The Marginal Impacts of Design, Traffic, Weather and Related Interactions on Roadside Crashes

VENKY N. SHANKAR [a]

SONGRIT CHAYANAN [b], SITTIPAN SITTIKARIYA [b], MING-BANG SHYU [b]

NAVEEN K. JUVVA [c]

JOHN C. MILTON [d]

[a] Assistant Professor
Department of Civil and Environmental Engineering
University of Washington, Box 352700
Seattle, WA 98195
Tel: (206) 616-1259
Fax: (206) 543-1543
vns@u.washington.edu

[b] Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Washington, Box 352700
Seattle, WA 98195
Tel: (206) 616-1259
songrit@u.washington.edu, sittipan@u.washington.edu, mbshyu@u.washington.edu

[C] Transportation Engineer
Parametrix, Inc.
5808 Lake Washington Blvd. N.E., Suite 200
Kirkland, WA 98033
Tel: (425) 822-8880
njuvva@parametrix.com

[d] Assistant Design Engineer
Washington State Department of Transportation
Maple Park East, Box 47329
Olympia, WA 98504
Tel: (360) 705-7299
miltonj@wsdot.wa.gov

Submitted for presentation and publication at the 83rd Annual Meeting of the National Transportation Research Board, Washington D.C., 2004.
Estimated Word Count: 4,143 text, 750 words in tables. 500 words in figures. Total 5,393 words.
December 13, 2003

ABSTRACT

A multivariate model that incorporates the effects of design, traffic, weather and related interactions with design variables on reported roadside crashes is presented in this paper. By providing for a framework that accounts for all measurable effects, the model minimizes the impact of omitted variable effects. Furthermore, the framework presented here accounts for partial observability effects that stem from fluctuations in environmental conditions as well as unobserved effects that contribute to heterogeneity in the traffic safety network. A sample of 318 one-mile sections is used for this study. These sections represent the state highway network in Washington State on the basis of environmental and road classification factors, and therefore were used for the collection of detailed precipitation, snowfall and temperature data in addition to roadway and roadside design and traffic parameters. The resulting model suggests that the marginal impact of weather is both in main effects and interactive form, suggesting that even after controlling for unobserved heterogeneity and partial observability, weather effects play a statistically significant role in roadside crash occurrence. In particular, it was found that in addition to precipitation, average monthly snowfall exceeding 4 inches as well as interactions between snow depths and horizontal curves were found to have a statistically significant effect on roadside crash frequency probabilities. The marginal effects of these variables were also statistically significant; furthermore, the contribution of weather and related interactions to the likelihood of roadside crash frequencies is approximately 19%, while design main effects contributed to 33% while traffic and design interactions contributed to 6%. Weather interactions with design contributed to approximately 6% of the overall likelihood. Traffic as a main effect contributed 36% to the overall roadside crash likelihood.

INTRODUCTION

Roadside crashes are crashes in which it was reported that the crash occurred off the roadway. By definition, this includes crashes occurring when vehicles leave the travel lane, encroach onto the shoulder and beyond, or hit one or more natural or man-made objects, such as utility poles, bridge walls, embankments, guardrails, parked vehicles, or trees. (1,2) In the United States every year, over 40,000 fatalities occur on the highway system, with over 44% of those fatalities occurring due to fixed-object and other roadside crashes. Roadside crashes account for nearly 25 % of all reported crashes, while the fatal consequences of those crashes are over-represented. No more than 0.7% of all reported crashes are fatal while nearly 1.3% of all roadside crashes are fatal. In Washington State for example, approximately 360 crashes are fatal every year on the state highway network. Of these crashes, 158 crashes were associated with the roadside. Furthermore, considering the need for mobility-related expansion and consequent constraints on adequate provision of the roadside as a forgiving environment, the importance of a safe roadside is heightened. From a state agency programming standpoint, significant efficiencies can be gained from a societal cost standpoint through the understanding of factors that statistically correlate with roadside crashes. In particular, understanding the role of weather and related interactions is useful since advance warning relating to significant effects in this regard from a geographical perspective could be designed to be more effective in preventing roadside crashes.

PREVIOUS RESEARCH

Previous crash research involved varied efforts to analyze different types of crashes for improved safety. In 1980, Haddon (3) developed a matrix of categories to assist researchers trying to systematically address injury prevention. The idea was to look at injuries in terms of causal factors and contributing factors, rather than just using a descriptive approach. The matrix divided these factors into human factors, vehicle factors, and environmental factors as shown in Table 1. However, significant attention has not been paid to the roadside crash frequency analysis from the standpoint of comprehensive multivariate frameworks that incorporate weather and related interactions with design and traffic variables. Table 1b presented an illustrative example of the importance of exposure and related impacts of key exposure-type factors on crash risk. Qualitative knowledge-based methodologies have been developed for selection and implementation of roadside safety improvements. Ramache (4) developed a system to experiment with different configurations of fixed roadside objects and select the configuration which minimizes the probability of collision. Most of the efforts have been focused on identifying the design improvements to avoid or reduce the severity of roadside crashes. Tracz and Gaca (5) developed a method of in-depth analysis in order to identify critical points, circumstances in which crashes occur and main, potential factors of road crash risk on the basis of computer crash data base. The research identified blackspots and crash circumstances, setting up together typical road design errors and characteristics of a roadside-favoring crash occurrence. These research efforts did not involve extensive statistical and econometric modeling of roadside crashes to identify the impacts of a wide range of roadway, roadside, traffic and weather factors.

