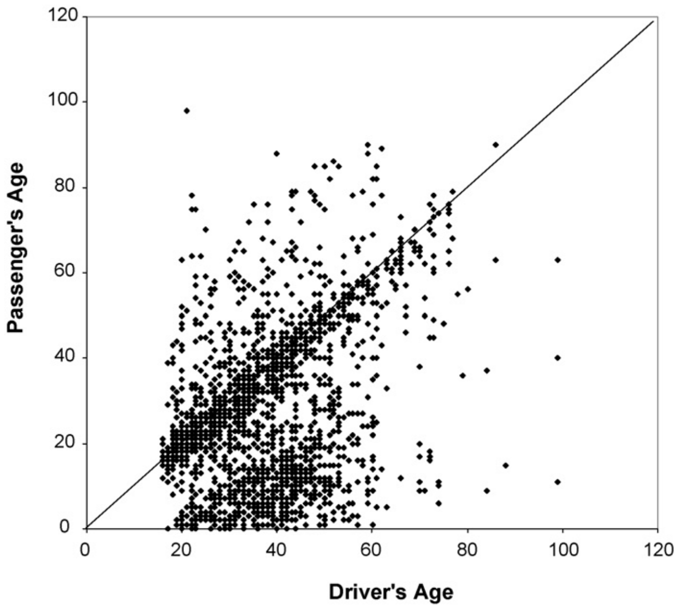

Fig. 1. Distribution of driver's age and pedestrian's age.

ate probit model identifies the factors affecting these two choices and also captures the correlation between them as follows:

$$\begin{aligned} z_1 &= \beta_1 x_1 + \varepsilon_1, \quad y_1 = 1 \quad \text{if } z_1 \geq 0, \ y_1 = 0 \ \text{otherwise} \\ z_2 &= \alpha z_1 + \beta_2 x_2 + \varepsilon_2, \quad y_2 = 1 \quad \text{if } z_2 \geq 0, \ y_2 = 0 \ \text{otherwise} \end{aligned} \tag{1}$$

where z_1 is a latent variable indicating a driver is accompanied by one or more passengers, y_1 is the observed first choice (0 = drive alone, 1 = drive with passengers), z_2 is a latent variable indicating a driver is cited for crash occurrence, y_2 is the observed second choice (0 = not cited, 1 = cited), x_1 and x_2 are explanatory variables related to the first and second choice, respectively, α, β_1 and β_2 are coefficients estimated by the model, and ε_1 and ε_2 are random error terms.

Since two choices are sequentially inter-related, the bivariate probit model predicts the possible outcome of the first choice (one of two values for a binary passenger-related variable) and then the possible outcome of the second choice (one of two values for a binary crash-related variable) based on the possible outcome of the first choice. Thus, the dependent variable in the first model (z_1) is defined as one of the explanatory variables in the second model. If two choices (z_1 and z_2) are inter-related, two error terms (ε_1 and ε_2) are correlated. The positive values of coefficients α, β_1 and β_2 represent higher likelihood of $y_1 = 1$ and $y_2 = 1$ as the values of explanatory variables x_1 and x_2 increase, and vice versa. Using these models, the association of passengers with crash potential was investigated from different perspectives. First, the following three binary choices of passenger characteristics were considered:

(1) *The presence of passengers*: Drive alone (=0) or drive with one or more passengers (=1).
(2) *The number of passengers*: Drive with one passenger (=0) or drive with more than one passengers (=1).
(3) *Younger driver/younger passenger combination*: Younger driver carrying only younger passenger(s) (=1) or otherwise (=0).

The third binary choice above was considered mainly because younger drivers with only younger passengers have been commonly recognized as a driver group with high crash potential. The above three choices were related to the following three binary choices of crash characteristics:

(1) *Driver citation for crash occurrence*: Driver is cited (=1) or not cited (=0) given that a crash occurs.
(2) *Crash type*: Single-vehicle crash (=1) or multi-vehicle crash (=0).
(3) *Driver's injury severity*: Fatal/severe injuries (=1) or otherwise (=0).

