Accident Analysis and Prevention 36 (2004) 447–456

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

Modeling rear-end collisions including the role of driver's visibility and light truck vehicles using a nested logit structure

Mohamed Abdel-Aty*, Hassan Abdelwahab

Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816-2450, USA

Received 23 April 2002; received in revised form 12 February 2003; accepted 13 February 2003

Abstract

This paper presents an analysis of the effect of the geometric incompatibility of light truck vehicles (LTV)—light-duty trucks, vans, and sport utility vehicles—on drivers' visibility of other passenger cars involved in rear-end collisions. The geometric incompatibility arises from the fact that most LTVs ride higher and are wider than regular passenger cars. The objective of this paper is to explore the effect of the lead vehicle's size on the rear-end crash configuration. Four rear-end crash configurations are defined based on the type of the two involved vehicles (lead and following vehicles). Nested logit models were calibrated to estimate the probabilities of the four rear-end crash configurations 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 CarTrk (a regular passenger car striking an LTV). Possibly, indicating a sight distance problem. A driver of a smaller car following an LTV, have a problem seeing the roadway beyond the LTV, and therefore would not be able to adjust his/her speed accordingly, increasing the probability of a rear-end collision. Also, the probability of a CarTrk rear-end crash increases in the case that the lead vehicle stops suddenly.

Keywords: Rear-end; Nested logit; Light truck vehicles; Vehicle compatibility; Crash configuration

1. Introduction

Rear-end is a common type of traffic crashes in the United States. In 2000, about a third of the US traffic crashes were rear-end collisions. James et al. (1997) reported that rear-end collisions are not frequently reported because most are of low overall crash severity. Therefore, the actual number of rear-end collisions is expected to be more than the reported crashes. Wang et al. (1999) reported that the most numerous crash category in the US is rear-end collisions. According to the Spinal Injury Foundation (2002), neck injuries can be the most severe outcome of rear-end collisions because of the fact that the head has nothing to stop it except a nonexistent or poorly adjusted head rest (which can actually cause more harm than good), which means that rear-end crashes are a real problem.

In general, a driver can minimize the likelihood of being involved in a rear-end crash by maintaining a space cushion that is appropriate for the driving conditions. A proper space cushion must provide a driver time to see and recognize a hazard and make a decision regarding what should be done. Then, there must be adequate space to bring the vehicle to a stop.

Kostyniuk and Eby (1998) explored rear-end crashes from the driver perspective to determine if knowledge of the circumstances of the crash as perceived by the driver could offer insight into the characteristics of potential countermeasures for rear-end crashes. The subjects were directly asked about the factors that contributed to the crash and to indicate the relative contribution of each factor. Action of the other driver in the lead vehicle was the dominant contributing factor according to the subjects. The drivers described these actions as the other car stopped unexpectedly and the other car did not move when it should have. The next most frequent set of responses to this question included personal inattention or distraction.

Graham (2000) reported that a major concern often voiced by the US motorists is that light truck vehicles (LTV) (light-duty trucks, vans, and sports utility vehicles) make it impossible for drivers in smaller vehicles to see the traffic ahead of them or to see the traffic flow when a driver is pulling out of a side street onto a major roadway. Therefore, driver's visibility significantly affects the chance

* Corresponding author. Tel.: +1-407-823-5657; fax: +1-407-823-3315.
E-mail address: mabdel@mail.ucf.edu (M. Abdel-Aty).

0001-4575/$ – see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0001-4575(03)00040-X

of being involved in a rear-end collision when the leading vehicle stops suddenly.

Sayer et al. (2000) examined the effect that the lead 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). The data were obtained from a random sample of licensed drivers who drove an instrumented passenger car, unaccompanied, as their personal vehicle for 2–5 weeks. Results showed that passenger car drivers followed LTVs at shorter distances than they followed passenger cars, but at the same velocities. Also, the results of this study suggested that knowing the state of traffic beyond the lead vehicle, even by only one additional vehicle, affects gap length. 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.

The aforementioned discussion about rear-end crashes and the substantial increase in the percentage of LTVs in the US traffic over the last few years, show the urgent need to study the association between LTVs and rear-end crash configurations. Rear-end crashes can be classified into four different classes based on the strike (lead vehicle) and stricken vehicles (follower vehicle). The four categories of the rear-end crashes are:

1. A regular passenger car striking another regular passenger car (CarCar).
2. A regular passenger car striking a light truck vehicle (CarTrk).
3. A light truck vehicle striking a regular passenger car (TrkCar).
4. A light truck vehicle striking another light truck vehicle (TrkTrk).

