

USING LOGISTIC REGRESSION TO IDENTIFY RISK FACTORS CAUSING ROLLOVER COLLISIONS

Essam Dabbour¹

¹ Abu Dhabi University, Abu Dhabi, P.O. Box 59911, United Arab Emirates

Received 28 July 2012; accepted 10 September 2012

Abstract: Rollover collisions are among the most serious collisions that usually result in severe injuries or fatalities. In 2009, there were 8,732 fatal rollover collisions in the United States of America that resulted in the death of 9,833 persons. Those numbers represent approximately 28% and 29% of the total numbers of fatal collisions and fatalities, respectively. The main objective of this paper is to examine the impact of different risk factors that may contribute to this type of serious collisions to help develop countermeasures that limit them. To avoid the bias that may be caused by interactions among different drivers, this analysis focuses on rollover related to single-vehicle collisions so that the behavior of the driver of the collided vehicle can be analyzed more effectively. Logistic regression technique is utilized to analyze single-vehicle rollover collisions that occurred on state and interstate highways in the states of Ohio and Washington in 2009. The results obtained from this analysis have the potential to help decision makers identify different strategies to limit the severity of this type of collisions.

Keywords: traffic accidents, rollover collisions, logistic regression analysis, risk management.

1. Introduction

It was previously founded that rollover collisions usually cause more fatalities and injuries than other types of collisions (Chang and Yeh, 2006; Harrison, 1997; Islam and Mannering, 2006; Shankar and Mannering, 1996). Statistically, single-vehicle collisions represented 46.3%, 46.7% and 46.9% of the total number of traffic-related fatal collisions in the United States in the years 2007, 2008, and 2009 respectively (NHTSA, 2012). Furthermore, since a single-vehicle collision is mainly related to a specific driver, who is usually identified, with no interactions with other road users, analyzing this type of collisions can provide valuable insights regarding the behavior of the specific driver involved in the collision, and when aggregated,

can provide more details about the effects of different risk factors related to specific driver groups.

More than 90% of vehicle rollovers occur around the longitudinal vehicle axis (Digges et al., 1991) where different kinematic scenarios for a rollover collision may take place, including:

1. Trip-over: when the lateral motion of the vehicle is suddenly slowed or stopped inducing a rollover. The opposing force may be produced by a curb, pot-hole, or pavement that the vehicle wheels dig into.
2. Fall-over: when the surface on which the vehicle is traveling slopes downward in the direction of vehicle movement so that the center of gravity (cg) becomes outboard

¹ Corresponding author: essam.dabbour@adu.ac.ae

of its wheels. A turn-over collision is when the vehicle is traveling upgrade under similar circumstances.

3. Flip-over: when a vehicle is rotated around its longitudinal axis by a ramp-like object such as a turned down guardrail or the back slope of a ditch. The vehicle may be in yaw when it comes in contact with a ramp-like object.
4. Bounce-over: when a vehicle rebounds off a fixed object and overturns as a consequence. The rollover usually occurs in close proximity to the object from which it is deflected.
5. Turn-over: when centrifugal forces from a sharp turn or vehicle rotation are resisted by normal surface friction (most common for vehicle with higher cg). The surface includes pavement surface and gravel, grass, dirt and there is no furrowing, gouging at the point of impact. If rotation and/or surface friction causes a trip, the rollover is classified as a turn-over.
6. Climb-over: when vehicle climbs up and over a fixed object (e.g., guardrail, barrier) that is high enough to lift the vehicle completely off the ground. The vehicle must roll on the opposite side from which it approached the object.
7. End-over-end: when a vehicle rolls primarily about its lateral axis (pitch motion).

Ponboon et al. (2010) identified driver behavior and roadside slope as two major factors that cause a rollover collision. McLean et al. (2005) found that rollover crashes in Australia tend to increase on weekends and during daylight; and they also found that sports utility vehicles are the most type of personal vehicles exposed to rollover crashes. Wright and Zador (1981)

found that sharp horizontal curves (especially left-handed) and grades are both major factors that lead to a rollover collision. Deleys and Parada (1986) found that the likelihood of rollover increases with the steepness and height of side slopes and the depth of ditches.

