The Methodology for Predicting Rollover Risk

Ellen S. Hertz, Ph. D.
National Highway Traffic Safety Administration, Washington DC
ehertz@nhtsa.dot.gov

Abstract

Rollover risk for passenger vehicles is studied using data from six states. The metric that is used is the number of rollovers per single vehicle accident. The analysis is performed with logistic regression and relevant factors such as weather and driver age are controlled for in the model. The use of categorical analysis permits the data to be used on the observation level, preventing the loss of information that occurs when aggregated data, such as an average rollover rate, are used. The static stability factor (SSF) is the ratio of one half the track width to the height of the center of gravity. The crashes from the state data in the analysis are restricted to those involving make/models for which the SSF has been measured, but this is a large share of the fleet. SSF turns out to be an excellent predictor of rollover rate. The scores of some dynamic tests are also considered, along with SSF, and these turn out to have some association with rollover but not as much as SSF. Predicted rollover rates are compared with actual rates. Also, hypothetical rollover rates for an "average" crash are computed as a function of SSF and performance in a dynamic maneuver test. The model that uses static stability factor in the form log(SSF-0.9) gives the best predictions, especially in the low SSF range.

Key Words: Traffic safety, static stability factor, logistic regression.

Introduction and Summary

While the number of highway fatalities has remained relatively stable over the past decade, the number of rollover deaths has increased substantially. In 2001, 10,138 people died in rollover crashes, a figure that represents 32 percent of the traffic fatalities for the year. In response to this trend, NHTSA has been evaluating rollover testing since 1993. In this study, data were used from Maryland, Florida, North Carolina, Missouri, Utah and Pennsylvania because they maintain electronic records that are available to NHTSA with vehicle identification numbers (VINs) from which it is possible to determine the make/model of each vehicle and with coding that makes it possible to determine whether or not a vehicle rolled over. Data from PA were from the years 1994-1997, data from NC were from 1994-1999, data from MD, MO and UT from 1994-2000 and data from FL from 1994-2001.

The static stability factor (SSF) is the ratio of one half the track width to the height of the center of gravity. The analysis is confined to single vehicle accidents involving vehicles for which the SSF is known. There are one hundred such vehicle groups and they capture a large share of the fleet. Additionally, four dynamic maneuver tests were performed on twenty-five of these vehicle groups and analysis is performed that uses the results of these dynamic tests. (See the Federal Register notice at 66 FR 3388 for a listing of the hundred vehicle groups and the SSF values, and see the Federal Register notice at 68 FR 59250 for a listing of the twenty-five vehicle group subset and details on the dynamic maneuver tests.) The method is logistic regression. An observation is a single vehicle crash in the state database and a positive response is rollover. The model includes driver gender, driver age (as three categories - under 25, over 70 or neither) and evidence of drug or alcohol use. It also includes weather, speed limit, curve, hill, darkness and bad surface conditions.

The dynamic tests are J-turn with a light load (JL), J-turn with a heavy load (JH), fishhook with a light load (FL) and fishhook with a heavy load (FH). These provide four 0/1 scores, each corresponding to the occurrence of wheel lift during the test. In each case, the heavy load, usually corresponding to five occupants, is a more stringent test than the light load, which corresponds to two occupants. For each load, the fishhook, which involves two steering inputs, is a more stringent test than the J-turn which involves one. That is, vehicles that trip up with J-turn at a given loading will trip with fishhook at the same load but not necessarily vice versa. Also, vehicles that trip up with a given maneuver, J-turn or fishhook with light load will do so with a heavy load.

