Final Research Report
Contract T2695, Task 09
Median Crossover

MEDIAN CROSSOVER ACCIDENT ANALYSES AND THE EFFECTIVENESS OF MEDIAN BARRIERS

by

Songrit Chayanan, Graduate Research Assistant, University of Washington
Seattle

Sittipan Sittkariya, Research Consultant, Bangkok, Thailand

Ming-Bang Shyu, Graduate Research Assistant, University of Washington
Seattle

Venky Shankar, Assistant Professor
University of Washington
Seattle

Gudmundur Ulfarsson, Assistant Professor, Washington University
St. Louis

Naveen Kumar Juvva, Transportation Engineer, Parametrix, Inc.
Kirkland, Washington

Washington State Transportation Center (TRAC)
University of Washington, Box 354802
University District Building
1107 NE 45th Street, Suite 535
Seattle, Washington 98105-4631

Washington State Department of Transportation Technical Monitor
John Milton, Assistant State Design Engineer
Design Office

Prepared for
Washington State Transportation Commission
Departmenst of Transportation
and in cooperation with
U.S. Department of Transportation
Federal Highway Administration

August 2004

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
WA-RD 591.1		

4. TITLE AND SUBTITLE	5. REPORT DATE
MEDIAN CROSSOVER ACCIDENT ANALYSES AND THE EFFECTIVENESS OF MEDIAN BARRIERS	August 2004
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
Venky Shankar, Songrit Chayanan, Sittipan Sittkariya, Ming-Bang Shyu, Gudmundur Ulfarsson, Naveen Kumar Juvva	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. WORK UNIT NO.
Washington State Transportation Center (TRAC) University of Washington, Box 354802 University District Building; 1107 NE 45th Street, Suite 535 Seattle, Washington 98105-4631	11. CONTRACT OR GRANT NO. Agreement T2695, Task 50
12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Research Office Washington State Department of Transportation Transportation Building, MS 47372 Olympia, Washington 98504-7372 Kathy Lindquist, Project Manager, 360-705-7976	Final Research Report 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. ABSTRACT

This research explored the use of count models to determine design and weather factors correlated with median crossover crashes on Washington State highways. The study 1) developed a roadside data system that can be consistently and systematically used in all six regions of Washington State; 2) developed a decision matrix comprising geometric, environmental, and traffic factors for estimating cross-over probability ranges; and 3) examined the impacts of barriering. Longitudinal data for the period 1990 to 1994 containing crash information on vehicle crossovers on non-barriered medians on Washington State highways were used as the dataset for this study. Two types of statistical models were examined: 1) a model that forecasts the mean number of yearly median crossovers, and 2) a model that examines the contribution of roadway geometrics, median widths, weather, traffic volumes and roadside characteristics to the annual societal cost of median crossovers.

Results of the study suggest these design policies: Barrier all medians less than or equal to 50 feet wide; do not install barriers for medians wider than 60 feet; consider case-by-case barriering medians in the 50-foot to 60-foot range.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Negative multinomial, median crossovers, count models, median width, barrier policy	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None		

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Washington State Department of Transportation. This report does not constitute a standard, specification, or regulation.

CONTENTS

FIGURES

TABLES

EXECUTIVE SUMMARY

This research explores the use of count models to determine design and weather factors correlated with median crossover crashes on Washington State highways. The Washington State Department of Transportation (WSDOT) currently relies on the relationship between average daily traffic (ADT) and median width when examining median barrier requirements. Such techniques, while being somewhat effective, are limited by their simplicity. As a result of this, median barrier requirements do not take into account fully multi-variate effects of roadway geometrics, traffic factors, and environmental conditions. The WSDOT follows established rules and procedures set forth in the design manuals of the American Association of State Highway Transportation Officials (AASHTO) when it considers median barrier placement to reduce or eliminate median-encroaching vehicle crashes. Recently, the WSDOT went one step further than the national standards. It examined the benefit/cost of median barrier installations to determine the optimal level of barrier placement. For example, it determined through a focused benefit/cost study that cable barriers installed on medians that are 50 feet wide or less would be optimal. A broader suggestion from this finding is that savings in crash-related societal costs would be maximized when medians that are less than or equal to 50 feet wide are barriered. Given the fact that median crossover accidents can be generally severe and result in a high cost to society, this finding motivates our current study.

The objectives of this study were threefold. The first was to develop a roadside data system that could be consistently and systematically used in all six regions of Washington State. Such a system could then be used for data inputs in median crossover model development and refinement. The second was to develop a decision matrix comprising geometric, environmental, and traffic factors so that WSDOT personnel could cross-check their subject locations within the matrix and estimate cross-over probability ranges. The third objective was to examine the impacts of barriering. Impacts could include less severe outcomes from fixed-object collisions or multi-vehicle collisions caused by the redirection of errant vehicles into the mainline. A before-and-after analysis would then allow WSDOT personnel to assess the cost effectiveness of barriering.

To pursue the stated objectives, we used longitudinal data from 1990 to 1994 containing crash information about vehicle crossovers on non-barriered medians on the Washington State highway network. The longitudinal study was especially useful for median crossovers because of the sporadic nature of median crossovers. The 1990-94 dataset consisted of 275 non-barriered highway sections over the entire Washington State highway network, totaling a length of nearly 670 center-line miles. The mean crossover frequency was 0.24 crashes per year, while per-lane average daily traffic was approximately 7,400 vehicles. The mean median width was 57 feet, with approximately 70 percent of all sections in the 40-foot to 75-foot median width range.

A variety of statistical approaches was employed to examine factors contributing to the median crossover problem. Data that provided correlation information to median crossovers included roadway geometrics, precipitation, traffic volumes, and roadside characteristics relating to shoulder and median portions of the highway. An integrated database consisting of 1,375 records of annual median crossovers on 275 sections was developed for the analysis. The electronic database contained 30-year averages for precipitation data for all median sections and median characteristics, along with conventional information such as geometrics and traffic volumes. Precipitation information was compiled from 30-year National Oceanic and Atmospheric Administration weather station profiles available for Washington State as a whole. A GIS-based matching procedure was used to identify weather stations closest to the median sections. Median characteristic information included qualitative information on several features of medians, such as median slopes, surface type and widths,

In developing the decision matrix on median barrier installation, two types of statistical models were examined. The first model examined the contribution of roadway geometrics, median widths, weather, traffic volumes, and roadside characteristics to annual median crossovers. In this model type, issues relating to unobserved effects and time trends were addressed. The integrated database consisting of weather, geometrics, volumes, and roadside characteristics was only a partial set of factors that were found to correlate with median crossovers. Unobserved effects relating to driver characteristics, local environmental fluctuations not captured completely in precipitation data, and vehicle speeds also contributed to the complexity of the statistical problem. Time trends occurred in the observed dataset because of the longitudinal nature of the data. As a result, the statistical framework was adapted to account for repeated information occurring annually in the 1990-1994 dataset. This type of model is referred to in this study as the “crossover frequency” model, i.e., a model that forecasts the mean number of yearly median crossovers.

The second type of statistical model examined the contribution of roadway geometrics, median widths, weather, traffic volumes, and roadside characteristics to the annual societal cost of median crossovers. The cost-level model in this study is referred to as the "median crossover societal cost" model, i.e., a model that forecasts the annual cost of median crossovers in 2002 dollars. The average cost of a single median crossover in 2002 dollars is approximately $445,000. The intent behind this model was to determine if the cost-level examination and frequency-level examination from the first model provided a common point of decision-making for barrier installation policy. For example, let us assume that the median crossover frequency model suggests that, controlling for all other factors, medians 61 feet or wider experience significantly fewer median crossovers that narrower medians. Let us also assume that the societal cost model suggests that medians 51 feet or wider are expected to contribute to significantly lower societal costs than those 50 feet or narrower. A common point for decision making begins to emerge here. One might consider the following design policy in this case:

a) Barrier all medians less than or equal to 50 feet wide.
b) Do not barrier medians wider than 60 feet.
c) Consider case-by-case assessments for barriering medians in the 50-foot to 60-foot range.

Our statistical analysis in fact concluded that this decision rule is most appropriate for Washington State.

Finally, the effectiveness of the installation of median barriers on the selected road sections was also tested. The 1990-94 accident data for sections without median barriers were used as the "without median barrier" data, and 1990-94 accident data for sections containing barriers (and similar in attributes to the sections without barriers) were used as "with median barrier" data. The median barrier sections were chosen so that they were physically near the sections without median barriers and hence very closely represented the behavior of the median barrier-less sections after median barriers were put in them. A preliminary contingency analysis of total crash counts determined that barrier installation did not necessarily uniformly increase overall crash counts on the highway section. Some sections reported lowered crash counts with barrier installation, whereas others reported higher crash counts. To examine the characteristics of sections that exhibited this difference in crash profiles, we estimated a statistical model for all reported crashes for sections with and without median barriers. In estimating the model, traffic volumes, precipitation, geometrics, and the presence of median barriers were controlled for. The examination determined that as the number of curves per mile increased in a section with barriers, the overall crash profile on those sections increased in comparison to similar sections with no barriers. It was also found that a section with two to five grade changes per mile and median barriers experienced fewer overall crashes than similar sections without median barriers. Injury profiles on sections with barriers were not significantly different from those without barriers, but that finding could be an artifact of the dataset used.

INTRODUCTION

Median crossover accidents are generally severe and result in a high cost to society. Such accidents also have a greater potential for creating liability, both because of their severity and because of their inherent link with design deficiencies, i.e., no or weak median barriers. The Washington State Department of Transportation (WSDOT) is creating a systematic process for determining median barrier requirements on state highways. Methodologies are also being developed to cost-effectively address roadway safety design issues related to divided highways. WSDOT is re-examining current median barrier installation guidelines, which are ad hoc and do not make use of current multivariate statistical techniques that can account for effects from roadway geometric factors, traffic, and the environment.

Studies on median barrier requirements do not currently benefit from advanced analysis of median crossover accidents. Limited safety analysis techniques have been researched and used in the area of median barrier accidents (see, for example, Graf and Winegard 1968; Ross 1974; and Bronstad, Calcote and Kimball 1976). The current version of the American Association of State Highways and Transportation Officials (AASHTO) Roadside Design Guide (AASHTO 2002) suggests the use of a simple bivariate analysis of average daily traffic (ADT) and median width as a guide in determining median barrier requirements. The relationship is a simple decision rule chart: if ADT is above a certain value and median width below a certain value, the chart shows the recommended type of median barrier. Currently, the WSDOT uses a related ADT vs. median relationship when examining median barrier requirements.

CURRENT GUIDELINES

The "median treatment study on Washington State Highways" report by Glad et al. (2002) presents a fairly clear picture of current WSDOT practice. The report states,

> WSDOT guidance for the installation of median barriers (Figure 700-7 in the WSDOT *Design Manual*) is essentially the same as that provided in the AASHTO Roadside Design Guide. AASHTO guidance was developed using a study conducted by the California DOT in 1968. This guidance provides criteria for median barrier installation based on the average daily traffic (ADT) and width of median. The criteria for barrier protection indicate that the designer should "evaluate the need for barrier" on all medians up to 32.8 feet in width when ADT is 20,000, or greater. Barrier is optional for all medians between 32.8 feet and 50 feet or when the median is less than 32.8 feet and the ADT is less than 20,000. AASHTO indicates "barrier not normally considered" for median widths greater than 50 feet. (Glad et al. 2002)

The report also suggests that there is significant variability in median barrier installation practice at the state level. It states that "the North Carolina Department of Transportation recommends median barrier installation for all new construction, reconstruction, and resurfacing projects with medians 70 feet or less in width. Cal-Trans has adopted more stringent warrants based on ADT for freeways with medians less than 75' in width."

