SEVERITY DISTRIBUTION FUNCTION FOR FREEWAY SEGMENTS

Srinivas Reddy Geedipally, Ph.D., P.E.¹
Assistant Research Engineer
Texas Transportation Institute
Texas A&M University System
Tel. (817) 462-0519
Email: srinivas-g@ttimail.tamu.edu

James A. Bonneson, Ph.D., P.E.
Senior Principal
Kittelson & Associates, Inc.
Tel. (979) 319-1886
Email: jbonneson@kittelson.com

Michael P. Pratt, P.E.
Assistant Research Engineer
Texas Transportation Institute
The Texas A&M University System
Tel. (979) 845-1907
Email: m-pratt@ttimail.tamu.edu

Dominique Lord, Ph.D., P.Eng.
Associate Professor
Zachry Department of Civil Engineering
Texas A&M University
Tel. (979) 458-3949
Email: d-lord@tamu.edu

Paper submitted to the 92nd Annual Meeting of the Transportation Research Board

Word Count: 5,492 + 2,000 (3 tables + 5 figures) = 7,492 words

November, 2012

¹ Corresponding author
ABSTRACT

To date, the focus of modeling efforts for freeway safety has been on developing safety prediction functions (SPFs) and crash modification factors (CMFs), with only limited consideration for crash severity distributions. As a result, relatively little is known about the safety effects of design elements, such as lane width, rumble strips, and longitudinal barriers on crash severity. In some cases, countermeasures are implemented with the intent to reduce fatal crashes, but the effect of these treatments on less severe crashes is not well understood.

Research was conducted to develop severity distribution functions (SDFs) to predict the proportion of crashes in each severity category as a function of roadway geometric design elements and traffic control features. The SDFs were calibrated using freeway segment data from California, Maine and Washington. The findings from this research show that barrier presence, increased traffic volume, increased lane width, and urban area type reduce the proportion of high-severity crashes. At the same time, rumble strip and horizontal curvature presence increase the proportion of high-severity crashes. These SDFs can be applied along with SPFs and CMFs to obtain more precise estimates of the safety effects of design decisions.

Keywords: Freeways, Crash Severity, Severity Distribution Function, Multinomial Logit Model.

INTRODUCTION

Up until now, the focus of modeling efforts for freeway safety has been on developing safety performance functions (SPFs) and crash modification factors (CMFs), with only a limited consideration for crash severity distributions. SPFs are intended to quantify the relationship between exposure and crash frequency, where exposure is represented by traffic count, segment length, and possibly other events that combine to form a chance set up for a crash (i.e., the necessary conditions). The crash frequency estimate from an SPF is intended to represent typical design and operating conditions. CMFs are intended to quantify the relative association between one or more design or operational elements and crash frequency. CMFs are used in a multiplicative manner with SPFs to adjust the estimate from the SPF to account for elements whose dimensions, conditions, or presence at the subject site are not consistent with those sites represented by the SPF. A table of severity distribution factors is often used with a safety prediction model to estimate the expected crash frequency for one or more severity categories. The factors are specified as average values for typical conditions. These values are not sensitive to the roadway geometric design or traffic control features.

This paper presents the modeling results of severity distribution functions (SDFs) for freeway segments. An SDF is represented by a discrete choice model that with independent variables that describe roadway geometric design elements and traffic control features. It is used to predict the proportion of crashes in each of the following severity categories: fatal (K), incapacitating injury (A), non-incapacitating injury (B), or possible injury (C). The property damage only (PDO) crashes were not considered in this analysis. The SDF can be used with the SPF to estimate the expected crash frequency for each severity category. The SDF includes various geometric and operational variables that allow the estimated proportion to be specific to an individual freeway segment. The SDF was developed using a highway safety database that
contained crash and road geometry data. It is envisioned that the SDF can be used with a safety prediction model to more accurately estimate the expected crash frequency by severity. Using these models, planners and engineers can better target and select countermeasures to address severe crash patterns. This information should lead to a better return on infrastructure investment and decision making.