Thus, a total of nine (3 × 3) different bivariate probit models were developed to investigate the inter-relationship of two binary choices and the parameters were estimated using LIMDEP (Greene, 1998). It was found that the variables included in 7 models among 9 models were statistically significant at a 90% confidence level (p-value < 0.1). The error term correlations were also statistically significant in these seven models at a 95% confidence level (p-value < 0.05), which clearly indicates that two choices are inter-related. In most other studies which analyzed the effect of passengers using logistic regression models, these important correlations could not be identified. The detailed results of each model are described as follows.

3.1. Presence of passengers—driver citation model

In this model, the association between the presence of passengers and driver citation was estimated. The first and second choices are described in the functions of variables as defined in Eq. (1). As shown in Table 3(a), it was found from the first model (presence of passengers) that drivers are more likely to carry passengers on weekends and at night. The results also indicate that drivers tend to show safer driving behavior such as wearing seatbelts and not being alcohol-impaired when passengers are present. These results are consistent with the findings in Geyer and Ragland's study (2005) using the U.S. nationwide crash database.

The results of the second model (driver citation) show that the presence of passengers is negatively correlated with driver citation. This implies that the presence of passengers partially contributes to reducing crash potential. Thus, this can be considered as a strong evidence of a positive impact of passengers on driver safety. However, apart from the presence of passengers, drivers are more likely to be cited when they are younger (16–24) and alcohol-impaired as indicated by positive coefficients of the model estimates. This suggests that the impact of the presence of passengers on driver behavior depends on driver's age and alcohol/drug use.

3.2. Number of passengers—driver citation model

In this model, the association between the number of passengers and driver citation was estimated. Thus, only the records of drivers carrying one or more passengers (1071 drivers) are used in this model. As shown in Table 3(b), it was found from the first model (number of passengers) that drivers carry more passengers on weekends and at night. This may be because there are more group activities (e.g. long-distance recreational trips) on weekends and at night. The results of the first model (number of passengers) also indicate that younger and alcohol-impaired drivers carry fewer passengers. In other words, younger drivers are less likely to drive in a large group on freeways and drivers are less likely to be intoxicated with more passengers. In particular, the latter finding seems to suggest that drivers feel more responsibility of safe driving when they carry more passengers.

The result of the second model (driver citation) shows a negative association of the number of passengers with driver citation. This implies that the presence of more passengers partially contributes to reducing crash potential. This result contradicts the findings of the earlier studies (Preusser et al., 1998; Keall et al., 2004; Hing et al., 2003; Lam et al., 2003; Williams, 2003) that higher number of passengers increases crash risk. However, as there are more passengers

Table 3
Estimated parameters of bivariate probit models (driver citation)

	Coefficient	*t*-Statistic	*p*-Value
(a) Correlation between the presence of passengers and driver citation			
First model: presence of passengers			
Constant	−0.71	−8.79	<.0001
Day of week (1 = weekends, 0 = weekdays)	0.35	7.11	<.0001
Night (1 = 7 pm–3 am, 0 = otherwise)	0.24	4.96	<.0001
Seatbelt (1 = seatbelt use, 0 = no seatbelt use)	0.31	4.00	0.0001
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	−0.42	−3.68	0.0002
Second model: driver citation (1 = cited)			
Constant	0.29	8.62	<.0001
Presence of passengers (1 = drive with passenger(s), 0 = drive alone)*	**−1.69**	**−39.36**	**<.0001**
Younger driver (1 = 16–24, 0 = otherwise)	0.12	3.22	0.0013
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	0.42	4.03	<.0001
Error-term correlation**	0.94	17.59	<.0001
Log-likelihood at convergence = −3546.40			
Number of observation = 2817			
(b) Correlation between the number of passengers and driver citation			
First model: number of passengers			
Constant	0.01	0.18	0.8583
Day of week (1 = weekends, 0 = weekdays)	0.13	1.92	0.0550
Night (1 = 7 pm–3 am, 0 = otherwise)	0.22	2.96	0.0031
Younger driver (1 = 16–24, 0 = otherwise)	−0.27	−3.51	0.0004
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	−0.93	−5.83	<.0001
Second model: driver citation (1 = cited)			
Constant	0.18	1.73	0.0832
Number of passengers (1 = more than one passengers, 0 = one passenger)*	**−1.35**	**−11.46**	**<.0001**
Error-term correlation**	0.92	13.58	<.0001
Log-likelihood at convergence = −1294.65			
Number of observation = 1071			
(c) Correlation between the younger driver/younger passenger combination and driver citation			
First model: younger driver/younger passenger combination			
Constant	−0.84	−19.00	<.0001
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	0.66	2.89	<.0039
Second model: driver citation (1 = cited)			
Constant	−0.96	−21.14	<.0001
Younger driver/younger passenger combination (1 = younger driver with only younger passenger(s), 0 = otherwise)*	**1.90**	**5.70**	**<.0001**
Error-term correlation**	−0.90	−4.93	<.0001
Log-likelihood at convergence = −1116.93			
Number of observation = 1071			

