JOINT TRANSPORTATION RESEARCH PROGRAM

FHWA/IN/JTRP-2006/21

Final Report

**SAFETY CONSCIOUS PLANNING IN INDIANA:
PREDICTING SAFETY BENEFITS IN CORRIDOR
STUDIES**

***VOLUME 1
RESEARCH REPORT***

**Andrew P. Tarko
Muhammad Asif Iqbal
Michal Inerowicz
Hong Liang
Gopal Panicker**

May 2007

Final Report

FHWA/IN/JTRP-2006/21

**Safety Conscious Planning in Indiana:
Predicting Safety Benefits in Corridor Studies**

***Volume 1
Research Report***

By

Andrew P. Tarko
Professor of Civil Engineering

Muhammad Asif Iqbal
Graduate Research Assistant

Michal Inerowicz
Graduate Research Assistant

Hong Liang
Graduate Research Assistant

and

Gopal Panicker
Graduate Research Assistant

School of Civil Engineering
Purdue University

Joint Transportation Research Program
Project No. C-36-59LL
File No. 8-5-38
SPR-2875

Conducted in Cooperation with the
Indiana Department of Transportation and the
U.S. Department of Transportation
Federal Highway Administration

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 Indiana Department of Transportation or the Federal Highway Administration at the time of publication. The report does not constitute a standard, specification, or regulation.

Purdue University
West Lafayette, Indiana
June 2007

1. Report No. FHWA/IN/JTRP-2006/21	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Safety Conscious Planning in Indiana: Predicting Safety Benefits in Corridor Studies		5. Report Date May 2007	
		6. Performing Organization Code	
7. Author(s) Andrew P. Tarko, Muhammad Asif Iqbal, Michael Inerowicz, Hog Liang, Gopal Panicker, Jorge Ramos		8. Performing Organization Report No. FHWA/IN/JTRP-2006/21	
9. Performing Organization Name and Address Joint Transportation Research Program 1284 Civil Engineering Building Purdue University West Lafayette, IN 47907-1284		10. Work Unit No.	
		11. Contract or Grant No. SPR-2875	
12. Sponsoring Agency Name and Address Indiana Department of Transportation State Office Building 100 North Senate Avenue Indianapolis, IN 46204		13. Type of Report and Period Covered Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes Prepared in cooperation with the Indiana Department of Transportation and Federal Highway Administration.			
INDOT units responsible for planning and pre-engineering analysis consider safety implications of studied alternative roadway and road network solutions to the extent allowed by the available tools. This project is meant to enhance the available tools of predicting crash frequencies in future road networks to improve the quality of safety consideration in intermediate and long-term road planning. Safety Performance Functions are widely used to estimate and predict road safety at intersections and on segments. Evaluation of the performance of the functions developed for regions other than Indiana indicated that only full re-calibration, which is in fact equivalent to developing own models, gives a chance of obtaining a useful prediction tool for Indiana conditions. Such Safety Performance Functions have been developed for Indiana and implemented in TransCAD by the means of two tools: Indiana Input to Safety In Planning tool (INSIP) and Safety In Planning tool (SIP). INSIP allows reformatting the data available in Indiana to the format that is required by the developed SPFs. INSIP creates intersection data tables that allow predicting crashes separately for segments and separately for intersections. In addition, the INSIP reformats the crash data to the format acceptable by the TransCAD and assigns crashes to segments and intersections. The second component – SIP calculates crashes for segments and intersections in studied road networks. It includes a convenient SPF editor that allows modifying the existing SPFs and entering new ones. Also, the INSIP facilitates calibration of the SPFs based on the user-defined partitioning of the road network. A user manual has been developed to assist analysts use the INSIP and SIP.			
17. Key Words safety prediction, transportation planning, corridor study, add-on tools		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 427	22. Price

Table of Contents

1	INTRODUCTION	13
1.1	Objectives	14
1.2	Scope	14
2	SAFETY IN PLANNING IN INDIANA AND RELEVANT RESEARCH	15
2.1	Corridor Planning Studies	17
2.2	Safety Consideration in Existing Planning Processes in Indiana	18
2.3	Safety performance functions (Past Studies)	20
2.4	An overview of crash reduction factors from previous studies	25
2.5	Issues related to the Estimation of CRF's	28
2.6	The effect of congestion on safety	29
2.7	Methodologies for obtaining CRF's	33
2.7.1	Historical Accident Data	33
2.7.2	Regression Models	33
2.7.3	Before-and-After Studies	34
2.7.4	Expert Judgment	34
2.8	TransCAD as a Planning Tool (short and long range planning)	36
2.9	Chapter summary	37
3	CONCEPT OF RESEARCH APPROACH	39
4	REVIEW OF EXISTING SAFETY PERFORMANCE FUNCTIONS	42
4.1	Rural Two-Lane Model 1	43
4.2	Rural Two-Lane Model 2	46
4.3	Rural Two-Lane Model 3	49
4.4	Rural Two-Lane Model 4	51
4.5	Rural Multilane Model 1	52
4.6	Rural Multilane Model 2	54
4.7	Rural Multilane Model 3	57
4.8	Rural Multilane Model 4	59
4.9	Rural Interstate Model 1	61

