**TRB No. 03-2017**

# Configuration Analysis of Two-Vehicle Rear-End Crashes

Mohamed A. Abdel-Aty*, Ph.D., P.E.
*Corresponding Author
Department of Civil & Environmental Engineering
University of Central Florida
Orlando, FL 32816-2450
Phone: (407) 823-5657
Fax: (407) 823-3315
Email: mabdel@mail.ucf.edu

Hassan T. Abdelwahab, Ph.D.
Department of Civil & Environmental Engineering
University of Central Florida
Orlando, FL 32816-2450
Phone: (407) 823-4552
Fax: (407) 823-3315
Email: hta76746@pegasus.cc.ucf.edu

Prepared for Presentation in the 82nd TRB annual meeting, and Possible Publication in the Journal of the Transportation Research Board

Number of words = 4,304
Number of tables and figures = 8
Total = 8 × 250 + 4,304 = 6,304

October 2002

**ABSTRACT**

Light tuck vehicles (LTV), including light-duty trucks, vans, minivans and sport-utility vehicles (SUV), are generally larger than common passenger cars and able to take on additional tasks. LTVs usually ride higher than other common passenger cars, which likely affect the visibility of passenger car drivers. This paper investigates the role of LTVs in rear-end crashes. Statistical models of unordered multiple categories are attempted, including multinomial logit (MNL), heteroscedastic extreme value (HEV), and bivariate probit (BVP) models. Four different rear-end crash configurations were defined based on the type (LTV or regular passenger car) of the two involved vehicles (lead and following vehicles). The General Estimates System (GES 2000) traffic crash data were used to calibrate the three suggested models (MNL, HEV, BVP). Modeling results showed that there are sight distance and discomfort problems when a driver in a regular passenger car is driving behind an LTV. The probability of a PC_LTV (a regular passenger car striking an LTV) rear-end crashes increases when the driver of the following vehicle is distracted. The analysis illustrated also that the probability of a regular car striking an LTV increases when the driver of the following vehicle had obscured view.

## INTRODUCTION

Rear-end collisions are the most common form of traffic crashes in the U.S. accounting for nearly a third of the 6 million crashes reported annually nationwide. Wang *et al.* (*1*) reported that the most numerous crash category in the U.S. is rear-end collisions. Data of the analysis were retrieved or derived from the General Estimates System (GES) for the five-year period 1989–93. James *et al.* (*2*) indicated that rear-end collisions are not frequently reported because most of them are of low crash severity. Therefore, the actual number of rear-end collisions is expected to be more than those reported.

In the past 2 years, the National Transportation Safety Board (*3*) investigated nine rear-end collisions in which 20 people died and 181 were injured. Common to all nine crashes was the rear following vehicle driver's degraded perception of traffic conditions ahead. Some of these collisions occurred because of atmospheric conditions, such as sun glare or fog and smoke, interfered with the driver's ability to detect slower moving or stopped traffic ahead. In other crashes, the driver did not notice that traffic had come to a halt due to congestion at work zones or to other crashes. Regardless of the individual circumstances, the drivers in these crashes were unable to detect slowed or stopped traffic and to stop their vehicles in time to prevent a rear-end collision.

Over the past decade, millions of consumers have discovered the extra utility available in light truck vehicles (LTV), such as greater seating capacity (vans & minivans) for larger families, more cargo space (SUVs, light-duty trucks and vans) for transporting luggage, more hauling and towing power, higher ground clearance (light-duty trucks and SUVs) for driving on snow and ice or on rough roads and off-roads, four-wheel-drive (many LTV models) for bad weather and rough terrain, and finally greater size and weight for occupants' safety and durability. However, concerns about the effect of these bigger vehicles on other small passenger cars are increasing, especially driver visibility. Graham (*4*) reported that a major concern often voiced by U.S. motorists is that light truck vehicles (LTV) including light-duty trucks, vans, minivans, and sports utility vehicles (SUV), make it impossible for drivers in smaller passenger cars to see the traffic ahead of them.

The objective of this paper is to investigate whether driving behind an LTV would increase the probability of being involved in a rear-end crash. We hypothesize that there might be a sight distance and discomfort problem. With the increased percentage of LTVs in traffic, traffic composition in the US is changing with potential safety implications. This paper attempts at addressing one of the issues that might result from this change. In short, the paper study the possible effect of the configuration of the two-vehicles involved in rear-end crashes. Rear-end crashes were classified into four different categories based on the lead and following vehicles as shown in Figure 1. In the analysis presented in this paper, light-duty trucks, vans, minivans, and sports utility vehicles are categorized as passenger vehicles of increased weight and dimensions (we refer to these car categories as light vehicle trucks-LTVs), as opposed to the common passenger cars (PC).