2. Models of unordered alternatives

In this paper we are modeling the probabilities of four unordered alternatives in two-vehicle rear-end crashes based on the type of these vehicles (LTV or regular car). The unordered multiple-choice model on N subjects facing J unordered (nominal or qualitative) alternatives assumes the relationship:

$$P_n(j) = F(\alpha_j + \beta_j X_n, \theta), \quad j = 1, \ldots, J-1,$$

$$P_n(J) = 1 - \sum_{j=1}^{J-1} P_n(j)$$

where $P_n(j)$ is the probability that subject n ($n = 1, \ldots, N$) belongs to category j, α_j an alterative specific constant, X_n a vector of measurable characteristics specific to subjects and alternatives, β_j a vector of estimable coefficients, and θ is a parameter that controls the shape of probability distribution F. The model is made operational by a particular choice of the distributional form of F. McFadden (1974) assumed a logistic distribution for F to derive the multinomial logit model (MNL) as follows:

$$P_n(j) = \frac{\exp(\beta_j X_n)}{\sum_{j=1}^{J-1}\exp(\beta_j X_n)}, \quad j = 1, \ldots, J-1,$$

$$P_n(J) = 1 - \sum_{j=1}^{J-1} P_n(j)$$

One of the most widely critique of the multinomial logit model is the independence from irrelevant alternatives property (IIA). The IIA property holds that for a subject n, the ratio of choice probabilities of any two categories (i and j) is entirely unaffected by any other alternative. That is $P_n(i)/P_n(j) = \exp(\beta_i - \beta_j)X_n$. McFadden (1977) investigated a wide range of computationally feasible tests to detect violations of the IIA property. This involves comparisons of logit models estimated with subsets of alternatives from the universal choice set J. If the IIA property holds for the full choice set, then the logit model also applies to a choice from any subset of alternatives. Thus, if the logit model is correctly specified, we can obtain consistent coefficient estimates of the same sub-vector of parameters from a logit model estimated with the full choice set and from a logit model estimated with a restricted choice set.

The IIA property of MNL comes from the assumption that F of all alternatives j, $j = 1, \ldots, J-1$, are independent identical distributed (IID). The IID assumption can be relaxed in one of three ways:

1. *Allowing non-identical distribution* (*different shape parameters—variances*), *but maintaining the independence assumption*. Bhat (1995) developed this model by using the concept of heterosedasticity in alternative error terms (shape parameter of a distribution). If the scale parameters of the random components of all alternatives are equal, then the probability expressions collapse to that of the MNL.
2. *Allowing F to be correlated while maintaining the assumption that they are identically distributed*. This specification defines a nested logit model. The nested logit model is currently the preferred extension to the simple MNL (Ben-Akiva and Lerman, 1985). The appeal of the nested logit model is its ability to accommodate differential degrees of similarity between subsets of alternatives in a choice set.
3. *Allowing non-identical and non-independent distributions of F*. Models of that type commonly use a multivariate normal distribution. The resulting model, referred to as the multinomial probit model, is difficult to handle because of introducing several additional parameters in the covariance matrix.

In this section, we demonstrate a general outline of nested logit models. The nested logit model requires a priori specification of homogenous sets of alternatives (i.e. partitioning the set J into several exclusive groups). To fix the idea of

Table 1
Crash distribution by manner of collision

Year	Manner of collision[a]								Total
	0	1	2	3	4	5	6	9	
1988	2208451 (32.11)	1622543 (23.59)	102438 (1.49)	4582 (0.07)	2439827 (35.48)	294011 (4.28)	87048 (1.27)	117880 (1.71)	6876780 (100)
1989	2189327 (32.95)	1609442 (24.22)	112899 (1.70)	3369 (0.05)	2352320 (35.40)	248426 (3.74)	68707 (1.03)	60059 (0.90)	6644549 (100)
1990	2082589 (32.23)	1516417 (23.47)	83983 (1.30)	2058 (0.03)	2364042 (36.58)	290984 (4.50)	60884 (0.94)	61169 (0.95)	6462126 (100)
1991	1967573 (32.20)	1476424 (24.16)	102170 (1.67)	527 (0.01)	2301756 (37.67)	173349 (2.84)	49373 (0.81)	38757 (0.63)	6109931 (100)
1992	1879664 (31.36)	1496875 (24.98)	96567 (1.61)	832 (0.01)	2166831 (36.16)	236413 (3.94)	59520 (0.99)	56236 (0.94)	5992938 (100)
1993	1926146 (31.60)	1547026 (25.38)	94536 (1.55)	1631 (0.03)	2240387 (36.76)	213611 (3.50)	50469 (0.83)	20965 (0.34)	6094772 (100)
1994	1953523 (30.10)	1753996 (27.03)	115356 (1.78)	1790 (0.03)	2333927 (35.97)	258237 (3.98)	50985 (0.79)	21308 (0.33)	6489122 (100)
1995	1946729 (29.10)	1845379 (27.58)	103940 (1.55)	2610 (0.04)	2465011 (36.85)	251099 (3.75)	50956 (0.76)	24337 (0.36)	6690061 (100)
1996	1992933 (29.48)	1882396 (27.84)	105677 (1.56)	2276 (0.03)	2464304 (36.45)	254560 (3.77)	46791 (0.69)	12115 (0.18)	6761051 (100)
1997	1946624 (29.44)	1871564 (28.31)	100127 (1.51)	1563 (0.02)	2386369 (36.09)	243174 (3.68)	47783 (0.72)	14702 (0.22)	6611906 (100)
1998	1803747 (28.52)	1871271 (29.58)	82986 (1.31)	1603 (0.03)	2264670 (35.80)	244994 (3.87)	45956 (0.73)	10015 (0.16)	6325242 (100)
1999	1840807 (29.35)	1848407 (29.47)	106314 (1.70)	2830 (0.05)	1941647 (30.96)	427307 (6.81)	85715 (1.37)	18498 (0.29)	6271524 (100)
2000	1964375 (30.74)	1892371 (29.62)	136941 (2.14)	2773 (0.04)	1931304 (30.23)	368568 (5.77)	80713 (1.26)	12265 (0.19)	6389310 (100)