BACKGROUND

Previous studies have documented how to estimate crash severity frequencies by jointly modeling the frequency and injury-severity of crashes using multivariate Poisson and mixed-Poisson models (for example 2). However, a recent study applied a two-stage frequency-severity model that linked an SPF with an SDF for identification of hotspots in England (3). The predictive method discussed in this paper applies an SDF with an SPF to estimate the expected crash frequency for each severity category. This approach is intended to minimize the frequency-severity indeterminacy problem described by Hauer (4). The SDF considers all severity categories together and can therefore predict the shift in crashes among categories due to a change in roadway conditions. The limitation with this method is that it uses only non-crash specific data (such as geometric, operation, and traffic data) and ignores post-crash data (such as driver behavior and environmental data). Using non-crash specific data result in a loss in the forecasting accuracy, and the statistical inferences that can be made with regard to the underlying injury-severity generating process (5).

Several statistical models are available to develop SDFs. The most common ones that have been used by transportation safety analysts include the ordered logit or probit, partially-ordered logit, ordered mixed logit, multinomial logit, nested logit, and random parameters (mixed) logit. The curious reader is referred to Savolainen et al. (6) for a thorough review of data and modeling issues as well as the availability of SDFs.

In this paper, the Multinomial Logit (MNL) model is used to predict the probability of crash severities. Due to the limitation of the MNL model, nested logit (NL) models are also used to evaluate the Independence from Irrelevant Alternatives (IIA) property. The IIA property of the MNL restricts the ratio of probabilities for any pair of crash severities to be independent of the existence and characteristics of other crash severities in the set of severities considered in the model. This restriction implies that the introduction of a new crash severity type in the set will affect all other severities proportionately. However, the nesting parameters /inclusive value parameter for the NL model was not significantly different from one. An inclusive value parameter equal to one indicates that there is no correlation in the unobserved factors within the nest, hence the NL model is not different than the MNL model. For an acceptable nesting structure, the inclusive values need to be between zero and one.

METHODOLOGY

This part of the paper describes the methodology used to calibrate the severity distribution functions. It is divided into two subsections. The first subsection gives the data collection process. The second subsection presents an overview of modeling approach.
Data Collection

The original research on which this paper is based used the data available through the Highway Safety Information System (HSIS) as the primary source of data for model calibration. The “HSIS” states California, Maine, and Washington were identified as including ramp volume data, which is of fundamental importance to all aspects of the original research project. These data were not available from the other HSIS states. Hence, the database assembly focused on these three states. The data collected include the width of road cross section elements, barrier presence and location, horizontal curvature, ramp configuration, ramp entrance location, median type, etc.

A total of 1,880 freeway segments are represented in the combined freeway segment database. These segments represent about 209, 101, and 200 miles of freeway in California, Maine, and Washington, respectively. The crash data considered in this study covered the time period 2005-2007 for California and Washington, and 2004-2006 for Maine. The characteristics of selected segments are provided in Table 17 of Bonneson et al. (7).

Modeling Approach

As discussed above, the MNL model was used to predict the probability of crash severities. An individual crash severity among the given severities was considered to be predicted if the crash severity likelihood function was maximum for that particular severity. Each crash severity likelihood function, which is a dimensionless measure of the crash likelihood, was considered to have a deterministic component and an error/random component. While the deterministic part is assumed to contain variables that can be measured, the random part corresponds to the unaccounted factors that impact injury severity. The deterministic part of the crash severity likelihood was designated as a linear function shown in the following equation:

\[
V_j = ASC_j + \sum_{k=1}^{K} b_{k,j} X_k
\]  

(1)

where,
- \(V_j\) = deterministic component of crash severity likelihood for severity \(j\) \((j = K, A, B)\);
- \(ASC_j\) = alternative specific constant for crash severity \(j\);
- \(b_{k,j}\) = regression coefficient for crash severity \(j\) and variable \(k\), \(k = 1, \ldots, K\);
- \(X_k\) = independent variable \(k\); and
- \(K\) = total number of independent variables included in the model.

The logit model was derived assuming that the error components are extreme value (or Gumbel) distributed and the probability for each crash severity is given by the following equation:

\[
P_j = \frac{e^{V_j}}{\sum_{j=1}^{J} e^{V_j}}
\]  

(2)
where, $P_j$ = probability of the occurrence of crash severity $j$; and $J$ = total number of crash severities to be modeled.