This paper examines the impact of different risk factors that may contribute to rollover collisions in order to help develop countermeasures that limit them. The factors investigated include driver-related, vehicle-related, environmental-related, and roadway-related factors. To avoid the bias that may be caused by interactions among different drivers, this analysis focuses on rollover related to single-vehicle collisions so that the behavior of the driver of the collided vehicle can be analyzed more effectively.

2. Data Collection and Analysis

This paper is based on logistic regression analysis to all rollover collisions that occurred on state and interstate highways in the states of Ohio and Washington in 2009. Logistic regression is a tool that has been widely used in road safety studies to identify and quantify different risk factors (Harb et al., 2008; Robertson and Aultman-Hall, 2001; Chang and Yeh, 2006). It is a generalized linear model that predicts the probability of occurrence of an event by fitting data to a logit function in the form of Eq. (1) (McCullagh and Nelder, 1989):

$$f(z) = e^z / (1 + e^z) \quad (1)$$

In the above equation z is the logit function, which is a measure of the total contribution of all the independent (explanatory) variables used in the model, and $f(z)$ is a dichotomous variable that is assumed to follow Bernoulli distribution. It represents the probability of a rollover collision; and therefore, it takes the

value of 1 if the collision type was rollover and 0 if not. The logit function has the following form (Eq. (2)):

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k \quad (2)$$

where β_0 is the intercept and $(\beta_1, \beta_2, \beta_3, \dots, \beta_k)$ are the regression coefficients of explanatory variables $(x_1, x_2, x_3, \dots, x_k)$, respectively. The explanatory variables were extracted from the data sets obtained from the Highway Safety Information System (HSIS) for all collisions occurred in Ohio and Washington states in 2009. Given the distinctive differences between two-wheel vehicles and four-wheel vehicles in respect to rollover characteristics, this research focuses on data related to four-wheel passenger cars. Analysis of rollover collisions for two-wheel vehicles (e.g. motorcycles) is adequately covered in previous research (Chang and Yeh, 2006; Suraji and Tjahjono, 2012). After imputing incomplete and irrelevant records, the following records were used to calibrate the logistic regression models for the two states:

- Ohio: total of 37,422 single-vehicle collisions (where 1739 collisions were rollover); and
- Washington: 28,830 single-vehicle collisions (where 1650 collisions were rollover).

A Wald test is used to test the statistical significance of each of the estimated coefficients $(\beta_1, \beta_2, \beta_3, \dots, \beta_k)$. The Wald statistic of asymptotic chi-square distribution is the squared value of the Z statistic and is computed as Eq. (3):

$$Wald = [\hat{\beta}_i / SE(\hat{\beta}_i)]^2 \quad (3)$$

Where $\hat{\beta}$ is the i^{th} estimated coefficient and $SE(\hat{\beta}_i)$ is the standard error of that coefficient. In this study, the odds ratio (OR) is used

to interpret the significance of different risk factors where an estimate of the odds ratio of a certain risk factor is $\exp(\hat{\beta}_i)$ while holding the other risk factors unchanged. Finally, the 95% confidence interval (CI) is also used in this study to describe the upper and lower limits of the odds ratio with 95% confidence level and is $[\hat{\beta}_i \pm Z_{0.95} \cdot SE(\hat{\beta}_i)]^2$.

3. Results and Discussion

Logistic regression coefficients for all explanatory variables tested are shown in Table 1, and the odds ratios (OR) associated with different explanatory variables (risk factors), including the lower and upper 95% confidence intervals (CI), are shown in Table 2. The overall model fit for both models were highly significant (p value < 0.001). The tables show that sports utility vehicles (SUVs) are the most vulnerable to rollover collisions with odds ratio 3.6176 (for Ohio) and 2.8754 (for California). This suggests that sports utility vehicles are 3.6176 (or 2.8754) more likely to be involved in a rollover collision than other types of personal-use vehicles.