## METHODOLOGY

### Models for Unordered Multiple Choices

An unordered multiple-choice model for $N$ subjects facing $J$ unordered (nominal or qualitative) alternatives assumes the relationship:

$$P_n(j) = F(\alpha_j + \beta_j X_n, \theta);\ j = 1,..., J-1$$
$$P_n(J) = 1 - \sum_{j=1}^{J-1} P_n(j) \qquad [1]$$

where $P_n(j)$ is the probability that subject $n$ ($n=1,\ldots,N$) belongs to category $j$, $\alpha_j$ is an alterative specific constant of category $j$, $X_n$ is a vector of measurable characteristics specific to subjects and alternatives, $\beta_j$ is a vector of estimable coefficients, and $\theta$ is a parameter that controls the shape of the probability distribution $F$. The model is made operational by a particular choice of the distributional form of $F$. McFadden (*5*) assumed a logistic distribution for $F$ to derive the multinomial logit (MNL) model as follows:

$$P_n(j) = \exp(\beta_j X_n) / \sum_{j=1}^{J-1} \exp(\beta_j X_n);\ j = 1,..., J-1$$
$$P_n(J) = 1 - \sum_{j=1}^{J-1} P_n(j) \qquad [2]$$

where $P_n(j)$, $X_n$, and $\beta_j$ are defined as in equation [1]. One of the most widely criticisms of the MNL 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 (i.e., $P_n(i)/P_n(j) = \exp(\beta_i - \beta_j)X_n$). The IIA property of the MNL model comes from the assumption that F of all alternatives is independent identical distributed (IID). Bhat (*6*) developed a more general MNL model by relaxing the assumption of identical distributions, but still maintaining the independence assumption. The resulting model is called heteroscedastic extreme value (HEV). If the scale parameters (variances) of the random components of all alternatives are equal, then the probability expressions collapse to that of the MNL model.

**Bivariate probit (BVP) model**

The bivariate probit model is a useful tool for analyzing interdependence between two binary endogenous variables. Suppressing observation subscripts, the bivariate probit model (*7*) is:

$$\begin{aligned} &z_1 = \beta x_1 + \varepsilon_1, \quad y_1 = 1 \text{ if } z_1 > 0,\ 0 \text{ otherwise,} \\ &z_2 = \gamma x_2 + \varepsilon_2, \quad y_2 = 1 \text{ if } z_2 > 0,\ 0 \text{ otherwise,} \\ &[\varepsilon_1, \varepsilon_2] \sim \text{bivariate normal (BVN) } [0,0,1,1,\xi] \end{aligned} \qquad [3]$$

where $\{z_1, z_2\}$ are latent variables, $\{y_1, y_2\}$ are the observed outcomes, $\{\beta, \gamma\}$ vectors of coefficients to be estimated, $\{x_1, x_2\}$ are vectors of explanatory variables influencing the outcomes, $\{\varepsilon_1, \varepsilon_2\}$ are random error terms, and $\xi$ is the correlation coefficient between the two error terms. Distributional assumptions need to be made on the random error terms $\varepsilon_1$ and $\varepsilon_2$ in order to express response probabilities. A bivariate probit (BVP) model assumes the joint distribution of $\varepsilon_1$ and $\varepsilon_2$ is given by the bivariate standard normal distribution $\phi(\varepsilon_1, \varepsilon_2, \xi)$. Then, the parameters $\{\beta, \gamma, \xi\}$ can be estimated simultaneously using the maximum likelihood estimation approach.

In short, both HEV and BVP are considered as extensions to the MNL in which the IIA property of the MNL model is relaxed. HEV allows non-identical distribution (different shape parameters--variances), but maintains the independence assumption for the stochastic terms. If the scale parameters of the random components of all alternatives are equal, then the probability expressions collapse to that of the MNL model. BVP Allows non-identical and non-independent distributions for the stochastic terms. Models of that type commonly use a multivariate normal distribution. The resulting model, referred to as the multinomial probit (MNP) model, is difficult to handle because of introducing several additional parameters in the covariance matrix. However, Weeks and Orme (8) showed that the BVP model could be considered as a special case of the more general multinomial probit (MNP). Therefore, the BVP is another extension to the MNL. The calibration process of a BVP model is much faster and straightforward than a nested MNL model (another extension to the MNL model), in which the analyst has to decide on the appropriate nesting structure, or multinomial probit (MNP) model (because of introducing several additional parameters in the covariance matrix).

Most traffic safety studies applied the above modeling techniques to study injury severity levels. Examples include Shankar and Mannering (*9*), McGinnis *et al.* (*10*), and Krull *et al.* (*11*). In this paper, these promising approaches were utilized to study the configuration of rear-end crashes.

**TRAFFIC CRASH DATA**

The General Estimates System (GES) data come from a nationally representative sample of police-reported traffic crashes of all injury levels, from minor to fatal. 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. The GES database consists of three files: *accident*, *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. For more detailed information, refer to the GES Analytical User's Manual (*12*).

Table 1 summarizes distributions of rear-end crash configurations blocked by year. Although the PC_PC rear-end crash category has the highest percentage because of the majority of the passenger vehicles are still regular passenger cars, it worth mentioning that the percentage of PC_LTV rear-end crashes is relatively high and accounting for second highest configuration (25.52% in 2000). This indicates a possible problem with this type of

rear-end collisions. Also, the percentage of PC_PC has a decreasing trend (61.98% in 1988 vs. 38.42% in 2000), while the percentages of the three other configurations have an increasing trend.
Table 2 shows the variables that are 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. Examples of driver's vision obstruction could include reflected glare, bright sunlight, and headlights, but also sight distance problems would be considered vision obstruction. In this paper we hypothesize that the driver of a smaller passenger car following an LTV might have his line of sight obstructed by the size of the LTV. Therefore, he will not be able to see ahead beyond the LTV, and adjust his/her speed according to the situation downstream. 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). Examples of driver distraction include eating or drinking, moving object in vehicle, talking or listening to phone, and adjusting radio cassette. This could be also be a result of the size of the LTV, and could be exacerbated by the inability to see beyond the LTV, which is related to the vision obstruction. This variable was added to the GES *vehicle/driver* file in 1990.