4.10	Rural Interstate Model 2	
	63	
4.11	Urban Two-Lane Model 1	64
4.12	Urban Two-Lane Model 2	66
4.13	Urban Two-Lane Model 3	69
4.14	Urban Multilane Model 1	70
4.15	Urban Multilane Model 2	71
4.16	Urban Multilane Model 3	74
4.17	Urban Multilane Model 4	77
4.18	Urban Multilane Model 5	80
4.19	Urban Interstate Model 1	81
4.20	Urban Interstate Model 2	84
4.21	Signalized Intersections Model 1	84
4.22	Signalized Intersections Model 2	86
4.23	Signalized Intersections Model 3	87
4.24	Signalized Intersections Model 4	89
4.25	Signalized Intersections Model 5	91
4.26	TWSC Intersections Model 1	93
4.27	TWSC Intersections Model 2	95
4.28	TWSC Intersections Model 3	96
4.29	TWSC Intersections Model 4	97
4.30	TWSC Intersections Model 5	99
4.31	TWSC Intersections Model 6	104
4.32	AWSC Intersections Model 1	110
4.33	Roundabouts Model 1	112
4.34	Roundabouts Model 2	113
4.35	Models Selected for Implementation	116
5	PRELIMINARY INDIANA MODELS	118
5.1	Basic Safety Performance Functions	118
5.1.1	Basic safety performance functions (BSPF) overview	118
5.1.2	Basic safety performance functions used.....	119

5.2	Crash Modification Factors	121
5.2.1	Crash modification factors overview	121
5.3	Preliminary Models	126
5.3.1	Segments	126
5.3.2	Intersection models	128
5.4	Model Variables	134
5.4.1	Model variables overview	134
5.4.2	Model variables definitions	135
5.5	Evaluating preliminary models for segments	140
5.5.1	140	
5.5.2	Overall Framework	140
5.5.3	Criteria for evaluation	140
5.5.4	Crash Predictions with BSPFs and SPFs	140
5.5.5	Comparison of BSPFs and SPFs	143
5.5.6	Summary	145
5.6	Evaluating preliminary models for intersections	145
5.6.1	Overall Framework	145
5.6.2	Intersection reclassification results	146
5.6.3	Conclusion	147
6	CALIBRATION OF PRELIMINARY MODELS	149
6.1	Data Used in the Calibration Task	150
6.2	Calibration Process	151
6.3	Comparison of the Preliminary and Final Recalibrated Models	152
6.4	Crash Prediction Models for Rural Two-Lane	152
6.5	Crash Prediction Models for Rural Multi-Lane	155
6.6	Crash Prediction Models for Rural Interstate	158
6.7	Crash Prediction Models for Urban Interstate	160
6.8	Crash Prediction Models for Urban Multi-Lane	162
6.9	Crash Prediction Models for Urban Interstate	165
6.10	Summary	167

7	INDIANA NEW MODELS	168
8	TRANSCAD TOOL	174
8.1	Developed tools in long term planning – concept	174
8.1.1	Crash prediction procedure for planning	174
8.1.2	Pre-Processing calibration input data.....	175
8.1.3	Updating crash predictions	175
8.1.4	Calibrating crash prediction models	175
8.1.5	Pre-processing prediction input data.....	176
8.1.6	Predicting crash frequencies for future network.....	178
8.2	Expanding segment table	178
8.2.1	Overview.....	178
8.2.2	Obtaining additional variables	179
8.3	Creating an intersection table	180
8.3.1	Overview.....	180
8.3.2	Generated intersection table structure.....	180
8.4	Default Values of Variables	181
8.4.1	Segments default values.....	182
8.4.2	Intersections default values.....	185
8.5	Procedure for Calibrating Safety Prediction Models used in Planning Road Networks	187
8.5.1	Calibration Problem.....	188
8.5.2	ML Method for Disaggregate Crash Counts.....	193
8.5.3	LS Method for counts aggregated in calibration subsets	199
9	CONCLUSIONS AND RECOMMENDATIONS.....	205
9.1	Conclusions	205
9.2	Recommendations	206
	Appendix A Redefining Variables and Combining Model Parameters	213
	Appendix B Linking Model Variables with TransCAD Variables	236
	Appendix C New Indiana Safety Performance Functions.....	267