It was also found that adverse weather (i.e. rain, snow and ice conditions) is another significant risk factor that leads to rollover collisions with odds ratios 1.9164 and 2.0428 for Ohio and Washington, respectively. This finding can be mostly attributed to the reduced friction between vehicle tire and road surface, which may lead to turn-over type of collisions where centrifugal forces exceed the friction between the tire and the road surface. Sports utility vehicles, that have higher center-of-gravity than other types of personal vehicles, may be more susceptible to this type of collisions especially on curved segments of highways. To test this hypothesis, combined risk factors (including sports utility vehicle, curved road and adverse weather)

Table 1
Logistic Regression Coefficients and their Standard Errors

| Risk Factor | Ohio | | Washington | | Overall | |
|------------------------|-------------|--------|-------------|--------|-------------|--------|
| | Coefficient | SE | Coefficient | SE | Coefficient | SE |
| Sports Utility Vehicle | 1.2858 | 0.0823 | 1.0596 | 0.0874 | 1.1758 | 0.0851 |
| Exceeding Speed Limit | 0.4621 | 0.1064 | 0.4283 | 0.0454 | 0.4483 | 0.0974 |
| Driver's Age | -0.0204 | 0.0027 | -0.0107 | 0.0018 | -0.187 | 0.0025 |
| Curved Road | 0.5028 | 0.0864 | 0.4038 | 0.0742 | 0.4687 | 0.0796 |
| Adverse weather | 0.6504 | 0.0843 | 0.4531 | 0.0513 | 0.5871 | 0.0618 |
| Winter Months | 0.5139 | 0.0798 | 0.1314 | 0.0204 | 0.3176 | 0.0581 |

Source: Author

Table 2
Odds Ratios Related with Different Risk Factors (with 95% Lower and Upper Confidence Intervals)

| Risk Factor | Ohio | Washington | Overall |
|------------------------|-----------------------------|-----------------------------|-----------------------------|
| Sports Utility Vehicle | 3.6176 (3.0788 – 4.2507) | 2.8754 (2.2154 – 3.1543) | 3.1875 (2.9572 – 3.8963) |
| Exceeding Speed Limit | 1.5874 (1.2886 – 1.9555) | 1.3591 (1.1855 – 1.8630) | 1.4858 (1.2984 – 1.8586) |
| Driver's Age | 0.9798 (0.9746 – 0.9851) | 0.9652 (0.9578 – 0.9793) | 0.9716 (0.9712 – 0.9843) |
| Curved Road | 1.6534 (1.3957 – 1.9585) | 1.6341 (1.4463 – 1.8455) | 1.6497 (1.4129 – 1.8732) |
| Adverse weather | 1.9164 (1.7462 – 2.1548) | 2.0428 (1.8759 – 2.2107) | 1.9428 (1.7852 – 2.1897) |
| Winter Months | 1.6718 (1.4296 – 1.9551) | 1.1482 (1.0168 – 1.3252) | 1.3837 (1.1759 – 1.6919) |

Source: Author

were analyzed using logistic regression. The logistic regression coefficients, along with associated standard errors, are presented in Table 3; and the odds ratios, along with 95% lower and upper CI's are presented in Table 4. Both tables show that combining the three risk factors together (i.e. SUV, curved road and adverse weather) significantly increases the odds ratios of a rollover collision to 6.7542 and 6.3127 for Ohio and Washington, respectively. Additionally, combining any two of the three risk factors also increases the odds ratios of a rollover collision than those associated with a single risk factor. Those findings suggest high level of interactions among the three risk factors (i.e. SUV, curved road and adverse weather) in case if they co-exist, which supports the hypothesis explained.