In 1999, extensive modifications were made to the codes of the driver distraction field in the GES *vehicle/driver* file. In addition, too many missing cases of this variable as well as the driver vision obstruction factor are found 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-vehicle (involving PC or LTV) rear-end collisions to be able to define the rear-end crash configuration. In the 2000 GES, the total number of reported crashes was 57,392, of them there were 11,933 rear-end crashes with two vehicles involved. Examining the whole two-vehicle rear-end crashes, 1,579 cases (13.2% of the total number of two-vehicle rear-end crashes) have complete information about all variables listed in Table 2, specially the driver distraction and vision obstruction fields. The crash distribution of these complete cases among the four configurations (37.71%, 25.89%, 22.04%, 14.36%) is similar to the full 11,933 cases. The chi-square statistical test was used to 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.

## MODELING RESULTS

The four different categories of the rear-end crashes were modeled using the MNL, HEV, and BVP approaches. All drivers have the same possibility of exposure for involvement in a rear-end crash with one of the four configurations. All models were calibrated to estimate probabilities of the four rear-end crash configurations shown in Figure 1 and can be described as follows:

1. A regular passenger car striking another regular passenger car (PC_PC)
2. A regular passenger car striking a light truck vehicle (PC_LTV)
3. A light truck vehicle striking a regular passenger car (LTV_PC)
4. A light truck vehicle striking another light truck vehicle (LTV_LTV)

To select the model that fits the data, all models were calibrated and tested using the same sets of data. The whole data set was randomly divided into two sets; one set for calibration and the other for testing. About 90% (1,421 cases) were used for calibration and the remaining 158 cases (10% of the total cases) were used for testing. The selection of the best model was based on the estimated classification accuracy of the test phase. We could not use the likelihood values to test non-nested models. For separate non-nested models, values of likelihood are not comparable. Therefore, a direct test is precluded (*7*). The estimated classification accuracy of a calibrated model was defined as:

$$\%CC(M) = \frac{1}{T}\sum_{t=1}^{T} I_{[Y(t),D(t)]} \qquad [4]$$

where $\%CC(M)$ denotes the percentage of correct classification of a model *M*, *T* is the number of patterns in the test set, $D(t)$ is the desired output of input pattern $X(t)$, and $Y(t)$ is the actual response of the calibrated model when activated by the input pattern $X(t)$, and the indicator function $I_{[Y(t),D(t)]} = 1$ if $Y(t) = D(t)$, 0 otherwise.

**MNL and HEV models**

Table 3 presents the estimates for the calibrated MNL model. The model has an estimated classification accuracy of 60.8%. Significant variables include driver age, gender, distraction, and lighting condition. In addition, the positive coefficient of the traffic control device and night time interaction term indicates an increase in probability of PC_LTV rear-end crash configuration. To decide appropriateness of the calibrated MNL model, the validity of the IID assumption must be checked. McFadden (*13*) investigated a wide range of computationally feasible tests to detect violations of the IID assumption. This involves comparisons of many MNL logit models estimated with subsets of alternatives from the universal choice set (*1*,…, *J*). If the IIA assumption 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. Many MNL models were calibrated with subsets of alternatives (e.g., PC_PC, PC_LTV, LTV_PC). The estimated coefficients of the developed MNL models were significantly different among the models, which indicate that the calibrated MNL model violates the IID assumption.

One way to relax the IID assumption is to use the HEV model as discussed before. Table 4 summarizes the calibration results of the HEV model. Since the probability comparisons are all pairwise, one of the variances is set to one. The scale parameter of extreme value distribution of the LTV_LTV rear-end crash configuration was set to one. The estimated classification accuracy of the calibrated HEV model is 61.2%, which is slightly higher than the MNL model. The HEV model has a better goodness-of-fit (likelihood ratio index) than the MNL model. The HEV model shows the significance of driver age, gender, distraction, traffic control, and light condition as the MNL. In addition, obscuring driver's vision of the following vehicle when the lead vehicle stopped increases the likelihood of PC_LTV rear-end collisions.

**BVP model**

A bivariate probit (BVP) model was calibrated to analyze the interdependence between the two binary outcomes (types of the following and lead vehicles). The calibration process of a BVP model is generally much faster and straightforward than a nested MNL model (another extension to the MNL model), in which the analyst has to decide on the appropriate nesting structure, or multinomial probit (MNP) model (because of introducing several additional parameters in the covariance matrix of alternatives' stochastic terms). To estimate a nested logit model, it is necessary to specify a nesting structure (i.e., partition the choice set into branches). For the case considered in this paper (four categories for the dependent variable), there are 26 different nesting structures for partitioning the four rear-end crash configurations. The 26 structures are one MNL (as shown in Table 3), 13 two-level, and 12 three-level nesting structures. Many studies in the literature present several sets of results based on different specifications of the tree structure. Furthermore, there is no well-defined testing procedure for discriminating among different tree structures, which is a problematic aspect of the model (*7*).