[a] 0: not collision with motor vehicle in transport; 1: rear-end; 2: head-on; 3: rear-to-rear; 4: angle; 5: sideswipe, same direction; 6: sideswipe, opposite direction; 9: other. The values in parentheses are in percent.

a nested logit model, suppose that J alternatives can be divided into G mutually exclusive subgroups. Logically, one may think of the choice process as that of choosing among G choice sets and then making the specific choice j within the chosen set g, $g = 1, \ldots, G$. Then, the mathematical form for a two-level nested logit model is as follows (Greene, 2000):

$$P_n(j, g) = P_n(g)P_n(j|g), \quad j = 1, \ldots, J-1,$$

$$P_n(j|g) = \frac{\exp(\beta_j X_n|g)}{\sum_{j\in g}\exp(\beta_j X_n|g)},$$

$$P_n(g) = \frac{\exp(\gamma_g M_g + \tau_g I_g)}{\sum_{g=1}^{G}\exp(\gamma_g M_g + \tau_g I_g)},$$

$$P_n(J) = 1 - \sum_{j=1}^{J-1} P_n(j)$$

where $P_n(j, g)$ is the probability of a subject n to belong to category j in group g, $P_n(j|g)$ the conditional probability of a subject n to belong to category j given that subject belongs to group g, $P_n(g)$ the probability of subject n to belong to group g, X_n a vector of measurable attributes, M_g a vector of groups' characteristics, $I_g = \ln\sum_{j\in g}\exp(\beta_j X_n|g)$ and is called the inclusive value of group g, β_j and γ_g are vectors of coefficients to be estimated using the maximum likelihood approach, and τ_g is the coefficient of the inclusive value of group g. If we restrict all inclusive value parameters to one, then the nested logit model will be reduced to the MNL. The nested logit model can be easily extended to higher levels (more than two levels). The complexity of the model increases geometrically with the number of levels. But the model has been found to be extremely flexible and is widely used in transportation modeling.

There are two ways to estimate the parameters of the nested logit model. A limited information maximum likelihood (LIML), can be done as follows: estimate β by treating the choice within branches as simple multinomial logit model, compute the inclusive values for all branches in the model, then estimate the parameters by treating the choice among branches as a simple multinomial logit model. The other approach of estimating a nested logit model is the full information maximum likelihood (FIML). In this approach, the entire model is estimated in a single phase. In general, the FIML estimation will be more efficient than the multi-step estimation. In this paper, all nested logit estimations were based on FIML. The goodness-of-fit of a nested logit model can be assessed by the likelihood ratio test, in which a restricted version of the model is tested against a full version of the model.

To gain a better understanding of the variables included in a calibrated nested logit model, marginal effects can be computed. The marginal effect formula (Greene, 2000) for two-level nesting structure of an attribute r of choice j in group g on the probability of alternative j^* in group g^* is,

$$\frac{\partial \ln P(j^*, g^*)}{\partial r_{(j,g)}} = \beta_{r(j,g)}\{1_{(g=g^*)}[1_{(j=j^*)} - P(j^*|g)] + \tau_g[1_{g=g^*} - P(g^*)]P(j|g)\}$$

where $\partial \ln P(j^*, g^*)/\partial r_{(j,g)}$ represents the marginal effect of attribute r of alternative j in a group g on probability of alternative j^* in a group g^*, $1_{(g=g^*)} = 1$, if $g = g^*$, 0, otherwise, $1_{(j=j^*)} = 1$, if $j = j^*$, 0 otherwise, $P(j)$ is probability of category j, $P(g)$ probability of group g, and $\beta_{r(j,g)}$ is the estimated coefficient corresponding to attribute r.

