Bayesian Formulations for Heterogeneity in Crash Count Models

Sittipan Sittikariya, PhD

Transportation Engineer

DKS Associates, 719 Second Avenue, Suite 1250, Seattle, WA 98104, USA

Tel: (1-206) 382-9800, Fax: (1-206) 587-0692, Email: sxs@dksassociates.com

Venkataraman N. Shankar*, PhD, PE

Associate Professor

Department of Civil and Environmental Engineering, Pennsylvania State University,

226C Sackett Building, University Park, PA 16802, USA

Tel: (1-814) 865-9434, Fax: (1-814) 863-7304, Email: shankarv@engr.psu.edu

*Corresponding Author

Abstract—In this paper, we present alternative formulations for capturing heterogeneity in traffic crash count models. Using median crossover crashes as illustrative examples, we present formulations involving alternative error structures. Predictions from the formulations are also presented providing some empirical insight into the relative capabilities of the error structures in capturing variations in crash counts over time using panel data.

Keywords—Median crossover crashes, Heterogeneity, Correlation, Hierarchical Bayes, Bayesian Heterogeneity Poisson model, Bayesian Cluster-Specific Effects Poisson model

1. Introduction

There are some common modeling problems in the estimation of crash count models. Unobserved heterogeneity is a common theme in crash occurrence, leading to the well-known overdispersion problem (Shankar et al., 1995). Two major sources causing overdispersion are high-valued crash counts or excess zeros due to under-reporting. The negative binomial (**NB**) model is suitable for overdispersed crash frequencies, as shown in prior studies (see for example Poch and Mannering, 1996; Milton and Mannering, 1996).

Another common issue that arises in crash data contexts is the issue of correlation among crash counts. Multiple years of cross-sectional data on highway crash occurrences are often available from public domains, including time series information on crash counts, traffic volumes, and roadway geometrics as well as roadside characteristics. In addition, weather information is also available from national databases (see for example Shankar et al., 2004). It is noted here that correlation among crash counts and unobserved heterogeneity can be viewed as a simultaneity problem. That is, due to explicit correlation among dependent variables, or implicit correlation among the error structures, a simultaneity issue that argues for joint density functions arises. Complicating the decomposition of this simultaneity issue is the "mathematical overlap" between correlation and heterogeneity. From an empirical standpoint, it is highly possible that serial correlation may be spuriously captured as heterogeneity; the converse is also true. In the presence of crash count correlation in multiple years of cross-sectional crash count data, the efficiency of parameter estimates comes into question. Similar to the classical linear regression model, one can expect parameter estimates to be inefficient in the presence of correlation in crash count models (Sittikariya et al., 2005). A method to adjust for repeated observation effects on parameter estimates is necessary to adjust for heterogeneity. One such method relates to the use of the random effects Poisson introduced by Hausman et al., 1984, where cluster-specific heterogeneity is accounted for. In the traffic crash context, the random effects approach has been shown to be reasonable (Shankar et al., 1998). The empirical evidence in the above-mentioned literature points heavily to the usefulness of characterizing parameter estimation in longitudinal crash datasets as fundamentally heterogeneity problems.

The highway crash literature has recently evidenced an increased interest in the use of Hierarchical Bayes techniques (see for example Ma et al., 2006). As an area of application in the area of forecasting as a whole, **HB** methods have shown significant promise. Bayesian analysis also allows researchers to use subjective prior information on the distributions of parameters in combination with information from the observed data. Furthermore, the advantage of Bayesian analysis in the assessment of key contributing factors on traffic crashes is that the robustness of results can be examined. As a result Bayesian formulations are introduced in this study to address critical modeling issues, namely heterogeneity and correlation. Median crossover crash frequencies and associated roadway geometrics, traffic volume and weather information in the Washington State are modeled as an illustrative example of such formulations. The hierarchical Bayesian heterogeneity Poisson (**BHP**) model is the base model. This primary model is extendable to a Bayesian cluster-specific effects Poisson (**BCEP**) structure to account for correlation among crash counts in panel data.

The rest of this paper presents in order, methodology, empirical setting, model estimations, and conclusions and recommendations.