The Glad et al. (2002) study also recommends that the 50-foot width requirement for median barriering is optimal from a benefit/cost analysis of observed crash histories on Washington State highways. This finding is based on the comparison of societal costs of median crossovers on sections with and without cable median barrier treatments. The study carefully notes that it did not account for "regression-to-the-mean" effects when considering the impacts of median barrier installation. That is, would median crossovers decrease naturally to a lifetime mean, even without cable barrier installation? If so, how would that affect the true impact of cable barriers? However, limitations exist in the Glad et al. study. The above methodologies do not facilitate accurate predictions of median crossover accidents. The study is strictly historical and is fairly sensitive to whether fatalities occurred in the observed time period. Because of the high cost of fatalities, the occurrence of a fatal collision can significantly improve the benefit/cost ratios for a given type of median barrier treatment.

There is much to gain from accurate predictions of median crossover frequencies. Installing median barriers effectively reduces median crossover rates to near zero, although there is always a small chance of median barrier penetration. However, it is not beneficial to install median barriers everywhere on the road network because the frequency of other types of accidents tends to increase in the presence of barriers even while they reduce the propensity for median crossover accidents. Median barriers reduce the area that vehicles have to recover, or escape, from an accident in the roadway, and they cause rebound accidents when vehicles strike the barrier and bounce back to strike another vehicle traveling in the same direction. Generally, median barriers reduce the frequencies of injury accidents, particularly severe accidents. However, contrary examples exist that show that the case is not that simple. For example, a before-and-after study (Seamons and Smith 1991) found a total increase of roughly 14 percent in all injury (including fatal injury) accidents when median barriers were installed at freeway locations. Median crossover accidents tend to be more serious, with a higher probability of fatalities, whereas barriers might sometimes increase the probability of some types of accidents but most at lesser severity. Median barriers also carry a maintenance cost that is threefold: direct monetary cost, traffic delays, and risk to road crews. Cost-benefit analysis that considers the whole project in addition to the predictive models is therefore an important part of developing a full picture to help plan effective measures. Departments of transportation attempt to strike a balance when scheduling sections for barrier installation, and the decision-making benefits from an accurate prediction of median crossover frequency. The trade-offs between economic and safety priorities and the possibility to strike a balance support the need for advanced median crossover accident analysis.

PREDICTION MODELS

With an objective to obtain accurate predictions of median crossover frequencies, this project sought to develop advanced predictive models for median crossover frequencies. There are some common modeling problems in the estimation of count models. Unobserved heterogeneity is a common occurrence in accident databases, leading to overdispersion. In accident databases such as median crossover frequencies, a significant number of zero accident counts exists, which suggests possible latent processes at work leading to spurious overdispersion. The other primary concern is possible serial correlation, which is of two types. The serial correlation among the observations from sections in the same time frame does not seriously affect the modeling assumptions and also is likely to be insignificant in a dataset with data aggregated over a long time period. The serial correlation between the observations for a single section over many periods (years, usually) violates the independence assumptions for unobserved error terms and, if left unaccounted for, causes the coefficient estimates to be inefficient and the estimated standard errors to be biased. Since median accident distributions occur in much lower numbers than for other accident types, and likely in a more sporadic fashion, a longitudinal history of median accident counts is used to examine fundamental propensities in the long-term for median crossovers. This leads to possible section-specific temporal correlation.

The objective of this research was to arrive at multivariate models that account for forms of unobserved heterogeneity and section-specific temporal correlation. This could be achieved by formulating variations of count models derived from the basic structure of the Poisson model to incorporate the effects of unobserved heterogeneity and section-specific temporal correlation. These models would help evaluate the effectiveness of barrier installations on roadway sections in Washington State in reducing the frequency of median crossovers and accidents of different severities. Also, from a programming standpoint, the efficiency and consistency of modeling median crossovers would be improved. Future research could be devoted to further improving the state-of-the-practice in Washington State for programming median crossover safety and devising better schemes for reducing median crossover frequency and severity.

In the next chapter, previous research to develop different modeling methodologies for accidents, median accidents in particular, is documented and discussed in the context of the current study. The following chapter presents the methodology employed in performing the study in detail. This includes the methodology for estimating different models for median crossover accidents and accidents of different severities, as well as the procedures performed to test the explanatory power of the factors included in the models. Next, the empirical setting for estimating the models and performing the tests for models and variables is discussed. The data mining efforts are then described, and the comprehensive data set is studied in terms of the descriptive statistics. Last, the modeling results and their bearing on the current study are elaborated.

LITERATURE REVIEW

A review of accident modeling literature reflects the variety of methods that have been used to model accidents. The conventional method for modeling accidents is to use linear regression to model accident rates, a continuous number (for example Mulinazzi and Michael 1969; Shah 1968). This is a straightforward method that models the number of accidents per million vehicle miles (known as the accident rate) for a given roadway segment. Research has identified that linear regression has many drawbacks, such as lack of distributional properties to describe random, sporadic, vehicle accident events on the road. Hence, accident frequency counts modeled with linear regression can result in inconsistent parameter estimates. The other alternative for modeling accident frequencies is to use count models such as Poisson and negative binomial (and their suitable variations) models.

More recent methods used for modeling accident frequencies include models such as the Poisson and negative binomial (see, for example, Shankar, Mannering and Barfield 1995; Poch and Mannering 1996; Milton and Mannering 1998) and the zero inflated Poisson and zero inflated negative binomial (for example, Shankar, Milton and Mannering 1997). The Poisson model, while possessing most of the above specified desirable statistical properties (that linear regression lacks), was found unsuitable for overdispersed data. Overdispersion occurs when unobserved heterogeneity in the data result in a scenario in which variance of the counts exceeds the mean, thereby violating the assumption of numeric equality between variance and the mean inherent in the Poisson model for count data. As a result, employment of the Poisson model for overdispersed data results in underestimation of coefficient variances and a likelihood of accidents. Shankar et al (1995) showed that the negative binomial model incorporates overdispersion because of unobserved heterogeneity and thus avoids the underestimation of coefficient variances and likelihoods. Shankar et al (1997) suggested that ZIP structure models are promising and have great flexibility in uncovering processes affecting accident frequencies on roadway sections observed with zero accidents and those with observed accident occurrences. The latent processes that determine the safety behavior of a roadway section are modeled by using a suitable count modeling structure that models the accident count probability as a sum of latent and non-latent count probabilities. This flexibility allows highway engineers to better isolate design factors that contribute to accident occurrence and also provides additional insight into variables that determine the relative accident likelihoods of safe versus unsafe roadways. The research revealed that the generic nature of the models and the relatively good power of the Vuong specification test (Green 1999) used in the non-nested hypotheses of model specifications offer roadway designers the potential to develop a global family of models for accident frequency prediction that can be embedded in a larger safety management system.

Previous research on median barrier requirements has been restricted to limited analyses of median cross-over accidents (for example Graf 1968; Ross 1964; and Bronstad 1976). The American Association of State Highways and Transportation Officials (AASHTO), in its current version of the Roadside Design Guide 1966, suggests that simplistic bivariate analysis of average daily traffic (ADT) and median width relationships be used as guidelines for examining median barrier requirements in the absence of site-specific data. The WSDOT uses a similar ADT-median width combination principle to examine its median barrier requirements. Previous research on accidents and their relationships to roadway geometrics, environmental factors, and functional class (see, for example, Seamons 1991, Shankar et al 1998, and Guo 1996) has indicated that several main effects and interactions of such factors play a critical role in determining accident causality. To attain accurate and reliable forecasts of median crossover frequencies and other accident types associated with medians, multivariate models incorporating the effects of roadway geometrics, weather, and traffic variables is necessary. Poisson and negative binomial models would be a good starting point for modeling median-related accidents.

Also, the time-series nature of multi-year data as used in the current study presents serial-correlation issues. In addition to over-dispersion, the model must handle the violation of the independent observations assumption (because of serial-correlation across time) made by both the Poisson and NB models for the estimation to remain in effect. Previous research in these matters has explored modeling techniques to handle section-specific serial correlation. Count model counterparts to the fixed effects and random effects least squares models were tested for their ability to account for temporal correlation in count data.

In 1984, Hausman, Hall, and Griliches examined fixed effects and random effects Poisson and negative binomial models for panel count data of patents filed and received by firms. The statistical models developed as part of this research were applications and generalizations of the Poisson model. After rewriting the Poisson model as a function of independent variables, two important issues had to be dealt with: (1) Given the panel nature of the data, how can one account for persistent individual (fixed or random) effects, and (2) How does one introduce the equivalent of disturbance-in-the-equation into the analysis of Poisson and other discrete variable models? The first problem was solved by conditioning on the total sum of outcomes over the observed years, while the second problem was solved by introducing an additional source of randomness, allowing the Poisson parameter to be itself randomly distributed, and compounding the two distributions. To incorporate the random effects into the Poisson model, the Poisson parameter was multiplied with a random, firm-specific effect. This way, the Poisson parameter became a random variable rather than a deterministic function of the exogenous variables. The fixed effects approach was to condition on the firm-specific effect and apply the conditional maximum likelihood techniques. The relevant likelihood functions and associated computational methods were analyzed in the study. The above approaches did not, however, solve the problem of overdispersion. Overdispersion caused by unobserved heterogeneity in the data was handled by letting each firm's Poisson parameter be randomly distributed, thus estimating the negative binomial model extensions for the data. This resulted in a fixed effects negative binomial (FENB) model and a random effects negative binomial (RENB) model. To arrive at the FENB specification, a convenient distribution was found for the sum of patents for a given firm, on which conditioning was done. Once the conditioning had been done on the firm-specific effect parameter, the Poisson parameter was deterministically specified. Thus, the FENB model did not allow for section-specific variation. The RENB model essentially layered a random "location and time" effect on the parent negative binomial by assuming that the overdispersion parameter was randomly distributed across groups. The key advantage of this approach is that the variance-to-mean ratio, which is likely to grow with the expected mean of accidents, is not constrained to be constant across locations, as it is in the case of the cross-sectional negative binomial. The RENB model allows for randomly distributed, section-specific variation. In the case of accident frequency, it is likely that section-specific effects will be important. The results indicated a huge improvement in likelihood from the conventional Poisson model to the Poisson model with the section-specific effects and from the Poisson to the NB model. But the increase in likelihood was not very high from the NB to the RENB model. Testing for the presence of fixed effects or random effects in the dataset and correlating between the unobserved section-specific effects and the explanatory variables were found to be helpful in modeling the data.