Recent methods used for modeling crash frequencies such as the Poisson and negative binomial (NB) models (see for example 6-8) have been shown to be insightful econometric modeling tools for crash frequencies. Extensions to these models in the form of zero-inflated

Poisson (ZIP) and zero-inflated negative binomial (ZINB) models (9,10) have been tested and were found to account for the possibility of zero-inflated counting processes. They model the crash frequency as the sum of latent and observed effects affecting the driving environment. The failure to identify and model the two different states results in biased estimates as there will be an over-representation of zero-crash counts, a majority of which do not follow the assumed behavior of crash frequencies. In the context of run-off-roadway crashes, a count model approach in the form of ZINB and ZIP was employed by Lee and Mannering (11) to provide additional insight into the impact that roadside features have on the frequency and severity of run-off-roadway crashes. The study was a highly localized one that focused on limited corridors and hence was not representative of statewide driving networks in the state of Washington. In addition, the study did not examine possible interaction or correlation between the roadside crash frequencies and weather effects. Other studies (for example; 12-18) explored the run-off-roadway and fixed object crashes, but the chronic lack of extensive data has hindered development of detailed statistical models and depreciated the effectiveness of the safety recommendations.

Another dimension of roadside crash research was explored by Chayanan et al (19) which provided useful information as the starting point to this research. The research proposed the possibility of interaction between the roadside and roadway crash, by exploring the nature of simultaneity in the relationship between section roadway and roadside crash rates. The goal in this study was to understand the role of design factors in safety programming efficiency. The potential roadway-roadside correlation in a seemingly unrelated fashion due to unobserved effects common to both was modeled using seemingly unrelated regression estimation (SURE). The research detected that there was no significant efficiency gain obtained by employing the simultaneous regression estimation. In particular, it noted that there potentially exists a forward-recursive structure between the roadway and the roadside, with the roadside influencing crash frequencies on the roadway. Given that consideration, it is important to identify the specification for the roadside prior to embarking on a complete model of crash frequencies as a whole.

STUDY MOTIVATION AND METHODOLOGICAL APPROACH

The present research attempts to explore the applicability of zero-altered counting processes to roadside crash frequencies to provide better insights into the impact of roadway, roadside, traffic and weather factors on the roadside crash frequencies. In particular, the impact of environmental fluctuations on the occurrence of roadside crashes is of interest. In addition, when one considers that not all crashes are reported, the partial observability that results from reporting crashes in the survey period alone along with any associated heterogeneities that occur due to spatial and temporal effects makes a zero-altered probability approach appealing. In the event partial observability and unobserved heterogeneity issues become significant, distribution modeled as the product of two latent processes offers a plausible correlational approach. For example, if "Z" represents the zero-crash count state of the roadside, and "Y^*" denotes the crash count state for that roadside section, neither "Z" nor "Y^*" is observed, but only the observed crash count "Y", such that $Y=Z*Y^*$. Determining the latent components can then be viewed as a mixing distribution problem, with "Z" being modeled as a dichotomous probability and "Y^*" modeled as a count probability. In vehicular crash contexts, such distributions have been found to be appropriate (10). In particular these studies have highlighted the importance of roadway design deviations as a motivator for partial observability effects. The effect of such deviations has been found to, at the least, cause partial observability, and in certain design situations, overdispersion as well. The choice of a mixture distribution, however, is not self-apparent. Apriori assumptions regarding the density generators are required in formulating the mixture distribution. Formally, let Y_i be the annual number of roadside crashes reported for corridor *i*, and let p_i be the probability that corridor *i* will exist in the zero-crash state over its lifetime. Thus 1- p_i is the probability that corridor *i* actually follows a true count distribution in the non-zero state. For our immediate purposes, we assume that this count state follows a negative binomial distribution, considering the prospect of heterogeneity in the roadside context. Given this,

$$Y_i = 0 \text{ with probability } \mathrm{p_i} + (1-\mathrm{p_i})[\frac{1}{1+\alpha\lambda_i}]^{1/\alpha} \tag{1}$$

and,

$$\mathrm{Y_i} = \mathrm{k} \text{ with probability } (1-p_i)\frac{\Gamma(n_i+\theta)}{\Gamma(\theta)n_i!}\left(\frac{\theta}{\theta+\lambda_i}\right)^{\theta}\left(\frac{\lambda_i}{\theta+\lambda_i}\right)^{n_i} \tag{2}$$

where *k* is the number of accidents (positive numbers starting from one), with λ_i being the mean, and $\theta = 1/\alpha$, with α as the dispersion parameter. The parameter "k" here refers to a particular non-zero value of "n_i" drawn from the distribution for site "I" with a reported number of crashes "n_i." Note that the dispersion parameter, a, relaxes the Poisson assumption that requires the mean to be equal to the variance by letting Var[Y_i]= E[Y_i]{1 + α[Y_i]}