To adjust for local conditions, Equation 2 is modified by considering the local calibration factor. The adjusted probability for each severity category is given as:

\[
P_K = \frac{e^{v_K}}{1 + e^{v_K} + e^{v_A} + e^{v_B}} \quad (3)
\]
\[
P_A = \frac{e^{v_A}}{1 + e^{v_K} + e^{v_A} + e^{v_B}} \quad (4)
\]
\[
P_B = \frac{e^{v_B}}{1 + e^{v_K} + e^{v_A} + e^{v_B}} \quad (5)
\]
\[
P_C = 1 - (P_K + P_A + P_B) \quad (6)
\]

where, $P_K$ = probability of severity level K (fatal); $P_A$ = probability of severity level A (incapacitating injury); $P_B$ = probability of severity level B (non-incapacitating injury); $P_C$ = probability of severity level C (possible injury); and $C$ = local calibration factor.

The probability of each severity category is obtained by combining Equations 3 to 6 with Equation 1.

The SAS (8) non-linear mixed modeling procedure (NLMIXED) was used for model calibration.

**MODELING RESULTS**

This section describes the modeling results and discussion. It is divided into three subsections. The first subsection describes the calibration data. The second subsection presents the predicted probabilities and discussion of results and the last subsection provides concluding remarks.

**Calibration Data**

The database assembled for calibration included crash severity level as the dependent variable. Geometric design features, traffic control features, and traffic characteristics were included as independent variables.
Table 1 presents a brief summary of the variables used for SDF development. The variables listed were those found to have an important influence on the crash severity level. For the summary of other variables, the reader is referred to Bonneson et al. (7).

### TABLE 1. Summary statistics of data used for SDF development

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of Barrier</td>
<td></td>
<td>0.58</td>
<td>0.24</td>
<td>0</td>
<td>1</td>
<td>8249</td>
</tr>
<tr>
<td>Proportion of AADT During High-Volume Hours</td>
<td></td>
<td>0.49</td>
<td>0.35</td>
<td>0</td>
<td>0.93</td>
<td>8249</td>
</tr>
<tr>
<td>Proportion of Rumble Strips</td>
<td></td>
<td>0.2</td>
<td>0.38</td>
<td>0</td>
<td>1</td>
<td>8249</td>
</tr>
<tr>
<td>Proportion of Horizontal Curve</td>
<td></td>
<td>0.28</td>
<td>0.33</td>
<td>0</td>
<td>1</td>
<td>8249</td>
</tr>
<tr>
<td>Lane Width, ft</td>
<td>Rural</td>
<td>11.99</td>
<td>0.59</td>
<td>10.1</td>
<td>14.91</td>
<td>8249</td>
</tr>
<tr>
<td>Area</td>
<td>Urban</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>5832</td>
</tr>
<tr>
<td>Severity</td>
<td></td>
<td>K</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>171</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>445</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2550</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>5083</td>
</tr>
</tbody>
</table>

The following model form was used for the deterministic component of the SDF during the regression analysis.

\[
V_K = ASC_K + b_{bar,K} \times P_b + b_{hv,K} \times P_{hv} + b_{rs,K} \times P_r + b_{hc,K} \times P_c + b_{l,K} \times W_l + b_{rural,K} \times I_{rural}
\]  
(7)

\[
V_A = ASC_A + b_{bar,A} \times P_b + b_{hv,A} \times P_{hv} + b_{rs,A} \times P_r + b_{hc,A} \times P_c + b_{rural,A} \times I_{rural}
\]  
(8)

\[
V_B = ASC_B + b_{bar,B} \times P_b + b_{hv,B} \times P_{hv} + b_{rs,B} \times P_r + b_{hc,B} \times P_c + b_{l,B} \times W_l + b_{rural,B} \times I_{rural}
\]  
(9)

\[
C = e^{b_{ca,I_{ca}}}
\]  
(10)

where,
- \(P_b\) = proportion of segment length with an outside and inside barrier;
- \(P_{hv}\) = proportion of AADT during hours where volume exceeds 1000 veh/h/ln;
- \(P_r\) = proportion of segment length with rumble strips present on the inside and outside shoulders;
- \(P_c\) = proportion of the segment length with curvature;
- \(W_l\) = lane width, ft;
- \(I_{rural}\) = rural indicator variable (=1.0 if area is rural, 0.0 if area is rural);
- \(I_{ca}\) = California indicator variable (=1.0 if segment in California, 0.0 otherwise); and
- \(b_{k,j}\) = calibration coefficient for variable \(k\) and crash severity \(j\) (see Table 2).
The final form of the regression model is described here, before the discussion of regression analysis results. However, this form reflects the findings from several preliminary regression analyses where alternative model forms were examined. The form that is described represents that which provided the best fit to the data, while also having coefficient values that are logical and constructs that are theoretically defensible and properly bounded.