In other research, Guo (1996) considered a negative multinomial model that made allowance for independent observations. Guo worked on a dataset containing quarterly counts of surgical procedures performed from 1988 to 1991 in a national sample of American hospitals. Guo started with a conventional Poisson regression model and subjected the multiple counts in the same cluster to a cluster of specific random effects representing the unobserved effects shared by all the counts of the cluster. A gamma distributed, cluster-specific effect in the formulation resulted in the negative multinomial model. Similar ideas were behind the strategies adopted for correlated observations in linear regression modeling (Searle 1987) and event history analysis (Clayton 1978, Cox and Oakes 1984, Guo and Rodriguez 1992). A comparison of the results from the Poisson, the negative binomial, and the negative multinomial models revealed that the negative multinomial model improved the fit to the data tremendously over the Poisson model, which was a strong indicator of the presence of hospital-specific effects. In comparison to the negative multinomial model, the Poisson model seriously underestimated the standard errors of the parameter estimates, thus overestimating the t-ratios. Although the negative binomial model improved the fit over the Poisson model and also accounted for overdispersion, the standard errors of some hospital-specific and time invariant parameters estimated by the NM model were found to be lower than those from the NB model, since the NM explicitly controlled for the unobserved time-invariant, cluster-specific effects.

Even though the presence of unobserved, time-invariant, cluster-specific effects allowed the NM model to be more efficient than the NB model, there is no strong evidence to show that the NM model incorporated excessive zero accident counts. It is possible that many zero accident counts were repeating in a cluster over time, and the NM model may have omitted this effect, leading to a wrongly specified model. Therefore, zero altered models did play a significant role in modeling median crossovers for the current study.

In the next chapter, the methodological approach applied to the dataset is discussed. The model formulations provided in the section are discussed with regard to median safety. Detailed analysis and examples for illustration can be found in econometric textbooks.

METHODOLOGICAL APPROACH

In developing the decision matrix for median barrier installation, two types of statistical models were examined. The first model examined the contribution of roadway geometrics, median widths, weather, traffic volumes, and roadside characteristics to annual median crossovers. In this model type, issues relating to unobserved effects and time trends were addressed. The integrated database consisting of weather, geometrics, volumes and roadside characteristics was only a partial set of factors that are observed to be correlated with median crossovers. Unobserved effects relating to driver characteristics, local environmental fluctuations not captured completely in precipitation data, and vehicle speeds also contributed to the complexity of the statistical problem. Time trends occurred in the observed dataset because of the longitudinal nature of the data. As a result, the statistical framework was adapted to account for repeated information occurring annually in the 1990-1994 dataset. This type of model is referred to in this study as the “crossover frequency” model, i.e., a model that forecasts the mean number of yearly median crossovers.

The second type of statistical model examined the contribution of roadway geometrics, median widths, weather, traffic volumes and roadside characteristics to the annual societal cost of median crossovers. The cost-level model in this study is referred to as the “median crossover societal cost” model, i.e., a model that forecasts the annual cost of median crossovers in 2002 dollars. The average cost of a single median crossover in 2002 dollars is approximately $445,000. The intent behind this model was to determine if the cost-level examination and frequency-level examination from the first model provided a common point of decision-making for barrier installation policy. For example, let us assume that the median crossover frequency model suggests that, controlling for all other factors, medians 61 feet or wider experience significantly fewer median crossovers than narrower medians. Let us also assume that the societal cost model suggests that medians 51 feet or wider are expected to contribute to significantly lower societal costs than those 50 feet or narrower. A common point for decision making begins to emerge here. One might consider the following design policy in this case:

- Barrier all medians less than or equal to 50 feet wide.
- Do not install barriers for medians wider than 60 feet.
- Consider case-by-case assessments for barriering medians in the 50-foot to 60-foot range.

In addition, the effectiveness of the installation of median barriers on the selected road sections was tested. The 1990-94 accident data for sections without median barriers were used as the "without median barrier" data, and 1990-94 accident data for sections containing barriers (and similar in attributes to the sections without barriers) were used as "with median barrier" data. The median barrier sections were chosen so that they were physically near the sections without median barriers and hence very closely represented the behavior of the non-barriered sections after median barriers had been placed in them. A preliminary contingency analysis of total crash counts determined that barrier installation did not necessarily uniformly increase overall crash counts on the highway section. Some sections reported lower crash counts as a result of barrier installation, whereas others reported higher crash counts. To examine the characteristics of sections that contributed to this difference in crash profiles, we estimated a statistical model for all reported crashes for sections with and without median barriers. In estimating the model, traffic volumes, precipitation, geometrics, and the presence of a median barrier were controlled for.

CROSSOVER FREQUENCY MODELING FRAMEWORK

The NM model is suggested as the proper means of estimation when serial correlations are present in the dataset. The NM model was used to fulfill the ultimate objective of this project, to forecast accident frequency variation by accident types. By employing the NM model, the probability of accident counts on non-barriered and barriered sections could be estimated from the Gauss program on the basis of the knowledge presented below:

The derivation of the NM model begins with the Poisson probability density function, which is an expression for the probability of the frequency equaling a particular count (Guo 1996):

$$\mathrm{P}(Y_{it} = y_{it}) = \frac{\mathrm{e}^{-\lambda_{it}} \lambda_{it}^{y_{it}}}{y_{it}!}; \qquad y_{it} = 0,1,2,\ldots, \tag{1}$$

where Y_{it} is the observed frequency of accidents in section i at time t, and λ_{it} is the mean of the number of accidents. The Poisson model is estimated by defining:

$$\ln \lambda_{it} = x_{it}\beta, \tag{2}$$

where x_{it} is a vector of section- and time-specific explanatory variables, and β is a vector of coefficients to be estimated. To account for the section-specific variation, the procedure is similar to that for the NB model. A random error term is added to the expression for the mean,

$$\ln \lambda_{it} = x_{it} \beta + \varepsilon_i, \tag{3}$$

where ε_i is a section-specific (not observation-specific as in the NB model) random error term, $\exp(\varepsilon_i)$ is assumed to be an independently and identically distributed gamma with mean 1 and variance $\alpha = 1/\theta$. The assumption of mean 1 does not cause a loss of generality if equation (3) includes an intercept term.

The conditional joint density function of all individual counts for a particular section i, given that the individual counts are distributed by equation (1) and conditioned on ε_i, can now be written as:

$$\mathrm{P}(Y_{i1} = y_{i1},...,Y_{it_i} = y_{it_i} \mid \varepsilon_i) = \prod_{t'=1}^{t_i} \mathrm{P}(Y_{it'} = y_{it'} \mid \varepsilon_i), \tag{4}$$

where t_i denotes the number of time periods observed for section i. This assumes that the accident counts in different sections are independent. This is not unreasonable because these sections are generally not next to each other and will therefore only share minimal unobserved effects. The unconditional joint density function for the negative multinomial distribution can now be derived by integrating equation (4) and by using the assumed distribution of $\exp(\varepsilon_i)$ to give:

$$\mathrm{P}(Y_{i1} = y_{i1},...,Y_{it_i} = y_{it_i}) = \frac{\Gamma(y_i + \theta)}{\Gamma(\theta) y_{i1}!...y_{it_i}!} \left(\frac{\theta}{\eta_i + \theta}\right)^{\theta} \left(\frac{\eta_{i1}}{\eta_i + \theta}\right)^{y_{i1}} ... \left(\frac{\eta_{it_i}}{\eta_i + \theta}\right)^{y_{it_i}}, \tag{5}$$

where Γ is the gamma function, $\eta_{it} = e^{x_{it} \cdot \beta}$, $\eta_i = \eta_{i1} + ... + \eta_{it_i}$, and $y_i = y_{i1} + ... + y_{it_i}$. Recall that the variance of $\exp(\varepsilon_i)$ is $\alpha = 1/\theta$. The degenerate case, when each section has only one observation, (i.e., there is no section-specific correlation) yields the negative binomial distribution.

The expected value, variance, and covariance for the NM model are:

$$\mathrm{E}(Y_{it}) = \eta_{it}, \qquad \mathrm{Var}(Y_{it}) = \mathrm{E}(Y_{it})[1+\alpha\,\mathrm{E}(Y_{it})], \qquad \mathrm{Cov}(Y_{it}, Y_{it'}) = \alpha\eta_{it}\eta_{it'}. \tag{6}$$

A likelihood function is written using equation (5) to give equation (*7*):

$$\mathrm{L}(\beta,\theta) = \prod_{i=1}^{n} \frac{\Gamma(y_i+\theta)}{\Gamma(\theta)} \left(\frac{\theta}{\eta_i+\theta}\right)^{\theta} \left(\frac{\eta_{i1}}{\eta_i+\theta}\right)^{y_{i1}} \cdots \left(\frac{\eta_{it_i}}{\eta_i+\theta}\right)^{y_{it_i}}, \tag{7}$$

where n is the total number of sections. The coefficients β and $\alpha = 1/\theta$ are estimated by maximizing the likelihood function (7) (see, for example, Green 1999).

Alternative count models for median crossover analyses were also examined. These models did not substantively change the findings for critical contributors to median crossovers. Appendix B contains a description of alternative models and associated results.

MEDIAN CROSSOVER SOCIETAL COST MODELING FRAMEWORK

In considering the impact of median characteristics on the annual societal cost of crossovers, FHWA-recommended cost estimates were used in valuating the societal costs of observed median crossovers. The average cost of median crossovers is approximately \$445,000 (in year 2002 dollars). Using this value, we estimated a regression model of societal cost as follows:

$$\ln C_{it} = Z_{it}\gamma + \zeta_i \tag{8}$$

where Z_{it} is a vector of section- and time-specific explanatory variables, γ is a vector of coefficients to be estimated, and C_{it} is the societal cost in year 2002 dollars of yearly median crossovers for section "I" in year "t."

PRELIMINARY CONTINGENCY TEST FOR COMPARISON OF OVERALL CRASH COUNTS ON BARRIERED AND NON-BARRIERED SECTIONS

The contingency test was performed to decide whether the characteristics of the non-barriered sections versus the corresponding barriered sections were significantly different. Before estimating the models, it was important to ensure that the median barrier sections chosen to represent the "after" data would act as appropriate surrogates. Because the barriered sections were chosen from the sections nearest to the non-barriered sections (within 2 miles of each end of a non-barriered section), other factors such as AADT or number of curves may have affected the number of accident counts. If such factors were significantly different for barriered versus non-barriered sections, the estimated models may not have provided any useful and accurate information.

In general, effects in a contingency table are defined as relationships between the row and column variables; i.e., it is tested if levels of the row variable are differentially distributed over levels of the column variables. Hypothesis tests on contingency tables are based on the Chi-squared statistic. Significance in this hypothesis test means that interpretation of the cell frequencies is warranted. Non-significance means that any differences in cell frequencies could be explained by chance. In other words, if the computed Chi-square value is less than the critical value from the table, the null hypothesis cannot be rejected. In this case, that would mean that the cell frequencies in the contingency table among severity types were the same across non-barriered and barriered sections.

Frequency tables of two variables presented simultaneously are called contingency tables. Contingency tables are constructed by listing all the levels of one variable as rows in a table and the levels of the other variables as columns, then finding the joint or cell frequency for each cell. The cell frequencies are then summed across both rows and columns. The sums are placed in the margins, the values of which are called marginal frequencies. The lower right hand corner value contains the sum of either the row or column marginal frequencies, which both must be equal to N (Stockburger 1996).