2. Methodology

Bayesian analysis produces a density function for the desired information rather than a point estimate given by frequentist or classical methods. The main difference between the Bayesian and frequentist methods is how probability is used. The Bayesian theory views probability as a confirmation of beliefs. Therefore, a numerical probability is the confidence the researcher has in various parameter values. In Bayesian analysis, the posterior distribution combines the likelihood and the prior distributions as well as the marginal distribution, and reflects knowledge about estimated parameters (θ) being updated after seeing the data. Bayesian inference is expressed as

$$P(\theta|y) = [P(y|\theta)* P(\theta)]/P(y) \tag{1}$$

where $P(\theta|y)$ is the posterior distribution; the likelihood, $P(y|\theta)$, is the expression for the distribution of the data to be observed given the set of parameters and the prior distribution of parameters, $P(\theta)$, represents beliefs about θ in the form of a probability distribution before seeing the data. The denominator, $P(y)$, is called the marginal distribution. A future observation $\tilde{y}$ may be predicted using a predictive distribution, $P(\tilde{y} \mid y)$, based on the posterior distribution as follows:

$$P(\tilde{y} \mid y) = \int P(\tilde{y} \mid \theta) P(\theta \mid y) d\theta \tag{2}$$

The advantage of Bayesian approaches is directly quantifying uncertainty when predictive models with many parameters are considered. This is especially the case in median crossover crashes where roadway geometrics, traffic factors, median characteristics and environmental conditions play a dominant role. Human factors are neither easily measurable nor maintainable over the long-term as self-sustainable sources of information. Hence, to develop a predictive paradigm that maximizes predictive power, the Bayesian paradigm gives freedom to set up complex models by supplying a conceptually simple method for coping with multiple parameters. Although a realistic model may require many parameters, interest usually focuses on a smaller number of parameters.

The **BHP** model specifies the distributions of the observed median crossover crash count vector "**y**" given the individual parameters λ_{it}'s as:

$$\mathbf{y_{it}} | \lambda_{it} \sim \text{Poisson}(\nu_{it}\lambda_{it}) \tag{3}$$

where the vector λ_{it} is the mean rate of median crossovers in section *i* at time *t*; λ_{it} is defined as $\log(\lambda_{it}) = \mathbf{X_{it}}\,\beta$ ($\mathbf{X_{it}}$ is vector of explanatory variables e.g. roadway geometrics, median characteristics and traffic as well as weather variables); β is a coefficient vector with a normal prior (mean 0 and variance 0.001) and ν_{it} is a multiplicative conjugate gamma form for heterogeneity. In this model, the individual Poisson parameters ν_{it} follows a gamma distribution as shown:

$$\nu_{it} \sim \text{Gamma}(\alpha, \alpha) \tag{4}$$

$$\alpha \sim \text{Gamma}(1, 1) \tag{5}$$

where α is the overdispersion parameter and is gamma distributed. By doing so, the structure of **BHP** model becomes a hierarchy from the product of individual heterogeneity and normal prior of the exogenous variables. A multilevel model with carefully chosen priors at various levels for the mean and variance is expected to be more flexible across observations and capture more heterogeneity in the data. The hierarchy of this structure is captured by allowing heterogeneity parameter to follow a gamma distribution. For this specification, the mean and the variance correspond to the classical **NB** form with $E[y_{it}] = \exp(\mathbf{X}_{it}\,\beta)$ and $Var[y_{it}] = E[y_{it}][1 + \alpha E[y_{it}]]$. Estimation algorithms corresponding to prior belief of data and different modeling structures of the Bayesian model are customized using the Gibbs Sampler available in WinBUGS 1.4.1 software. The Markov Chain Monte Carlo (**MCMC**) algorithms involved 10,000 iterations and three chains to ensure stationary distributions.

The **BHP** model in equation 3 assumes that the unobserved effects or heterogeneity for each individual segment vary across time and space. One may argue that the differences in heterogeneity in the same roadway section across time are minimal because the geometrics of roadway and roadside remain the same and the traffic volumes across the year are proximate with usual differences equal to the nominal growth rate of 3 to 4 percent. The statistically significant differences are mainly group-specific, in this case, specific to the group of years for each segment. To address group-specific effects, a Bayesian cluster-specific effects Poisson model is developed. The **BCEP** model can capture the correlations among crash counts in the same cluster by constraining cluster heterogeneity to be identical in the same roadway section. By doing so, the correlations among the crash counts can be absorbed through one single parameter for each roadway section over multiple years.