Exceeding speed limit was also found to be a significant risk factor that leads to rollover collisions with odds ratios 1.5874 and 1.3591 for Ohio and Washington, respectively. This finding reinforces the current trend from law makers and law-enforcement authorities, backed by other organizations, to enforce drivers' obedience to speed limits. From Tables 3 and 4, it is evident that exceeding speed limit on curved segments of roadways will have more significant effect with odds ratios 3.4182 and 3.2047 for Ohio and Washington, respectively. This finding suggests that exceeding speed limit will increase centrifugal forces on curved roadways until the centrifugal forces exceed the friction between the tire and the road surface, which will result in a turn-over type of collisions.

Driver's age was found to be another risk factor contributing to rollover collisions where it was found that the odds of a rollover collision decreases for every year of driver's age by 0.0202 and 0.0348 for Ohio and Washington, respectively. This means that a 20-years-old

driver is more likely to be involved in a rollover collision than a 60-years-old driver by 1.808 (in Ohio) or 2.392 (in Washington). This is explained in the light of the fact that the age may increase driver's experience and therefore increasing the age may lead to taking more proper evasive measures (in terms of proper steering and/or braking) that may reduce the probability of a rollover collision. However, it should also be noted that increasing the age for very old drivers may increase the probability of a rollover collision due to the deteriorating health conditions that may increase driver's perception-reaction time and therefore slow driver's evasive measures. To test for this hypothesis, a new dichotomous variable was created that accounts to whether the driver is in the older age group (above 65 years), where the variable takes a value of 1 if the driver was above the age of 65 years and value of 0 otherwise. It was found that the odds ratios of being involved in a rollover collision increase by 1.8452 in Ohio (or 1.7896 in Washington) if the driver was above the age of 65 years. It should be noted that there were other potential risk factors that were tested and found to be insignificant in increasing the risk of rollover collisions. Those risk factors that were found to be insignificant include driver's gender (male vs. female), highway setting (urban vs. rural), day of the week (weekends vs. weekdays), time of the day (daylight vs. night) and highway lighting conditions.

4. Conclusion

In this study, logistic regression was used to identify and quantify the effects of different risk factors that may increase the risk of a rollover collision. The risk factors identified include the type of the vehicle, where it was found that sports utility vehicles (SUVs) are more susceptible to rollover collisions than other types of personal vehicles. The odds

Table 3

Logistic Regression Coefficients and Standard Errors for Combined Risk Factors

| Combined Risk Factors | Ohio | | Washington | | Overall | |
|--|-------------|--------|-------------|--------|-------------|--------|
| | Coefficient | SE | Coefficient | SE | Coefficient | SE |
| Sports Utility Vehicle + Curved Road + Adverse Weather | 2.0558 | 0.1227 | 1.9365 | 0.1024 | 1.9870 | 0.1128 |
| Sports Utility Vehicle + Curved Road | 1.9257 | 0.1134 | 1.8879 | 0.1086 | 1.9068 | 0.1102 |
| Sports Utility Vehicle + Adverse Weather | 1.8754 | 0.0967 | 1.9147 | 0.0878 | 1.8905 | 0.0913 |
| Curved Road + Adverse Weather | 0.8157 | 0.0254 | 0.8697 | 0.0337 | 0.8491 | 0.0294 |
| Exceeding Speed Limit + Curved Road | 1.1056 | 0.0631 | 0.9486 | 0.0487 | 1.0367 | 0.0571 |

Source: Author

Table 4

Odds Ratios Related with Combined Risk Factors (with 95% Lower and Upper Confidence Intervals)

| Combined Risk Factors | Ohio | Washington | Overall |
|--|-----------------------------|-----------------------------|-------------------------------|
| Sports Utility Vehicle + Curved Road + Adverse Weather | 6.7542 (6.1028 – 7.1240) | 6.3127 (6.0328 – 6.7145) | 6.5952 (6.0893 – 6.9923) |
| Sports Utility Vehicle + Curved Road | 5.4285 (5.9842 – 6.7892) | 5.2382 (5.8926 – 6.6921) | 5.3980 (5.9105 – 6.7054) |
| Sports Utility Vehicle + Adverse Weather | 5.1028 (4.4895 – 5.5207) | 5.3189 (4.6983 – 5.6851) | 5.1873 (4.5632 – 5.5607) |
| Curved Road + Adverse Weather | 2.7483 (2.4318 – 3.0781) | 3.0574 (2.7832 – 3.2073) | 2.9364 (2.6087 – 3.1076) |
| Exceeding Speed Limit + Curved Road | 3.4182 (3.0254 – 3.7568) | 3.2047 (2.8637 – 3.6395) | 3.3893 (2.9680 – 3.3.7108) |

