

INSURANCE INSTITUTE FOR HIGHWAY SAFETY

May 13, 2008

The Honorable Nicole R. Nason
Administrator
National Highway Traffic Safety Administration
1200 New Jersey Avenue, SE, West Building
Washington, DC 20590

Supplemental Notice of Proposed Rulemaking; 49 CFR Part 571, Federal Motor Vehicle Safety Standards, Roof Crush Resistance; Docket No. NHTSA-2008-0015

Dear Administrator Nason:

The Insurance Institute for Highway Safety (IIHS) has conducted a study that demonstrates a direct relationship between roof strength and injury risk reduction in rollover crashes (Brumbelow et al., 2008). We included this study in our previous comment to the docket (IIHS, 2008) because of its relevance to the National Highway Traffic Safety Administration's (NHTSA) rulemaking under Federal Motor Vehicle Safety Standard (FMVSS) 216.

Finding that stronger roofs reduce the risk of injury in rollover crashes, the IIHS study contradicts two previous studies on the topic (Moffatt and Padmanaban, 1995; Padmanaban et al., 2005). Two authors of these earlier studies have submitted a comment and additional analysis to NHTSA (Padmanaban and Moffatt, 2008), questioning the IIHS study and concluding that "stronger roofs are not safer roofs."

The comments by Padmanaban and Moffatt (2008) contain misleading statements about the IIHS study that are detailed in item 6 of the attached document, "Logical and Statistical Errors in Comments by Padmanaban and Moffatt on the Insurance Institute for Highway Safety Study, 'Roof Strength and Injury Risk in Rollover Crashes.'" In addition, the analytical tactics recommended and used by Padmanaban and Moffatt depart in fundamental ways from appropriate use and interpretation of statistical results (see item 4). Of most concern is their insistence on including ejection, belt use, and alcohol use as control variables in their analysis when, in fact, these variables are either direct outcomes of roof crush strength or affected by the dependent variable, injury risk. Inclusion of them in the analysis obfuscates the real effects of roof strength on injury risk (see items 1-3).

These concerns are detailed in the attachment. We would be happy to discuss the issues further if NHTSA has questions.

Sincerely,



Adrian K. Lund, Ph.D.
President

cc: Docket Clerk, Docket No. NHTSA-2008-0015

Logical and Statistical Errors in Comments by Padmanaban and Moffatt on the Insurance Institute for Highway Safety Study, “Roof Strength and Injury Risk in Rollover Crashes”

1. Ejection is an outcome of rollover and is influenced by roof strength. Including ejection as a predictor of death or serious injury in a rollover crash masks a major benefit of roof strength.

Padmanaban and Moffatt argue that IIHS should have included a number of additional variables in the predictive model of injuries and deaths in rollovers. One of these variables is ejection. Their argument is that ejection greatly increases the risk of injury while “ejection is...likely to be unrelated to roof strength” (pg. 1).

- a. This argument is illogical. Roof strength may not affect injury risk once a person is ejected, but a strong roof may prevent occupants from being ejected in the first place. Preventing an occupant compartment from collapsing obviously can reduce ejection risk by preventing broken glazing and deformed structure, which create ejection paths.
- b. This argument is testable. Using the midsize SUVs in the IIHS study, IIHS researchers investigated the relationship between roof strength and ejection risk with an additional analysis. The risk of ejection was 31 percent lower for each 1-unit increase in peak roof strength-to-weight ratio (SWR) measured within 5 inches of plate displacement (p-value of 0.004). Appendix A reports details of this analysis. Clearly, ejection risk is not “unrelated to roof strength.”
- c. By treating ejection as a risk factor unrelated to roof strength, when reduced ejection risk is one of the benefits of stronger roofs, Padmanaban and Moffatt bias their analysis against finding a relationship between roof strength and injury risk.
- d. Padmanaban and Moffatt’s concern about ejection implies that roof strength does not matter if ejected occupants are not counted. However, a new IIHS analysis limited to drivers coded by police as not having been ejected reveals that stronger roofs reduced injury risk among these drivers. Many of the fatal and incapacitating injuries in the overall analysis were sustained by ejected drivers, but risk reductions for drivers not ejected were statistically significant and very similar to the overall analysis. Appendix B reports the full results.