The formulation of the **BCEP** distribution begins with the **BHP** density as defined in equation 3. To account for the section-specific variation, we modify the multiplicative heterogeneity term in the **BHP** model (by dropping the time subscript) as follows:

$$y_{it}|\lambda_{it} \sim \text{Poisson}(\nu_i\lambda_{it}) \tag{6}$$

$$\nu_i \sim \text{Gamma}(\alpha, \alpha) \tag{7}$$

$$\alpha \sim \text{Gamma}(1, 1) \tag{8}$$

where ν_i is a cluster-specific heterogeneity (not observation-specific as in the **BHP** model) for in roadway section *i*. This cluster-specific heterogeneity is also gamma distributed. It is also noted here that the degenerate case, when each roadway section has only single year of observation, yields the **BHP** distribution.

In Bayesian analysis, the deviance information criterion (**DIC**) is suggested as a common measure of overall model fit. The **DIC** is intended to be a hierarchical modeling generalization of the **AIC** (Akaike information Criterion) and **BIC** (Bayesian information Criterion) and it is useful for Bayesian model selection purpose. The **DIC** is a combination of the likelihood function and the effective number of parameters in the model. In general, the likelihood function tends to decrease as the number of parameters in a model increases. However, the effective number of parameters will compensate for this effect. The models with smaller **DIC** should be preferred to models with larger **DIC**. The reader is urged to consult Spiegelhalter et al., 2002, for in-depth details of **DIC**.

Both models were estimated using the same set of specifications as a benchmark to identify the robustness of the explanatory variables. To determine this set of specifications, several hundred modeling specifications were experimented to determine the most meaningful and robust specifications from the database. In other words, this approach allows us to probe for variables that are robust to specification assumptions. By doing so, the intent of this research is to identify "common denominator variables" that would appear to be consistently statistically significant under various treatments of heterogeneity.

3. Empirical Setting of Median Crossover Crash in Washington State

In this discussion, information regarding the data for model analysis and evaluations is presented. The Washington State highway system contains over 7,000 miles of state highways. In order to pursue the stated objectives, longitudinal data for the period 1990 to 1994 containing crossover crash information on unbarriered medians on the Washington State highway network was used. The longitudinal study is especially useful for median crossovers due to the sporadic nature of median crossovers. The 1990-1994 dataset consisted of 275 unbarriered highway sections over the entire Washington State highway network, totaling a length of nearly 670 center-line miles. Mean crossover frequency was 0.24 crashes per year, while per-lane average daily traffic was approximately 7,400 vehicles. Mean median width was 57 feet, with approximately 70 percent of all sections in the 40-foot to 75-foot median width range.

All crash records responded by the Washington State Patrol on the Washington State highway system were recorded in the police report. The Washington State Department of transportation (**WSDOT**) developed the crash database, the Washington State Master Crash Record System (**MARS**), based on these police records. The crash records from this database are the primary source of data for the current median crossover crash study. The panel data in this research consists of five years (1990–1994 inclusive) of annual median crossover crash counts for 275 roadway sections in Washington State. As a result, the total number of observations in the database is 1,375 observations. The panel data is balanced, with all sections having a full five-year history. This panel data represents all sections (longer than 2,624 ft) without median barriers on divided state highways. The reasons why only sections longer than 2,624 ft are selected are that about 95 percent of shorter sections on divided highways have barriers, and that the shorter sections are more affected by access controls and intersections (Ulfarsson and Shankar, 2003). An objective in part in this paper is to identify in a robust manner, factors associated with median crossover crashes.

In addition to median crossover crash information, other components of data extracted from the database included roadway geometrics, median characteristics and traffic volumes. Since section definitions in this analysis are based on the continuity of section without barrier, section lengths are unequal. The geometric and traffic data was aggregated using a weighted average from the section listed in the database. The geometric data contain roadway widths, lane widths, number of lanes, shoulder widths, horizontal curve information, legal speed limit, surfacing type, terrain, median widths and so on. The traffic data include average annual daily traffic (**AADT**) and truck volume as percentage of **AADT**.