The procedure used to test the significance using contingency tables is similar to that of all other hypothesis tests. The test begins as follows:

The first step is to develop a contingency table as shown in Table 1, which presents the roadway section on State Route 2 from 1990 to 1994 as an example. The table contains the accident counts by types: property damage only, possible injury, and evident injury-involved accidents.

Table 1. Non-barriered and barriered sections of State Route 2

Observed Cell Frequency (O)										
BARRIER						**NON-BARRIER**				
Year, ID	PDO	P-INJ	INJ	Total		Year, ID	PDO	P-INJ	INJ	Total
1990, 500	7	2	1	10		1990, 1	0	0	1	1
1991, 500	5	0	2	7		1991, 1	1	0	0	1
1992, 500	3	0	0	3		1992, 1	0	1	0	1
1993, 500	8	4	0	12		1993, 1	0	0	1	1
1994, 500	2	1	0	3		1994, 1	0	0	0	0

The next step in computing the Chi-squared statistic is the computation of the expected cell frequency for each cell. This is accomplished by multiplying the marginal frequencies for the row and column (row and column totals) of the desired cell and then dividing by the total number of observations. The formula for computation can be represented as follows:

$$\text{Expected Cell Frequency} = (\text{Column Total} \times \text{Row Total}) \div N$$

Table 2 shows the expected cell frequency calculated from the values in Table 1, row 1 only.

Table 2. Expected cell frequency table for row 1 in Table 1.

1990	PDO	P-INJ	INJ	Total
Barrier	7	2	1	10
Non-barrier	0	0	1	1
Total	7	2	2	11

Expected Cell Frequency (E)			
1990	**PDO**	**P-INJ**	INJ
Barrier	**6.36**[1]	**1.82**	1.82
Non-barrier	0.64	0.18	0.18[2]

Example:

[1]6.36 = (7 × 10)÷11

[2]0.18 = (2 × 1)÷11

The next step is to subtract the expected cell frequency from the observed cell frequency for each cell (O-E) and then divide the squared difference by E. The result is shown in Table 3.

Table 3. The squared differences divided by the expected cell frequency for each cell

$(O\text{-}E)^2/E$					
BARRIERED			**NON- BARRIERED**		
PDO	P-INJ	INJ	PDO	P-INJ	INJ
0.064	0.018	0.368	0.636	0.182	3.682

The last step is to sum the values to get the computed Chi-square value. Thus, for this section of SR 2, the Chi-square value is 4.950. The degree of freedom (df) is equal to (number of rows - 1) × (number of columns -1). In this case, the df is 2. The critical value from the Chi-square table is 5.99 at $\alpha = 0.05$. Since 4.95 is less than 5.99, the interpretation is that the accident counts are the same in barriered and non-barriered sections. In other words, we can conclude that the two sections involved in testing the effectiveness of median barriers are not different because of other factors that might contribute to the number of accident counts in these sections.

The interpretation of the cell frequencies may be guided by the amount each cell contributes to the Chi-squared statistic, as seen in the $(O\text{-}E)^2/E$ value. In general, the larger the difference between the observed and expected values, the greater this value. If the value of the observed Chi-square statistic is less than the expected value, then the model of no effects cannot be rejected and the table is not significant. It can be said that no effects were discovered. In this case an interpretation of the cell frequencies is not required because the values could have been obtained by chance alone.

COUNT MODEL FRAMEWORK FOR COMPARISON OF BARRIERED AND NON-BARRIERED SECTIONS

In addition to the contingency test, a test for identifying factors that contribute to differences in barriered and non-barriered sections was conducted. This test employed the familiar negative binomial framework, with the median barrier indicator as a dummy effect interacting with explanatory variables. Formally, this test is denoted as follows:

$$\ln \lambda'_{it} = Z_{it}\gamma + W_{it}D\theta + \zeta_i \tag{9}$$

where Z_{it} is a vector of section- and time-specific explanatory variables, and γ is a vector of coefficients to be estimated, λ'_{it} is the overall yearly crash count for section "I" in year "t," W_{it} is a vector of section- and time-specific explanatory variables interacted with the median barrier dummy "D", and θ is a vector of coefficients to be estimated for that interaction. If θ is significant, then it can be said that certain contributing factors in median barrier sections are significant over and above the baseline non-barrier effect.

EMPIRICAL SETTING

This chapter presents information regarding the data, framework, and criteria for model analysis and evaluations. The Washington State highway system contains over 7,000 centerline miles of state highways. The annual records in the database developed by WSDOT contain almost every roadway and roadside incident that occurs during the year. The panel data in this research consisted of five years (1990–1994 inclusive) of annual accident counts for 275 roadway sections in Washington State. The panel was balanced, with all sections having a full five-year history. This panel represented all sections (longer than 2,624 ft) without median barriers on divided state highways. The reason that only sections longer than 2,624 ft were selected is that about 95 percent of shorter sections on divided highways have barriers, and those without medians are more affected by access controls and intersections (Gudmundur and Shankar 2003). In addition to accident information, the other items extracted from the database included roadway geometrics, median characteristics, and traffic volumes. The database did not provide any of the weather information required for the study. A GIS-based technique was used to correlate weather information with roadway sections. The GIS program ArcView 3.2 was used to match each roadway section to its weather attributes stored in the historical weather database provided by Western Regional Climate Center. The mapping criteria involved linking the non-median barrier roadway sections to the nearest corresponding weather stations. Each weather station collected climate data that include daily, monthly, and annual temperature; precipitation; and snowfall, including snow depth, with records dating back to 1948. If the weather data of any nearest station were unavailable, the data for the second nearest weather station were selected instead. More information about the procedures of mapping weather stations with roadway sections is provided in Appendix A. Table 4 provides descriptive statistics of key variables in the without median barrier dataset.

Table 4. Key statistics of without-median barrier data

Variable	Mean	Std. Dev.	Min	Max
Number of crossover accidents in section	0.24	0.65	0	7
Number of total accidents in section	16.45	21.87	0	182
Number of property damage accidents in section	9.21	12.05	0	96
Number of possible injury accidents in section	3.54	6.53	0	64
Number of injury accidents in section	3.70	4.56	0	35
AADT	37,355	36,975	3,350	172,560
AADT per the number of lanes	7,445	5,830	835	28,690
Single truck percentage	4.20	1.22	1.90	10
Double truck percentage	7.76	4.62	0.55	17.80
Truck-train percentage	2.21	1.60	0	7
Percentage of AADT in the peak hour	10.72	7.314	0	19.40
Speed limit	60	5.5	35	65
Maximum median shoulder width in feet	5.31	2.49	0	18
Minimum median shoulder width in feet	4.48	1.6833	0	10
Percentage of medians narrower than 40 feet	32.36	46.80		
Percentage of medians between 40 and 50 feet wide	11.64	32.08		
Percentage of medians between 50 and 60 feet wide	5.81	23.42		
Percentage of medians wider than 60 feet	50.18	50.02		
Percentage of 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	4
The number of horizontal curves in section	2.75	2.86	0	29
The number of horizontal curves per mile	1.44	0.96	0	5
Maximum horizontal central angle in degrees	30.29	23.88	0	111.49
Minimum radius of horizontal curve in feet	4267.24	4875.08	0	38,400
Average annual snow depth in inches	19.44	45.66	0	652*
Average annual precipitation in inches	29.86	21.98	4.53	131.74
Number of grade changes	3.865	4.089	0	28
Average roadway width	57.42	15.47	24	121

* Weather station data for mountainous section

Ideally, to assess the impact of median barriers, comparing crash counts on the same section with and without barriers would have provided greater certainty about the relative impact of barriers. However, data on before-and-after characteristics for 275 section were unavailable on a consistent basis. Hence, an alternative method for assessing the relative impact of barriers was employed. By controlling for traffic volumes, weather, and geometric and roadside effects on similar sections that were proximate to non-barriered sections, a comparison of overall crash counts would provide some insight into the relative impact of barriers. Proximate sections were defined as sections that were within 2 miles and had comparable ADT and weather characteristics. Accordingly, 31 median barrier sections from 1990 to 1994 were selected and compared with 31 non-barriered proximate sections. The mean values of key variables for this 31-section dataset are shown in Table 5.

Table 5. Mean values of key variables of the comparison dataset.

Variable	**Mean**	
	With Barrier	Without Barrier
Number of total accidents in section	18.30	17.10
Number of property damage accidents in section	10.70	9.65
Number of possible injury accidents in section	4.54	4.43
Number of injury accidents in section	3.10	3.00
AADT	56,225	50,250
Per-lane AADT	10,375	9,035
Speed limit	61	57.90
Average roadway width	63.38	62.19
Length of the roadway section in miles	0.87	1.24
The number of interchanges in section	0.85	0.77
The number of horizontal curves per mile	2.26	1.60
Minimum radius of horizontal curve in feet	5,557	2,937
Average annual precipitation in inches	36.83	31.34
Average snowfall in inches	16.98	18.13
Number of grade changes	1.29	2.03

A t-test of sample means was conducted to verify whether the means of the key variables shown in Table 5 between the non-barriered and barriered median datasets were statistically different. The results showed that, between the two datasets, the means of these selected variables were insignificantly different from each other.

MODEL ESTIMATION RESULTS

MEDIAN CROSSOVER FREQUENCY MODEL

Table 6 presents the results on significant variables correlated to yearly median crossovers.

Table 6. Negative multinomial model for yearly median crossovers.

Variable	Estimated Coefficient	Standard Error	T-statistic
Constant	-1.73	0.19	-8.95
Percentage of AADT per the number of lanes <= 5000	-0.85	0.22	-3.83
Length of section in miles	0.18	0.06	3.25
Indicator for sections where medians are wider than 60 feet	-0.78	0.18	-4.24
The number of interchanges in the section	0.30	0.12	2.55
Median surface indicator (1 if median is paved, 0 otherwise)	0.64	0.34	1.85
Interaction between low average annual precipitation and the number of horizontal curves per mile (1 if average annual precipitation <=18 inches and the number of horizontal curves per mile <= 0.5, 0 otherwise)	-0.90	0.44	-2.03
Interaction between high average annual precipitation and the number of horizontal curves per mile (1 if average annual precipitation >30 inches and the number of horizontal curves per mile <= 0.5, 0 otherwise)	0.51	0.28	1.83
α	0.46	3.03	0.72
Restricted log-likelihood (All parameters = 0, α ~ 0)	-6253.34		
Log-likelihood at convergence	-725.23		
ρ^2	0.88		
Number of observations	1,375		

Table 6 shows that the model convergence is fairly good, with a rho-squared measure of 0.885. The predictions for the mean median crossovers per year for the entire

275-section dataset indicate that an average of 0.25 crossovers is expected per year. The observed mean crossover per year is 0.24.

The results show that all the variables included in the model are highly significant and thus have a high level of confidence (exceeding 90 percent). Factors positively correlating with median crossover accidents include the number of interchanges in the section, paved median surface, length of section, and the interaction between high average annual precipitation and the number of horizontal curves per mile. Sections with average annual precipitation in excess of 30 inches and fewer than 0.5 horizontal curves per mile are likely to experience higher median crossovers.