2. Belt use cannot be used in a model evaluating roof strength and injury likelihood because information about belt use in crashes is inaccurate, incomplete, and subject to influence by the injury outcomes.

Another variable that Padmanaban and Moffatt argue should be included as a control (predictor) variable in the IIHS study is police-reported belt use. According to Padmanaban and Moffatt, “It is well known that the majority of rollover KA injuries and fatalities are to unbelted occupants, mostly ejectees” (pg. 2) and, later, “... 56% of the fatalities and 28% of the serious/fatal injuries were unbelted and completely ejected” (pg. 5). As a result, Padmanaban and Moffatt conclude that belt use should have been a predictor variable. However, because this variable is difficult to know with precision, inclusion as a predictor variable can bias any analysis of roof strength.

- a. The principal source of bias in belt use codes is that police-coded belt use is subject to distortion by crash outcomes. No official typically is present to observe belt use prior to a crash. Instead, police must *judge* belt use based on information gathered after the crash including statements by occupants about their own belt use, statements by witnesses to the crash and, significantly, the presence of injuries and whether police believe they are consistent or inconsistent with belt use. In other words, Padmanaban and Moffatt include in their analysis a variable that is itself subject to

influence by the outcome (injury severity and pattern) to be predicted. In addition, occupant statements about belt use are influenced by the fact that it is illegal in most states to be unbelted. A result of these twin biases is that belt use in crashes can be overestimated, especially for occupants with lesser injuries whose claims of belt use are more believable (Schiff and Cummings, 2004). Models including belt use as a predictor of injury severity not only introduce general inaccuracy but also overestimate the effect of belt use on reducing injury, simultaneously masking the effects of any other variables.

Evidence of the bias toward overestimating belt use in the dataset used in the IIHS study is provided by comparisons with NHTSA's National Occupant Protection Use Survey (NOPUS), which records rates of belt use for the general population observed during daylight hours. During the calendar years of the IIHS study, NOPUS data show driver belt use averaging 70-75 percent, which is lower than the 83 percent recorded by police for drivers in the rollover crashes in the IIHS study. It is unlikely that drivers involved in single-vehicle rollover crashes, many of which occur at night when belt use rates are lower (NHTSA, 2005, 2007), were wearing belts more often than the general population during daylight hours.

- b. Because of these problems, IIHS did not include belt use as a predictor. However, IIHS did examine whether the effects varied by coded belt use. As reported in the study, additional statistical models were run for occupants coded as belted (83 percent), for those coded as unbelted (10 percent), and for those coded as unknown (7 percent).
 - i. For those coded as belted, the pattern of effects of roof strength varied little from the overall analysis. This is not surprising because most drivers in the study were coded as belted. In addition, if belt use is miscoded, as argued above, then many of the drivers actually were unbelted, again meaning that this analysis is very similar to the overall analysis.
 - ii. For those coded as unknown, the pattern also was quite similar to the overall analysis. Again, this is not surprising because the unknown group also included both belted and unbelted occupants.
 - iii. Effects estimated for those coded as unbelted were much smaller, but this would be expected from the twin biases noted in item 2.a. It is likely many of those coded as unbelted received their codes because their injuries were serious and inconsistent with belt use. This bias would occur for both weak and strong roofs, masking the effect of roof strength by assigning higher weight to the (overestimated) effect of belt use.

The conclusion from these separate analyses is that coded belt use does not affect the estimated effect of roof strength on injury severity, except in a way that would be expected from the biases and inaccuracies inherent in police-coded belt use.