The ArcGIS program was used to match the roadway sections to their weather attributes stored in the historical weather database provided by the Western Regional Climate Center. The mapping criteria involved linking the non-median barrier roadway sections to the nearest corresponding weather stations. Each weather station provided climate data including daily, monthly and annual measurements of temperature, precipitation and snowfall including snow depth, with the records dating back to 1948. The selected roadway sections and all weather stations in Washington State are illustrated in figure 1. Median sections that are unbarriered represent nearly all unbarriered sections in the state system. Hence, only weather data proximate to the identified sections was used in the analysis. In this research, the 30-year average monthly precipitation depth and the 30-year average monthly snow depth are selected as key weather variables. The reason why the average values are selected is that the monthly precipitation and the monthly snow depth in each year are not significantly different from the 30-year average values. Furthermore, preliminary analysis indicated that parameter effects are statistically similar; on the other hand, the predictive usefulness of historical averages for weather provides an advantage from a forecasting standpoint.

Table 1 provides aggregate descriptive statistics of key variables in the median crossover crash dataset for the entire Washington State highway network. The median crossover crash frequency is the dependent variable with traffic variables, roadway geometric variables, median variables, weather variables and interaction variables being used as explanatory variables.

4. Model Estimations

4.1 Empirical Findings

In the development of both Bayesian models, WinBUGS 1.4.1 software is used. In order to compare the robustness of variables with different modeling issues and the accuracy of predictions, the same set of specifications is used.

As results of both Bayesian models presented in table 2 show, factors positively correlating with median crossover crashes include the interaction variables between length of median and three categories of median widths and in addition the number of interchanges variable. The three categories of the median width variable interacted with the length of the section variable are 1) less than or equal to 40-foot median width, 2) 40 to 60-foot median width, and 3) greater than 60-foot median width. Different ranges of the median width were experimented with but the result showed that these three categories, when interrelated with the section length, had the greatest impact on the crossover crash likelihood. The magnitudes of the coefficients of the length variables suggest that for a given length, the likelihood of median crossover crashes increases the greatest on sections with median widths between 40 and 60 feet wide in comparison to the other width categories. On the other hand, sections with median width wider than 60 feet have the least contribution to the likelihood of median crossover frequency. This would imply that decisions to barrier should strongly consider median widths up to 60 feet.

Three factors negatively correlate with median crossover crash frequencies. These include the traffic volume indicator variable (if average annual daily traffic was less than 5,000), the interaction variable between the number of horizontal curves less than or equal to 0.5 per mile in the section and the average monthly precipitation being less than or equal 1.5 inches, and the interaction variable between the number of horizontal curves greater than 0.5 per mile in the section and the average monthly precipitation being greater than 4 inches.

The weather effect appearing in the form of the interaction variables played a significant role in median crossover likelihood. In a section where the average number of horizontal curves was less than or equal to 0.5 per mile, the median crossover counts were expected to decrease if the average monthly precipitation was less than or equal to 1.5 inches. Median crossover crash frequency also decreases when average monthly precipitation exceeds 4 inches on sections with greater than 0.5 horizontal curves per mile. Both effects point to the range of interactions between precipitation and horizontal curves on median crossover frequencies. When precipitation amount is not significant, the direction of the effect of its interaction with horizontal curvature is the same as its interaction when precipitation is significant, both being negative.

Table 2 shows the comparison of estimated parameters from the **BHP** and the **BCEP** models of median crossover crash frequency. In comparison with the **BHP** model, the coefficient values of the **BCEP** structure are similar but not identical. Compared to the **BCEP** model, the **BHP** model underestimated the standard errors of the parameter estimates, therefore overestimating the t-statistic. It is noted that the **BHP** model is the special case of the **BCEP** model if there is one year of observation in the roadway section. Without the section-specific effects, all estimated parameters and standard errors would be exactly the same as those from the **BHP** model. The deviance information criterion (goodness of fit penalized for degrees of freedom) in the **BHP** models is noticeably lower in comparison with the **BCEP** model. A lower value of deviance information criterion from **BHP** model suggests that this model is preferred in terms of overall model fit.