* The parameters in bold represent the dependent variable of the first model used in the second model as an explanatory variable.
** The statistically significant error-term correlation indicates that two choices are inter-related.

in the car, drivers are more likely to carry passengers of different ages. In fact, Regan and Mitsopoulos (2003) reported based on the focus group discussion that drivers (particularly younger drivers) have a greater sense of responsibility and drive more cautiously when they carry older or younger passengers. This was also partially reflected by driver's less alcohol-drug use with higher number of passengers in the first model. From this perspective, higher number of passengers may help reduce driver's crash potential rather than increasing it.

3.3. *Younger driver/younger passenger combination—driver citation model*

In this model, the association between the younger driver–younger passenger combination and driver citation was estimated. It was found from the first model (younger driver/younger passenger combination) that younger drivers are more likely to be intoxicated when they carry only younger passengers as shown in Table 3(c). This reflects that younger drivers tend to be careless and display unsafe driving behavior when they are accompanied by their peers. In the second model (driver citation), it was found that younger drivers carrying only younger passengers have higher likelihood of being cited. This result suggests that younger drivers are more distracted by younger passengers, which is more likely to lead to crashes. This result agrees with the finding in Cooper et al. (2005). It is also possible that younger drivers are more likely to purposely display risk-taking behaviors to show off in front of their young peers.

3.4. *Presence of passengers—crash type model*

In this model, the association between the presence of passengers and crash type was estimated. To identify the characteristics of drivers who cause certain type of crashes, the records of only at-fault drivers (978 drivers) were used. Similar to the driver citation model, the presence of passengers are positively correlated with weekends and night, and negatively correlated with driver's alcohol-drug use in the first model (presence of passengers) as

Table 4
Estimated parameters of bivariate probit model (crash type): correlation between the presence of passengers and crash type

	Coefficient	*t*-Statistic	*p*-Value
First model: presence of passengers			
Constant	−0.80	−15.11	<.0001
Day of week (1 = weekends, 0 = weekdays)	0.61	5.88	<.0001
Night (1 = 7 pm – 3 am, 0 = otherwise)	0.24	2.21	0.0273
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	−0.34	−2.18	0.0294
Second model: crash type (1 = single-vehicle crash)			
Constant	−1.58	−21.39	<.0001
Presence of passengers (1 = drive with passenger(s), 0 = drive alone)*	**1.17**	**3.43**	**0.0006**
Younger driver (1 = 16–24, 0 = otherwise)	0.27	2.64	0.0083
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	1.18	7.63	<.0001
Error-term correlation**	−0.72	−4.79	<.0001
Log-likelihood at convergence = −853.33			
Number of observation = 978			

* The parameters in bold represent the dependent variable of the first model used in the second model as an explanatory variable.
** The statistically significant error-term correlation indicates that two choices are inter-related.