Analyses and Results

The static stability factor (SSF) was available for 100 vehicle groups, that is, make/model/model year ranges. For 25 of these, the four dynamic tests were performed. For each of these vehicle groups , SSF had been measured and all four dynamic tests were performed. Seven of these tipped up on J-turn heavy, 3 on J-turn light, 11 on fishhook heavy and 6 on fishhook light. Analyses that involved only SSF used crashes involving any of the 100 vehicle groups; analyses with dynamic tests as covariates, of course, were limited to crashes of the twenty-five vehicles for which we had scores. The first model used SSF as the only vehicle metric. That is,

Logit (Pr(Rollover))=
SSF STORM FAST HILL CURVE BADSURF MALE YOUNG OLD DRINK
DUMMYFL DUMMYMD DUMMYNC DUMMYPA DUMMYUT

The variables DUMMY<state> represent the change in Logit(Pr(Rollover)) due to the crash's taking place in that state as compared to an otherwise similar crash in MO. They are included to control for differences in traffic patterns and reporting practices that effect rollover rates between the states. For each value of SSF, one can compute the proportion of rollovers predicted by the model by summing the predicted probabilities of rollover for all of the cases with that SSF and dividing by the number of cases with that SSF. Since the model controls for the types of driver and terrain under which the crash took place, this predicted probability of rollover is not monotonically decreasing with SSF.

Figure 1 displays these predicted proportions and compares them to the actual rollover rates by SSF. It should be mentioned that since more than one vehicle group can have the same SSF, some of the dots represent more than one vehicle group. Figure 1 shows that the predictive power of the model with SSF only is fairly good in general, but is low for the low SSF vehicles. This is a serious liability because those are the vehicles most at risk of rollover.

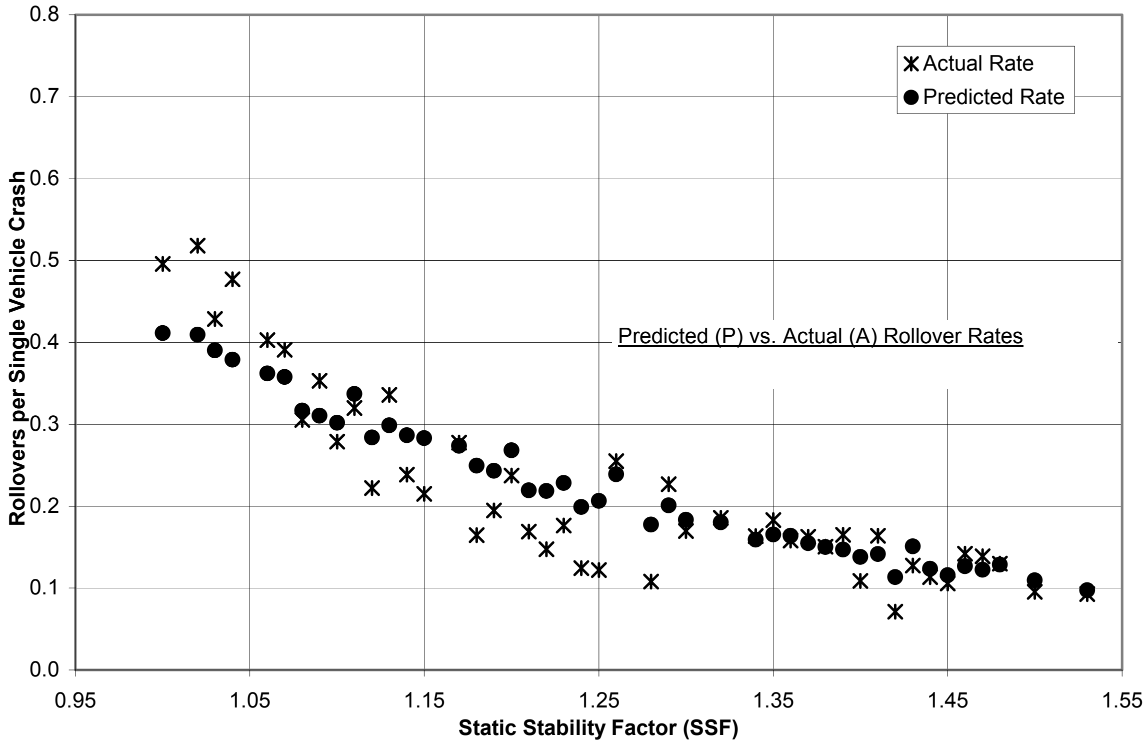

Figure 1: Logistic regression model operating on SSF (w/o transformation)
100 Vehicle Database – SSF Only