Table 2 summarizes the estimation results of MNL model. An examination of the coefficient values and their implication on the corresponding crash severity levels are documented in a subsequent section. In general, the sign and magnitude of the regression coefficients in Table 2 are logical and consistent with previous research findings.

The t-statistics indicate a test of the hypothesis that the coefficient value is equal to 0.0. Those t-statistics with an absolute value that is larger than 2.0 indicate that the hypothesis can be rejected with the probability of error in this conclusion being less than 0.05. For those few variables where the absolute value of the t-statistic is smaller than 2.0, it was decided that the variable was important to the model and its trend was found to be consistent with previous research findings (even if the specific value was not known with a great deal of certainty as applied to this database).

Indicator variables were included for the states of California and Maine. However, only the coefficient for California was statistically significant. The coefficient for this variable is shown in the last row of Table 2.

### Table 2. Parameter estimation for SDF

<table>
<thead>
<tr>
<th>Variable</th>
<th>Inferred Effect of...</th>
<th>Fatality (K)</th>
<th>Incapacitating injury (A)</th>
<th>Non-Incapacitating injury (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Value</td>
<td>t-statistic</td>
<td>Value</td>
</tr>
<tr>
<td>ASC</td>
<td>Alternative specific constant</td>
<td>-0.1705</td>
<td>-0.09</td>
<td>-2.3929</td>
</tr>
<tr>
<td>$b_{bar}$</td>
<td>Proportion of barrier</td>
<td>-0.3883</td>
<td>-1.06</td>
<td>-0.3253</td>
</tr>
<tr>
<td>$b_{hv}$</td>
<td>Proportion volume during high-volume hours</td>
<td>-0.9239</td>
<td>-3.03</td>
<td>-0.8528</td>
</tr>
<tr>
<td>$b_{rs}$</td>
<td>Proportion of rumble strips</td>
<td>0.3868</td>
<td>1.67</td>
<td>0.3906</td>
</tr>
<tr>
<td>$b_{hc}$</td>
<td>Proportion of horizontal curves</td>
<td>0.2079</td>
<td>0.88</td>
<td>0.2427</td>
</tr>
<tr>
<td>$b_{l}$</td>
<td>Lane width, ft</td>
<td>-0.2608</td>
<td>-1.72</td>
<td>--</td>
</tr>
<tr>
<td>$b_{rural}$</td>
<td>Added effect of rural area type</td>
<td>0.4919</td>
<td>2.47</td>
<td>0.4302</td>
</tr>
<tr>
<td>$b_{Ca}$</td>
<td>Added effect of California</td>
<td>0.349</td>
<td>6.45</td>
<td>0.349</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of observations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$AIC$</td>
<td>Akaike information criterion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$BIC$</td>
<td>Bayesian information criterion</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: possible injury is the base scenario with coefficients restricted at zero. – means highly insignificant
The coefficients in Table 2 were combined with Equations 7 to 9 to obtain the deterministic component of each crash severity for freeway crashes. The form of each model is:

\[ V_K = -0.1705 - 0.3883 \times P_b - 0.9239 \times P_{hv} + 0.3868 \times P_r + 0.2079 \times P_c - 0.2608 \times W_i + 0.4919 \times I_{rural} \]  
\[ (11) \]

\[ V_A = -2.3929 - 0.3253 \times P_b - 0.8528 \times P_{hv} + 0.3906 \times P_r + 0.2427 \times P_c + 0.4302 \times I_{rural} \]  
\[ (12) \]

\[ V_B = 0.0732 - 0.2499 \times P_b - 0.872 \times P_{hv} + 0.1347 \times P_r + 0.1312 \times P_c - 0.0464 \times W_i + 0.2079 \times I_{rural} \]  
\[ (13) \]

The probability of each severity category is obtained by combining Equations 3 to 6 with Equations 11 to 13.

**Predicted Probabilities**

**Barrier Presence**

Two variables that define the existence of barrier include the proportion of segment length with a barrier present in the median \( P_{ib} \), and the proportion of segment length with a barrier present on the roadside \( P_{ob} \). Initially, the proportion of median barrier and roadside barrier were considered separately during the model calibration. However, both variables showed similar effect and thus the average of those two variables \( P_b = \frac{P_{ib} + P_{ob}}{2} \) is represented in the final model. Barrier is defined herein to be any combination of cable barrier, concrete barrier, guardrail, or bridge rail.