List of Figures

Figure 2.1 Long Range Plan Development Process.....	16
Figure 2.2 Minimum Accident Rate when $v/c=0.78$	31
Figure 2.4 Flow Diagram of the Accident Prediction Algorithm for a Single Roadway Segment or Intersection	36
Figure 3.1 Flow Diagram Illustrating Proposed Safety Conscious Technical Planning.....	41
Figure 7.1 Tools in planning	174
Figure 7.2 Updating crash predictions.	175
Figure 7.3 Calibrating crash prediction models.	176
Figure 7.4 Pre-processing prediction input data.	177
Figure 7.5 Predicting crash frequencies for future network.	178
Figure 7.6 An example road network.	191
Figure 7.7 Indiana state road network.....	196

List of Tables

Table 4.1 Parameter Estimates for total crashes for Rural-Two-Lane Model 1	44
Table 4.2 Estimates for PDO crashes for Rural-Two-Lane Model 1	45
Table 4.3 Parameter Estimates for Fatal/Injury crashes for Rural-Two-Lane.....	45
Table 4.4A Parameter Estimates for total crashes for Rural-Two-Lane Model 2	47
Table 4.6 Parameter Estimates for total crashes for Rural-Two-Lane Model 3	50
Table 4.7 Parameter Estimates for PDO crashes for Rural-Two-Lane Model 3	50
Table 4.8 Parameter Estimates for Fatal/Injury Rural-Two-Lane Model 3.....	51
Table 4.9 Parameter Estimates for total crashes for Rural Multi-lane Model 1	53
Table 4.10 Parameter Estimates for total crashes for Rural Multi-lane Model 2	55
Table 4.11 Parameter Estimates for PDO crashes for Rural Multi-lane Model 2	56
Table 4.12 Parameter Estimates for Fatal/Injury crashes for Rural Multi-lane Model 2	56
Table 4.13 Parameter Estimates for PDO crashes for Rural Multi-lane Model 3	58
Table 4.14 Parameter Estimates for Fatal/Injury crashes for Rural Multi-lane Model 3	58
Table 4.15 Parameter Estimates for total crashes for Rural Multi-lane.....	58
Table 4.16 Parameter Estimates for PDO crashes for Rural Interstate Model 1	61
Table 4.17 Parameter Estimates for Fatal/Injury crashes for Rural Interstate.....	62
Table 4.18 Parameter Estimates for total crashes for Rural Interstate Model 1	62
Table 4.19 Parameter Estimates for total crashes for Urban Two-Lane Model 1	65
Table 4.20 Parameter Estimates for PDO crashes for Urban Two-Lane Model 1	65
Table 4.21 Parameter Estimates for Fatal/Injury crashes for Urban Two-Lane Model 1.....	66
Table 4.22 Parameter Estimates for PDO crashes for Urban Two-Lane Model 2	68
Table 4.23 Parameter Estimates for Fatal/Injury crashes for Urban Two-Lane Model 2.....	68
Table 4.24 Parameter Estimates for Total crashes for Urban Two-Lane Model 2	68
Table 4.25 Parameter Estimates for Total crashes for Urban Multi-Lane Model 1	70
Table 4.26 Parameter Estimates for Total crashes for Urban Multi-Lane Model 2	73
Table 4.27 Parameter Estimates for PDO crashes for Urban Multi-Lane Model 2.....	73
Table 4.28 Parameter Estimates for Fatal/Injury crashes for Urban Multi-Lane Model 2	74
Table 4.29 Parameter Estimates for Total crashes for Urban Multi-Lane Model 3	76
Table 4.30 Parameter Estimates for PDO crashes for Urban Multi-Lane Model 3	76

Table 4.31 Parameter Estimates for Fatal/Injury crashes for Urban Multi-Lane Model 3	77
Table 4.32 Parameter Estimates for Total crashes for Urban Multi-Lane Model 4	79
Table 4.33 Parameter Estimates for PDO crashes for Urban Multi-Lane Model 4	79
Table 4.34 Parameter Estimates for Fatal/Injury crashes for Urban Multi-Lane Model 4	80
Table 4.35 Estimates	81
Table 4.36 Estimates	82
Table 4.37 Estimates	83
Table 4.38 Estimates	83
Table 4.39 Parameter Estimates	85
Table 4.40 Signalized Intersection Model 2	87
Table 4.41 Parameter Estimates	88
Table 4.42 Variables and units	89
Table 4.43 Variable description and units	90
Table 4.44 Parameter Estimates	90
Table 4.45 Variable description and units	92
Table 4.46 Parameter Estimates	92
Table 4.47 Parameter Estimates	94
Table 4.48 Parameter Estimates	95
Table 4.49 Parameter Estimates	97
Table 4.50 Parameter Estimates\	98
Table 4.51 Variable description and units	100
Table 4.52 Parameter Estimates	101
Table 4.53 Parameter Estimates	102
Table 4.54 Parameter Estimates	103
Table 4.55 Parameter Estimates	104
Table 4.56 Variable description and units	106
Table 4.57 Parameter Estimates	107
Table 4.58 Parameter Estimates	108
Table 4.59 Parameter Estimates	109
Table 4.60 Parameter Estimates	110
Table 4.61 Parameter Estimates	111