Table 2
Rear-end crash configuration distribution blocked by year

Year	CarCar	CarTrk	TrkCar	TrkTrk	Total
1988	716502 (61.98)	209778 (18.15)	172492 (14.92)	57173 (4.95)	1155945 (100)
1989	688946 (60.18)	210017 (18.35)	186057 (16.25)	59706 (5.22)	1144727 (100)
1990	585958 (57.31)	204999 (20.05)	168748 (16.50)	62732 (6.14)	1022437 (100)
1991	592936 (54.14)	240953 (22.00)	182621 (16.68)	78581 (7.18)	1095091 (100)
1992	610071 (54.00)	242791 (21.49)	186931 (16.55)	89965 (7.96)	1129757 (100)
1993	610294 (51.24)	265669 (22.30)	213529 (17.93)	101585 (8.53)	1191076 (100)
1994	684904 (50.06)	308368 (22.54)	256744 (18.76)	118247 (8.64)	1368263 (100)
1995	691666 (48.10)	328482 (22.84)	279739 (19.45)	138101 (9.60)	1437987 (100)
1996	676169 (46.22)	350041 (23.93)	280182 (19.15)	156546 (10.70)	1462939 (100)
1997	650522 (45.04)	353907 (24.50)	280720 (19.44)	159121 (11.02)	1444270 (100)
1998	623139 (43.02)	367374 (25.36)	286539 (19.78)	171382 (11.83)	1448433 (100)
1999	545404 (38.22)	376794 (26.40)	300829 (21.08)	204110 (14.30)	1427138 (100)
2000	565173 (38.42)	375249 (25.51)	321020 (21.82)	209510 (14.24)	1470952 (100)

CarCar: a regular passenger car striking another regular passenger car; CarTrk: a regular passenger car striking a light truck vehicle; TrkCar: a light truck vehicle striking a regular passenger car; TrkTrk: a light truck vehicle striking another light truck vehicle.

3. GES traffic crash database

The general estimates system (GES) databases were used in the analysis. The GES data come from a nationally representative sample of police-reported motor vehicle crashes of all types, from minor to fatal. The GES database is a relational database consisting of three files: *crash*, *vehicle/driver*, and *person*. Each file deals with a specific aspect of traffic crashes. Files may be linked as needed to combine the information contained in each. The *crash* file contains general information about the crash characteristics and circumstances. The *vehicle/driver* file contains information about the vehicles and drivers' actions in traffic crashes. The *person* file contains information about individuals involved in traffic crashes.

The GES system began operation in 1988, and was created to identify traffic safety problem areas, provide a basis for regulatory and consumer initiatives, and form the basis for cost and benefit analyses of traffic safety initiatives. In order for a crash to be eligible for the GES sample, a police crash report (PCR) must be completed, it must involve at least one motor vehicle traveling on a traffic way, and the result must be property damage, injury, or death. These crash reports are chosen from 60 areas that reflect the geography, roadway mileage, population, and traffic density of the United States. For more detailed information, refer to the GES Analytical User's Manual (National Highway Traffic Safety Administration, 2000).

Table 1 shows the crash distribution by manner of collision. Single vehicle (not collision with motor vehicle in transport), angle, and rear-end collisions are the three major crash types. For the year 2000, the percentages of single vehicle, angle, and rear-end collisions were 30.94, 30.23, and 29.62%, respectively. The yearly rear-end crash percentage has an increasing trend. Table 2 summarizes distributions of rear-end crash configurations blocked by year. In 2000, the rear-end crash distribution by configuration was 38.4, 25.6, 21.8, 14.2% for CarCar, CarTrk, TrkCar, and TrkTrk, respectively. Although CarCar has the highest percentage because the majority of the vehicles on the road are still regular cars, it is worth mentioning that the percent of CarTrk is relatively high and accounting for the second highest configuration. Also this type has an increasing trend, possibly due to the increasing percentage of LTVs in traffic. Tables 1 and 2 illustrate the increasing percentage of rear-end crashes and of the CarTrk configuration type. This indicates a possible problem with this type of rear-end collisions.

Table 3 shows the variables that were used in analyzing rear-end crash configurations. Most of the variables are self explanatory and usually coded in any traffic crash database. Among the interesting variables are driver's vision obstruction and driver distraction. Driver's vision obstruction identifies visual circumstances that may have contributed to the cause of the crash. Driver distraction attempts to capture distractions which may have influenced driver performance and contributed to the cause of the crash. The distractions can be both inside the vehicle (internal) and outside the vehicle (external). This variable was added to the *vehicle/driver* file in 1990. In 1999, extensive modifications were made to the codes of the driver distraction field. In addition, too many missing cases of this variable as well as driver vision obstruction are in the 1999 GES vehicle/driver file. Therefore, the 2000 GES crash database is used in this analysis, because of better and complete reported cases. The analysis was limited to two-vehicles rear-end collisions. In the 2000 GES, the total number of reported crashes was 57,392 crashes, and there were 11,933 rear-end crashes with two vehicles involved.

Examining the whole two-vehicle rear-end crashes, 1579 cases (13.2% of the total number of two-vehicle rear-end crashes) have complete information about all variables listed in Table 3, especially the driver distraction and vision obstruction fields. The crash configuration distribution of these complete cases (37.7, 25.9, 22.0, 14.4%) is similar to the full 11,933 cases. The chi-square statistical test was used to