shown in Table 4. More importantly, the presence of passengers was found to be positively correlated with single-vehicle crashes in the second model (crash type). Since single-vehicle crashes more frequently occur due to driver's error during uncongested traffic conditions than congested conditions as opposed to multi-vehicle crashes, the result implies that passengers are less helpful in reducing driver's crash potential in uncongested conditions, i.e. they are more likely to distract drivers. On the other hand, this result can be interpreted in a way that passengers can help reduce driver's crash potential in congested conditions (e.g. relieving driver's impatience (Vollrath et al., 2002) or warning drivers of an impending queue). However, the models relating the number of passengers and younger driver/younger passenger combination to crash type were not found to be statistically significant and did not yield practically reasonable results.

3.5. *Presence of passengers—driver's injury severity model*

In this model, the association between the presence of passengers and driver's injury severity was estimated. It should be noted that driver's injury severity is originally categorized in 5 levels in the crash records and these 5 levels are aggregated into the following two levels: (1) fatal/severe injury: fatal, incapacitating and non-incapacitating evident injuries and (2) non-severe injury: possible and no injuries. In the first model (presence of passengers), as shown in Table 5(a), it was again found that the presence of passengers is positively correlated with weekends and driver's seatbelt use, and negatively correlated with driver's alcohol-drug use.

The result of the second model (driver's injury severity) indicates lower likelihood of driver's fatal/severe injuries when passengers are present. Given that crash injury severity is greatly affected by the speed of a vehicle at the time of a collision, it can be speculated that driver's speed tends to be lower (i.e. cautious driving) with passengers when they are involved in crashes and the impact of the collision on drivers is likely to be weaker. However, it should be noted that since only one driver's information was randomly selected from each crash, the presence of passengers in the other vehicles (and its consequent effect on the speed of those vehicles) involved in multi-vehicle crashes cannot be captured in this analysis. In fact, the impact of the collision is also affected by the speed of the other vehicle(s) involved in the crash. However, it is expected that the impact of the collision for lower speed vehicles are likely to be weaker than that of higher speed vehicles regardless of hitting the other vehicle(s) or being hit by the other vehicle(s). Finally, it was also found that alcohol-impaired and female drivers are more likely to be fatally/severely injured than non-impaired and male drivers, respectively.

3.6. *Number of passengers—driver's injury severity model*

In this model, the association between the number of passengers and driver's injury severity was estimated. As shown in Table 5(b), the results of the first model (number of passengers) indicates higher likelihood of more passengers on weekends and at night, but less passengers for younger drivers and impaired drivers. It was found in the second model (driver's injury severity) that higher number of passengers reduces the likelihood of driver's fatal/severe injuries. This result contradicts the finding of the study by Chen et al. (2000). However, as mentioned earlier, this result indicates that drivers tend to drive more cautiously with more passengers at lower speed than driving with only one passenger and the consequence of crashes is likely to be less severe. The result also indicates that driver's injury severity generally increases with driver's age.

3.7. *Younger driver/younger passenger combination—driver's injury severity model*

In this model, the association between the younger driver/younger passenger combination and driver's injury severity was estimated. The result of the first model (younger driver/younger passenger combination) is similar to the result of driver citation model—younger drivers are more likely to be impaired when they are accompanied by only younger passengers as shown in Table 5(c). The result of the second model (driver's injury severity) shows that drivers are more likely to be fatally/severely injured when both drivers and passengers are younger. This reflects that younger drivers tend to be speeding with only younger passengers and high impact of collision results in more severe driver's injuries.

However, the effect of traffic flow parameters was not considered in the above bivariate probit models due to a limited set of loop detector data. Instead, a binary logit model which is a simpler form of discrete choice model was used to estimate this effect. The results of the binary logit model are presented in the next section.