The second model also uses SSF as the only vehicle metric but in the form log(SSF). Figure 2 shows some improvement but the proportion of rollovers is still slightly underestimated in the low SSF range.

It was conjectured that since reductions in SSF make more difference in the low ranges, a model with log(SSF-margin) would produce a better fit. The margin had to be big enough to make a difference but small enough to keep the argument of log far enough from zero. A margin of 0.9 was selected.
Figure 3 is a plot of actual versus predicted rollover rates for the model that uses SSF in the form log(SSF-0.9) The predictions are vastly improved for the lower values of SSF. They are also good for the upper ranges. There are some points in the mid SSF range that suggest that the expected rollover rate is affected by factors other than SSF and the covariates in the model.

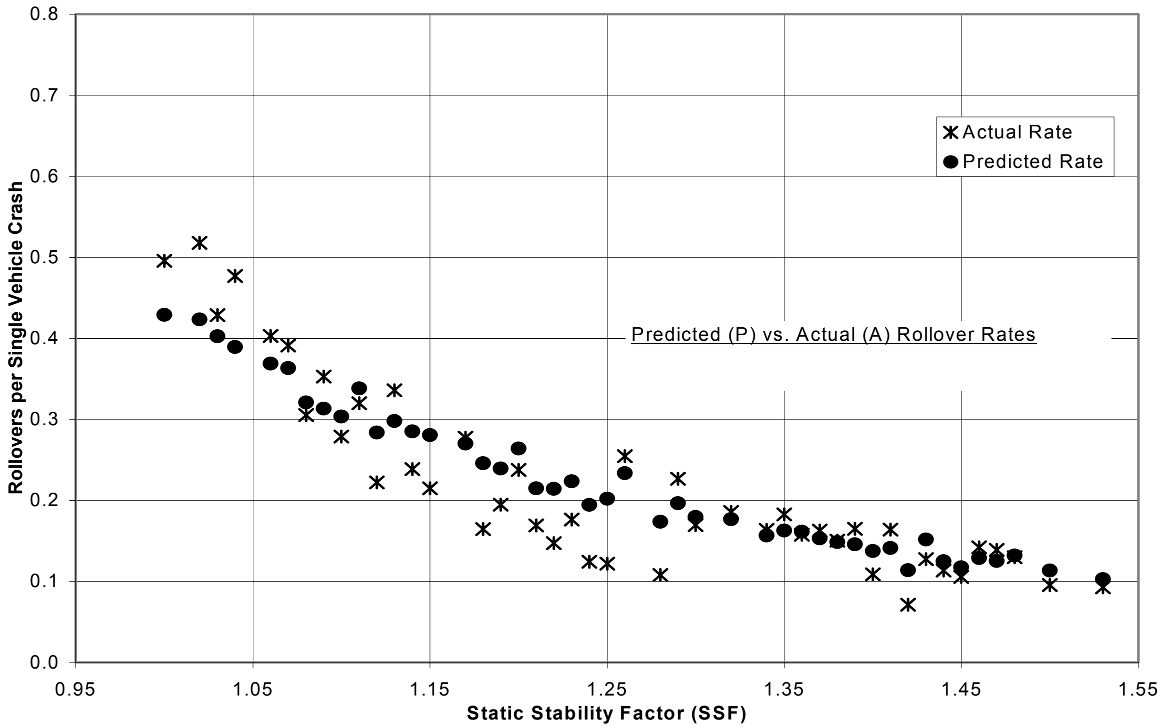

Figure 2: Logistic regression model operating on the LOG(SSF)
100 Vehicle Database – SSF Only