In equations 1 and 2, the probability of being in the zero-accident state p_i is formulated as a logistic distribution such that $\log\left(p_i/1-p_i\right)=\mathbf{G_i}\,\gamma$ and λ_i satisfies $\log(\lambda_i)=\mathbf{H}_i\,\beta$, where $\mathbf{G}_i$ and $\mathbf{H}_i$ are covariate vectors, and γ and β are coefficient vectors. The covariates that affect the mean λ_i of the Poisson state may or may not be the same as the covariates that affect the zero-accident state probability (i.e., p_i). Alternatively, vectors $\mathbf{G}_i$ and $\mathbf{H}_i$ may be related to each other by a single, real-value shaped parameter τ. In such a case, a natural parameterization is $\log(\mathrm{p_i}/1-\mathrm{p_i})=\tau \mathrm{H_i}\,\beta$. Theoretically, one must be confirm that this restriction is valid before assuming the "tau" parameterization. It is probable that some variables in $\mathbf{H}_i$ may actually be insignificant when estimated as separate parameters in the vector $\mathbf{G}_i$. Then, it is more useful to classify $\mathbf{G}_i$ as $\mathbf{B}_i$ where $\mathbf{B}_i$ differs from $\mathbf{H}_i$ in that some covariates that were significant in the count model (i.e. in the vector $\mathbf{H}_i$) may be excluded from the model determining the probability of the zero-accident state because they are insignificant. Thus vector $\mathbf{B}_i$ can be equal to or a subset of vector $\mathbf{H}_i$ as well as other variables not significant in $\mathbf{H}_i$. If τ is insignificantly different from zero, then the corridor is equally likely to be in the zero or non-zero lifetime state. Equations 1 and 2 combined provide the zero-inflated negative binomial (ZINB) model of roadside crash frequency. The formulations shown in equations 1 and 2 follow established methods in Greene (20).

In statistically validating the ZINB model, one has to distinguish between the base count model (such as the negative binomial model) from the zero-inflated probability model (such as the ZINB). A statistical test for this has been proposed by Vuong (21). The Vuong test is a t-statistic-based test with reasonable power in count-data applications (20). The Vuong statistic (V-statistic) is computed as

$$V = \frac{\overline{m}\sqrt{N}}{S_m} \tag{3}$$

where $\overline{m}$ is the mean with m = $\log\left[\frac{f_1(.)}{f_2(.)}\right]$, (with $f_1(.)$ being the density function of the ZINB distribution and $f_2(.)$ is the density function of the parent-negative binomial distribution), and S_m and N are the standard deviation and sample size respectively. The advantage of using the Vuong test is that the entire distribution is used for comparison of the means, as opposed to just the excess zero mass. A value greater than 1.96 (the 95 percent confidence level for the t-test) for the V-statistic favors the ZINB while a value less than -1.96 favors the parent-negative binomial (values in between 1.96 and -1.96 mean that the test is indecisive). The intuitive reasoning behind this test is that if the processes are statistically not different, the mean ratio of their densities should equal one. To carry out the test, both the parent and zero-inflated distributions need to be estimated and tested using a t-statistic. Studies (20) have shown that a Vuong statistic has reasonable power and hence is quite reliable.

EMPIRICAL SETTING

A random sample of one-mile sections from Washington State highway system was used to estimate the roadside crash frequency model. Crash data for the year 1995 was used to estimate the model. Sections with construction related crashes during 1995 were not included as accurate geometric and traffic data during the construction was not available. Highway sections that are routinely closed during the year were also not included. Sections with incomplete geometric or traffic data were also not included in the dataset.

The geometric and traffic data for each highway section was taken from the Washington State Department of Transportation highway geometric/traffic database. This database contains information for all the state highways divided into varying highway sections lengths with each section having homogenous characteristics. Since modeling varying section lengths could introduce unnecessary heteroskedasticity, this project uses one-mile sections. Due to data collection constraints, approximately 500 one-mile segments were sampled from the 7,000-mile system. Of these sections, 318 sections had complete roadway geometric, traffic and weather data. A sample target that ensured adequate representation of all functional classes, including interstates, principal, minor and collector arterials resulted in approximately 36 one-mile interstate sections, 110 principal arterial sections, 95 minor arterial sections and 77 collector arterial sections. The geometric and traffic data for the one-mile sections was aggregated using a weighted average from the section lengths listed in the database. Minimum, maximum, and weighted averages were recorded geometric data such as roadway widths, shoulder widths, traffic volumes, number of lanes, grades, curvature and speed limits.