Table 5 summarizes the estimation results of the BVP model for probabilities of the four rear-end crash configurations. The model has an estimated classification accuracy of 64.5%, which is higher than the other two models (MNL and HEV), and likelihood ratio index ($\rho^2$) of 0.10, which is also higher than the other two models. The first part of the calibrated BVP model estimates the probability of the following vehicle in a two-vehicle rear-end crash is a regular passenger car. The second part of the model estimates the probability of the lead vehicle in a two-vehicle rear-end crash is a regular passenger car. The correlation coefficient ($\xi$) measures the correlation between the disturbances in the two equations. The significance of the correlation between the two error terms underscores the importance of accounting for cross-equation correlation. Failure to take into account the association of the responses (types of the following and lead vehicles) will result in inefficient estimates (by "inefficient" we mean that the parameter estimates will have smaller variance if the correlation is included).

All variables included in the calibrated BVP model are self-explanatory and their coefficients are readily interpretable. For example, a positive coefficient of the stop-sign variable indicates that a regular passenger car is more likely to be the following vehicle, when involves in a two-vehicle rear-end collision at night at a stop sign. Since, all explanatory variables included in the calibrated BVP model were all dummy variables, then their coefficients and scaled (t-statistic) values represent a clear approach to assess impacts of these variables on rear-end crash configurations. The type of following vehicle (1 is passenger car, 0 if LTV) has highest scaled-value among all explanatory variables included in the model. In addition, the negative coefficient of this variable in the second equation indicates that if the following vehicle in rear-end crash is a regular passenger car then it is most likely to have an LTV as the lead vehicle.

Results of the calibrated BVP for driver gender and age are consistent with the MNL and HEV. For example, the MNL and HEV indicated that a male driver in the lead vehicle of a rear-end crash has a lower representation in rear-end crashes of type PC_PC and PC_LTV, which is similar to the negative sign of driver gender in the first equation of the calibrated BVP model. Also, older driver age indicator of the following vehicle in the MNL and HEV has positive coefficient specific to PC_PC and PC_LTV rear-end crash configuration, which is similar to the positive coefficient of same variable in the first equation of the calibrated BVP model. Impacts of light condition on rear-end crash configuration among the three calibrated models are consistent. Negative coefficients for this variable for PC_PC and LTV_PC rear-end crash configuration coincides with negative coefficient of the same variable in the second equation of the calibrated BVP model.

The positive coefficient of "driver of the following vehicle had obscured view when lead vehicle stopped" in the first equation indicates an increase in the likelihood of a regular passenger car being the following vehicle and its driver's vision obscured. While the negative coefficient of "driver in the following vehicle was under distraction when the lead vehicle suddenly stopped" interaction term in the second equation suggests that the lead vehicle is less likely to be a regular passenger car when the driver in the following vehicle was distracted. If we also consider that the lead vehicle is less likely to be a passenger car when the following vehicle is a passenger car (instrumentation of $y_1$ in the second equation), then when a driver in a passenger car following another passenger vehicle has a higher probability of having PC_LTV rear-end category when his/her vision is obscured and he/she was under distraction.

Marginal effects of the calibrated BVP model were calculated for all explanatory variables included in the model. There are several marginal effects of the event probabilities in the calibrated BVP model as summarized in Table 6. The first two rows and columns of Table 6 indicate the joint distribution: for example $\Phi_{11}$ is the probability of a PC_PC rear-end collision, while $\Phi_{10}$ is the probability of a PC_LTV rear-end collision. Row 3 and column 3 report the marginal probabilities: $\Phi_1$ is the marginal probability of the following vehicle to be a regular passenger car, while $\Phi_1^*$ is the marginal probability of the lead vehicle to be a regular passenger car. Rows 4 to 5 and columns 4 to 5 indicate the relevant conditional probabilities: for example $\Phi_{0|1}$ is the probability of the lead vehicle to be an LTV given that the following vehicle is a regular passenger car, while $\Phi_{1|1}^*$ is the probability of the following vehicle to be a regular passenger car given that the lead vehicle is also a regular passenger vehicle. The marginal effects of variable $x_k$ on probability of the event PC_LTV, denoted by $\Pr[y_2 = 0 | y_1 = 1, x]$, where $x$ is defined as the union vector of vectors $x_1$ and $x_2$ (i.e., $x = x_1 \cup x_2$), are given by (14, 15):

$$\begin{aligned}
&\Pr[y_2 = 0 | y_1 = 1, x] = \Pr[y_1 = 1, y_2 = 0 | x] / \Pr[y_1 = 1 | x] = \Phi_{10} / \Phi_1, \\
&\partial \Pr[y_2 = 0 | y_1 = 1, x] / \partial x_k = \Phi_1^{-2} \left( \Phi_1 \partial \Phi_{10} / \partial x_k - \Phi_{10} \partial \Phi_1 / \partial x_k \right), \\
&\partial \Phi_{10} / \partial x_k = \Phi_{0|1} \phi_1 \beta_k + \Phi_{1|0}^* \phi_0^* \gamma_k, \\
&\partial \Phi_1 / \partial x_k = \phi_1 \beta_k + (\Phi_{1|1}^* \phi_1^* - \Phi_{1|0}^* \phi_0^*) \gamma_k
\end{aligned} \qquad [5]$$

where, $\beta_k$ is the coefficient of variable $x_k$ in the following-vehicle equation, and $\gamma_k$ is the coefficient of variable $x_k$ in the lead-vehicle equation. The $\partial \Phi_{10} / \partial x_k$ equation describes the marginal effects of changes in variable $x_k$ on the probability of a PC_LTV rear-end collision as the weighted sum of two terms. The first term accounts for the effect of variable $x_k$ on the marginal probability of the following vehicle to be a regular passenger car weighted by the conditional probability ($\Phi_{0|1}$) of the lead vehicle to be an LTV. The second term ($\Phi_{1|0}^* \phi_0^* \gamma_k$) accounts for the effect of variable $x_k$ on the marginal probability of the lead vehicle being an LTV weighted by the conditional probability ($\Phi_{1|0}^*$) of the following vehicle to be a passenger vehicle.