4. Analysis using binary logit model

In this analysis, the association among two crash characteristics (i.e. crash type and driver's injury severity), three passenger characteristics (i.e. the presence of passengers, the number of passengers and the younger driver–younger passenger combination)

Table 5
Estimated parameters of bivariate probit models (driver's injury severity)

	Coefficient	*t*-Statistic	*p*-Value
(a) Correlation between the presence of passenger and driver's injury severity			
First model: presence of passengers			
Constant	−0.86	−9.07	<.0001
Day of week (1 = weekends, 0 = weekdays)	0.42	7.49	<.0001
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	−0.30	−2.58	0.0099
Seatbelt (1 = seatbelt use, 0 = no seatbelt use)	0.50	5.24	<.0001
Second model: driver's injury severity (1 = fatal/severe injury)			
Constant	−0.86	−5.12	<.0001
Presence of passengers (1 = drive with passenger(s), 0 = drive alone)*	**−0.70**	**−2.81**	**0.0049**
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	0.55	4.25	<.0001
Driver's gender (1 = female, 0 = male)	0.20	3.75	0.0002
Error-term correlation**	0.47	2.97	0.0029
Log-likelihood at convergence = −3172.04			
Number of observation = 2817			
(b) Correlation between the number of passengers and driver's injury severity			
First model: number of passengers			
Constant	0.00	−0.07	0.9447
Day of week (1 = weekends, 0 = weekdays)	0.18	2.20	0.0275
Night (1 = 7 pm – 3 am, 0 = otherwise)	0.16	1.82	0.0689
Younger age (1 = 16–24, 0 = otherwise)	−0.23	−2.69	0.0071
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	−0.85	−4.07	0
Second model: driver's injury severity (1 = fatal/severe injury)			
Constant	−0.45	−1.62	0.1049
Number of passengers (1 = more than one passengers, 0 = one passenger)*	**−1.00**	**−2.71**	**0.0067**
Driver's age (continuous)	0.01	1.81	0.0699
Error-term correlation**	0.60	2.53	0.0114
Log-likelihood at convergence = −1249.10			
Number of observation = 1071			
(c) Correlation between the younger driver/younger passenger combination and driver's injury severity			
First model: younger driver/younger passenger combination			
Constant	−0.83	−18.86	<.0001
Alcohol (1 = alcohol-drug use, 0 = no alcohol-drug use)	0.42	1.96	0.0504
Second model: driver's injury severity (1 = fatal/severe injury)			
Constant	−1.02	−21.97	<.0001
Younger driver/younger passenger combination (1 = younger driver with only younger passenger(s), 0 = otherwise)*	**1.66**	**2.83**	**0.0046**
Error-term correlation**	−0.87	−3.60	0.0003
Log-likelihood at convergence = −1065.48			
Number of observation = 1071			

* The parameters in bold represent the dependent variable of the first model used in the second model as an explanatory variable.
** The statistically significant error-term correlation indicates that two choices are inter-related.

and traffic flow parameters collected prior to crash occurrence were investigated. Traffic flow parameters include 5-min average speed, volume and occupancy, and variations in speed, volume and occupancy during a 5-min period (measured in coefficients of variation) near the crash site before the reported time of a crash.

In the analysis of crash type, since the purpose of this analysis is to identify typical traffic conditions when drivers are more likely to cause certain type of crashes, the data for at-fault drivers were only used. In the analysis of driver's injury severity, the data for both at-fault and not-at-fault drivers were used because the traffic conditions are more important than driver citation. However, due to the missing data at some detector stations on certain time periods, the traffic data are only available for 783 crashes among 978 crashes with at-fault drivers and 2225 crashes among 2817 crashes with both at-fault and not-at-fault drivers. Since the missing data occur due to random occurrence of detector malfunction, it was assumed that these crashes were randomly selected from the total crashes.