Table 3
Variables' coding for the rear-end collision analysis

Variable	Code
Crash related variables	
Rear-end crash configuration	1. A passenger car strike another passenger car (CarCar) 2. A passenger car strike an LTV (CarTrk) 3. An LTV strike a passenger car (TrkCar) 4. An LTV strike another LTV (TrkTrk)
Posted speed limit	Mile per hour
Roadway alignment	1 straight, 0 curve
Roadway profile	1 level, 0 grade
Roadway surface condition	1 dry, 0 not dry
Light condition	1 daylight, 0 not daylight
Atmospheric condition	1 no adverse condition, 0 adverse conditions
Day of the week	1 weekend, 0 weekday
Traffic control device	(1, 0, 0) no controls, (0, 1, 0) traffic signal, (0, 0, 1) traffic stop sign, (0, 0, 0) other
Relation to junction	1 intersection related, 0 otherwise
Striker vehicle	
Speed at the time of the crash	Mile per hour
Driver's violation	1 driver made a violation, 0 no violation
Driver age	Actual age
Driver gender	1 male, 0 female
Alcohol involvement	1 drunk driver, 0 sober driver
Driver distraction	1 driver under distraction, 0 no distraction
Driver's vision obstruction	1 vision was obstructed, 0 no vision obstruction
Event initiated by the other vehicle	1 stopped, 0 otherwise
Stricken vehicle	
Speed at the time of the crash	Mile per hour
Driver's violation	1 driver made a violation, 0 no violation
Driver age	Actual age
Driver gender	1 male, 0 female
Alcohol involvement	1 drunk driver, 0 sober driver

test the difference in rear-end crash distribution for the complete case and the full set. Results indicated no significant difference between the two sets.

4. Modeling results

The four different categories of the rear-end crashes can be modeled using models of un-ordered alternatives (i.e. nested logit model). All drivers have the same possibility of exposure for involvement in any rear-end crash configuration. The data was randomly divided into two sets. One set for calibration and the other for testing. About 90% (1421 cases) were used for calibration and the remaining 158 cases (10% of the total cases) were used for testing.

To select the best nesting structure, all possible structures that examine possible correlation among the unobserved effects of the four rear-end crash categories were considered. Table 4 lists all possible nesting structures for grouping the four rear-end crash configurations. We limited our analysis to the five reasonable and intuitive structures shown in Fig. 1 including the simple multinomial logit model.

The first structure deals with the four configurations as independent categories (simple multinomial logit structure). The second structure divides the four configurations into two groups: (i) same-size group (CarCar and TrkTrk) and (ii) different-size group (TrkCar and CarTrk). The third model groups the four categories by the type of the striker vehicle (i.e. CarCar and CarTrk in one group, TrkCar and TrkTrk in

Table 4
All possible nesting structures of rear-end crashes' models

Model	Nesting structure
1	(CarCar), (CarTrk), (TrkCar), (TrkTrk)
2	(CarCar, TrkTrk), (CarTrk, TrkCar)
3	(CarCar, CarTrk), (TrkCar, TrkTrk)
4	(CarCar, TrkCar), (CarTrk, TrkTrk)
5	(CarCar), (CarTrk, TrkCar, TrkTrk)
6	(CarCar, CarTrk), (TrkCar), (TrkTrk)
7	(CarCar), (CarTrk), (TrkCar, TrkTrk)
8	(CarCar), (TrkCar), (CarTrk, TrkTrk)
9	(CarCar, TrkCar), (CarTrk), (TrkTrk)
10	(CarCar), (TrkTrk), (CarTrk, TrkCar)
11	(CarCar, TrkTrk), (CarTrk), (TrkCar)
12	(CarTrk), (CarCar, TrkCar, TrkTrk)
13	(TrkCar), (CarCar, CarTrk, TrkTrk)
14	(TrkTrk), (CarCar, CarTrk, TrkCar)
15	(CarCar), ([CarTrk, TrkCar], TrkTrk)
16	(CarCar), (CarTrk, [TrkCar, TrkTrk])
17	(CarCar), ([CarTrk, TrkTrk], TrkCar)
18	(CarTrk), ([CarCar, TrkCar], TrkTrk)
19	(CarTrk), (CarCar, [TrkCar, TrkTrk])
20	(CarTrk), ([CarCar, TrkTrk], TrkCar)
21	(TrkCar), ([CarCar, CarTrk], TrkTrk)
22	(TrkCar), (CarCar, [CarTrk, TrkTrk])
23	(TrkCar), ([CarCar, TrkTrk], CarTrk)
24	(TrkTrk), ([CarCar, CarTrk], TrkCar)
25	(TrkTrk), (CarCar, [CarTrk, TrkCar])
26	(TrkTrk), ([CarCar, TrkCar], CarTrk)