Similar equations can be generated for the marginal effect of variable $x_k$ on probabilities of other events such as LTV_PC rear-end collisions (i.e., $\Pr[y_2 = 1 | y_1 = 0, x]$). Table 7 summarizes the marginal effects of the all the explanatory variables included in the calibrated bivariate probit model. Values of these marginal effects represent change in conditional probability of the lead vehicle to be a light truck vehicle given the type of following vehicle (PC or LTV). In case of binary explanatory variables, a marginal effect represents the change in the probability due to a change of that variable from zero to one. Driver distraction in the following vehicle has the highest marginal effect when the following vehicle is a regular passenger car. Driver's vision of the following vehicle was obscured when the lead vehicle stopped has the second highest marginal effect on a rear-end crash, when the following vehicle was a regular passenger car. This suggests that LTVs may prevent drivers in cars behind them from being aware of the traffic situation ahead. 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 PC_LTV. 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 LTVs than females (*16*). Hence, male drivers have a higher percentage of LTV_PC rear-end crashes.

## CONCLUSIONS

The popularity of LTVs has caused safety researchers to become concerned about size and height incompatibilities and their effects on driver visibility in rear-end crashes. Currently in the United States, rear-end crashes are the most numerous crash category. In this paper, rear-end crashes were classified into four categories depending on the lead and follower vehicles. Then, these four events were modeled using multinomial logit (MNL), heteroscedastic extreme value (HEV), and bivariate probit (BVP) models. Both HEV and BVP models relax the irrelevant alternatives property of the MNL model, which is a crucial consideration. Also, the HEV and BVP models do not require a priori identification of mutually exclusive partitions, as does the nested MNL structure (another extension to the MNL).

All models had similar results on the significance of driver age, gender, and visibility. The BVP model has the highest estimated classification accuracy (64.5%) as compared to the MNL and HEV models. Results of the calibrated BVP model indicated that 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 PC_LTV (a regular passenger car striking an LTV). Both driver's vision and distraction of the following vehicle in the calibrated BVP model indicated that there are sight distance and discomfort problems when a driver in a regular passenger car is driving behind an LTV. Presence of a stop sign at an intersection positively affects the probability of having a PC_LTV rear-end crash at night. Finally, results if the calibrated BVP and HEV models highlight the problem of geometric incompatibility between LTVs and other passenger cars on drivers' visibility of passenger cars involved in rear-end collisions. This emphasizes the risk of driving a passenger car too closely behind an LTV.

There is an indication from the results of this study that there is a higher chance of rear-end crashes when a regular passenger car follows an LTV due to driver distraction and limited sight distance. This implies that two phenomena are taking place at the same time with the increase of the percentage of LTVs in traffic; more rear end crash probability of a regular passenger car following an LTV, and less rear-end crash probability between two LTVs, two regular cars, and an LTV following a car. Therefore, there is likely changes occurring in rear-end crash patterns due to the changes in passenger car composition in traffic that result from the increase in sales and registration of LTVs. While there might be an early increase in rear-end crashes, it is likely that eventually there will be some benefits from the increased percentage of LTVs. Safety awareness campaigns that warn drivers of smaller passenger cars not to follow too closely an LTV because of the safety hazards that have been shown in this research might be one of the suggestions of this study. Future research should address the implications of the change in vehicle composition, particularly the substantial increase of the percentage of LTVs in traffic, on traffic safety.

## REFERENCES

1. Wang, J., Knipling, R., and Blincoe, L., The Dimensions of Motor Vehicle Crash Risk, *Journal of Transportation and Statistics*, Vol. 2, 1999, pp. 19-43.
2. James, M., Decker, R., and Warner, C., Severe and Fatal Injuries in Rear Impacts, *41st Annual Association for the Advancement of Automotive Medicine Conference*, Orlando, Florida, 1997, pp. 291-305.
3. National Transportation Safety Board, *Vehicle- And Infrastructure-Based Technology for the Prevention of Rear-End Collisions*, Special Investigation Report, report number PB2001-917003, Washington, D.C. 20594, 2001.
4. Graham, J., Civilizing the Sport Utility Vehicle, *Issues in Science and Technology*, Vol. 17, pp. 57-62, 2000.
5. McFadden, D., Conditional Logit Analysis of Qualitative Choice Behavior: Economic Theory and Mathematical economics, *Frontiers in Econometrics*, Academic Press, 1974, pp. 105-142.
6. Bhat, C., Heteroscedastic Extreme Value Model of Intercity Mode Choice, *Transportation Research Part B: Methodological*, Vol. 29, 1995, pp. 471-483.
7. Greene, W., *Econometric Analysis*, 4th ed., Prentice Hall, Upper Saddle River, New Jersey, 2000.
8. Weeks M. and Orme C., The Statistical Relationship Between Bivariate and Multinomial Choice Models, manuscript, Department of Applied Economics, University of Cambridge, UK, http://www.econ.cam.ac.uk/faculty/weeks/bvpnewx.pdf, 1998.