Furthermore, the useful results from the Bayesian analysis are the credibility intervals associated with coefficient values. As shown in tables 2, the trend in parameter signs and magnitudes as their credibility levels increase from 2.5 percent to 97.5 percent can be noted. The mean value is the value of the parameter typically reported in most studies. As expected, some variation exists between the 2.5^{th} percentile credibility and 97.5^{th} percentile credibility estimates of the coefficients in every model. All variables including constants maintained their “sign” as credibility levels increase. The trend in estimated parameters’ credibility levels suggests that all variables remain fairly robust in response to common modeling issues such as heterogeneity and correlations among crash counts.

4.2 Model Predictions

To assess the predictive abilities of both models, model predictions are presented in this section. In the general case of crash count models, two measures of predictive effectiveness are used: a) mean absolute deviation (**MAD**) and b) root mean square error (**RMSE**). The **MAD** is the average of the absolute values of the prediction errors. It is appropriate when the cost of forecast errors is proportional to the absolute size of the forecast error. This criterion is also called mean absolute error (**MAE**). The **RMSE** is the square root of the average of the squared values of the prediction errors and is appropriate to situations in which the cost of an error increases as the square of that error. In this study, percent change of predicted count from observed count is not appropriate due to the significant presence of zero counts in the database. The **MAD** and **RMSE** used in this research are formulated as follows:

$$\mathbf{MAD} = \left(\sum_{i=1}^{275} \sum_{t=1}^{5} |O_{it} - P_{it}|\right) / 1375 \tag{9}$$

$$\mathbf{RMSE} = \sqrt{\left(\sum_{i=1}^{275} \sum_{t=1}^{5} (O_{it} - P_{it})^2\right) / 1375} \tag{10}$$

where O_{it} and P_{it} are observed and predicted median crossover crash counts in section *i* at time *t*. Both models were developed using 5 years of data (1990-1994) and then the predictions were tested against within-sample observations. In addition, both models were developed using 3 years of data (1990-1992) and 4 years of data (1990-1993) and then tested against the extra-sample observations (i.e. the predicted crossover frequencies in the fourth and the fifth years (1993-1994) for the first case and the predicted crossover frequencies in the fifth year (1994) for the second case). If the predicted count is close to observed count, the **MAD** and the **RMSE** will be minimized. In other words, the lesser the **MAD** and the **RMSE**, the more accurate the prediction provided by the model.

The predictions of both models are presented in table 3. For sub-sample predictions, the **BHP** model shows a lower **MAD** of 0.30 compared to a **MAD** of 0.33 in the **BCEP** model. In addition, the **BHP** model arrives at a lower **RMSE** of 0.41 in comparison with a **RMSE** of 0.51 in the **BCEP** model. Furthermore, the findings from extra-sample predictions (1 year) confirm that the **BHP** model provides lower or the same **MAD** and **RMSE** in comparison with the **BCEP** structure. It is noted here that the **BHP** structure experienced a convergence issue while performing extra-sample predictions (2 years) due to lack of sufficient panel data. The **DIC** value for the **BHP** model is 1389.82 compared to 1420.71 for the **BCEP** model.

5. Conclusions and Recommendations

5.1 Conclusions

This paper intends to present alternate methodologies for the estimation of traffic crash count models. In general, heterogeneity and correlation among crash counts are major modeling problems in the estimation of highway crash models using cross-sectional panel data. We established the **BHP** model with individual unobserved effects first and then we extended to the **BCEP** structure with cluster-specific heterogeneity to account for correlation in panel data. In this study, multiple years of median crossover crash frequency were used as case studies for the formulation of the hierarchical Bayesian structures.

The main effects such as average daily traffic, and the number of interchanges were found to have statistically significant impact on the frequency of median crossover crashes. The interaction variables of length of the roadway section and median width as well as the interaction between monthly precipitation and the number of horizontal curves per mile are also found to be important contributors to median crossover likelihoods.