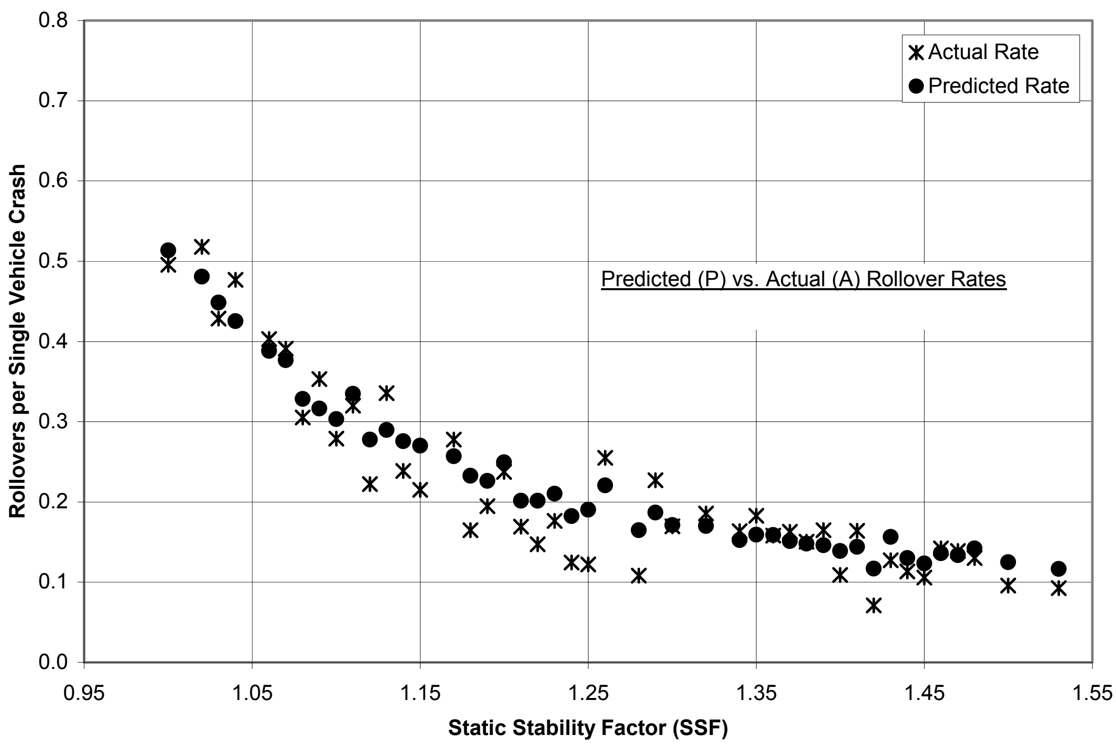

Figure 3: Logistic regression model operating on the LOG(SSF-0.9)
100 Vehicle Database – SSF Only

Analyses that included the dynamic tests in the model were of the form
Logit (Pr(Rollover)) =

Log(SSF-0.9) FH STORM FAST HILL CURVE BADSURF MALE YOUNG OLD DRINK DUMMYFL DUMMYMD DUMMYNC DUMMYPA DUMMYUT

where FH is a dummy variable that is scored as a 1 if the vehicle experienced tip-up during the fishhook test with heavy loading, 0 otherwise. FL, JH and JL are analogous dummy variables for the fishhook test with light loading and the J-turn test with heavy and light loadings, respectively.

Models of this form were run for each of FH, FL, JH and JL. Models that included two or more or the dynamic tests did not turn out to be useful as the weaker of the two tests would either be non-significant or have a coefficient with a negative sign suggesting that tip up on the test was protective. In fact, while FH and JH have positive coefficients, FL and JL had negative ones even when alone. Since it is not possible that tipping up on any test is indicative of a lower rollover risk, it is most likely that too few vehicles tipped up on these tests and the results are driven by a small number of vehicle groups. Figure 4 is a graph of the predicted rollover rate for an average crash as a function of SSF and FH.