3. Like police-coded belt use, police-coded alcohol involvement in crashes is incomplete, inaccurate, and may be related to the injury severity. Besides, Padmanaban and Moffatt offer no justification other than the empirical relationship, which could be spurious, for including alcohol use codes in the prediction equation.

- a. Results of blood alcohol concentration (BAC) tests are the most objective measures of the presence of alcohol, but only a small percentage of crash-involved drivers typically are tested. Queries of the state databases used in the IIHS study show that about 11 percent of the drivers studied were tested. Padmanaban and Moffatt report using a combination of BAC test results and "had been drinking" codes. They do not specify in their comments to NHTSA what

percentage of the codes resulted from actual BAC tests, what codes were used for those not tested, or the extent of missing data. In response to an IIHS inquiry, they provided this additional information:

- i. Of drivers identified in their analysis as positive for alcohol use, about 18 percent were tested. About 13 percent tested positive, and 5 percent were coded as having positive alcohol use despite negative BAC tests. Thus 5 percent were coded as positive for alcohol despite chemical tests to the contrary.
 - ii. For drivers without BAC test results, Padmanaban and Moffatt determined alcohol use from a variety of codes regarding police judgment of alcohol use or factors contributing to the crashes. When alcohol was not listed as a factor, alcohol use was coded as negative.
- b. It is incorrect to assume that all of the drivers not tested were alcohol-free based on police not listing alcohol as a contributing factor to the crashes. According to Moskowitz et al. (1999), police most often cite breath odor in determining alcohol involvement in traffic offenses, but the ability to detect this odor is unreliable even under controlled laboratory conditions.
 - c. It is likely that reported alcohol use is spuriously related to injury outcome because more seriously injured people are more likely to undergo close examination. About half of the states included in the IIHS study mandate BAC testing of fatally injured drivers (NHTSA, 2004), creating inherent reporting bias because the likelihood of testing is correlated with injury outcome. Padmanaban and Moffatt do not report or account for this bias.
 - d. It is likely that factors such as crash severity, vehicle damage, and driver age and gender have some influence on whom police choose to test for alcohol as well as which crashes they judge to be influenced by alcohol. Previous research has found that driver age and gender affect which drivers at sobriety checkpoints are judged not drinking (Wells et al., 1997).
 - e. Although alcohol clearly increases crash likelihood, Padmanaban and Moffatt offer no explanation of how alcohol increases the likelihood of K/A injury, given that a crash already has occurred. Absent convincing evidence that alcohol increases the susceptibility of human tissue and bones to injury, the primary determinants of whether an injury occurs to alcohol-impaired or sober occupants are the forces experienced during the crash. It might be argued that sober drivers' rollover crashes would be more severe, and their injurious forces greater, than those of drinking drivers because more extreme circumstances would be required for the sober drivers to lose control of their vehicles or leave the road. But this argument leads to the opposite of the effect claimed by Padmanaban and Moffatt. Any empirical relationship to the contrary observed between alcohol and K/A injury likelihood is likely to be spurious and related to the absence of objective evidence of alcohol involvement after a crash has occurred.

4. Padmanaban and Moffatt's docket submission is based on unsound and inconsistent statistical treatment. It contains numerous misstatements and omissions that undermine its conclusions.

- a. They either misunderstand or misconstrue the fundamental concepts of statistical estimation and significance testing. The object of a study of roof strength is to obtain the best estimate permitted by the data. In this context, statistical significance is only a way of representing how often one expects to be wrong in concluding that the observed estimate is indicative of a real non-zero effect. Padmanaban and Moffatt claim that if the estimated effect of roof strength on injury risk is found to be "not significant, then the lives saved [by strengthening roofs] could just as well be

zero or negative” (pg. 2). This trivializes the process of statistical estimation in a way that is fundamentally misleading.