Factors that decrease median crossover frequencies include median widths in excess of 60 feet, per-lane ADT of less than 5,000, and the interaction between high average annual precipitation and the number of horizontal curves per mile. Sections with average annual precipitation of less than 18 inches and fewer than 0.5 horizontal curves per mile are likely to experience fewer median crossovers. The median width variable has a significant impact on median crossover frequency. Sections with a median width of less than 60 feet are expected to experience twice the frequency of crossovers, controlling for other factors, than those with median widths in excess of 60 feet.

MEDIAN CROSSOVER SOCIETAL COST MODEL

Table 7 shows the results for the median crossover societal cost model. As mentioned previously, the societal cost model of median crossovers assumed an annual average cost of $445,000 (year-2002 dollars) per median crossover. This represents a significantly higher severity distribution than that of common types of crashes. The individual observation costs used in the model estimation, however, are costs assigned to reported severities as a weighted measure for any given year. Approximately 235 observations out of the 1,375-observation sample resulted in median-crossover-related costs. The statistical model is based on an ordinary least squares estimation of the natural logarithm of the per-mile societal cost, using per-lane ADT, median width in excess of 50 feet, paved median surface, and interactions between precipitation and horizontal curvature.

Table 7. An OLS model of the natural logarithm of median crossover societal cost per mile

Variable	Estimated Coefficient	Standard. Error	T-statistic
Constant	2.11	0.19	11.38
Per-lane AADT (1 if less than 5,000; 0 otherwise)	-1.05	0.22	-4.88
Indicator for sections where medians are wider than 50 feet	-0.73	0.21	-3.51
Median surface indicator (1 if median is paved, 0 otherwise)	3.01	0.50	6.0
Interaction variable between low precipitation and the number of horizontal curves per mile (1 if average annual precipitation < 18 inches and the number of horizontal curves per mile <= 0.5; 0 otherwise)	-0.51	0.36	-1.41
Interaction variable between high precipitation and the number of horizontal curves per mile (1 if average annual precipitation > 30 inches and the number of horizontal curves per mile <= 0.5;0 otherwise)	0.71	0.44	1.62
Adjusted R^2	0.07		
Number of observations	1,375		

Note that the societal cost model is based only on reported severities. As a result, information related to sections that reported no median crossovers in a given year was not used. While most sections reported a median crossover in at least one of the five years of the study, selectivity bias stemming from our ignoring information on zero crash sections may still have occurred. Factors that positively contribute to median crossover societal cost include paved median surfaces, and the interaction between high precipitation and horizontal curvature. Factors that decrease societal cost include median widths in excess of 50 feet, per-lane ADTs less than 5,000, and the interaction between low precipitation and horizontal curvature.

CONTINGENCY TEST

This test was intended to determine whether factors besides median barriers installed in median barrier sections contributed their effect the same way they contributed to non-barriered sections along the same route. The contingency table shown in Table 1 was extended to cover 31 barriered sections to be compared with the 31 nearest non-barriered sections. When this test was applied, we examined the overall crash profile of sections in terms of injury categories such as property damage, possible injury, and injury. The total number of rows in this table was 155 (31 sections multiplied by 5 years). The results showed that 148 out of 155 observations fell in the accepted region (a computed Chi-square value of less than 5.99 at a 95 percent confident level with 2 degrees of freedom). Thus, the assumption that the effect should be the same across the barriered and non-barriered sections was valid. Since seven observations fell in the rejected regions, these sections may have contained some properties that made the median barriered and non-median barriered sections' crash profiles different. Figure 1 shows differences in crash profiles for the seven observations that rejected similarity. Four state and Interstate routes contributed to this pattern: I-5, I-90, SR 205 and SR 410. Table 8 presents the injury profiles for these seven observations.

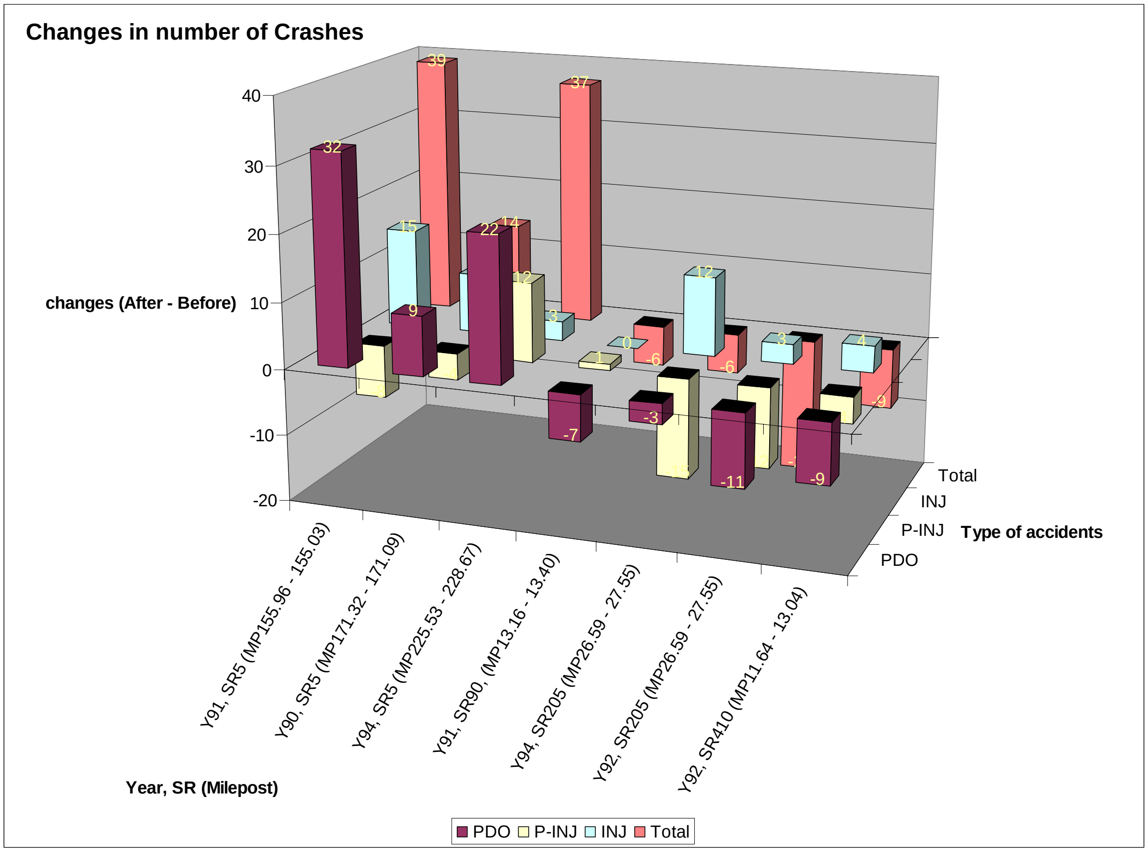

Figure 1. Differences in crash profiles between median barriered and non-barriered sections.

Table 8. List of the seven observations falling in the rejected-the-null-hypothesis region

	OBSERVED CRASHES								
	WITH BARRIER					WITHOUT BARRIER			
Route	Year	PDO	P-INJ	INJ	Total	PDO	P-INJ	INJ	Total
5	91	79	30	25	134	47	38	10	95
5	90	24	8	11	43	15	12	2	29
5	94	29	12	9	50	7	0	6	13
90	91	0	1	1	2	7	0	1	8
205	94	25	5	17	47	28	20	5	53
205	92	16	4	14	34	27	16	11	54
410	92	5	2	5	12	14	6	1	21

COUNT MODEL RESULTS FOR COMPARISON OF BARRIERED AND NON-BARRIERED SECTIONS

The dummy variable-based test represented in equation 9 was employed to examine the characteristics of sections that contributed to this difference in crash profiles. We estimated a statistical model for all reported crashes for sections with and without median barriers. In this test, a larger dataset was used with greater variation in traffic volumes and weather effects than that used in the contingency study. In estimating the model, traffic volumes, precipitation, geometrics, and the presence of median barrier were controlled for. The test determined that as the number of curves per mile increased in a section with barriers, the overall crash profile on those sections increased in comparison to similar sections with no barriers. The test also found that sections with two to five grade changes per mile and median barriers experienced fewer overall crashes than similar sections without median barriers. Injury profiles on sections with barriers were not significantly different from those without barriers, but that finding could be an artifact of the dataset used.

CONCLUSIONS AND RECOMMENDATIONS

A panel dataset consisting of five years (1990–1994 inclusive) of annual crash counts for 275 roadway sections in Washington State was used to examine the contribution of median widths to annual crossover crashes. The panel was balanced, with all sections having a full five-year history. Key findings from the study result in the following conclusions and recommendations:

CONCLUSIONS

- A statistical model of yearly median crossovers determined that, all factors controlled for, median widths in excess of 60 feet do not require barriering.
- A statistical model of weighted societal costs of median crossovers determined that, all factors controlled for, median widths under 50 feet may require barriering to minimize the societal costs of those crashes.
- A comparison of sections with and without barriers indicated that as the number of curves per mile increases in a section with barriers, the overall crash profile on those sections increases in comparison to similar sections with no barriers. Analysis also found that sections with two to five grade changes per mile and median barriers experience fewer overall crashes than similar sections without median barriers.

RECOMMENDATIONS

- It is recommended that WSDOT consider median widths narrower than 50 feet for mandatory barriering. In the 50- to 60-foot range, it is recommended that

WSDOT evaluate sections on a case-by-case basis. It is also recommended that WSDOT not install barriers for medians that are wider than 60 feet.

- It is recommended that WSDOT continue to evaluate these policy recommendations with the latest available crash histories. The likelihood that median crossovers will escalate exists, and it is prudent to ensure that such escalations are not significant to maintain policy robustness.

REFERENCES

American Association of State Highway and Transportation Officials (2002). *Roadside Design Guide*.

Bronstad, M.E., Calcote, L.R., and Kimball, C.E. (1976) *Concrete Median Barrier Research*. Report No. FHWA-RD-73-3. FHWA, June.

Clayton, D.G. (1978) A model for association in bivariate life tables and its application in epidemiological studies of family tendency in chronic disease incidence. *Biometrika* 65, pp.141-151.

Cox D.R., and Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall, London, UK.

Glad R.C., Albin R., Macintosh D. and Olson D. "*Median treatment study on Washington State Highways,"* WSDOT Research Report WA RD 516.1, 2002.

Graf, V.D., and Winegard, N.C. (1968). *Median Barrier Warrants*. Traffic Department of the State of California.

Greene, W.(1999) *Econometric Analysis*, 4th Ed. Prentice Hall.

Greenwood, M., and Yule, G.U.(1920). An Inquiry into the Nature of Frequency Distribution of Multiple Happenings. *Journal of the Royal Statistical Society, A*, Vol. 83, pp. 255–279.

Gudmunder, U.F., and Shankar, V.N.(2003). An Accident Count Model Based on Multi-Year Cross-Sectional Roadway Data with Serial Correlation, presented at the 82nd Annual Meeting of the Transportation Research Board, Washington D.C.

Guo, G.A. and Rodriguez G.(1992). Estimating a Multivariate proportional Hazards Model for Clustered Data Using the EM Algorithm with an Application to Child Survival in Guatemala. *Journal of the American Statistical Association* 87(420), pp. 969-976.