The type of information available from this database is geometric data such as roadway widths, lane widths, number of lanes, shoulder widths, horizontal curve information, median widths, and barrier types, traffic data such as average annual daily traffic (AADT), truck volume as a percentage of AADT, and peak hour volumes as a percentage of AADT, and other data such as roadway classification, surfacing type, terrain, access restrictions and legal speed limit.

Historical weather data was collected for each one-mile roadway section. The data was taken from 240 National Oceanic and Atmospheric Administration (NOAA) weather stations located around Washington State. The location of each one-mile section was determined and matched to the nearest weather station. For some stations, data had been collected since the 1930's. In general, thirty-year histories were available for all stations. The data gathered was the average maximum monthly temperature, average minimum monthly temperature, average monthly total precipitation, average monthly snowfall, and average monthly snow depth for all twelve months and the annual averages for temperatures, precipitation, snowfall, and snow depth. It should be noted here that 1995 weather data did not show significant differences from the longer thirty-year historical averages. This offers an advantage from the standpoint of prediction. From a prediction standpoint, this is advantageous since the time dummies would otherwise drop out for future years.

Monthly weather data were too disaggregate to be of any help to the models. In particular, consecutive months had similar values for temperature and precipitation which showed up as significant in the models, but with parameters that are statistically indifferent from each other. For this reason, the data was aggregated to clarify the variability at the seasonal level. Maximum average temperature, minimum average temperature, total precipitation, total snowfall and total snow depth for winter (November – March), spring (April – June), summer (July and August) and autumn (September and October) were computed for possible specification in the crash frequency model. These combinations were chosen based on their similar values for each category, i.e., when variation within that period was minimal.

Table 2 shows descriptive statistics for key variables potentially affecting the roadside crash context. It should be noted in Table 2, statewide measures are provided for roadway classification and weather-related information to support our sample representativeness. The mean number of roadside crash counts for the 318 one-mile crash context was approximately 1.09 reported roadside crashes per year which is consistent with the equivalent roadside crash rate of 0.693 per million vehicle miles. Figure 1 shows the observed crash distribution for the sample. Approximately 50% of the sample reported zero roadside crashes. The maximum reported roadside crash count was ten. Statewide on the Washington State highway system, the roadside crash rate is approximately 0.69 crashes per million vehicle miles. Table 3 shows the distribution of roadside crash counts in the dataset.

MODEL ESTIMATION

At the early stage, estimated variables obtained from the previous study (18) were used as the starting variables. Estimation was carried out by standard maximum likelihood procedures using

Limdep 7.0, the econometric software. The negative binomial model was used as the starting point. Insignificant variables were rejected and new variables introduced to ensure that the best zero-inflated negative binomial model was achieved. The Vuong statistic value of 2.53 was greater than 1.96 (the 95% confidence level for the t-test) which implies that the zero-inflated model structure is preferable. The ZINB model was validated as the most plausible model, with a "tau" parameter of -22.34 and a p-value of 0.0134. The overdispersion parameter "alpha" is also significant. All key variables in the model are significant. Table 3 presents p-values. The "tau" variable is a parameter for computational simplicity. One can hypothesize three basic state variable effects for the zero state. A) a pure constant, B) a tau effect, and C) a separate vector of covariates. One will find in crash contexts that the constant effect is nothing but a "tau" effect in a different form. In addition, one will also find very few intuitive variables in a separate vector additional to that found in the count state. We present the "tau" effect to capture the basic essence of the roadside crash problem. And that is, when tau is insignificantly different from zero, no matter what the specification is (A, B, or C as explained above), the zero-state probability is near 50-50. That appears very intuitive. In our case, the behavior of "tau" is along the negative real line. As it approaches negative infinity, the zero state probability approaches zero. The Poisson or negative binomial do not capture this; in fact they significantly overstate the probability. And it does not resonate with locations where it is overstated. This finding strengthens the validation of the zero-inflated model, because it allows us to accommodate the less-than 50-percent probability of a zero state, which is entirely within the realm of reality for roadside crash contexts. This appears to be the case even more so, when weather effects AND interactions are taken into account. Roadside encroachments are far more likely in the modern context due to deviation effects, with narrower roadsides created by mobility-related widening improvements.

Key Variables

The results show that all the variables included in the model were highly significant and thus have a high level of confidence (exceeding 90%). It is also noted that the negative sign of the tau parameter indicates that the variables found to correlate with the count state as shown in table 3 correlate with the zero crash state in opposite sign. For example, per-lane average daily traffic positively correlates with roadside crash counts, while it negatively correlates with the probability of a section being a zero count state in its lifetime. The signs of the marginal effects and the coefficients are found to be the same and their significances were similar in value. The main effects of the variables can be summarized into three groups: design, traffic and weather variables. There are also two other groups: the interaction variables between the traffic and design, and interaction between weather and design factors.

Design factors positively correlating with roadside crash frequencies included interchange presence, median presence, shoulder widths less than 5 feet. In addition, horizontal curve sections including intersections were positively correlated with roadside crash frequencies. High-speed sections, (exceeding 45 miles per hour speed limits) with horizontal curvature positively correlated with roadside crash frequencies, as truck volume on those sections linearly increased as a percentage.