- i. It is misleading to treat any estimate with a p-value slightly above 0.05 as if it were drastically different from estimates with p-values slightly below 0.05. For example, among the effects estimated for reductions in the likelihood of driver death with increased roof strength, the p-value for SWR within 5 inches of crush was slightly greater than 0.06. This means that if one were to conclude that an effect this large is different from zero, one would expect to be wrong about 6 times out of 100 (a p-value of 0.05 would lower the error risk only slightly, to 5 times in 100). This 6 percent error risk also means that the likelihood of seeing effects as large as that estimated for roof strength when the true effect is zero or negative is only about 3 in 100. Padmanaban and Moffatt misrepresent the logic of statistical estimation and misconstrue the implications of significance testing.
 - ii. This illogical approach leads them to ignore the overwhelming consistency of the results of the IIHS study. Their docket submission suggests that a single IIHS estimate for injury risk reduction that was not significant at the 0.05 level contradicts and invalidates the overall finding that stronger roofs reduce injury risk. Of the 12 estimates for K/A injury risk related to roof strength measured in 4 different ways and at 3 different crush distances, all were significant at $p < 0.0001$. For the 12 estimates for K injury risk, 9 were significant at $p < 0.0001$, 2 at $p < 0.05$, and 1 at $p < 0.07$. Robustness of an empirical pattern when measured in different ways is much more important than the fact that 1 of 24 tests did not meet an arbitrary level of $p < 0.05$.
- b. The docket submission does not include sample sizes for any of Padmanaban and Moffatt’s 7 statistical models. In response to subsequent requests by IIHS, they indicated sample sizes ranging from 1,352 to 20,010. These details should have been included in the discussion of their statistical modeling, especially given their ill-advised reliance on levels of statistical significance for interpretation of results. For example, they emphasize that odds ratios in the IIHS study were not statistically significant for the subset of drivers that police coded as unbelted, asserting that this means roof strength is not beneficial for these occupants. However, these drivers account for only 10 percent of the total sample, limiting the power to detect statistically significant effects.
 - c. Padmanaban and Moffatt do not give parameter estimates for the predictors of injury risk they chose to include in their comment. Without these, it is unknown whether the effects being estimated by their models are consistent or realistic relative to some underlying reasonable theory. Subsequent IIHS inquiries produced some, but not all, of the parameter estimates (see item 5.a.i. below).
 - d. Padmanaban and Moffatt do not present p-values for their additional parameters in the model that looked at fatality risk, saying only that roof SWR was not significant at a p-value of 0.10. It is possible that some variables previously claimed to be major factors (alcohol, belt use, ejection status) in injury outcome were not significant in this model.

5. Padmanaban and Moffatt’s docket submission is based on questionable engineering judgment.

- a. They stress the importance of aspect ratio (height divided by track width) in previous research and criticize IIHS for excluding it. In their reproduction of the IIHS study, they find it statistically significant. This is problematic for 4 reasons:

- i. Based on data provided to IIHS, their models predict greater injury risk in SUVs with larger aspect ratios. This directly contradicts their previous studies, which reported decreased injury risk for vehicles with larger aspect ratios. Padmanaban and Moffatt do not explain or even disclose this fact in their submission to NHTSA.
 - ii. They do not offer a hypothesis for how the shape of these SUVs, as defined by aspect ratio, would affect injury risk. This also is true of their previous research, although they have stated that it is unrelated to differences in headroom. If the small geometric differences between these midsize SUVs are important in the rollover crash dynamics, more meaningful measurements would include maximum vehicle width or vehicle width at the height of the roof.
 - iii. The range of aspect ratios given for these vehicles is very small. Height and track width vary by up to only about 2 inches.
 - iv. There is enough variation in the specified height and track width measurements between model years of several of the study vehicles to invalidate whatever data were used.
- b. Padmanaban and Moffatt do not seem to understand the IIHS motivation for including static stability factor (SSF) in the statistical models, stating that “the purpose of the IIHS study and of ours was to evaluate the likelihood of serious/fatal injuries given a rollover and not the likelihood of rollovers.” The IIHS study clearly explains why SSF may be correlated to crash severity: By definition, more stable vehicles require more severe events to cause them to roll over.
 - c. Padmanaban and Moffatt do not explain why vehicle weight should be included in two different places in their statistical models. They include it both as an independent variable and in the calculation of SWR.