Guo, G. (1996). Negative Multinomial Regression Models for Clustered Event Counts. *Sociological Methodology*, Vol. 26, pp.113–132.

Hausman, J.A., Hall, B.H., and Griliches, Z.(1984). Econometric Models for Count Data with an Application to the Patents-R&D Relationship. *Econometrica*, Vol. 52, No. 4, pp. 909–938.

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.

Milton, J. and Mannering, F. (1998). The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies, *Transportation* 25, 395-413.

Mulinazzi, T.E. and Michael, H.L.(1969). Correlation of Design Characteristics and Operational Controls with Accident Rates on Urban Arterials, *Purdue University Road School Proc*.

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

Ross Jr., H.E. (1974). *Impact Performance and Selection Criterion for the Texas Median Barriers*. Research Report 140-8. Texas Transportation Institute, Texas A&M University.

Seamons, L. L., and Smith, R.N. (1991). *Past and Current Median Barrier Practice in California.* Final Report, CALTRANS-TE-90-2. California Department of Transportation.

Searle, S.R.(1987). *Linear Models for Unbalanced Data.* New York: John Wiley.

Shah, K. (1968). A Methodology to Relate Traffic Accidents to Highway Design Characteristics, Ohio University, 31 August.

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*, Vol. 27, No.3, pp. 371-389.

Shankar, V.N., Milton, J.C. and Mannering, F.L.(1997). Modeling Statewide Accident Frequencies as Zero-Altered Probability Processes: An Empirical Inquiry. *Accident Analysis and Prevention*, Vol. 29, No.6, pp. 829–837.

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. In *Transportation Research Record 1635*, TRB, National Research Council, Washington, D.C., pp 44–48.

Shankar, V.N and Chayanan, S. (2004). The Marginal Impacts of Design, Traffic, Weather, and Related Interactions on Roadside Crashes, presented at the 82nd Annual Meeting of the Transportation Research Board, Washington D.C.

Stockburger, W.D. (1996). Chi-square and Tests of Contingency Tables, *Introductory Statistics: Concepts, Models, and Applications*, Retrieved December 9, 2003, from http://davidmlane.com/hyperstat/chi_square.html, 1996.

Voung, Q.(1989). Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. *Econometrica* 57 no.2, pp. 307-333.

Western Regional Climate Center (2002). Historical Climate Information, Desert Research Institute, Retrieved March 2002, from http://www.wrcc.dri.edu/index.html

APPENDIX A –
GIS-Weather-Highway Network Integration Protocol
Procedures for Mapping Weather Stations to Washington State Highways

Install ArcView 3.2 or later
    Put extensions into
        ESRI/AV_GIS30/ARCVIEW/EXT32
            MadogGPS.avx
            MadogLRS.avx
            MadogMap.avx

            Coordinate.avx
            nearstneighbor2.avx

Put WA State GIS files into C:/gisosc in the WA directory structure beginning with /Geodata.

Run ArcView

Load the above extensions in File – Extensions.
Create New View
Select View
Change Tool Kit to Base Map
Add State with Base Maps – Add State – Select C: drive – OK

State appears in the view, and its name becomes Washington State.

You can turn features on/off by clicking on the checkbox next to them in the left window in the view window.

Create a table with the SR route information
To form the table Attributes of State Routes. It must be exported out.
    Change Tool Kit to Basic Arcview.
    Select the State Route theme in the view
    Theme – Table
        this opens an attribute table that now appears in the Table list
    Select the table Attributes of State Routes . (it should be open)
    File – Export as delimited text, and save this table in a good location.

Use this table along with the roadway section table to map SR-IDs to SR numbers in the roadway section table. Do this for example in MS Access. You should then have a table with at least SR#, SR-ID, SR, beginning-SRMP, end-SRMP. Call it RoadSections.txt.

Add tables into project by clicking Add button and browsing to table locations. Choose delimited text files if appropriate.

Add weather station table, it should have columns: weather station ID, lon, lat in degrees decimal.

Add roadway section table, it should have columns: SR#, SR-ID, SR, beginning-SRMP, ending-SRMP. Make sure the roadway section table is sorted by SR and by beginning SRMP.

Change Tool Kit to LRS Tools
Select Route Theme – Select – State Routes (LRS) – OK
Place Events – Map Your Data

Click on the Right button of the top two (not the Y X button)
Route theme – State Routes (LRS)
Route field – Sr-ID
Points
Table: Roadway Section table
Event field: Secid (ID of your road section)
Location field: Beginning SRMP
Offset field: <none>
OK

This creates a new theme in the view window, called Roadsections.txt (the name of the newly created input roadway section table)
Select it to plot your roadway sections on the State Route highway map in the View.

Change Tool Kit to GPS Tools
Load GPS Table

Add Table
Select the Weather Station table which has X (lon), Y (lat), in degrees decimal.

This creates a new theme in the view.
Select it to plot the weather stations on the State map in the View.

Change Tool Kit to Basic Arcview
Change table text themes to Shapefiles in

Theme – Convert to shapefile

Select the shapefile theme.
Theme - Table

to view the attribute table for the theme.

Select a point in the attribute table and zoom in on selection to view the selected area on the State map.

To create Lat Lon coordinates for the SRMP located road sections use: Coordinate Utility extension. To activate it the distance units must be set.

Select View
View – Properties – Distance units – Miles

Then select the chosen theme in the View window and convert it to a shapefile if needed.

Theme – Convert to shapefile

Select

Coordinate – Coordinate utility

Choose a shape theme for conversion, planar X, Y coordinates, choose the center of the shape. This adds X,Y columns to the shape Theme Table Attributes.

Select Nearest Neighbor

1 Nearest Neighbor

select road section shape theme for source

select weather shape theme for target

select (by shift clicking or ctrl-clicking) the two X, Y coordinates in the road section shape file, for source coordinates

select similarly from weather shape for target X,Y coordinates

select COOP as the ID field from the weather shape to transfer to road section shape.

Select distance range for nearest neighbor search

Wait for program to finish

Say OK to load new shape

Create Layouts to print graphs.

To change fonts in Layout select textbox and go to Window – Show Symbol Window

To get more options on points, select corner at the lower right on buttons with the mouse, and hold it to get a drop-down menu of other buttons.

APPENDIX B –
Alternative Statistical Models of Median Crossovers

ZERO INFLATED POISSON AND ZERO INFLATED NEGATIVE BINOMIAL MODELS

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 opposite direction accident, and “Y^*” denotes the crash count state for that roadway section, neither “Z” nor “Y^*” is observed, but only the observed opposite direction 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 (Shankar et al, 1997). 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. A priori assumptions regarding the density generators are required in formulating the mixture distribution. Formally, let Y_i be the annual number of accident counts reported for section *i*, and let p_i be the probability that section *i* will exist in the zero-count state over its lifetime. Thus 1-p_i is the probability that section *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,

$$Y_i = 0 \text{ with probability } p_i + (1 - p_i)e^{-\lambda_i} \tag{8}$$

and,

$$\mathrm{Y_i} = \mathrm{k} \text{ with probability } (1 - p_i)\frac{\Gamma(n_i + \theta)}{\Gamma(\theta)n_i!}\left(\frac{\theta}{\theta+\lambda_i}\right)^{\theta}\left(\frac{\lambda_i}{\theta+\lambda_i}\right)^{n_i} \tag{9}$$

where *k* is the number of accidents (positive numbers starting from one), with λ_i being the mean, and $\theta = 1/\alpha$, with α is the dispersion parameter. 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]}. If α is zero (t-stat less than 1.96), the ZIP model may be used and (9) reduces to

$$\mathrm{Y_i} = \mathrm{k} \text{ with probability } (1 - p_i)(\frac{e^{-\lambda i}\lambda_i^{y_i}}{y!}) \tag{10}$$

In (8) and (9), 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 $\log(\lambda_\mathrm{i}) = \mathbf{H}_\mathrm{i}\,\beta$,

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 $\log\left(p_i / 1 - p_i\right) = \tau \mathbf{B}_i \beta$, 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 $\mathbf{H}_i$. If τ is insignificantly different from zero, then the corridor is equally likely to be in the zero or non-zero lifetime state. The (8) and (9) combined provide the zero-inflated negative binomial (ZINB) model. (for example see Shankar and Chayanan 2004). The formulations shown in (8) and (9) follow established methods in Greene (1999).

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 in 1989. The Vuong test is a t-statistic-based test with reasonable power in count-data applications (Green 2000). The Vuong statistic (V-statistic) is computed as

$$V = \frac{\overline{m}\sqrt{N}}{S_m} \tag{11}$$

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. It should be noted that in case of the ZIP model, $f_1(.)$ being the density function of the ZIP distribution and $f_2(.)$ is the density function of the parent Poisson distribution. 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 (Greene 1999) have shown that a Vuong statistic has reasonable power and hence is quite reliable.

ESTIMATION

After the new set of variables was found in the base NM model, we continued the estimation using the ZIP(τ) and ZINB(τ). The ZIP(τ) model was first estimated and computed the Vuong statistic. The result revealed that the Vuong statistic of ZIP(τ) was greater than 1.96, suggesting that ZINB(τ) should also be tested. The Vuong statistic of ZINB(τ), however, fell between -1.96 and 1.96. Therefore, we can concluded that the specification estimated by the ZIP(τ) model was hold.

We continued tested a ZIP full model to be compared with the ZIP(τ) model. The ZIP full model is different from the ZIP(τ) model at which the independent variables in the zero state are not constrained to be the same as the that of the non-zero state. Finally, the results still indicated that the ZIP(τ) model was proper than the ZIP full model.

Since the serial correlation problem could still be present in the ZIP(τ) model, the NM estimation was applied for the adjustment of the standard errors in ZIP(τ) model. The process of the adjustment started from getting the proportion of the standard error of the parameter estimated in the NM model and that of NB model with when these two models employed the same set of variables. After the proportion or the load factor for each variable was calculated, the standard errors in the ZIP(τ) model were adjusted by multiplying these proportions. This adjustment technique was derived from the fact that, in the NM model, the standard error of a parameter would usually be larger than the standard error of the same parameter in the NB model. By comparison to account for the serial correlation, the ZIP(τ) model should also have its standard errors larger than the actual standard errors from the estimation. Table 8 presents the NB, NM, and ZIP(τ) models with the adjusted standard errors for the median crossover accident count model.