Geographic Effects

There are differences in classification effects. However, much of this difference is absorbed by the constant and the alpha parameter. This is intuitive. The constant picks up the mean effect of crash counts when the vector of variables is set to zero. The alpha parameter captures the heterogeneity due to spatial overdispersion occurring across functional classifications. As a result, when functional classes are introduced as "dummies" the nature of the model does not change. The model structure remained a ZINB specification in the presence of functional class dummies. Therefore, for simplicity, we leave the dummies out (in fact the dummies are not entirely separable) and present a single ZIP model for roadside crashes across functional classifications.

Marginal Contribution of Weather and Weather-related Interactions

The main effects relating to weather include those involving the impact of precipitation in autumn and spring months, as well as high snowfall months. Autumn precipitation was found to positively correlate with crash frequencies while spring month precipitation was found to negatively correlate with roadside crash frequencies. The interpretation of this finding was that precipitation during September through October increased the probability of roadside crash occurring, however, the probability reduced due to precipitation during April through June. A possible reason was that early runoff on roadways where gasoline, oil, and dust accumulated on the surface during summer season induced the chance of roadside crashes. Another reason combined with the runoff may be subjected to drivers' unfamiliarity in weather change. Drivers may still be familiar to summer driving since they drove at or higher than the posted speed limits. Precipitation in spring would not induce the similar effect to driver's behaviors; therefore, the probability of roadside crashes was lower. In addition, the combined precipitation variable of autumn and spring seasons provided insignificant impact to the crash count, confirming that these two indicators had different reasons to support the finding.

The finding on high snowfall indicator was found appears counter-intuitive; however when one examines the impact of high-snowfall exposure (13%) on the entire sample and the impact of snow as a visual deterrent and associated driving characteristics in those regions especially along I-90 and SR-2 in Washington state, the expectation is that roadside crash likelihoods are tempered. Furthermore, snow depth interactions with horizontal curvature are expected to reduce roadside crash frequencies as well. This finding is consistent with design treatments on locations with horizontal curvature and high snowfall. Advance warning signs that enhance driver awareness of adverse conditions temper roadside crash likelihoods.

Discussion of counter-intuitive effects

Autumn precipitation was found to be not of the same effect as Spring precipitation. Combining them would be an undue and incorrect parametric restriction. The correlation of variables with AADT is an interesting one. Surely, precipitation variables are correlated with AADT but not to the extent where inclusion of both effects cancels out the significance. In fact the marginal contribution of both AADT and weather are significant as shown in the paper. In addition, it must be noted that when a variable is omitted and is correlated with the included variables, the bias is fatal for model specification. The signing effect relating to weather was examined as well. Some adverse weather locations have advance warning signs, and some do not. However, the critical issue here is the impact of endogeneity of those locations with advance warning signs. That is, endogeneity as a result of increased probability of advance warning sign placement due to increased roadside crashes. When the precipitation variable is instrumented as a surrogate for advance warning sign placement, it was found that the parameter values are not that biased as would be expected due to endogeneity. There is movement in the parameter values however. But the effect is significant enough to categorize the weather variable as endogenous. This suggests that a partial effect may be present.

Marginal Contribution to Crash Likelihood

Interesting findings arose from the examination of design, traffic, weather and interactions to the likelihood of roadside crash frequencies. In particular, as noted in figure 2 and table 3, the marginal effects of these variables were also statistically significant. Marginal effects of variables are computed as the first partial derivate of the expected value of roadside crash count with respect to the particular variable. The presented value of the marginal effect for key variables in the model as shown in Table 3 is computed at the sample means. In addition to marginal effects, the marginal contribution of key variables to the overall roadside crash likelihood was also computed. Incrementally, each variable category (traffic, weather, design and interactions) were assessed for their marginal contribution to the likelihood value. Proceeding in this sense, it was found that the contribution of weather and related interactions to the likelihood of roadside crash frequencies was approximately 19%, while design main effects contributed to 33% while traffic and design interactions contributed to only 6%. Weather interactions with design contributed to approximately 6% of the overall likelihood. Traffic as a main effect contributed 36% to the overall roadside crash likelihood.

CONCLUSIONS AND RECOMMENDATIONS

This research has developed a count model for roadside crashes for Washington State highways. Three main ideas were tested in this research. The first area of interest in this research is the inclusion of historical weather data in the model specification. Temperature and precipitation values were combined for the four seasons since monthly values were shown to be too disaggregate. Weather variables were significant in both main effects and interaction form. Main effects relating to precipitation and high snowfall were statistically significant. Second, the impact of partial observability on modeling roadside crashes. The validation of the ZINB structure statistically indicates that partial observability is an issue in modeling roadsides, even after one accounts for unobserved heterogeneities. Unobserved heterogeneity in the form of overdispersion was also accounted for in the modeling framework and found to be significant. The third idea tested was the marginal contribution of weather-related effects to the likelihood of roadside crash frequencies. The contribution of weather to the understanding of roadside crash frequency serves as the good knowledge base for designing systems that can provide strategic direction for advance warning. Furthermore, design deviation situations such as those occurring at interchange and median locations emphasize the importance of design requirement enforcement to minimize roadside crash frequencies.