By constraining the set of specifications to be the same in both models, the impact of critical modeling issues in the analysis of panel crash counts was addressed. The results show that the estimated parameters from both models are similar but not identical while the standard errors from the **BHP** are downward biased and therefore lead to overestimation of the variable significance. Overestimating the t-statistic potentially results in the identification of irrelevant variables as significant effects. From an information-based safety programming standpoint, irrelevant variables place an unnecessary data collection burden. Despite the fact that the **BCEP** models have higher standard errors; leading to lower t-statistics, high levels of confidence in the **BCEP** models still remain (exceeding 95 percent). From a prediction standpoint, compared to the **BCEP** model, the **BHP** structure appears to hold valid in the median crossover crash context.

To this extent, while the **BHP** model, on the one hand, helps provide better predictions, it may also introduce the identification of irrelevant variables as significant effects. The trade-off between assessments of parameter uncertainty, predictions, and model data requirements appears to be a prominent area of research for years to come. While it is not surprising that the **BHP** model would provide a lower **DIC** due to unconstrained error relationships within sections, one has to consider the impact of downward biased standard errors on variable identification as well. In our context, this does not appear to have an impact on variable significance; however it is a critical enough issue considering downward biases in standard errors exceed 20 percent.

5.2 Recommendations

The data used in this research is somewhat limited in its coverage of geographic, environmental and geometric effects. The question raised from this research is that "will the common variables identified in this study still remain robust in models based upon geography, environment and geometric design different from those in Washington State?" One possible outcome is that all common variables would remain at a high level of significance in the model. However, the estimated magnitudes may differ. Furthermore, the group specific effect as well as the overdispersion parameter may vary in both significance and magnitude. It is fair to say that the common variables developed from this research provide a reasonable "starter dataset" for other geographic, environmental and geometric design contexts. As future research, one can use common variables developed by this study as naïve specifications to explore median crossover crash frequency in greater depth.

Methodologically, modeling issues remain. Bayesian analysis allows the modeler to use subjective prior information on the distribution of parameters in combination with information from the observed data. In this sense, an alternate approach is to assume normal heterogeneity instead of gamma distributed heterogeneity. The effect of Poisson-normal overdispersion may be to influence parameter standard errors differently from the gamma-heterogeneity assumption. Our preliminary analysis in this regard indicates that normal heterogeneity structures may not be as well behaved, but this may be a finding that arises as an artifact of the dataset at hand.

References

Hausman, J., Hall, B. and Griliches, Z. (1984) ‘Econometric-models for count data with an application to the patents R and D relationship’, *Econometrica*, **52**(4), pp. 909-938

Ma, J. and Kockelman, K.M. (2006) ‘Bayesian multivariate Poisson regression for models of injury count, by severity’, *Transportation Research Record*, **1950**, pp. 24-34

Milton, J.C. and Mannering, F.L. (1996) ‘The relationship between highway geometrics, traffic related elements and motor vehicle accidents,’ Final Research Report, WA-RD 403.1, Washington State Department of Transportation, Washington

Poch, M. and Mannering, F.L. (1996) ‘Negative binomial analysis of intersection-accident frequencies’, *Journal of Transportation Engineering*, **122**(3), pp. 105-113

Shankar, V.N., Chayanan, S., Sittikariya, S., Shyu, M.B., Juvva, N.K. and Milton J.C. (2004) ‘The marginal impacts of design, traffic, weather, and related interactions on roadside crashes’, *Transportation Research Record*, **1897**, pp. 156-163

Shankar, V.N., Albin, R.B., Milton, J.C. and Mannering, F.L. (1998) ‘Evaluating median cross-over likelihoods with clustered accident counts: an empirical inquiry using the random effects negative binomial model’, *Transportation Research Record*, **1635**, pp. 44–48

Shankar, V.N., Mannering, F.L. and Barfield, W. (1995) ‘Effect of roadway geometrics and environmental conditions on rural accident frequencies’, *Accident Analysis and Prevention*, **27**(3), pp. 371-389

Sittikariya, S., Shankar, V.N., Shyu, M.B. and Chayanan, S. (2005) ‘Accounting for Serial Correlation in Count Model of Traffic Safety’, *Journal of Eastern Asia Society for Transportation Studies*, **6**, pp. 3645-3657