6. Padmanaban and Moffatt misrepresent the IIHS study.

- a. They say they “agree [with the IIHS study] that SWR within 5 inches is the most useful and universally accepted roof strength metric,” but the IIHS study makes no such claim. Its calculations of lives saved use this metric simply because FMVSS 216 uses the same metric. SWR within 5 inches of plate displacement is 1 of 12 roof strength metrics IIHS evaluated, and several of the other metrics predict greater reductions in injury risk across the range of tested vehicles. Even with their problematic predictors, it is possible that Padmanaban and Moffatt would have found statistically significant results with different roof strength metrics.
- b. Padmanaban and Moffatt claim that the regression line in Figure 1 of the IIHS study is the “primary finding” and later in their submission to NHTSA dedicate much time to discussing this line. However, they separately state their understanding that the plot is included “solely to present a visual representation of their raw data. [IIHS does] not rely upon it in any way for their conclusions.” This second statement is correct, and it is disingenuous to criticize the statistical fit of a plot presented for visualization and understood to be uncorrected for known confounding factors.
- c. They claim IIHS used the estimate for the reduction of fatal and incapacitating injury in the lives-saved calculations because the fatality estimate alone was not statistically significant (see items 4.a.i. and 4.a.ii. above). However, the former estimate was used because it is based on more observations (of injuries) and therefore likely to be more accurate. For the other 11 roof strength

metrics, little variation was observed between effect estimates for K/A injury and for fatal injury, so the choice was well founded.

- d. Padmanaban and Moffatt say their analysis does not “differ significantly from [IIHS] raw data counts” but do not give any details. Responses to subsequent requests from IIHS indicate their analysis includes 2,807 fewer drivers overall and 100 more drivers with fatal or incapacitating injuries. These differences are not explained. Padmanaban and Moffatt fail to demonstrate that their data and analysis replicate the IIHS study before including additional predictor variables. If their initial analysis cannot replicate IIHS’s, then none of their subsequent claims are applicable to the current discussion.

7. Padmanaban and Moffatt’s docket submission and associated analysis cannot be fully evaluated due to the lack of detailed information about data sources, methods, and results.

In contrast, IIHS methods and findings are fully described in the study. IIHS staff further assisted JP Research in understanding the construction of the statistical models used in the study. All information necessary to reconstruct the IIHS study is available to the public.

- a. For some additional predictor variables, unexplained discrepancies exist between the data counts in the state files and the counts JP Research reported to IIHS. For example, JP Research reports that ejection status was known for all but 2,198 drivers, whereas IIHS observed that ejection status was coded as unknown or completely missing for 8,713 drivers in the state data files. It would be useful to know how JP Research obtained the ejection status for their analyses.
- b. The docket submission includes statements about the methods used in their two previous studies that were not disclosed in that research. For example, the submission claims that both earlier studies controlled for ejection and rural/urban land use, but their 2005 study mentions neither among the factors included in the logistic regression models. The docket comment says “all our previous models also controlled for states, though it was not explicitly stated in the reports” (pg. 3). It is impossible to judge the credibility of any study when important details are omitted about how the research was conducted.
- c. Padmanaban and Moffatt report access to the results of other roof strength tests of the IIHS study vehicles that differ substantially from the IIHS results. These other results are not public, so it is impossible to determine their relevance. Previous research by Padmanaban and Moffatt included confidential tests conducted by vehicle manufacturers on non-production vehicles (Moffatt and Padmanaban, 1995; Padmanaban et al., 2005), and we do not know the nature of any additional test data on IIHS study vehicles.
- d. As detailed above, Padmanaban and Moffatt exclude several important facts that were revealed to IIHS only after follow-up inquiries to JP Research (see items 3.a., 4.b., 4.c., 5.a.i., 6.d., and 7.a.).