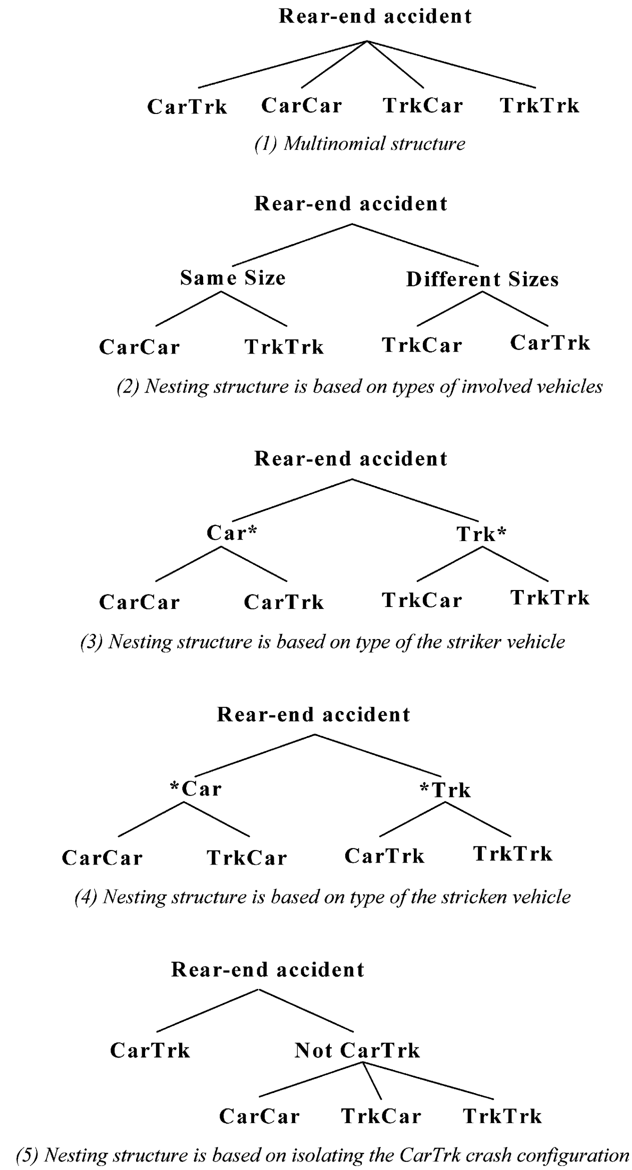

Fig. 1. Five nesting structures of rear-end crashes' models.

the other group). The fourth structure categorizes the four rear-end configuration based on the type of the stricken vehicle (lead vehicle). In this model, CarCar and TrkCar are grouped in one nest and the other configurations (CarTrk and TrkTrk) are in the other nest.

The fifth nesting structure shown in Fig. 1 isolates the CarTrk rear-end crash collision and deal with other rear-end configurations as one group. The reason of this nesting structure is the hypothesis that we are testing that there are sight distance and discomfort problems when a driver in a passenger car is driving behind an LTV. Finally, any of the nested logit models shown in Fig. 1 can be expanded (complicated) or reduced (simplified) as needed. The best-nested logit model can be then selected based on the significance of the nesting structure coefficients, likelihood ratio index, and percent correct classification of the test set.

Table 5
Multinomial logit model (structure 1) for rear-end crash configuration

Variable	Estimated coefficient	*t*-statistics
CarCar specific constant	1.517	7.33
CarTrk specific constant	0.583	4.18
TrkCar specific constant	0.972	4.09
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarCar	−0.539	−3.18
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarTrk	−0.837	−4.85
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to TrkCar	0.502	2.56
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarCar and CarTrk	0.963	7.62
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarCar	1.287	3.09
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarTrk	1.048	2.33
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to CarCar	−0.849	−6.76
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to TrkCar	−1.115	−7.56
Light condition (1 daylight, 0 otherwise), specific to CarCar	−0.490	−3.01
Light condition (1 daylight, 0 otherwise), specific to TrkCar	−0.580	−3.12
Crash occurred at traffic signal and night time, specific to CarTrk	0.818	2.93
Driver of the striker vehicle was distracted when the stricken vehicle stopped, specific to TrkCar	−0.399	−2.07
Goodness-of-fit measures		
Likelihood ratio index (ρ^2)	0.06	
Classification accuracy of the testing data	60.8%	

Tables 5–9 summarize the calibration results of the five nesting structures shown in Fig. 1. Based on the goodness-of-fit measures (likelihood ratio index and classification accuracy of the test set), the fifth nesting structure shown in Fig. 1 was found to be the best-nested logit model. The model structure is a two-level nesting structure that groups the CarTrk in one group and the other categories (CarCar, TrkCar, TrkTrk) in another group. The inclusive value coefficient is significantly different from zero and one, which provides a statistical validation of using this nesting structure.

The significant variables of the calibrated nested logit model include; driver age, gender, and traffic control device. In addition, two two-way interaction terms were found to significantly affect the probability of a rear-end crash of the CarTrk configuration. The first interaction term describes a rear-end crash situation when the driver of the striker vehicle was distracted when the stricken vehicle stopped. The positive coefficient of this variable indicates that it increases the probability of CarTrk rear-end crash configuration given that a rear-end collision has occurred. The other interaction term describes a situation when the driver's vision of the striker vehicle (following vehicle) was obscured and the stricken vehicle (lead vehicle) stopped. It also increases the probability of the CarTrk category. These interaction terms suggest that LTV obscure drivers' visibility of other passenger cars. Also, LTV may prevent drivers in cars behind them from being aware of traffic situation ahead.