Additional work in the area of roadside inventory is needed as well. Minimal data is available regarding roadside conditions. As more detailed surveillance data on the roadside becomes available, it is likely that unobserved heterogeneity resulting from the roadside will be minimized. Such an effort is highly data-intensive, and in the absence of that information, the model presented here offers a portable framework that can be extended to the statewide system as a whole without additional data requirements. In particular, it is noted here that weather information statewide is readily available from NOAA weather networks and GIS-mapping algorithms that match weather and roadway locations increase model portability.

REFERENCES

1. Miaou, S.-P., Estimating Roadside Encroachment Rates with Combined Strength of Accidents and Encroachment-based Approaches: Final Report. *Federal Highway Administration*, FHWA-RD-01-124, 2001.
2. Taylor, H. and Meczkowski L., Safer Roadside. *Public Roads*, 2003.
3. Haddon, W., Advances in the Epidemiology of Injuries as a Basis for Public Policy. *Public Health Reports* (95)5,1980, pp 411-421.
4. Ramache, A., Roadside Safety: a Knowledge-based System for Alternative Configurations of Fixed Roadside Objects. *University of Newcastle Upon Tyne*, Ph.D. Thesis, 1991.
5. Tracz, M., and Gaca, S. Effects of Roadway and Roadside Design on Road Safety. *Road Safety in Europe and Strategic Highway Research Program*, Lille, France, 1995, pp 231-245.
6. Shankar, V., Mannering, F., and Barfield, W., Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequency. *Accident Analysis and Prevention* 27 (3),1995, pp 371-389.
7. Poch, M. and Mannering, F., Negative Binomial Analysis of Intersection-accident Frequencies. *Journal of Transportation Engineering*, 122(3),1996, 105-113.
8. Milton, J. and Mannering, F. The Relationship Among Highway Geometrics, Traffic-related Elements and Motor-vehicle Accident Frequencies. *Transportation* 25,1998, pp 395-413.
9. Miaou, S.-P., The Relationship between Truck Accidents and Geometric design of Road Sections: Poisson Versus Negative Binomial Regressions. *Accident Analysis and Prevention* 26 (4),1994, pp 471-482.
10. Shankar, V., Milton, J., Mannering, F., Modeling Accident Frequencies as Zero-Altered Probability Processes: an Empirical Inquiry. *Accident Analysis and Prevention* 29 (6),1997, pp 829-837.
11. Lee, J. and Mannering, F., Impact of Roadside Features on the Frequency and Severity of Run-off-roadway Accidents: an Empirical Analysis. *Accident Analysis and Prevention* 34, 2002, pp 149-161.
12. Wright, P.H; Mak, K.K., Single Vehicle Accident Relationships, *Traffic Engineering* 1976/01, 46(1), 1976, pp 16-21.
13. Turner, D., Prediction of Bridge Accident Rates. *Journal of Transportation Engineering* 110 (1),1984, pp 45-54.
14. Michie, J., Bronstad, M., Highway Guardrails: Safety Feature or Roadside Hazard? *Transportation Research Record* 1468,1994, pp 1-9.
15. Zegeer, C., Council, F., Safety Relationships Associated with Cross-sectional Roadway Elements. *Transportation Research Record* 1512,1995, pp 29-36.
16. Kennedy, J., Effect of Light Poles on Vehicle Impacts with Roadside Barriers. *Transportation Research Record* 1599,1997, pp 32-39.

17. Reid, J., Sicking, D., Faller, R., Pfeifer, B., Development of a New Guardrail System. *Transportation Research Record* 1599,1997, pp 72-80.
18. Ray, M., Impact Conditions In Side-impact Collisions with Fixed Roadside Objects. *Accident Analysis and Prevention* 31 (1),1999, pp 22-30.
19. Chayanan, S., Shankar, V., Nebergall, M., Juvva, N., and Ouyang, Y., Interaction Between the Roadway and Roadside - An Econometric Analysis of Design and Environmental Factors Affecting Segment Accident Rates, Research Report T1803, Task 31, *Washington State Department of Transportation*, 2003
20. Greene, W., *Econometric Analysis*, Fourth Edition, Prentice Hall, 2000.
21. Voung, Q., Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. *Econometrica* 57, 1989, pp 307-334.

LIST OF TABLES

LIST OF FIGURES

TABLE 1a. An Example Haddon Matrix.

TIME	HUMAN FACTORS	VEHICLE FACTORS	ENVIRONMENT FACTORS
Pre-Crash	Drinking	Poor maintenance	Horizontal Curvature
During-Crash	Inattention	Tire Failure	Black Ice
Post-Crash	Ejected	Event Data Records	Emergency Response

TABLE 1b. An Example Exposure Matrix.