The relationship between the proportion-of-segment-with-barrier variable and severity level is shown in Figure 1. The negative value of the associated coefficient (in Table 2) indicates that, as the proportion of barrier increases, the likelihood of severity levels K, A, and B decreases. The trends in Figure 1 indicate that the fatal crash percentage changes from 7.1 percent without a barrier, to 5.7 percent with a continuous barrier. A similar trend is shown for severity levels A and B.

The trends shown in Figure 1 are consistent with the findings of other researchers. For example, Bligh et al. (9) found that locations where barrier was not present had cross-median and other median-related crashes that were more likely to be of the K, A, or B severity, relative to locations with longitudinal barrier. Donnell and Mason (10) found that the installation of median barriers decreases the probability of fatal and injury crashes. In another study, Tarko et al. (11) noted that installing a concrete barrier decreased the most severe head-on crashes but increased the other non-severe crash types.
High-Volume

This variable indicates the proportion of AADT during hours where volume exceeds 1000 veh/h/ln. It is computed using the average hourly volume distribution associated with the subject segment. This distribution will typically be computed using the data obtained from the nearest continuous traffic counting station (on a freeway of similar character). The freeway station nearest to each segment in the study state database was identified and the hourly volume distribution for the average day acquired. This distribution was then used to compute the hourly traffic volume for each hour of the average day for each segment. The volume of 1,000 veh/h/ln corresponds to an average vehicle headway of 3.6 s. The average speed tends to drop as flow rates increase beyond 1,000 veh/h/ln. This trend suggests that drivers are reducing their speed to improve their comfort and safety as the headway gets shorter than 3.6 s. For more details about this variable, the reader is referred to Bonneson et al. (7). The variable is positively correlated with the volume to-capacity ratio experienced by the segment on an hourly basis.

The relationship between the proportion of AADT during high-volume hours and severity level is shown in Figure 2. The negative value of the associated coefficient (in Table 2) indicates that, as the proportion of high-volume hours increases, the likelihood of severity levels K, A, and B decreases. The trends in Figure 2 indicate that the fatal crash percentage changes from 7.1 percent when the proportion equals 0.0, to 4.4 percent when the proportion equals 1.0. A similar trend is shown for severity levels A and B.

It is rationalized that the trends in Figure 2 are due to the correlation between the proportion variable and speed. During high-volume conditions, the running speed of the vehicles decreases and, thus, the chance to be involved in a severe crash decreases. Martin (12) investigated the relationship of traffic flow and crash severity on French motorways and concluded that lower traffic flow is associated with an increase in crash severity. In a study of
London highways, Noland and Quddus (13) found that congestion is less likely to be associated with severe crashes in urban conditions. The authors stated that the mobility benefits of reduced traffic congestion might be offset by the occurrence of more severe crashes. In another study, Quddus et al. (14) explored the relationship between crash severity and congestion. They found that the level of traffic congestion does not affect crash severity, which contradicts the 2005 study.

![Figure 2. Freeway severity distribution based on the proportion of AADT during high-volume hours.](image)

**Rumble Strips**

The presence of shoulder rumble strips was found to have some association with the crash severity distribution. This presence is quantified as the proportion of the segment with rumble strips. Two variables that represent the presence of rumble strips are \( P_{ir} \) and \( P_{or} \). The proportion \( P_{ir} \) represents the proportion of the segment length with rumble strips present on the inside shoulders. It is computed by summing the length of roadway with rumble strips on the inside shoulder in both travel directions and dividing by twice the segment length. The proportion \( P_{or} \) represents the proportion of the segment length with rumble strips present on the outside shoulders. It is computed by summing the length of roadway with rumble strips on the outside shoulder in both travel directions and dividing by twice the segment length. An initial regression model contained both the inside and outside rumble strip proportions. However, the two variables yielded a similar relationship with crash severity level and thus the average of these two variables (\( P_r = \frac{P_{ir} + P_{or}}{2} \)) is represented in the final model.
The relationship between the proportion-of-segment-with-rumble-strips variable and severity level is shown in Figure 3. The positive value of the associated coefficient (in Table 2) indicates that, as the proportion of rumble strips increases, the likelihood of severity levels K, A, and B also increases. It is important to note that this variable is significant at 5% level for severity level A but at 10% level for severity levels K and B. The trends in Figure 3 indicate that the fatal crash percentage changes from 7.1 percent when the proportion equals 0.0, to 9.2 percent when the proportion equals 1.0. A similar trend is shown for severity levels A and B.