Given that the crash characteristics are defined as binary variables, the relationships between the crash characteristics and explanatory variables were estimated using a binary logit model. The binary logit model describes the probability of a response variable being classified into one of two categories in the function of explanatory variable(s). For instance, the probability of single-vehicle crashes relative to multi-vehicle crashes can be described as follows:

$$P(Z=1)=\frac{\exp(\alpha+\beta' x+\gamma' y)}{1+\exp(\alpha+\beta' x+\gamma' y)},$$
$$P(Z=0)=\frac{1}{1+\exp(\alpha+\beta' x+\gamma' y)} \quad (2)$$

where $P(Z=i)$ is the probability of crash type z being i(1 = single-vehicle crash or 0 = multi-vehicle crash), α is the constant, x is the vector of passenger characteristics, y is the vector of traffic flow parameters, and β and γ are the vectors of coefficients.

Similarly, in terms of crash injury severity, driver's non-severe injuries and fatal/severe injuries are defined as categories 0 and 1, respectively. Before estimating these binary logit models, the correlation between passenger characteristics and traffic flow parameters was examined. It was found that only the presence of passenger is strongly correlated with the variation in speed. More specifically, crashes tend to occur during higher variation in speed when drivers carry passenger(s) than driving alone. This implies that drivers have more difficulty with handling traffic turbulence if passengers are present due to distraction. However, the number of passengers and younger driver–younger passenger combination were not strongly correlated with any traffic flow parameter. Thus, it is reasonable to assume that passenger characteristics are independent from traffic flow parameters and they can be both used as explanatory variables in the binary logit models.

As a result of the binary logit model estimation, it was found that both crash type and driver's injury severity were strongly correlated with average speed and volume. The results indicate higher likelihood of single-vehicle crashes and driver's fatal/severe injuries as average speed increases and average volume decreases. However, it was found that only crash type was correlated with both passenger characteristics and traffic flow parameters if at-fault drivers carry passengers. Based on the functional relationship of the binary logit model, the ratio of the probability of single-vehicle crashes ($Z=1$) to the probability of multi-vehicle crashes ($Z=0$) (or the odds of single-vehicle to multi-vehicle crashes) is described as follows:

$$\ln\left(\frac{P(Z=1)}{P(Z=0)}\right) = -3.39 + 0.99 \times \underset{(p\text{-value}=0.046)}{\text{YOUNG}} - 0.83 \times \underset{(0.096)}{\text{NOPASS}} + 0.06 \times \underset{(0.045)}{\text{AS}} - 0.17 \times \underset{(0.006)}{\text{AV}}$$

where YOUNG is the younger driver–younger passenger combination (1 = yes, 0 = no), NOPASS is the number of passengers (1 = carry more than one passenger, 0 = carry one passenger), AS is 5-min average speed before crash (miles/h), and AV is 5-min average volume before crash (vehicles).

As shown in p-values, all the explanatory variables are statistically significant at a 90% confidence level. The signs of the parameters indicate that single-vehicle crashes are more likely to occur than multi-vehicle crashes when both drivers and passengers are younger, the number of passengers is one, average speed is higher and average volume is lower. In fact, higher likelihood of single-vehicle crashes during low traffic volume is the expected result; however, it is worthy to note high involvement of younger drivers with younger passengers in single-vehicle crashes than other combinations of driver–passenger age groups. This implies that when traffic volume is low (i.e. when single-vehicle crash is more likely to occur), particularly younger drivers with younger passengers are more likely to make errors leading to crashes due to their risky and distracted driving.

It should also be noted that the above model with both passenger characteristics and traffic flow parameters yields better model fit accounting for more number of explanatory variables (indicated by lower Akaike Information Criterion) than the model with only traffic flow parameters (average speed and volume). This clearly indicates the importance of considering the impact of passengers along with traffic conditions on single-vehicle crashes.

5. Discussion

The results of bivariate probit models and binary logit models showed that the presence of passengers is strongly correlated with crash characteristics. In general, it was found that drivers generally displayed safer driving behavior when passengers accompany drivers—higher likelihood of seatbelt use and lower likelihood of alcohol use, driver citation and driver's fatal/severe injury (potentially due to lower speed in the collision).