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Van der Linde, A. (2002) ‘Bayesian measures of model complexity and fit’, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, pp. 583-639

Ulfarsson, G. F. and Shankar V.N. (2003) ‘An accident count model based on multi-year cross-sectional roadway data with serial correlation’, *Transportation Research Record*, **1840**, pp. 193-197

Table 1 Descriptive Statistics of Key Median Crossover Crash Related Variables

Variable	Mean	Std. Dev.	Min	Max
Number of crossover crashes in section	0.24	0.65	0.00	7.00
AADT	37,355.00	36,975.00	3,350.00	172,560.00
AADT per lane	7,445.00	5,830.00	835.00	28,690.00
Single truck percentage	4.20	1.22	1.90	10.00
Double truck percentage	7.76	4.62	0.55	17.80
Truck-train percentage	2.21	1.60	0.00	7.00
Percentage of AADT in the peak hour	10.72	7.31	0.00	19.40
Speed limit	60.00	5.50	35.00	65.00
Maximum median shoulder width in feet	5.31	2.49	0.00	18.00
Minimum median shoulder width in feet	4.48	1.68	0.00	10.00
Percent medians narrower than 40 feet	32.36	46.80		
Percent medians between 40 feet and 50 feet in width	11.64	32.08		
Percent medians between 50 feet and 60 feet in width	5.81	23.42		
Percent medians wider than 60 feet in width	50.19	50.02		
Percent medians that are paved	4.36	20.44		
Length of the roadway section in miles	2.43	2.69	0.50	19.30
The number of interchanges in section	0.85	0.84	0.00	4.00
The number of horizontal curves in section	2.75	2.86	0.00	29.00
The number of horizontal curves per mile	1.44	0.96	0.00	5.00
Maximum horizontal central angle in degrees	30.29	23.88	0.00	111.49
Minimum radius of horizontal curve in feet	4267.24	4875.08	0.00	38,400.00
Average annual snow depth in inches	19.44	45.66	0.00	652.00[*]
Average annual precipitation in inches	29.86	21.98	4.53	131.74
Number of grade changes	3.87	4.09	0.00	28.00
Average roadway width	57.42	15.47	24.00	121.00

[*] Weather station data for mountainous section

Table 2 Parameter Comparison between the Bayesian Heterogeneity Poisson and the Bayesian Cluster-Specific Effects Poisson Models of Median Crossover Crash Frequency

Variable	**Bayesian Heterogeneity Poisson Model**				**Bayesian Cluster-Specific Effects Poisson Model**			
	β* σ ** t***	**Credibility Percentiles of Coefficient**			**β* σ ** t*****	**Credibility Percentiles of Coefficient**		
		2.5	**50**	**97.5**		**2.5**	**50**	**97.5**
Constant	-1.897 0.139 -13.638	-2.165	-1.894	-1.613	-1.981 0.183 -10.831	-2.357	-1.983	-1.615
Per-lane AADT indicator (1 if per-lane AADT <= 5000 vehicles, 0 otherwise)	-0.953 0.182 -5.242	-1.324	-0.950	-0.609	-0.974 0.200 -4.877	-1.383	-0.979	-0.578
Length of the roadway section on median widths less than or equal to 40 feet	0.323 0.043 7.565	0.230	0.322	0.406	0.357 0.051 7.065	0.262	0.357	0.458
Length of the roadway section on median widths greater than or equal to 41 feet and less than or equal to 60 feet	0.431 0.064 6.781	0.305	0.431	0.555	0.451 0.079 5.744	0.301	0.446	0.611
Length of the roadway section on median widths greater than 60 feet	0.112 0.026 4.305	0.063	0.112	0.162	0.122 0.033 3.725	0.057	0.123	0.183
The number of interchanges in the section	0.234 0.083 2.823	0.057	0.234	0.384	0.253 0.113 2.240	0.023	0.245	0.477
Interaction 1 between average monthly precipitation and the number of horizontal curves per mile (1 if average monthly precipitation <= 1.5 inches and the number of horizontal curves per mile <= 0.5, 0 otherwise)	-1.050 0.355 -2.962	-1.786	-1.046	-0.382	-1.032 0.398 -2.596	-1.845	-1.004	-0.306
Interaction 2 between average monthly precipitation and the number of horizontal curves per mile (1 if average monthly precipitation > 4.0 inches and the number of horizontal curves per mile > 0.5, 0 otherwise)	-0.558 0.222 -2.515	-1.000	-0.550	-0.136	-0.469 0.225 -2.082	-0.908	-0.460	-0.034
The inverse of dispersion parameter ($\theta = 1/\alpha$)	1.206 0.288 4.192	0.769	1.159	1.934	2.097 0.537 3.906	1.249	2.029	3.364
Deviance Information Criterion	1389.820				1420.710			