The trends shown in Figure 3 are consistent with the findings of other researchers. For example, Griffith (15) found this trend in his before-after study of the effect of continuous shoulder rumble strips on freeways. He concluded that, for impaired drivers, there could be a possible transfer from single-vehicle run-off-the-road crashes to multiple-vehicle crashes, the latter crash type being more severe. Hu and Donnell (16) concluded that an increase in severity could be caused by over-steering, which results in the driver leaving the traveled-way in order to avoid possible multi-vehicle rear-end or sideswipe crashes with vehicles in adjacent lanes. It is possible that some drivers who encounter rumble strips on curves may respond by over steering.

Marvin and Clark (17) conducted a before-after study to evaluate the safety effectiveness of shoulder rumble strips. They concluded that, in certain situations related to terrain or demographic of driver population, rumble strips increased the severity of rollover crashes. They hypothesized that the increase could be through rumble strip deployment or other undefined factors. Smith and Ivan (18) speculated that an increase in the proportion of multiple-vehicle crashes is possible with the addition of rumble strips. They stated that a possible reason for the increase in the multiple-vehicle crashes is because of a driver hitting the rumble strip and subsequently panicking, which causes the driver to swerve and hit another vehicle. In contrast, Sayed et al. (19) reported a reduction in crash severity after the installation of shoulder rumble strips.

The association of rumble strip treatment with increased crash severity could be the result of confounding variables. That is, it is possible that rumble strips are being installed at locations where crash severity is relatively high. However, the application of rumble strips on freeways is a fairly routine practice among state transportation agencies, such that rumble strips tend to be installed as a matter of policy rather than as a problem-site treatment. Moreover, at least two of the studies cited in the previous paragraphs used a before-after design, which would control for some confounding elements.

**Horizontal Curve**

The “proportion of segment with horizontal curve” variable \( P_c \) is computed as the ratio of the length of all curves on the segment to the length of the segment. For example, consider a segment that is 0.5 mi long and has only one curve that is 0.2 mi long. If one-half of the curve is on the segment, then \( P_c = 0.20 (= 0.1/0.5) \). In fact, this proportion is the same regardless of the curve’s length (provided that it is 0.1 mi or longer and 0.1 mi of this curve is located on the segment). In addition to the curve proportion, different combinations of curve radius and length were considered during the model calibration. However, the proportion-of-segment-with-curve variable \( P_c \) was only found to significantly affect the crash severity.
Figure 3. Freeway severity distribution based on the proportion of segment with rumble strips.

The relationship between the proportion of segment with horizontal curve and severity level is shown in Figure 4. The positive value of the associated coefficient (in Table 2) indicates that, as the proportion of the segment with horizontal curvature increases, the likelihood of severity levels K, A, and B also increases. The trends in Figure 4 indicate that the fatal crash percentage increases from 7.1 percent for a segment with no curve to 8.0 percent for a segment located fully on a horizontal curve. A similar trend is shown for severity levels A and B.

Hu and Donnell (16) found a similar result when analyzing median barrier crashes on curved and uncurved freeway segments. They concluded that the trend is likely due to the fact that vehicles on horizontal curves impact longitudinal barriers at higher impact angles, relative to those on tangent segments. Shankar et al. (20) found that road sections with frequent horizontal curves tend to have a larger proportion of injury crashes, relative to road sections with infrequent curves. Abdel-Aty (21) also found that roadway curves contribute to higher probability of injuries on roadway sections. Donnell and Mason (10) analyzed the effect of curved sections on crash severity. They found the likelihood that a crash is designated as fatal or injury is higher for curved sections than for tangent sections.
Figure 4. Freeway severity distribution based on the proportion of segment with horizontal curve.

Lane Width

The lane width used in this research is an average for all through lanes on the segment. Shoulder width was also considered during model calibration, but it was not found to have a correlation with crash severity.

The relationship between lane width and severity level is shown in Figure 5. The negative value of the associated coefficient (in Table 2) indicates that, as the lane width increases, the likelihood of severity levels K, A, and B decreases. The trends in Figure 5 indicate that the fatal crash percentage decreases from 5.7 percent at 10-ft lane width, to 2.3 percent for a 14-ft lane width. A similar trend is shown for severity levels A and B. Geedipally et al. (22) found that if a crash occurs on rural two-lane highways with a wider lane width, it is less likely to be classified as a rear-end collision. With wider lane widths, it is possible that drivers have more opportunity to avoid rear-end and head-on collisions, which in turn reduces the likelihood of high crash severity.