Appendix A – Relationship between roof strength and ejection risk

To address Padmanaban and Moffatt’s claim that ejection is “likely to be unrelated to roof strength,” IIHS conducted a logistic regression analysis of ejection likelihood based on roof strength. Vehicle and crash data were the same as in IIHS’s analysis of vehicle roof strength and injury risk (Brumbelow et al., 2008). Figure 1 shows the relationship in the raw data between peak roof SWR within 5 inches of plate displacement and ejection rate before adjusting for any potentially confounding factors. Of 22,817 rollover crashes of study vehicles, police coded 13,086 drivers as not ejected, 1,018 as fully or partially ejected, and the rest were coded as unknown or had missing values. Only the drivers with known ejection status were included in this analysis. Table 1 presents results of the logistic regression model controlling for the effects of state, driver age, and vehicle SSF. For a 1-unit increase in peak SWR, ejection risk was reduced 32 percent. For each 10-year increase in driver age, there was an 11 percent decrease in ejection risk. Both of these results are statistically significant at the 0.05 level. An increase in SSF of 0.1 was predicted to increase ejection risk by 4 percent, but this result was not statistically significant at the 0.05 level.

Figure 1 – Rates of full or partial driver ejection by peak SWR within 5 inches of plate displacement

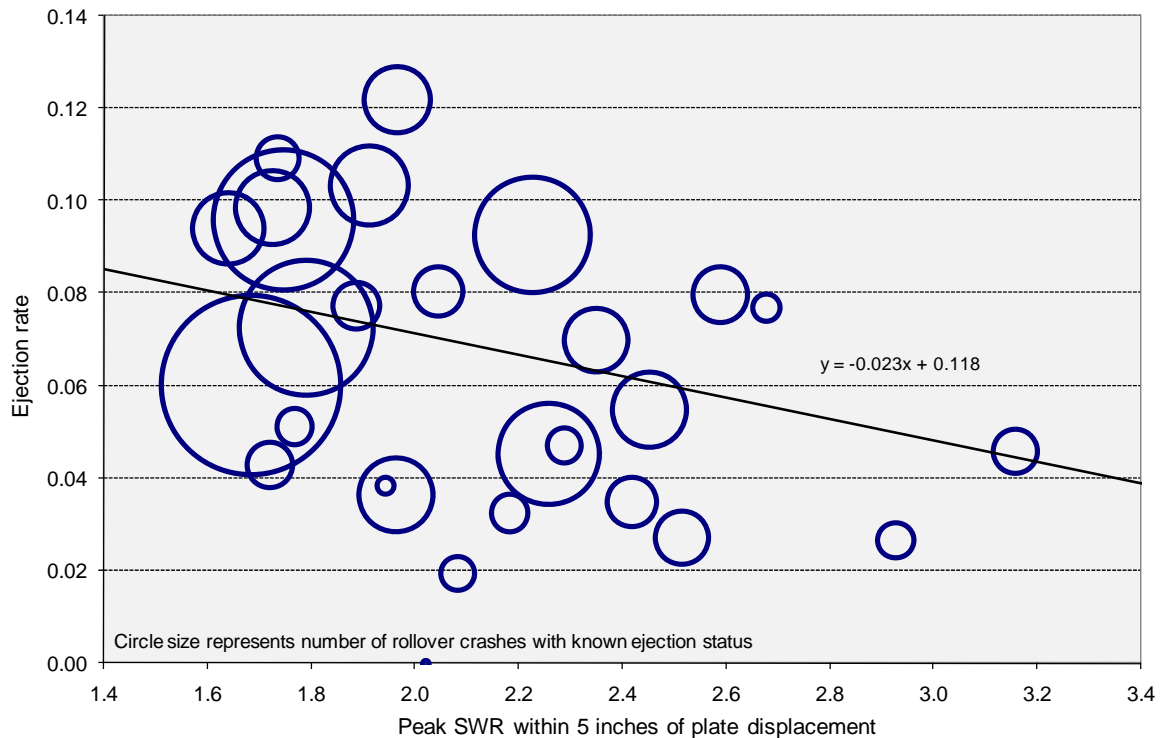


Table 1 – Results of logistic regression model for risk of ejection

Parameter	Odds ratio
Roof SWR within 5 inches (1-unit increase)	0.68*
Driver age (10-year increase)	0.89*
SSF (0.1-unit increase)	1.04