Table B-1. NB, NM, and ZIP(τ) Models with the Adjusted Standard Errors for Yearly Median Crossovers

Variable	Negative Binomial Model			Negative Multinomial Model		
	Estimated Coefficient	Standard. Error	T-statistic	Estimated Coefficient	Standard Error	T-statistic
Constant	-2.1108	0.1343	-15.7140	-2.1906	0.2099	-10.4387
Per-lane AADT indicator (1 if AADT per the number of lanes is <=5000, 0 otherwise)	-0.8745	0.1821	-4.8030	-0.8789	0.2211	-3.9743
Length of section where medians are less than 40 feet wide	0.2991	0.0370	8.0950	0.3323	0.1018	3.2633
Length of section where medians are between 40 feet and 60 feet wide	0.4307	0.0519	8.2990	0.4434	0.0988	4.4884
Length of section where medians are wider than 60 feet	0.1111	0.0293	3.7890	0.1198	0.0415	2.8864
The number of interchanges in the section	0.2571	0.0809	3.1770	0.2858	0.1245	2.2952
Median surface indicator (1 if median is paved, 0 otherwise)	0.8142	0.3032	2.6860	0.8307	0.3311	2.5088
Interaction between low average annual precipitation and the number of horizontal curves per mile (1 if average annual precipitation <=18 inches and the number of horizontal curves per mile <= 0.5, 0 otherwise)	-0.8704	0.3499	-2.4880	-0.9069	0.4731	-1.9169
Interaction between high average annual precipitation and the number of horizontal curves per mile (1 if average annual precipitation >30 inches and the number of horizontal curves per mile <= 0.5, 0 otherwise)	0.5657	0.2180	2.5950	0.5275	0.2871	1.8375
α	0.7225	0.2159	3.3470	0.3844	4.1563	0.6259
Restricted log-likelihood (All parameters = 0, α almost equals to 0)				-6253.34		
Log-likelihood at convergence	-722.8549			-719.648		
ρ^2	-			0.8849		
Number of observations	1375[1]			1375		

[1] Calculated by multiplying 5 years with the total 275 non-median barrier sections.

Table B-1 (continued). NB, NM, and ZIP(τ) Models with the Adjusted Standard Errors for Yearly Median Crossovers

Variables	Zero-Inflated Poisson Model			Load Factors[2]	Adjusted Standard Error	Adjusted T-Statistic
	Estimated Coefficient	Standard. Error	T-statistic			
Zero accident probability state as logistic function and non-zero accident Poisson probability state (regressors constrained to be same)						
Constant	-0.7295	0.1273	-5.7310	1.5622	0.1988	-3.6688
Per-lane AADT indicator (1 if AADT per the number of lanes is <=5000, 0 otherwise)	-0.4221	0.1318	-3.2020	1.2144	0.1601	-2.6367
Length of section where medians are less than 40 feet wide	0.1509	0.0191	7.9030	2.7558	0.0526	2.8679
Length of section where medians are between 40 feet and 60 feet wide	0.2186	0.0447	4.8930	1.9031	0.0850	2.5708
Length of section where medians are wider than 60 feet	0.0560	0.0187	3.0010	1.4157	0.0264	2.1196
Number of interchanges in section	0.1393	0.0497	2.8010	1.5388	0.0765	1.8204
Median surface indicator (1 if median is paved, 0 otherwise)	0.4504	0.1899	2.3720	1.0922	0.2074	2.1715
Interaction between low average annual precipitation and the number of horizontal curves per mile (1 if average annual precipitation <=18 inches and the number of horizontal curves per mile <= 0.5, 0 otherwise)	-0.4053	0.1875	-2.1620	1.3522	0.2535	-1.5987
Interaction between high average annual precipitation and the number of horizontal curves per mile (1 if average annual precipitation >30 inches and the number of horizontal curves per mile <= 0.5, 0 otherwise)	0.3340	0.1366	2.4460	1.3170	0.1799	1.8570
τ	-1.4792	0.4588	-3.2240	-	-	-
Number of observations	1375					
Restricted log-likelihood (constant only)	-889.7221[3]					
Log-likelihood at convergence	-732.0256					
Vuong statistic	12.5486					

[2] The load factor was the proportion of the S.E. of NM over the S.E. of NB.
[3] This restricted log-likelihood was obtained from the restricted log-likelihood computed by the Poisson model.

APPENDIX C –
Alternative Statistical Tests for Comparing Crash Profiles for Sections with and without Median Barriers

LOG-LIKELIHOOD RATIO TEST

A likelihood-ratio test (LR) is a statistical test relying on a test statistic computed by taking the ratio of the maximum value of the likelihood function under the constraint of the null hypothesis to the maximum with that constraint relaxed. The general form of the test is presented below:

$$\lambda = -2(\text{LL restricted} - \text{LL unrestricted}),$$

where LL represents the computed log-likelihood value.

There are a number of ways to set a priori restriction. An example of A priori restriction can be the restriction of estimated parameter from one model on the same set of data. If the a priori restriction is valid, the restricted and unrestricted (log) likelihood should not be different. In this case, λ should be close to zero. If the sample size is large, it can be shown that the test statistic λ follows the chi-square (χ^2) distribution with degree of freedom equal to the number of restrictions imposed by the null hypothesis. In the current study, the LR test is used to test whether the coefficients of the non-barrier model are the same as the coefficients of the barrier model. To accomplish this, three types of log-likelihood ratio tests are performed:

a) -2[LL(*coefficients of the non-median barrier model applied to the w/median barrier dataset*) - LL(coefficients of the median barrier model)]
b) -2[LL(*coefficients of the median barrier model applied to the w/o median barrier dataset*) - L(coefficients of w/o median barrier model)]

c) -2[LL(*coefficients of the non-median barrier model and coefficients of the median barrier model in the combined dataset*) - LL(coefficients of the non-median barrier model in the w/o median dataset) -LL(coefficients of the median barrier model in the w/ median dataset)

The number of degrees of freedom is equivalent to the number of constrained coefficients in the restricted model (*italic*).

A brief outline of the methodology employed is in model estimation and validation is outlined below (this does not include data mining, which is discussed in the next chapter):

1. Perform the contingency test to determine whether characteristics of factors beside median barrier installed in median barrier section contribute their effect the same way they contribute to non-median barrier sections in the same route. The details of the contingency test were outlined in the methodological approach chapter.
2. Perform t-test to verify whether the means of a random variable between the w/o median barrier and w/median barrier datasets are statistically different from each other. Several continuous variables including weather variables were tested. The hypothesis testing can be made by comparing the computed t-value and the value from the t table with the degree of freedom which is equal to n1 + n2 - 2.
3. Estimate the median crossover accident count model by using the explanatory variables obtained from Shankar et al in 1998 as the starting point in the NB, ZIP, and ZINB models, and then compared with the NM model. ZIP and ZINB were used to determine whether the partial observability and unobserved heterogeneity

issues were significantly present in the non-median barrier dataset. Finally, estimate the predicted probability of median crossover accidents from the model.

4. Estimate the Negative Multinomial model for the non-median sections for the following accident types: all accidents, property damage only, possible injury, and evident-injury types. The dataset used is the 1990-94 accident data for the non-median barrier sections.
5. Estimate the median barrier models by using the NM estimator on property damage only, possible injury, and evident-injury types. Compare the coefficients and their standard deviations of a similar set of variables estimated across the two models. For example, to model the PDO accident counts in the w/barrier dataset, the same set of variables presented in the non-median barrier section model were included in estimating the NM model for the median barrier sections. The new coefficients and standard errors were observed.
6. Perform the likelihood ratio (LR) test to evaluate whether the coefficients of the non-barrier model are the same as the coefficients of the barrier model. Three types of log-likelihood (LL) ratios mentioned earlier are tested.

Estimate the probability of different accident counts based on the estimated models. The probability would provide important information on particular sections of roadways whether or not median barriers were required. As mentioned earlier, installing barriers on some sections might increase the frequency of some accidents. Hence, the roadway sections should be studied for the effectiveness of barrier installation on reducing the frequencies of different accident severities. For the sections where barrier installation is deemed appropriate, corresponding design improvements could also be derived from the estimated models.

THE NEGATIVE MULTINOMIAL ESTIMATION FOR NON-MEDIAN BARRIER MODELS IN ALL ACCIDENTS, PDO, PINJ, AND EINJ

According to the distributions of accident means over 5 years for each accident type listed in Figure C-1, the trend of all accident counts (total) has a similar trend that of PDO. The number of all accident counts was derived from the sum of accident counts in PDO, PINJ, and EINJ. These trends were quite useful for model predictions and measuring how model output should look like. Since all accident counts contained more accident information than any specific accident type, we firstly started modeling this type of accident by using NM estimator.

Once the non-median barrier model for all accidents was estimated, the same specification of variables was applied to the estimation in the PDO model. Several best sets of the PDO model were created; however, we selected one that passed the LR test. After the best PDO model was chosen, we employed the same set of variables in the PDO model as the starting point of the PINJ and EINJ models.

The results of the models estimation for non-median barrier models in PDO, PINJ, and EINJ were presented in tables C-1 to C-3.

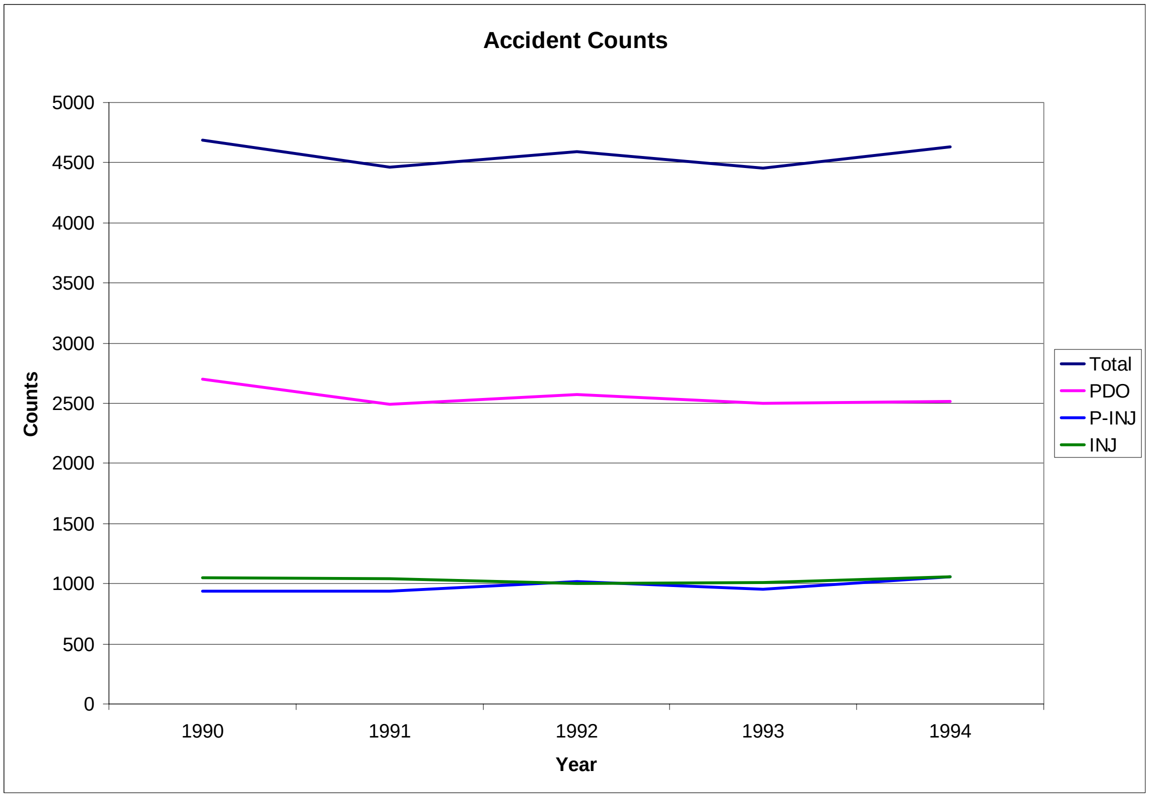

Figure C-1. Mean distribution of accident counts by type

However, the parameters estimated and the output from the predictions of the all accident model and the PDO model for non-median barrier sections were projected almost the same results. This provided no sense when considering that the prediction from the PDO model could be used to predict the total number of accidents. It was likely possible that PDO counts were the majority of the all accident counts which may be serving as a surrogate of the all accident counts. Therefore, we selected to present only the PDO model in this report.