Source: Author

ratio of a SUV being involved in a rollover collision significantly increase with curved roadway and/or adverse weather conditions. Exceeding speed limit was also found to be a significant risk factor that leads to rollover collisions, especially on curved roadways where odds ratio of rollover collisions significantly increase. Driver's age was also a significant risk factor where it was found that the odds of a rollover collision decreases for every year of driver's age so that a 20-years-old driver is more likely to be involved in a rollover collision than a 60-years-old driver by 1.808 (in Ohio) or 2.392 (in Washington). This is attributed to the more driving experience gained with age so that an older driver is more likely to take more proper evasive measures (in terms of proper steering and/or braking) that may reduce the probability of a rollover collision. However, it was also found that odds ratios of being involved in a rollover collision significantly increase if the driver was above the age of 65 years due to the deteriorating health conditions that may increase driver's perception-reaction time and therefore slow driver's evasive measures. There were other potential risk factors that were tested and found to be insignificant in increasing the risk of rollover collisions, including driver's gender (male vs. female), highway setting (urban vs. rural), day of the week (weekends vs. weekdays), time of the day (daylight vs. night) and highway lighting conditions.

The findings of this research provide the scientific evidence to support the efforts made by various law-making and law-enforcement agencies to enforce speed limits. Furthermore, this research can serve as a scientific evidence to educate the public about the disadvantages of sports utility vehicles, which are incorrectly thought by many people to be safer than passenger cars. Furthermore, the findings of this research

provide scientific evidence to support the efforts made by several government bodies to impose extra taxes on sports utility vehicles in order to discourage the public from using them given their negative impacts on safety as well as on the environment due to their high rates of fuel consumption. In general, the findings of this research also have the potential to help decision makers identify the more significant risk factors that increase the risk of rollover collisions so that they can allocate the resources to reduce those factors in order to increase road safety and ultimately improve public safety and the overall quality of life.

References

- Chang, H.; Yeh, T. 2006. Risk factors to driver fatalities in single-vehicle crashes: comparisons between non-motorcycle drivers and motorcyclists, *ASCE Journal of Transportation Engineering*. DOI: [http://dx.doi.org/10.1061/\(ASCE\)0733-947X\(2006\)132:3\(227\)](http://dx.doi.org/10.1061/(ASCE)0733-947X(2006)132:3(227)), 13(3): 227-236.
- DeLeys, N.J.; Parada, L.O. 1986. Rollover potential of vehicles on embankments, sideslopes and other roadside features. *Report FHWA/RD-86/164*, Federal Highway Administration, USA. 246 p.
- Digges, K.H.; Malliaris, A.C.; Ommaya, A.K.; McLean, A.J. 1991. Characterization of rollover casualties. In *Proceedings of the 17th International IRCOBI Conference*, University of Virginia and DeBlois Associates. 309-319.
- Harb, R.; Radwan, E.; Yan, X.; Pande, A.; Abdel-Aty, M. 2008. Freeway work-zone crash analysis and risk identification using multiple and conditional logistic regression, *ASCE Journal of Transportation Engineering*. DOI: [http://dx.doi.org/10.1061/\(ASCE\)0733-947X\(2008\)134:5\(203\)](http://dx.doi.org/10.1061/(ASCE)0733-947X(2008)134:5(203)), 134(5): 203-214.
- Harrison, W.A. 1997. An exploratory investigation of the crash involvement of disqualified drivers and motorcyclists, *Journal of Safety Research*. DOI: [http://dx.doi.org/10.1016/S0022-4375\(96\)90042-0](http://dx.doi.org/10.1016/S0022-4375(96)90042-0), 28(3): 213-219.