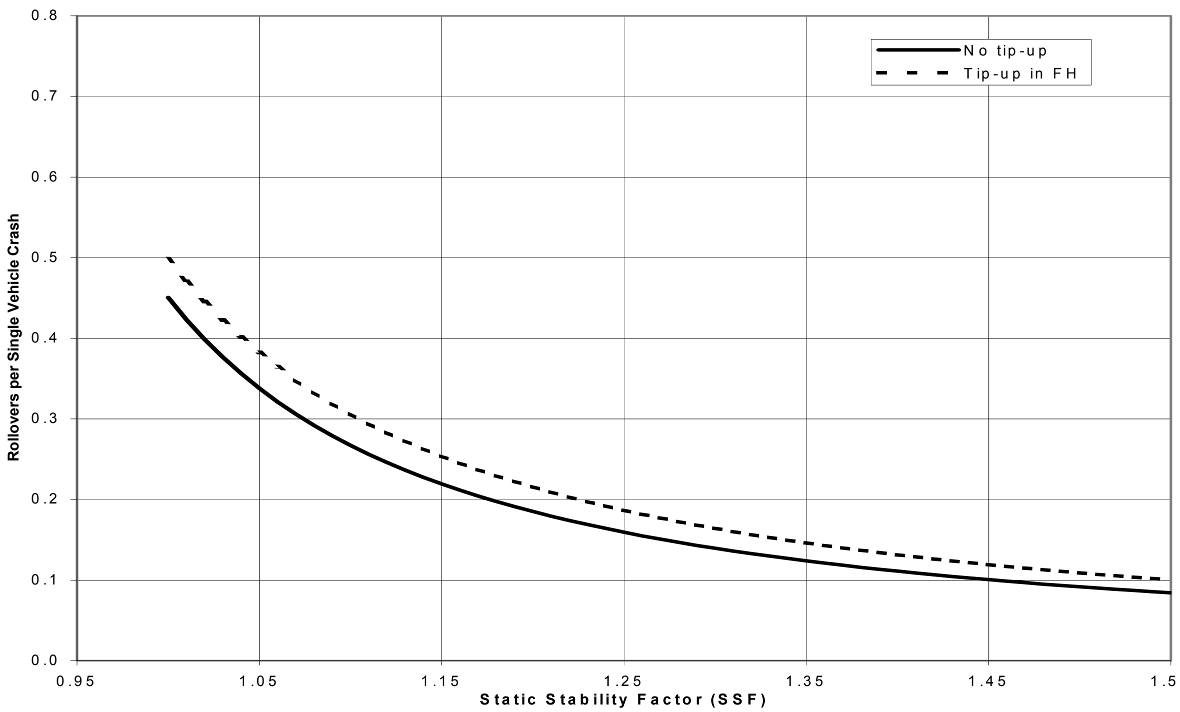

Figure 4: Model with Single Dynamic Variable – Fishhook Heavy

To see how much improvement in predictive power FH adds to the model, a plot displaying predicted roll rates and actual rates was created, where the predicted roll rates are based on the model with only log(SSF-0.9). This plot (designated by solid circles) is similar to that shown in Figure 3 except that the model is based on only the 25 vehicle groups for which dynamic tests were performed rather than the 100 vehicle groups used in Figure 3. This was to permit a fair comparison with the model that also used the FH dynamic maneuver variable that was calculated using the same 25 vehicle groups. The later model is the same one plotted in Figure 4 for average crash conditions, but the open circles in Figure 5 show it applied to the individual conditions of each crash involving vehicles in the twenty-five vehicle groups. Note that the SSF only model in Figure 5 is not the same as the predictions of the SSF plus FH model with FH = 0 (no tip-up in the test). Plots for the two models along with the plot of actual observed rollover rates for the twenty-five vehicle groups are displayed in Figure 5.