Hypothesized Marginal Effects	At-Risk Locations	Not At-Risk Locations
Traffic-Human	Very significant	Very significant
Design	Very significant	Very significant
Environment	Significant	Maybe significant
Interactions	Possibly significant	Not significant

TABLE 2 Key Sample Statistics.

	Mean	Std. Dev.	Min	Max
Roadside crash count	1.09	1.55	0	10
Roadway Classification Proportions*				
Percent interstate sections (statewide mean percentages in parentheses)	11.32 (10.84)			
Percent principal arterial sections	34.59 (38.02)			
Percent minor arterial sections	29.87 (26.79)			
Percent collector arterial sections	24.21 (24.35)			
Weather Variables				
Average annual maximum temperature (statewide in parentheses)	60.57 (59.21)	2.83	45.10	66.00
Average annual minimum temperature (in Fahrenheit)	39.20 (38.45)	3.63	30.10	46.70
Total precipitation in inches	32.66 (42.60)	22.67	6.90	119.30
Average annual snowfall in inches	22.51 (21.30)	24.07	0.10	217.20
Traffic and Roadway Geometric Variables				
Per-lane AADT	2,649	3,256	97	29,907
Percent sections with peak hour volume exceeding 12% of ADT	12.89			
Truck percentage	14.66	9.42	2.10	71.50
Average right shoulder width in feet	5.55	2.83	0.00	12.00
Percent sections with narrow right shoulder (right shoulder ≤ 5 ft.)	50.94			
Percent sections with full/highly restrictive access control	27.04			
Percent sections with partial access control	66.70			
Percent sections with varying speed limits	15.41			
Average speed limit in miles per hour	54.75	8.61	25	70
Percent sections with horizontal curves	77.04			
Percent sections with intersections	70.75			
Percent sections with change in lanes in cross section	4.72			
Percent sections with interchanges	5.03			
Percent sections with level or level/rolling terrain	22.96			
Percent sections with divided medians	16.04			
Percent sections with special use lanes (e.g., climbing and auxiliary	6.66			
Interaction Variables				
Percent sections with horizontal curves and posted speeds > 50 mph	56.92			
Percent sections with horizontal curves and intersections in section	51.26			
Percent sections with measurable snow depth and horizontal curves	46.86			
Percent sections with intersections and posted speeds > 50 mph	43.08			

* Percentage was calculated by the proportion of the category over the total multiplied by 100

TABLE 3 Zero–Inflated Negative Binomial Model and Marginal Effects for Roadside Crashes.

Variables	Zero–Inflated Negative Binomial Model					
	Coefficients			Marginal Effects*		
	Estimated coefficient	Standard error	p-value	Estimated coefficient	Standard error	p-value
Constant	-0.1017	0.0369	0.0059			
Main Effects						
The change of speed limit indicator (1 if speed limit changes in section; 0 otherwise)	-0.1439	0.0472	0.0023	-0.4408	0.1712	0.0100
Per-lane AADT in 1000s	0.0690	0.0112	0.0000	0.2114	0.0407	0.0000
Right shoulder indicator (1 if right shoulder is ≤ 5 feet; 0 otherwise)	0.1376	0.0384	0.0003	0.4213	0.1394	0.0025
Total precipitation for autumn months (Sep-Oct)	0.0451	0.0191	0.0181	0.1380	0.0693	0.0465
Total precipitation for Spring months (Apr-Jun)	-0.0389	0.0160	0.0148	-0.1191	0.0580	0.0402
High monthly snowfall indicator (1 if monthly snowfall is higher than 4 inches; 0 otherwise)	-0.0840	0.0424	0.0474	-0.2572	0.1536	0.0940
Interchange indicator (1 if one or more interchanges in section; 0 otherwise)	0.3902	0.2167	0.0717	1.1950	0.7851	0.1280
Median indicator 1 (1 if all divided; 0 otherwise)	0.3459	0.1531	0.0238	1.0593	0.5547	0.0562
Interaction Effects						
Interaction between horizontal curves and intersections indicator (1 if there are horizontal curves and intersections in section; 0 otherwise)	0.1133	0.0409	0.0056	0.3471	0.1483	0.0193
Interaction variable between snow depth and horizontal curve (1 if snow depth > 0 for any month and horizontal curve; 0 otherwise)	-0.0899	0.0416	0.0309	-0.2752	0.1508	0.0680
Interaction variable among number of trucks, horizontal curves and average speed (1 if number of trucks ≥ 200, horizontal curve indicator = 1 and average speed limit ≥ 45; 0 otherwise)	0.0753	0.0384	0.0498	0.2305	0.1393	0.0979
α	0.1525	0.0833	0.0671			
τ	-22.3413	9.0341	0.0134			
Number of Observations		318				
Restricted log-likelihood (Constant only)		-470.9807				
Log-likelihood at convergence		-383.6697				
Vuong Statistic		2.5269				

* Partial derivatives of expected value with respect to the vector of characteristics. They are computed at the means of the independent variables.