9. Shankar, V., and Mannering, F., An Exploratory Multinomial Logit Analysis of Single-Vehicle Motorcycle Accident Severity, *Journal of Safety Research*, Vol. 27, pp. 183-194, 1996.
10. McGinnis, R., Wissinger, L., Kelly, R., and Acuna, C., Estimating the Influences of Driver, Highway, and Environmental Factors on Run-off Road Crashes Using Logistic Regression, *The 78th Annual Meeting of the Transportation Research Board*, Washington, D.C., 1999.
11. Krull, K., Khattak, A., and Council, F., Injury Effects of Rollovers and Events Sequence in Single-Vehicle Crashes, *The 80th Annual Meeting of the Transportation Research Board*, Washington, D.C., 2000.
12. National Highway Traffic Safety Administration, (NHTSA), *GES Analytical User's Manual 1988-1999*, U.S. Department of Transportation, 2000.
13. McFadden, D., An Application of Diagnostic Tests for the Irrelevant Alternatives Property of The Multinomial Logit Model, *Transportation Research Record*, No. 637, 1977, pp. 39-46.
14. Christofides, L., Stengos, T., and Swidinsky, R., On the Calculation of Marginal Effects in the Bivariate Probit Model, *Economics Letters*, 1997, Vol. 54, pp. 203-208.
15. Christofides, L., Stengos, T., and Swidinsky, R., (2000), Corrigendum to: on the Calculation of Marginal Effects in the Bivariate Probit Model, *Economics Letters*, 2000, Vol. 68, pp. 339.
16. Kockelman, K., and Zhao, Y., Behavioral Distinctions: The Use of Light-Duty Trucks and Passenger Cars, *Journal of Transportation and Statistics*, Vol. 3, 2000, pp. 47-60.

**List of Tables**



**List of Figures**

**Table 1** Rear-end crash configuration distribution blocked by year

| Year | Rear-end crash configuration | | | | Total |
|---|---|---|---|---|---|
| | PC_PC | PC_LTV | LTV_PC | LTV_LTV | |
| 1988 | 716,502 | 209,778 | 172,492 | 57,173 | 1,155,945 |
| | 61.98% | 18.15% | 14.92% | 4.95% | 100% |
| 1989 | 688,946 | 210,017 | 186,057 | 59,706 | 1,144,727 |
| | 60.18% | 18.35% | 16.25% | 5.22% | 100% |
| 1990 | 585,958 | 204,999 | 168,748 | 62,732 | 1,022,437 |
| | 57.31% | 20.05% | 16.50% | 6.14% | 100% |
| 1991 | 592,936 | 240,953 | 182,621 | 78,581 | 1,095,091 |
| | 54.14% | 22.00% | 16.68% | 7.18% | 100% |
| 1992 | 610,071 | 242,791 | 186,931 | 89,965 | 1,129,757 |
| | 54.00% | 21.49% | 16.55% | 7.96% | 100% |
| 1993 | 610,294 | 265,669 | 213,529 | 101,585 | 1,191,076 |
| | 51.24% | 22.30% | 17.93% | 8.53% | 100% |
| 1994 | 684,904 | 308,368 | 256,744 | 118,247 | 1,368,263 |
| | 50.06% | 22.54% | 18.76% | 8.64% | 100% |
| 1995 | 691,666 | 328,482 | 279,739 | 138,101 | 1,437,987 |
| | 48.10% | 22.85% | 19.45% | 9.60% | 100% |
| 1996 | 676,169 | 350,041 | 280,182 | 156,546 | 1,462,939 |
| | 46.22% | 23.93% | 19.15% | 10.70% | 100% |
| 1997 | 650,522 | 353,907 | 280,720 | 159,121 | 1,444,270 |
| | 45.04% | 24.50% | 19.44% | 11.02% | 100% |
| 1998 | 623,139 | 367,374 | 286,539 | 171,382 | 1,448,433 |
| | 43.02% | 25.37% | 19.78% | 11.83% | 100% |
| 1999 | 545,404 | 376,794 | 300,829 | 204,110 | 1,427,138 |
| | 38.22% | 26.40% | 21.08% | 14.30% | 100% |
| 2000 | 565,173 | 375,249 | 321,020 | 209,510 | 1,470,952 |
| | 38.42% | 25.52% | 21.82% | 14.24% | 100% |

PC_PC: A regular passenger car striking another regular passenger car
PC_LTV: A regular passenger car striking an LTV
LTV_PC: An LTV striking a regular passenger car
LTV_LTV: An LTV striking another LTV

**Table 2** Variables' coding for the rear-end collision analysis

| Variable | Code |
|---|---|
| *Crash related variables* | |
| 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) traffic signal, (0,1) traffic stop sign, (0, 0) no or any other control (e.g., advisory speed sign, railroad grade crossing) |
| Relation to junction | 1 intersection related, 0 otherwise |
| *Following 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 |
| *Lead 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 |