It is important to note that driver's crash potential can even be reduced further when more than one passenger accompany drivers—lower likelihood of driver citation and driver's fatal/severe injuries. This finding indicates that the positive effect of driver's increased responsibility with carrying more passengers and their consequent safer driving behavior surpasses the negative effect of driver distraction by more passengers.

It was also found that younger drivers accompanied by only younger passengers were more likely to cause crashes than the other combinations of driver–passenger age groups. Similarly, the result of the binary logit model indicated that younger drivers with younger passengers were more likely to be involved in single-vehicle crashes than other driver age groups during uncongested traffic conditions. This is because younger drivers were distracted by young peers or were more likely to take risk to show off in front of young peers.

6. Conclusions and recommendation

This study examines the correlation between the presence of passengers and freeway crash potential. The study and the approach used show that there is an interplay between the presence of passengers and the crash characteristics, and that this issue is not as simple as looking into one dimension or with simple statistics. The bivariate probit modeling approach would identify this issue easily, producing efficient and unbiased estimators. There is a possibility of having similar results to studies that used simpler approaches, but this does not mean that these studies were "more correct". It is a known fact in any statistical analysis that multivariate modeling and full information maximum likelihood are more superior. The point of whether the presence of passengers should be an independent or dependent variable, can easily be resolved using our approach of instrumenting it before it could be used as an explanatory variable. This approach guarantees that we do not run into the risk of potential endogeneity.

This study developed a set of bivariate probit models that relate three passenger characteristic variables such as the presence of passenger, the number of passengers and the younger driver/younger passenger combination to three crash characteristic variables such as driver citation, crash type and driver's injury severity. The models show that the presence of passengers is correlated with lower prevalence of risky behavior (i.e. non-seatbelt use and alcohol use) that ultimately reduces crash potential. The bivariate probit models also identified statistical significance of the error-term correlation between inter-related responses of passenger and crash characteristics. However, the correlation of two response variables has been overlooked in most other studies which analyzed the association of passengers with crash risk using logistic regression models. While this study provided identical results as the other studies, it also provided the new results that were not observed in the past. Considering the correlation of passenger and crash characteristics, the bivariate probit model provides additional insights such as decreased driver citation and crash injury severity with more passengers.

However, it is possible that there exist some other factors that could not be measured in this study, but could affect individual's driving behavior. For instance, driver distraction could have been a cause of crashes. Drivers are likely to be involved in distraction-related crashes while they are adjusting radios and distracted by other passengers and outside objects (Stutts et al., 2001). The risk of crash injury can also be increased by the distraction caused by the use of a hand-held phone particularly for younger drivers (Lam, 2002).

Based on the results, it is strongly recommended that younger drivers are accompanied by one or more older passengers. It is also recommended that younger drivers should take more caution and drive slower particularly when they drive with only younger passengers in high-speed and low-volume traffic conditions when they are more likely to be distracted and potentially make errors. In this regard, various potential crash countermeasures can be implemented. For instance, warning messages are displayed in the roadside signs to drivers temporarily during free-flow conditions only. It is expected that this temporary or variable messages will attract driver's attention more effectively than fixed warning messages. The graduate licensing program can also help reduce crashes by reducing the number of younger drivers who drive with younger passengers under high-speed conditions. Although all drivers should take caution and drive more safely, younger drivers should take more cautions as their driving tends to be more aggressive than older driver's driving. Clearly, the education programs for younger drivers are needed to advice them of driving slowly during uncongested conditions. However, more detailed work is needed to investigate how we can effectively enforce any such measures related to passengers in order to reduce younger drivers' crash potential.

The limitations of this study are that the study was conducted only for freeways and the data were collected from one local area. Thus, in order to generalize the findings, it is worthwhile to test the transferability of the results to other roadway types and freeways in other areas. It is also worthwhile to observe actual driving behavior of individual drivers with and without passengers in instrumented vehicles under various traffic/environmental/road geometric conditions. This observation will help identify the specific conditions when the presence of passengers is more likely to affect driver behavior.