Islam, S.; Mannering, F.L. 2006. Driver aging and its effect on male and female single-vehicle accident injuries: Some additional evidence, *Journal of Safety Research*. DOI: <http://dx.doi.org/10.1016/j.jsr.2006.04.003>, 37(3): 267-276.

McCullagh, P.; Nelder, J.A. 1989. *Generalized linear models*, 2nd Ed. Chapman and Hall, USA. 532 p.

McLean, A.J.; Kloeden, C.N.; Ponte, G.; Baldock, M.R.J.; Lindsay, V.L.; Van den Berg, A.L. 2005. *Rollover crashes*. Centre for Automotive Safety Research, Report CASR026. The University of Adelaide, Australia. 28p.

National Highway Traffic Safety Administration (NHTSA). 2012. Fatality Analysis Reporting System (FARS). Available from Internet: <<http://www-fars.nhtsa.dot.gov>>.

Ponboon, S.; Boontob, N.; Tanaboriboon, Y.; Bin Islam, M.; Kanitpong, K. 2010. Contributing factors of road crashes in Thailand: evidences from the accident in-depth study, *Journal of the Eastern Asia Society for Transportation Studies*, 8(3): 1958-1970.

Robertson, S.; Aultman-Hall, L. 2001. Impact of road conditions on elderly drivers, *ASCE Journal of Transportation Engineering*. DOI: [http://dx.doi.org/10.1061/\(ASCE\)0733-947X\(2001\)127:3\(244\)](http://dx.doi.org/10.1061/(ASCE)0733-947X(2001)127:3(244)), 127(3): 244-246.

Shankar, V.; Mannering, F. 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity, *Journal of Safety Research*. DOI: [http://dx.doi.org/10.1016/0022-4375\(96\)00010-2](http://dx.doi.org/10.1016/0022-4375(96)00010-2), 27(3): 183-194.

Suraji, A.; Tjahjono, N. 2012. A confirmatory factor analysis of accidents caused by the motorcycle aspect in urban area. *International Journal for Traffic and Transport Engineering*, 2(1): 60-69.

Wright, P.H.; Zador, P. 1981. Study of fatal rollover crashes in Georgia, *Transportation Research Record*, 819(1): 8-17.

PRIMENA LOGISTIČKE REGRESIJE U IDENTIFIKACIJI FAKTORA RIZIKA KOJI UZROKUJU SUDARE SA PREVRTANJEM

Essam Dabbour

Sažetak: Sudari sa prevrtanjem spadaju među najozbiljnije sudare koji obično rezultiraju teškim povredama ili smrtnim ishodom. U SAD je 2009. godine zabeleženo 8732 sudara sa prevrtanjem sa smrtnim ishodom u kojima je poginulo 9833 osoba. Ovi podaci predstavljaju približno 28% i 29% od ukupnog broja sudara sa smrtnim ishodom i smrtnih slučajeva, respektivno. Osnovni cilj ovog rada je da se ispita uticaj različitih faktora rizika koji mogu doprineti pojavi ove vrste teških sudara što bi pospešilo razvoj zaštitnih mera. Da bi se izbegli mogući uticaji do kojih može doći usled interakcije između različitih vozača, ova analiza je usmerena na prevrtanje pojedinačnog vozila kako bi se ponašanje vozača vozila koje je učestvovalo u sudaru moglo efikasnije analizirati. Pri tome je korišćena logistička regresija za analizu sudara sa prevrtanjem pojedinačnog vozila koji su se desili na državnim i međudržavnim autoputevima u državama Ohajo i Vašington 2009. godine. Rezultati do kojih se došlo ovom analizom mogu poslužiti donosiocima odluka u cilju identifikacije različitih strategija za ograničenje težine ovog tipa sudara.

Ključne reči: saobraćajne nesreće, sudari sa prevrtanjem, logistička regresija, upravljanje rizikom.