**Table 3** Estimates of a MNL model for rear-end crash configuration

| Variable | Estimated coeff. | t-stat. |
|---|---|---|
| PC_PC specific constant | 1.517 | 7.33 |
| PC_LTV specific constant | 0.583 | 4.18 |
| LTV_PC specific constant | 0.972 | 4.09 |
| Driver gender (1 if driver is male, 0 if female) of the following vehicle, specific to PC_PC | -0.539 | -3.18 |
| Driver gender (1 if driver is male, 0 if female) of the following vehicle, specific to PC_LTV | -0.837 | -4.85 |
| Driver gender (1 if driver is male, 0 if female) of the following vehicle, specific to LTV_PC | 0.502 | 2.56 |
| Young driver age indicator (age between 15 and 24) of the following vehicle, specific to PC_PC and PC_LTV | 0.963 | 7.62 |
| Old driver age indicator (age 75 and older) of the following vehicle, specific to PC_PC | 1.287 | 3.09 |
| Old driver age indicator (age 75 and older) of the following vehicle, specific to PC_LTV | 1.048 | 2.33 |
| Driver gender (1 if driver is male, 0 if female) of the lead vehicle, specific to PC_PC | -0.849 | -6.76 |
| Driver gender (1 if driver is male, 0 if female) of the lead vehicle, specific to LTV_PC | -1.115 | -7.56 |
| Light condition (1 daylight, 0 otherwise), specific to PC_PC | -0.490 | -3.01 |
| Light condition (1 daylight, 0 otherwise), specific to LTV_PC | -0.580 | -3.12 |
| Crash occurred at traffic signal and night time, specific to PC_LTV | 0.818 | 2.93 |
| Driver of the following vehicle was distracted when the lead vehicle stopped, specific to LTV_PC | -0.399 | -2.07 |
| Goodness of fit measures (no. of cases = 1,421) | | |
| Likelihood ratio index ($\rho^2$) at convergence | 0.06 | |
| Overall Classification accuracy | 60.8% | |
| Specific classification accuracy by configuration | | |
| PC_PC | 64.5% | |
| PC_LTV | 68.2% | |
| LTV_PC | 54.2% | |
| LTV_LTV | 48.1% | |

**Table 4** Estimates of a HEV model for rear-end crash configuration

| Variable | Estimated coeff. | t-stat. |
|---|---|---|
| PC_PC specific constant | 1.706 | 8.19 |
| PC_LTV specific constant | 0.989 | 5.04 |
| LTV_PC specific constant | 1.423 | 6.61 |
| Driver gender (1 if driver is male, 0 if female) of the following vehicle, specific to PC_PC | -0.671 | -4.35 |
| Driver gender (1 if driver is male, 0 if female) of the following vehicle, specific to PC_LTV | -0.903 | -5.10 |
| Young driver age indicator (age between 15 and 24) of the following vehicle, specific to PC_PC and PC_LTV | 0.748 | 4.76 |
| Old driver age indicator (age 75 and older) of the following vehicle, specific to PC_PC | 1.088 | 2.66 |
| Old driver age indicator (age 75 and older) of the following vehicle, specific to PC_LTV | 0.867 | 2.14 |
| Driver gender (1 if driver is male, 0 if female) of the lead vehicle, specific to PC_PC | -0.702 | -3.69 |
| Driver gender (1 if driver is male, 0 if female) of the lead vehicle, specific to LTV_PC | -0.895 | -5.21 |
| Light condition (1 daylight, 0 otherwise), specific to PC_PC | -0.420 | -2.69 |
| Light condition (1 daylight, 0 otherwise), specific to LTV_PC | -0.489 | -3.04 |
| Crash occurred at traffic signal and night time, specific to PC_LTV | 0.644 | 2.53 |
| Driver of the following vehicle was distracted when the lead vehicle stopped, specific to LTV_PC | -0.262 | -1.69 |
| Driver of the following vehicle had obscured view when lead vehicle stopped, specific to PC_LTV | 0.513 | 2.45 |
| Scale parameter of extreme value distribution of PC_PC | 1.330 | 2.31 |
| Scale parameter of extreme value distribution of PC_LTV | 1.246 | 3.55 |
| Scale parameter of extreme value distribution of LTV_PC | 1.345 | 4.04 |
| Goodness of fit measures (no. of cases = 1,421) | | |
| Likelihood ratio index ($\rho^2$) at convergence | 0.08 | |
| Overall Classification accuracy | 61.28% | |
| Specific classification accuracy by configuration | | |
| PC_PC | 66.1% | |
| PC_LTV | 68.2% | |
| LTV_PC | 54.2% | |
| LTV_LTV | 48.1% | |

**Table 5** Estimates of a BVP model for rear-end crash configuration

| Variable | Estimated coeff. | t-stat. |
|---|---|---|
| *Equation of the following vehicle* | | |
| Constant | 0.279 | 4.69 |
| Driver gender (1 if driver is male, 0 if female) of the following vehicle | -0.431 | -6.53 |
| Young driver age indicator (age 15-24) of the following vehicle | 0.538 | 7.22 |
| Old driver age indicator (age 75+) of the following vehicle | 0.813 | 3.33 |
| Lead vehicle slowed down its speed | 0.240 | 2.19 |
| Driver of the following vehicle had obscured view when lead vehicle stopped | 0.702 | 3.59 |
| Crash occurred at a stop sign and at night | 0.360 | 2.83 |
| *Equation of the lead vehicle* | | |
| Constant | 1.238 | 15.96 |
| Type of the following vehicle (1 if passenger car, 0 if LTV) | -1.346 | -15.79 |
| Driver gender of the lead vehicle (1 male, 0 female) | -0.422 | -7.09 |
| Light condition (1 daylight, 0 not daylight) | -0.230 | -3.62 |
| Driver in the following vehicle was under distraction when the lead vehicle suddenly stopped (1 yes, 0 no) | -0.137 | -2.13 |
| Young driver age indicator (age 15-24) of the lead vehicle | 0.314 | 4.54 |
| Old driver age indicator (age 75+) of the lead vehicle | 0.474 | 2.56 |
| Disturbances correlation ($\xi$) | 0.849 | 10.25 |
| Goodness of fit measures (no. of cases = 1,421) | | |
| Likelihood ratio index ($\rho^2$) at convergence | 0.10 | |
| Overall Classification accuracy | 64.5% | |
| Specific classification accuracy by configuration | | |
| PC_PC | 71.0% | |
| PC_LTV | 70.5% | |
| LTV_PC | 54.2% | |
| LTV_LTV | 55.0% | |