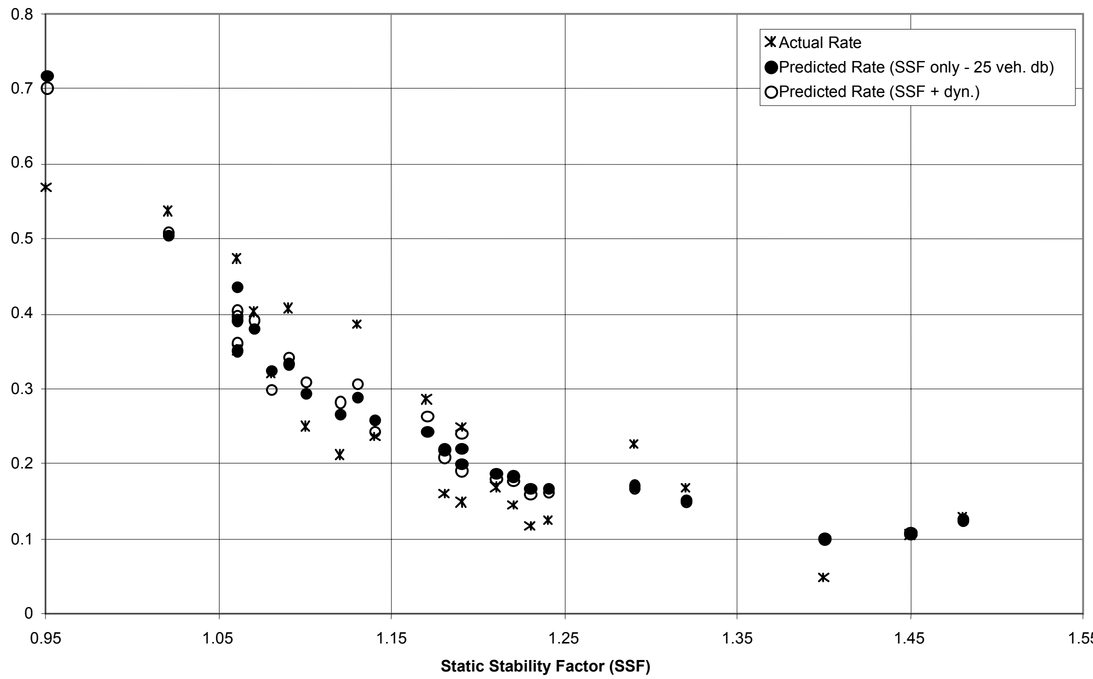

Figure 5: Predicted Versus Actual Rollover Rates for the 25VGs With Only Log(SSF-0.9) in the Model With and Without Fishhook Heavy

It should be mentioned that in the logistic regression with both log(SSF-0.9) and FH, log(SSF-0.9) had a chi square of over 1,848 and FH had a chi square of around 76.

Conclusion

Of the vehicle metrics that have been considered, static stability factor is, by far, the best single predictor of rollover risk. The model with both SSF and FH has the best predictive power, but the improvement in predictive power gained by the dynamic maneuver test variable is modest. This is reasonable outcome because the dynamic maneuver test measures the susceptibility of vehicles to rolling over as a result of tire traction forces on the pavement alone (on-road untripped rollover) rather than as a result of the higher tripping forces induced by striking a curb, soft soil, a guard rail or other tripping mechanism off the roadway. The dynamic maneuver test represents a relatively rare type of rollover crash but one that creates a high level of public interest.

References

The National Highway Traffic Safety Administration State Data System
The SAS® System
Hosmer and Lemeshow, Applied Logistic Regression
The Federal Register: 68 FR 59250 Final Rule Notice on Rollover Resistance Consumer Information