References

Chen, L.H., Baker, S.P., Braver, E.R., Li, G., 2000. Carrying passengers as a risk factor for crashes fatal to 16- and 17-year-old drivers. JAMA 283 (12), 1578–1582.

Cooper, D., Atkins, F., Gillen, D., 2005. Measuring the impact of passenger restrictions on new teenage drivers. Accident Analysis and Prevention 37, 19–23.

Doherty, S.T., Andrey, J.C., MacGregor, C., 1998. The situational risks of young drivers: the influence of passenger, time of day and day of week on accident rates. Accident Analysis and Prevention 30, 45–52.

Geyer, J.A., Ragland, D.R., 2005. Vehicle occupancy and crash risk. Transportation Research Record 1908, 187–194.

Glass, R.J., Graham, J.D., 1999. Kids at risk: where American children sit in passenger vehicles. Journal of Safety Research 30, 17–24.

Greene, W.H., 1998. LIMDEP Version 7.0 User's Manual Revised Edition. Econometric Software Inc., Plainview, NY.

Hing, J.Y., Stamatiadis, N., Aultman-Hall, L., 2003. Evaluating the impact of passengers on the safety of older drivers. Journal of Safety Research 34, 343–351.

Isaac, N.E., Kennedy, B., Graham, J.D., 1995. Who's in the car? Passengers as potential interveners in alcohol-involved fatal crashes. Accident Analysis and Prevention 27 (2), 159–165.

Keall, M.D., Frith, W.J., Patterson, T.L., 2004. The influence of alcohol, age and number of passengers on the night-time risk of driver fatal injury in New Zealand. Accident Analysis and Prevention 36, 49–61.

Lam, L.T., 2002. Distractions and the risk of car crash injury: the effect of drivers' age. Journal of Safety Research 33, 411–419.

Lam, L.T., Norton, R., Woodward, M., Connor, J., Ameratunga, S., 2003. Passenger carriage and car crash injury: a comparison between younger and older drivers. Accident Analysis and Prevention 35, 861–867.

Lin, M., Fearne, K.T., 2003. The provisional license: Nighttime and passenger restrictions—A literature review. Journal of Safety Research 34, 51–61.

Preusser, D.F., Ferguson, S.A., Williams, A.F., 1998. The effect of teenage passengers on the fatal crash risk of teenage drivers. Accident Analysis and Prevention 30, 217–222.

Regan, M.A., Mitsopoulos, E., 2003. Understanding Passenger Influences on Driver Behaviour: Implications for Road Safety and Recommendations for Countermeasure Development. Monash University Accident Research Centre, Report #180.

Rueda-Domingoa, T., Lardelli-Claret, P., Luna-del-Castillo, J., Jiménez-Moleón, J.J., García-Martín, M., Bueno-Cavanillas, A., 2004. The influence of passengers on the risk of the driver causing a car collision in Spain: analysis of collisions from 1990 to 1999. Accident Analysis and Prevention 36, 481–489.

Simons-Morton, B.G., Lerner, N., Singer, J., 2005. The observed effects of teenage passengers on the risky driving behavior of teenage drivers. Accident Analysis and Prevention 37, 973–982.

Stutts, J.C., Reinfurt, D.W., Staplin, L., Rodgeman, E.A., 2001. The role of driver distraction in traffic crashes. AAA Foundation for Traffic Safety. www.aaafoundation.org.

Vollrath, M., Meilinger, T., Krüger, H., 2002. How the presence of passengers influences the risk of a collision with another vehicle. Accident Analysis and Prevention 34, 649–654.

Williams, A.F., 2003. Teenage drivers: patterns of risk. Journal of Safety Research 34, 5–15.

Williams, A.F., Wells, J.K., 1995. Deaths of teenagers as motor-vehicle passengers. Journal of Safety Research 26 (3), 161–167.

Williams, A.F., Shabanova, V.I., 2002. Situational factors in seat belt use by teenage drivers and passengers. Traffic Injury Prevention 3, 201–204.