* Estimated Coefficient, ** Standard Error, *** t-statistic

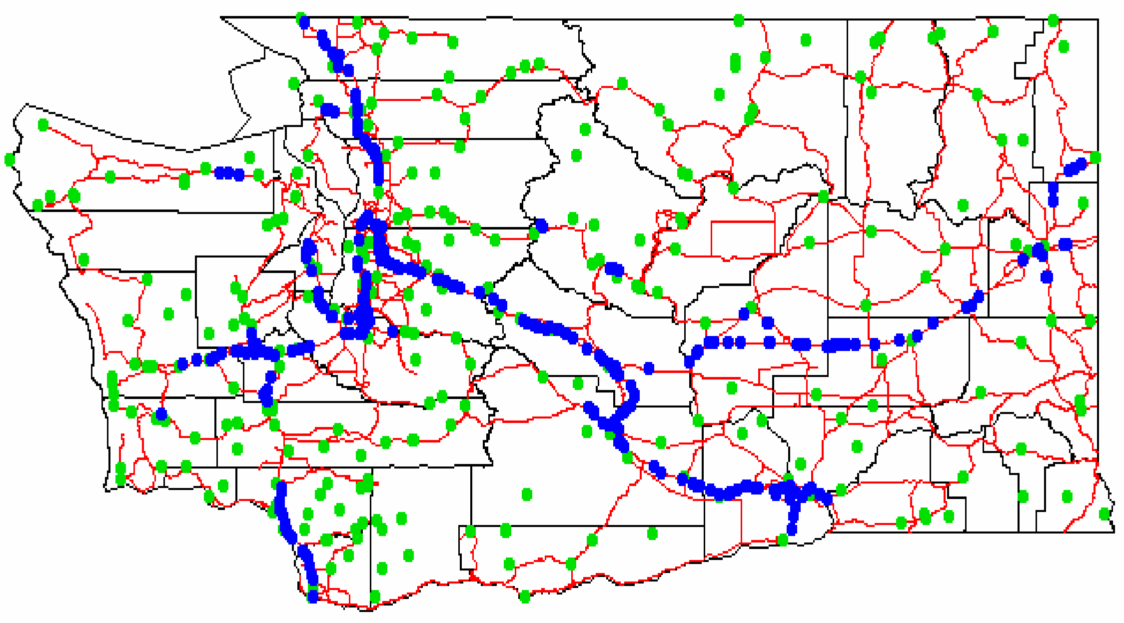
Table 3 The Predictions of the Bayesian Heterogeneity Poisson and the Bayesian Cluster-Specific Effects Poisson Models of Median Crossover Crash Frequency

Model	**Errors in Forecasting Observations**		
	Sub-Sample 5 years (1990-1994)	**Extra-Sample 1 year (1994)**	**Extra-Sample 2 years (1993-1994)**
The **BHP** Model	0.30* (0.41)**	0.30 (0.45)	Not Applicable
The **BCEP** Model	0.33 (0.51)	0.30 (0.46)	0.31 (0.51)

*Mean Absolute Deviation (**MAD**), ** Root Mean Square Error (**RMSE**)

LIST OF FIGURE

Figure 1 Selected Roadway Sections and Weather Stations in Washington State



*Blue (darker color) = The Selected Roadway Sections, **Green (lighter color) = The Weather Stations

Figure 1 Selected Roadway Sections and Weather Stations in Washington State