**Table 6** Event probabilities in the calibrated BVP model

| Following vehicle ($y_1$) | Lead vehicle ($y_2$): Passenger car ($y_2 = 1$) | Lead vehicle ($y_2$): LTV ($y_2 = 0$) | Marginal distribution of $y_1$ | Conditional distribution of $y_2$ on $y_1$ | |
|---|---|---|---|---|---|
| Passenger car ($y_1 = 1$) | $\Phi_{11}$ | $\Phi_{10}$ | $\Phi_1$ | $\Phi_{1\|1}$ | $\Phi_{0\|1}$ |
| LTV ($y_1 = 0$) | $\Phi_{01}$ | $\Phi_{00}$ | $\Phi_0$ | $\Phi_{1\|0}$ | $\Phi_{0\|0}$ |
| Marginal dist. of $y_2$ | $\Phi_1^*$ | $\Phi_0^*$ | 1 | | |
| Conditional distribution of $y_1$ on $y_2$ | | | | | |
| Passenger car ($y_1 = 1$) | $\Phi_{1\|1}^*$ | $\Phi_{1\|0}^*$ | | | |
| LTV ($y_1 = 0$) | $\Phi_{0\|1}^*$ | $\Phi_{0\|0}^*$ | | | |

**Table 7** Marginal effects of the calibrated BVP model

| Variable | Following vehicle | |
|---|---|---|
| | PC | LTV |
| Driver gender (1 if driver is male, 0 if female) of the lead vehicle | 0.185 | 0.146 |
| Light condition at the time of the crash Light condition (1 daylight, 0 not daylight) | 0.101 | 0.082 |
| Driver in the following vehicle was distracted and the lead vehicle stopped | 0.401 | 0.044 |
| Young driver age indicator (age 15-24) of the following vehicle | 0.019 | 0.038 |
| Old driver age indicator (age 75+) of the following vehicle | 0.029 | 0.040 |
| Driver gender (1 if driver is male, 0 if female) of the following vehicle | -0.126 | -0.117 |
| The lead vehicle slowed down its speed | 0.070 | 0.061 |
| Driver of the following vehicle had obscured view when lead vehicle stopped | 0.250 | 0.175 |
| Crash occurred at stop sign and at night | 0.153 | 0.092 |

* Values represent change in conditional probability of the lead vehicle is an LTV

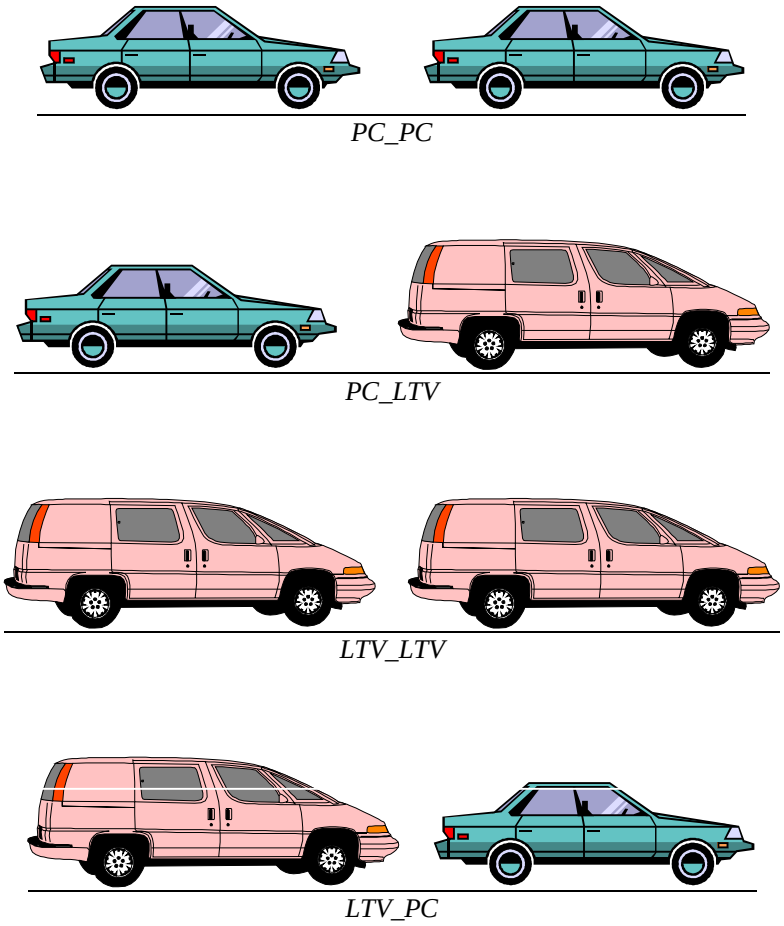


**Figure 1** Rear-end crash configurations