*Statistically significant at 0.05 level

Appendix B – Relationship between roof strength and injury risk for drivers coded as not ejected

The logistic regression model described in Appendix A demonstrates that reducing the risk of driver ejection is one benefit of stronger roofs. Also of interest is how stronger roofs benefit drivers who remain inside a vehicle during a rollover crash. Police coded 13,086 drivers in the IIHS study as not ejected. Figure 2 shows the relationship between the rate of fatal or incapacitating injury among the nonejected drivers and the peak roof SWR measured within 5 inches of plate displacement for each of the vehicles. The figure plots the raw data before adjusting for any confounding factors. Controlling for state effects, SSF, and driver age, a logistic regression model estimated a 27 percent reduction in the risk of fatal or incapacitating driver injury for a 1-unit increase in peak SWR within 5 inches of plate displacement. Nearly identical to the risk reduction estimated for all drivers in the IIHS study (see Table 2), this result is not surprising because nonejected drivers represent 93 percent of all drivers with known ejection status. A 10-year increase in driver age was predicted to increase the risk of K/A injury by 18 percent. A 0.1-unit increase in SSF was associated with a 6 percent increase in K/A injury risk. The odds ratios for SWR and driver age were significant at the 0.05 level, but the odds ratio for SSF was not.

Figure 2 – Rates of fatal or incapacitating driver injury by peak SWR within 5 inches of plate displacement