Area Type

The relationship between area type and crash severity level was also considered during model calibration. Previous studies have documented differences in crash severity between urban and rural roadways (23). It is generally recognized that probabilistic models should be developed separately for urban and rural crashes. However, separate models were not developed in this project due to the limited sample size.
The relationship between area type and severity level is shown in Table 3. The positive value of the associated coefficient (in Table 2) indicates that a crash in a rural area is likely to be more severe than a crash in an urban area, when all other variables are controlled. The percentages in Table 3 indicate that the fatal crash percentage on a rural freeway is 7.1 percent, and it is 5.1 percent on an urban freeway. A similar trend is shown for severity levels A and B.

The trend with area type in Table 3 can be attributed to higher operating speeds on rural freeways. A crash that occurs at higher speed typically has higher severity than a crash at lower speed.

**State**

In addition to the roadway variables, an indicator variable for each state was included in the calibrated model to account for differences between states that could not be explained with the other variables in the model. Indicator variables for the states of California and Maine were initially included in the regression model. However, the coefficient for Maine was very small and not statistically significant. This finding suggests that the state effect is very similar between Maine and Washington. The Maine indicator variable was removed as a result.

The coefficient for California is relatively large and statistically significant. Its positive sign indicates that a crash on freeways in California is more severe than a crash on freeways in Washington or Maine, when all other variables are controlled. This difference may be explained by different crash reporting practices, highway design practices (e.g., use of different roadside design features, etc.), terrain, weather for the various states, or other unknown factors. The severity distribution for California is compared with that of the other states in Table 3.
### TABLE 3. Freeway severity distribution based on area type and state

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Severity Level Distribution, percent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Fatal (K)</td>
</tr>
<tr>
<td>Area type</td>
<td>Rural</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>5.1</td>
</tr>
<tr>
<td>State</td>
<td>Calif.</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>6.1</td>
</tr>
</tbody>
</table>

### CONCLUSIONS

This paper has examined the severity distribution of crashes on freeway segments in three states. The findings from this research indicate that as the proportion of barrier increases, the likelihood of high severity crashes decreases. In other words, installing a barrier decreases the most severe crash types (such as head-on crashes) but it increases other less severe crash types (such as single-vehicle crashes). It is important to note that the individual coefficients in the model may not always accurately represent the incremental effects of the independent variables. This is because of the inherent limits in modeling, such as the possibility of confounded or correlated variables. However, the model as a whole can be used to accurately estimate the crash severity proportions.

The current national focus on the elimination of fatal crashes (i.e., “toward zero deaths”) and the recently reported decline in fatalities may be partially attributed to increased traffic safety efforts such as installing barriers, educating drivers, enforcing strict laws, etc. However, the results of this research indicate that the likelihood of fatal crashes also decreases with an increase in the proportion of hours each day that experience congestion (see Figure 2). Hence, some of the reported decline in fatalities may be partly due to increased congestion on the nation’s freeway system (24).

The results of this research indicate that as the proportion of rumble strips increases, the likelihood of fatal and injury crashes also increases. The main purpose of installing rumble strips is to reduce run-off-the-road crashes. However, the research findings suggest that a crash occurring on a freeway segment with rumble strips tends to be more severe, relative to a crash on a segment without rumble strips. One reason for this trend could be that the rumble strips trigger a shift from single-vehicle run-off-the-road crashes to multiple-vehicle crashes, the latter crash type being more severe (25). Additional research is needed to quantify the effect of rumble strips on the overall reduction in single-vehicle crashes and confirm the possible increase in multi-vehicle crashes.

The results also show that crashes are likely to be more severe on horizontal curves than on tangent segments. They also indicate that as lane width increases, the proportion of high severity crashes decreases. This trend is possibly due to the reduction in head-on and rear-ends collisions associated with the increase in lane width, which are typically severe. The research results indicate that a crash occurring in a rural area tends to more severe than a crash in urban area, which is likely due to the higher operating speeds in rural areas. In addition, a crash on
freeways in California tends to be more severe than a crash on freeways in Washington or Maine, when all other variables are statistically controlled. This difference may be explained by different crash reporting practices, highway design practices (e.g., use of different roadside design features, etc.), terrain, or weather for the various states.

REFERENCES