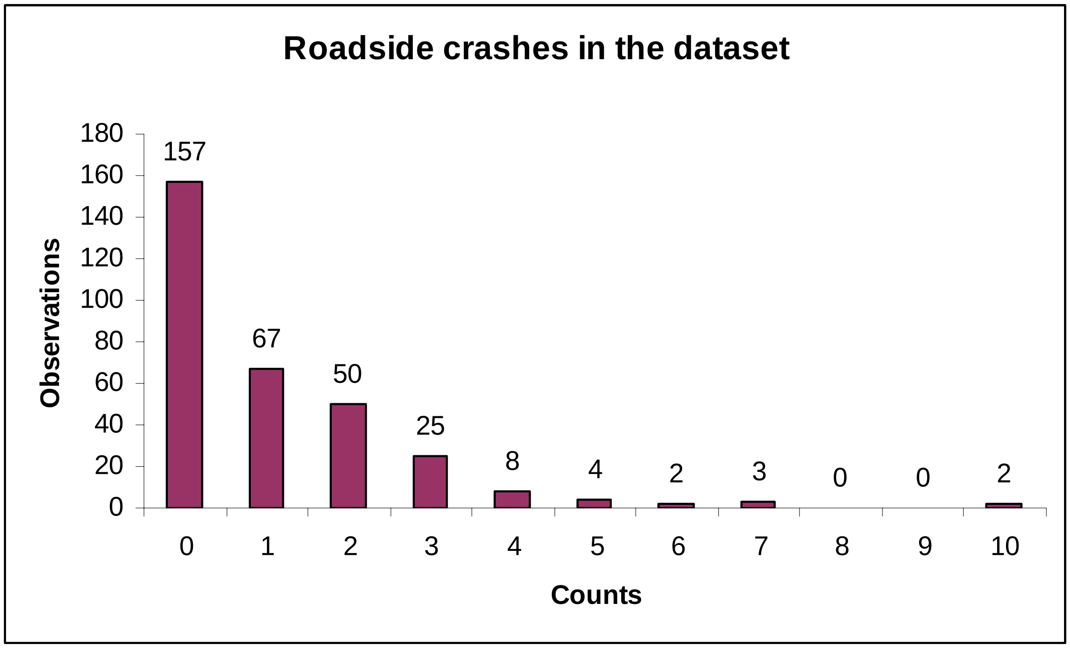

Figure 1. Observed Distribution of Roadside Crashes.

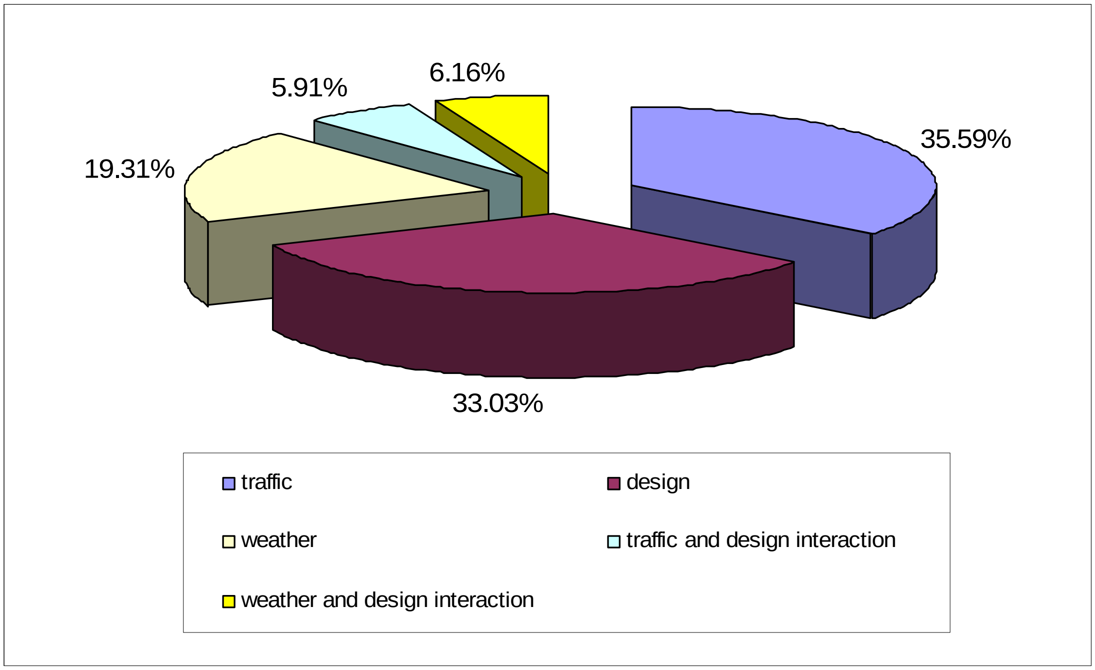

Figure 2. Contribution of Key Parameters to Roadside Crash Frequency Likelihood.