Table 6
Nested logit model (structure 2) for rear-end crash configuration

Variable	Estimated coefficient	*t*-statistics
CarCar specific constant	1.550	2.90
Same-group specific constant	−1.209	−3.75
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarCar	−0.587	−2.20
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarTrk	−1.022	−2.78
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to TrkCar	0.577	2.03
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarCar and CarTrk	1.017	2.88
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarCar	1.375	2.18
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarTrk	1.105	1.71
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to CarCar	−0.897	−3.04
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to TrkCar	−1.158	−2.86
Light condition (1 daylight, 0 otherwise), specific to CarCar and TrkCar	−0.429	−2.02
Crash occurred at traffic signal and night time, specific to CarTrk	0.732	2.13
The stricken vehicle stopped suddenly, specific to same-group	1.116	9.70
Nesting coefficient of the same-group	0.923	2.88
Nesting coefficient of the different group	0.914	2.76
Goodness-of-fit measures		
Likelihood ratio index (ρ^2)	0.10	
Classification accuracy of the testing data	63.1%	

Table 7
Nested logit model (structure 3) for rear-end crash configuration

Variable	Estimated coefficient	t-statistics
CarTrk specific constant	−1.589	−4.58
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to TrkCar	1.966	7.86
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarCar and CarTrk	0.959	7.41
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarCar	1.360	2.52
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to CarCar	−1.456	−4.01
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to TrkCar	−1.869	−4.76
Light condition (1 daylight, 0 otherwise), specific to CarCar	−0.389	−1.55
Light condition (1 daylight, 0 otherwise), specific to TrkCar	−0.335	−1.83
Crash occurred at traffic signal and night time, specific to CarTrk	1.029	2.50
Driver of the striker vehicle was distracted when the stricken vehicle stopped, specific to TrkCar	1.500	3.81
Natural logarithm of speed limit at crash location, specific to Car* group	0.364	3.65
The stricken vehicle stopped suddenly, specific to Car* group	0.973	7.80
Driver vision was bad, specific to Car* group	0.527	2.63
Nesting coefficient of the Car* group	0.533	3.86
Nesting coefficient of the Trk* group	0.442	4.81
Goodness-of-fit measures		
Likelihood ratio index (ρ^2)	0.09	
Classification accuracy of the testing data	63.5%	

Table 10 summarizes direct marginal effects of the variables included in the calibrated nested logit model. This table shows estimates of the effect of the explanatory variables included in the model on the probability of a certain rear-end crash configuration while taking into account the impact of the other explanatory factors. For example, the driver's vision of the striker vehicle when obstructed and the stricken vehicle stopped suddenly has the biggest effect of having a rear-end collision of configuration CarTrk. This emphasizes that there are sight distance and discomfort problems when a driver in a passenger car is driving behind an LTV.

In terms of driver age, young drivers (age between 15 and 24) and old drivers (age 75 and older) are more likely to be involved in rear-end crashes of type CarTrk. This could be attributed to that young and older drivers are more likely to drive regular passenger cars. Furthermore, for younger drivers, lack of driving experience and aggressive attitude may make them follow too closely other vehicles. Male drivers are more likely to drive LTV than females (Kockelman and Zhao, 2000). Hence, male drivers have a higher percentage of TrkCar rear-end crashes. Presence of a traffic signal positively affects the probability of having a CarTrk rear-end crash.

Table 8
Nested logit model (structure 4) for rear-end crash configuration

Variable	Estimated coefficient	t-statistics
CarCar specific constant	0.609	3.40
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarTrk	−0.295	−2.95
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to TrkCar	0.940	5.34
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarCar	0.852	3.96
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarTrk	0.899	3.71
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarCar	0.794	2.29
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to CarCar	−0.878	−6.91
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to TrkCar	−1.121	−7.63
Light condition (1 daylight, 0 otherwise), specific to CarCar	−0.429	−2.68
Light condition (1 daylight, 0 otherwise), specific to TrkCar	−0.473	−2.92
Driver of the striker vehicle was distracted when the stricken vehicle stopped, specific to CarTrk	0.302	2.11
Crash occurred at traffic signal and night time, specific to CarTrk	0.781	2.73
Driver's vision of the striker vehicle was obstructed when the stricken vehicle stopped, specific to CarTrk	0.574	2.60
The stricken vehicle stopped suddenly, specific to *Car group	−0.765	−6.26
Crash happened in a weekday (1 if weekend, 0 if weekday)	−0.269	−1.93
Nesting coefficient of the *Car group	0.874	3.87
Nesting coefficient of the *Trk group	0.978	2.83
Goodness-of-fit measures		
Likelihood ratio index (ρ^2)	0.09	
Classification accuracy of the testing data	62.5%	

Table 9
Nested logit model (structure 5) for rear-end crash configuration

Variable	Estimated coefficient	t-statistics
CarCar specific constant	0.346	1.87
CarTrk specific constant	0.565	2.29
TrkCar specific constant	0.432	2.34
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarCar	−0.472	−2.49
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarTrk	−0.818	−4.63
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to TrkCar	0.523	2.42
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarCar	0.987	4.73
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarTrk	0.950	5.49
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarCar	1.388	2.79
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarTrk	1.078	2.25
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to CarCar	−0.908	−5.86
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to TrkCar	−1.12	−5.85
Crash occurred at traffic signal and night time, specific to CarTrk	0.436	2.61
Driver of the striker vehicle was distracted when the stricken vehicle stopped, specific to CarTrk	0.532	2.23
Driver's vision of the striker vehicle was obstructed when the stricken vehicle stopped suddenly, specific to CarTrk	1.533	6.74
Nesting coefficient of the other group	0.812	5.29
Goodness-of-fit measures		
Likelihood ratio index (ρ^2)	0.18	
Classification accuracy of the testing data	67.7%	