The all accident estimation was useful in term of finding the right combination of variables in the other models. There was possibility that some variables in the other models could not be found if the all accident model were not firstly estimated.

Table C-1. Estimation of negative multinomial for non-median barrier model on property damage only accident probabilities.

Variable	Estimated Coefficient	Coefficients Standard Error	T-statistic
Constant	-5.176	1.104	-4.691
Natural Log of (AADT/ the number of lanes)	0.789	0.147	5.353
The number of interchanges in the section	0.283	0.085	3.318
The number of curves per mile	-0.166	0.076	-2.174
Maximum median shoulder width in feet	0.066	0.028	2.395
The number of absolute grade changes exceeding 0.5%	0.097	0.022	4.516
The number of total trucks in 10^{-6}	-0.209	0.103	-2.019
The square of the number of total trucks	0.016	0.007	2.102
The number of absolute grade changes exceeding 0.5% per mile indicator (1 if The number of absolute grade changes exceeding 0.5% per mile indicator is > 5, 0 otherwise)	-0.556	0.339	-1.638
Interaction variable between average precipitation and average speed (1 if average annual precipitation > 30 inches and average speed limit <= 50 miles/hour; 0 otherwise)	-1.531	0.656	-2.334
The number of absolute grade changes exceeding 0.5% per mile indicator (1 if The number of absolute grade changes exceeding 0.5% per mile indicator > 2 but <= 5, 0 otherwise)	-0.458	0.139	-3.284
Percentage of AADT in the peak hour indicator (1 if the percentage > 13, 0 otherwise)	0.315	0.184	1.714
Minimum radius of horizontal curve in feet indicator (1 if minimum radius > 0 foot but <= 2000 feet, 0 otherwise)	0.339	0.147	2.308
θ	2.418	0.971	2.489
Restricted log-likelihood (Constant only)		-14400.095	
Log-likelihood at convergence		-3689.472	
ρ^2		0.744	
Number of observations		1,375	

Table C-2. Estimation of negative multinomial for non-median barrier on possible injury accident probabilities.

Variable	Estimated Coefficient	Coefficients Standard Error	T-statistic
Constant	-10.739	1.195	-8.990
Natural Log of (AADT/ the number of lanes)	1.316	0.155	8.519
The number of interchanges in the section	0.274	0.091	3.011
The number of curves per mile	-0.296	0.087	-3.403
Maximum median shoulder width in feet	0.066	0.031	2.151
The number of absolute grade changes exceeding 0.5%	0.087	0.022	3.880
The number of total trucks in 10^{-6}	-0.346	0.096	-3.616
The square of the number of total trucks	0.023	0.007	3.470
Interaction variable between average precipitation and average speed (1 if average annual precipitation > 30 inches and average speed limit is <= 50 miles/hour; 0 otherwise)	-1.348	0.775	-1.741
Minimum radius of horizontal curve in feet indicator (1 if minimum radius > 0 foot but <= 2000 feet, 0 otherwise)	0.341	0.163	2.090
The number of absolute grade changes exceeding 0.5% per mile	-0.097	0.0467	-2.075
Interaction variable between the number of curves and average precipitation (1 if the number of curves > 0 and average annual precipitation > 0 inch, 0 otherwise)	0.566	0.275	2.055
θ	2.082	0.861	2.412
Restricted log-likelihood (Constant only)		-12,687.730	
Log-likelihood at convergence		-2,547.021	
ρ^2		0.799	
Number of observations		1,375	

Table C-3. Estimation of negative multinomial for non-median barrier on evident injury accident probabilities.

Variable	Estimated Coefficient	Coefficients Standard Error	T-statistic
Constant	-3.612	1.063	-3.398
Natural Log of (AADT/ the number of lanes)	0.512	0.138	3.722
The number of interchanges in the section	0.171	0.086	1.985
The number of curves per mile	-0.407	0.082	-4.961
Maximum median shoulder width in feet	0.063	0.030	2.073
The number of absolute grade changes exceeding 0.5%	0.107	0.025	4.228
The number of total trucks in 10^{-6}	-0.156	0.088	-1.771
The square of the number of total trucks	0.011	0.006	1.783
Minimum radius of horizontal curve in feet indicator (1 if minimum radius > 0 foot but <= 2000 feet, 0 otherwise)	0.414	0.152	2.718
The number of absolute grade changes exceeding 0.5% per mile	-0.135	0.043	-3.125
Interaction variable between the number of curves and average precipitation (1 if the number of curves > and average annual precipitation > 0 inch, 0 otherwise)	0.500	0.263	1.898
The number of absolute grade changes exceeding 0.5% per mile indicator (1 if The number of absolute grade changes exceeding 0.5% per mile indicator > 2 but <= 3, 0 otherwise)	-0.292	0.157	-1.860
θ	3.014	1.767	1.706
Restricted log-likelihood (Constant only)		-13,314.917	
Log-likelihood at convergence		-2,800.676	
ρ^2		0.790	
Number of observations		1,375	

THE NEGATIVE MULTINOMIAL ESTIMATION FOR MEDIAN BARRIER MODELS IN PDO, PINJ, AND EINJ.

The results of the NM estimation for median barrier models were shown in tables C-4 to C-6. The variables appeared to be less significant in median barrier models were highlighted in the tables.

Table C-4. Estimation of negative multinomial for median barrier model on property damage only accident probabilities.

Variable	Estimated Coefficient	Coefficients Standard Error	T-statistic
Constant	-7.463	1.564	-4.771
Natural Log of (AADT/ the number of lanes)	1.027	0.166	6.178
The number of interchanges in the section	0.091	0.044	2.062
The number of curves per mile	-0.066	0.088	-0.746
Maximum median shoulder width in feet	0.004	0.001	2.783
The number of absolute grade changes exceeding 0.5%	0.561	0.184	3.052
The number of total trucks in 10^{-6}	-0.001	0.0015	-0.602
The square of the number of total trucks	0.0001	0.00014	0.741
The number of absolute grade changes exceeding 0.5% per mile indicator (1 if The number of absolute grade changes exceeding 0.5% per mile indicator is > 5, 0 otherwise)	-2.666	0.440	-6.055
Interaction variable between average precipitation and average speed (1 if average annual precipitation > 30 inches and average speed limit <= 50 miles/hour; 0 otherwise)	-0.202	0.807	-0.250
The number of absolute grade changes exceeding 0.5% per mile indicator (1 if The number of absolute grade changes exceeding 0.5% per mile indicator > 2 but <= 5, 0 otherwise)	-1.392	0.361	-3.854
Percentage of AADT in the peak hour indicator (1 if the percentage > 13, 0 otherwise)	0.024	0.088	0.273
Minimum radius of horizontal curve in feet indicator (1 if minimum radius > 0 foot but <= 2000 feet, 0 otherwise)	0.260	0.462	0.563
θ	1.864	0.867	2.149
Restricted log-likelihood (Constant only)		-5,882.811	
Log-likelihood at convergence		-2,387.527	
ρ^2		0.594	
Number of observations		440	

Table C-5. Estimation of negative multinomial for median barrier on possible injury accident probabilities.

Variable	Estimated Coefficient	Coefficients Standard Error	T-statistic
Constant	-14.063	2.238	-6.284
Natural Log of (AADT/ the number of lanes)	1.560	0.234	6.655
The number of interchanges in the section	0.106	0.042	2.499
The number of curves per mile	-0.187	0.131	-1.426
Maximum median shoulder width in feet	0.0040	0.002	2.479
The number of absolute grade changes exceeding 0.5%	0.330	0.211	1.562
The number of total trucks in 10^{-6}	-0.0003	0.001	-0.252
The square of the number of total trucks	0.0001	0.0004	0.239
Interaction variable between average precipitation and average speed (1 if average annual precipitation > 30 inches and average speed limit is <= 50 miles/hour; 0 otherwise)	-0.005	0.311	-0.016
Minimum radius of horizontal curve in feet indicator (1 if minimum radius > 0 foot but <= 2000 feet, 0 otherwise)	-0.008	0.245	-0.033
The number of absolute grade changes exceeding 0.5% per mile	-0.349	0.098	-3.569
Interaction variable between the number of curves and average precipitation (1 if the number of curves > 0 and average annual precipitation > 0 inch, 0 otherwise)	1.379	0.593	2.325
θ	1.675	1.125	1.488
Restricted log-likelihood (Constant only)		-4,971.783	
Log-likelihood at convergence		-1,858.948	
ρ^2		0.626	
Number of observations		440	

Table C-6. Estimation of negative multinomial for median barrier on evident injury accident probabilities.

Variable	Estimated Coefficient	Coefficients Standard Error	T-statistic
Constant	-9.522	1.751	-5.438
Natural Log of (AADT/ the number of lanes)	1.074	0.188	5.726
The number of interchanges in the section	0.059	0.042	1.416
The number of curves per mile	-0.180	0.107	-1.684
Maximum median shoulder width in feet	0.003	0.001	2.734
The number of absolute grade changes exceeding 0.5%	0.612	0.193	3.166
The number of total trucks in 10^{-6}	-0.0005	0.002	-0.295
The square of the number of total trucks	0.0001	0.0002	0.556
Minimum radius of horizontal curve in feet indicator (1 if minimum radius > 0 foot but <= 2000 feet, 0 otherwise)	-0.133	0.451	-0.296
The number of absolute grade changes exceeding 0.5% per mile	-0.443	0.082	-5.393
Interaction variable between the number of curves and average precipitation (1 if the number of curves > and average annual precipitation > 0 inch, 0 otherwise)	0.788	0.511	1.542
The number of absolute grade changes exceeding 0.5% per mile indicator (1 if The number of absolute grade changes exceeding 0.5% per mile indicator > 2 but <= 3, 0 otherwise)	-0.381	0.375	-1.016
θ	2.536	2.519	1.006
Restricted log-likelihood (Constant only)		-4,294.705	
Log-likelihood at convergence		-1,124.191	
ρ^2		0.738	
Number of observations		440	

LOG-LIKELIHOOD RATIO TEST

The LR test, -2[LL(*coefficients of the median barrier model applied to the w/o median barrier dataset*) - L(coefficients of w/o median barrier model)] was conducted. Table 15 shows the log-likelihood values of PDO, PINJ and EINJ and the result of the LR test.

Table C-7. Log-likelihood values of PDO, PINJ and EINJ and the test results.

Model	LL(restricted)	LL(unrestricted)	$-2[LL_{restricted} - L_{unrestricted}]$	Chi-squared Value
PDO	-6959.55	- 3689.47	6540.16	22.362 (df =13)
PINJ	-2696.3	- 2547.02	298.56	21.026 (df =12)
EINJ	-3111.92	-2800.6	622.48	21.026 (df =12)

The results suggested that the coefficients of the non-median barrier models cannot be transferred to the median barrier models. In the other words, the different set of variables should be estimated for the median barrier models.