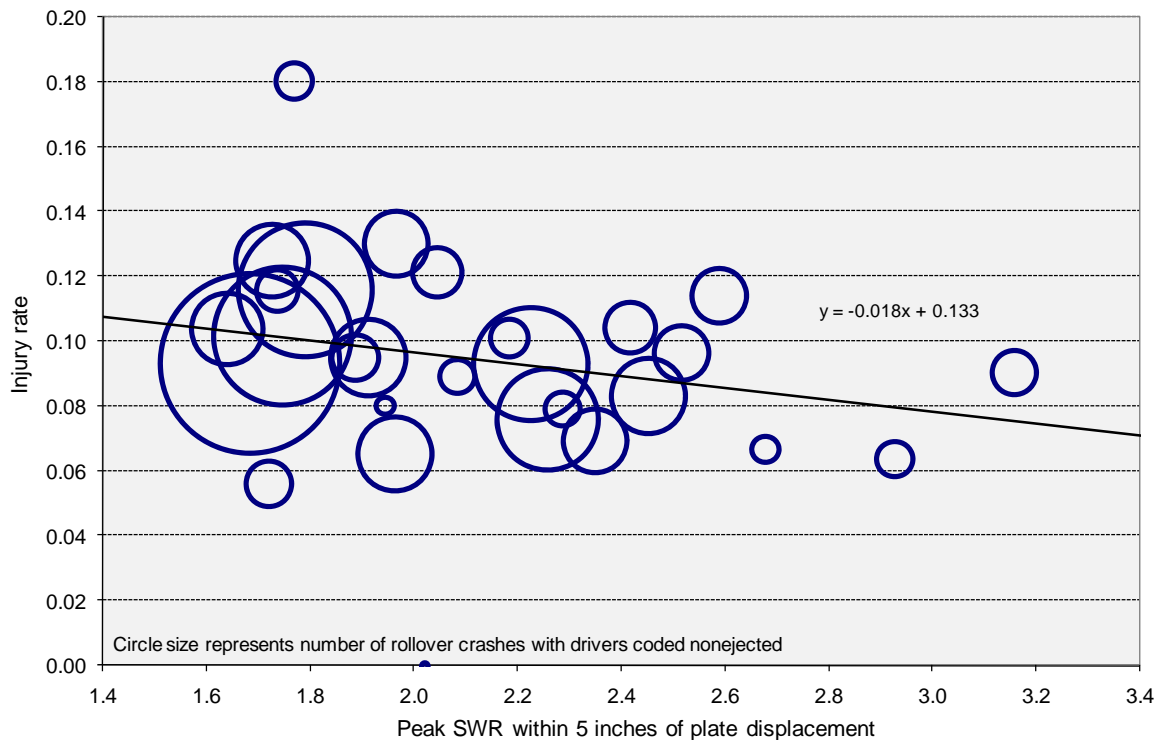


Table 2 – Results of logistic regression model for risk of fatal or incapacitating injuries for drivers coded as nonejected by police and for all drivers

Parameter	Odds ratio for drivers coded as nonejected	Odds ratio for all drivers
Roof SWR within 5 inches (1-unit increase)	0.73*	0.72*
Driver age (10-year increase)	1.18*	1.12*
SSF (0.1-unit increase)	1.06	0.96

*Statistically significant at 0.05 level

Nicole Nason
May 13, 2008
References

References

Brumbelow, M.L.; Teoh, E.R.; Zuby, D.S.; and McCartt, A.T. 2008. Roof strength and injury risk in rollover crashes. Arlington, VA: Insurance Institute for Highway Safety.

Insurance Institute for Highway Safety. 2008. Comment to the National Highway Traffic Safety Administration concerning proposed changes to Federal Motor Vehicle Safety Standard 216, Roof Crush Resistance; Docket No. NHTSA-2008-0015, March 27. Arlington, VA.

Moffatt, E.A. and Padmanaban, J. 1995. The relationship between vehicle roof strength and occupant injury in rollover crash data. *Proceedings of the 39th Annual Conference of the Association for the Advancement of Automotive Medicine*, 245-67. Des Plaines, IL: Association for the Advancement of Automotive Medicine.

Moskowitz, H.; Burns, M.; and Ferguson, S. 1999. Police officers' detection of breath odors from alcohol ingestion. *Accident Analysis and Prevention* 31:175-80.

National Highway Traffic Safety Administration. 2004. State laws and practices for BAC testing and reporting drivers involved in fatal crashes. Report no. DOT HS-809-756. Washington, DC: US Department of Transportation.

National Highway Traffic Safety Administration. 2005. Connecticut's day and night safety belt use. Report no. DOT HS-809-954. Washington, DC: US Department of Transportation.

National Highway Traffic Safety Administration. 2007. Daytime and nighttime seat belt use at selected sites in New Mexico. Report no. DOT HS-810-705. Washington, DC: US Department of Transportation.

Padmanaban, J. and Moffatt, E.A. 2008. Comment to the National Highway Traffic Safety Administration concerning proposed changes to Federal Motor Vehicle Safety Standard 216, Roof Crush Resistance; Docket No. NHTSA-2008-0015, March 27. Mountain View, CA.

Padmanaban, J.; Moffatt, E. A.; and Marth, D.R. 2005. Factors influencing the likelihood of fatality and serious/fatal injury in single-vehicle rollover crashes. SAE Technical Paper Series 2005-01-0944. Warrendale, PA: Society of Automotive Engineers.

Schiff, M.A. and Cummings, P. 2004. Comparison of reporting of seat belt use by police and crash investigators: variation in agreement by injury severity. *Accident Analysis and Prevention* 36:961-65.

Wells, J.K.; Greene, M.A.; Foss, R.D.; Ferguson, S.A.; and Williams, A.F. 1997. Drinking drivers missed at sobriety checkpoints. *Journal of Studies on Alcohol* 58:513-17.