Table 10
Marginal effects of the calibrated nested logit model for rear-end crashes

Variable	Direct marginal effect
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarCar	−0.148
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to CarTrk	−0.153
Driver gender (1 if driver is male, 0 if female) of the striker vehicle, specific to TrkCar	−0.124
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarCar	0.172
Young driver age indicator (age between 15 and 24) of the striker vehicle, specific to CarTrk	0.177
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarCar	0.195
Old driver age indicator (age 75 and older) of the striker vehicle, specific to CarTrk	0.180
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to CarCar	−0.103
Driver gender (1 if driver is male, 0 if female) of the stricken vehicle, specific to TrkCar	−0.124
Crash occurred at traffic signal and night time, specific to CarTrk	0.078
Driver of the striker vehicle was distracted when the stricken vehicle stopped, specific to CarTrk	0.097
Driver's vision of the striker vehicle was obstructed when the stricken vehicle stopped suddenly, specific to CarTrk	0.201

5. Conclusions

During the past decade, a profound shift in the composition of the passenger vehicle fleet has been realized in the United States. Registrations of light truck vehicles (light trucks, vans, sports utility vehicles) currently account for over a third of all light vehicle registrations (Goodman, 1999), and are a growing component of the US fleet. Although the rapid growth in LTV sales has not been associated with an overall increase in traffic crashes in the United States, there is a need to determine the effect of these LTVs on other passenger cars. In this paper, nested logit models were developed to study the role of LTV on rear-end collisions.

To develop an appropriate nesting structure, many possible nesting structures were considered and calibrated. Using the likelihood ratio index and classification accuracy of the test set as measures of goodness-of-fit, the final model was selected. The structure of the final model is a two-level nesting structure that groups the CarTrk in one group and the other categories (CarCar, TrkCar, TrkTrk) in another group. The model shows the significant variables to be driver's age, gender, traffic control device, action initiated by the lead vehicle, and inattention and vision obstruction of the driver in the following vehicle. Results of the calibrated nested logit model suggested that LTV obscure drivers' visibility of other passenger cars. LTVs may prevent drivers in cars behind them from being aware of traffic situation ahead, therefore more susceptible to collide with the LTV in case of sudden application of the breaks. Also, the probability of CarTrk rear-end crashes increases when the driver of the striker vehicle is distracted when the stricken vehicle stopped. The marginal effects of the model indicate that the vision obstruction of the rear vehicle, in case a regular car is striking an LTV, has the most important effect on rear-end crashes.

References

Ben-Akiva, M., Lerman, S., 1985. Discrete Choice Analysis: Theory Application to Travel Dem. MIT Press, Cambridge, MA.

Bhat, C., 1995. Heteroscedastic extreme value model of intercity mode choice. Transport. Res. B 30 (1), 16–29.

Goodman, R., 1999. 1999 Ward's Automotive Yearbook. Ward's Communication, Detroit, MI.

Graham, J., 2000. Civilizing the sport utility vehicle. Issues Sci. Technol. 17 (2), 57–62.

Greene, W., 2000. Econometric Analysis, fourth ed. Prentice Hall, Upper Saddle River, New Jersey.

Kostyniuk, L., Eby, D., 1998. Exploring Rear-End Roadway Crashes from the Driver's Perspective, Technical Report. Human Factors Division, Transportation Research Institute, Michigan University, Ann Arbor. Sponsor: Industry Affiliation Program for Human Factors in Transportation Safety, Michigan University, Ann Arbor.

McFadden, D., 1974. Conditional Logit Analysis of Qualitative Choice Behavior: Economic Theory Mathematical Economics. Frontiers in Econometrics, Zarembka ed. Academic Press, New York, pp. 105–142.

McFadden, D., 1977. An Application of Diagnostic Tests for the Irrelevant Alternatives Property of The Multinomial Logit Model. Transportation Research Record, No. 637, pp. 39–46.

National Highway Traffic Safety Administration, 2000. GES Analytical User's Manual 1988–1999. US Department of Transportation.

Sayer, J., Mefford, M., Huang, R., 2000. The Effect of Lead-Vehicle Size on Driver Following Behavior, Technical Report. Human Factors Division, Transportation Research Institute, Michigan University, Ann Arbor. Sponsor: Industry Affiliation Program for Human Factors in Transportation, Safety Michigan University, Ann Arbor. Report No. UMTRI-2000-15. UMTRI-93805.

Spinal Injury Foundation, 2002. http://www.whiplash101.com/driver%27s2.htm, accessed on 11 March 2002.

Wang, J., Knipling, R., Blincoe, L., 1999. The dimensions of motor vehicle crash risk. J. Transport. Stat. 2 (1), 19–43.