Table 4.62 Crash Modification Factors	112
Table 4.62 Models with Selected Variables for Implementation	116
Table 5.1 Basic Safety Performance Functions: segment and intersections.....	119
Table 5.2 Basic Safety Performance Functions roundabouts	120
Table 5.3 Crash modification factors (CMF).....	123
Table 5.4 Parameter Estimates for RTL	126
Table 5.5 Parameter Estimates for RML	126
Table 5.6 Parameter Estimates for RI	127
Table 5.7 Parameter Estimates for UTL	127
Table 5.8 Parameter Estimates for UML	127
Table 5.9 Parameter Estimates for UI.....	128
Table 5.10 Parameter Estimates for Type 1 - TWSC Intersection	128
Table 5.11 Parameter Estimates for Type II - TWSC Intersection.....	129
Table 5.12 Parameter Estimates for Type III - TWSC Intersection	129
Table 5.13 Parameter Estimates for Type IV - TWSC Intersection	130
Table 5.14 Parameter Estimates for Type V - All-Way Stop Controlled Intersection	130
Table 5.15 Parameter Estimates for Type VI – Signalized Intersection.....	130
Table 5.16 Preliminary equations for segments.....	131
Table 5.17 Preliminary equations for intersections (Total crashes)	132
Table 5.18 Preliminary equations for intersections (Injury crashes)	133
Table 5.19 Results for Injury/Fatal Predicted Crashes Using BSPF	141
Table 5.20 Results for PDO Predicted Crashes Using BSPF	141
Table 5.21 Results for Injury/Fatal Predicted Crashes Using Safety Performance Functions	142
Table 5.22 Results for PDO Predicted Crashes Using Safety Performance Functions	143
Table 5.23 Calibration Factors for BSPFs and SPFs	144
Table 5.24 Summary of TWSC intersection classification checking.	148
Table 6.1 Sample Size Per Facility and Crash History.....	150
Table 6.2 Injury and Fatal Crash Model for Rural Two-Lane Roads.....	153
Table 6.3 PDO Crash Model for Rural Two-Lane Roads	153
Table 6.4 Injury and Fatal Crash Model for Rural Multi-Lane Roads	155
Table 6.5 PDO Crash Model for Rural Multi-Lane Roads.....	156

Table 6.6 Injury and Fatal Crash Model for Rural Interstates	158
Table 6.7 PDO Crash Model for Rural Interstates.....	158
Table 6.8 Injury and Fatal Crash Model for Urban Two-Lane Roads.....	160
Table 6.9 PDO Crash Model for Urban Two-Lane Roads	160
Table 6.10 Injury and Fatal Crash Model for Urban Multi-Lane	162
Table 6.11 PDO Crash Model for Urban Multi-Lane Roads.....	163
Table 6.12 Injury and Fatal Crash Model for Urban Interstates.....	165
Table 6.13 PDO Crash Model for Urban Interstates.....	165
Table 7.1 Linking between segment network table and generated intersection network table	181
Table 7.2 Mean values	185
Table 7.3 Coincidence matrix Δ for example network elements and five calibration factors (resulted subsets s are shown in the last column).	192
Table 7.4 Basic safety performance functions for Indiana (Tarko and Kanodia, 2005).....	197
Table 7.5 Calibration subsets.....	197
Table 7.6 Calibration results for injury/fatal safety performance functions.....	201
Table 7.7 Calibration results for PDO safety performance functions.....	202
Table A.1 Redefined shoulder variables and their slopes for UTL	215
Table A.2 Combining and converting variables for RTL facility.....	219
Table A.3 Combining and converting variables for RML facility.....	220
Table A.4 Combining and converting variables for RI facility	221
Table A.5 Combining and converting variables for UTL facility	222
Table A.6 Combining and converting variables for UML facility	223
Table A.7 Combining and converting variables for UI facility	224
Table A.8 Variables for Signalized Intersections (Total Accidents)	225
Table A.9 Variables for Signalized Intersections (Injury Accidents).....	226
Table A.10 Variables for TWSC Intersections (Total Accidents).....	227
Table A.11 Variables for TWSC Intersections (Injury Accidents).....	228
Table B.1 Model variables and TransCAD variables	236
Table B.2 Curves classes used in the HPMS	237
Table B.3 Grade classes used in the HPMS.....	238

Table B.4 Curb presence.....	239
Table B.5 Item 58 — Shoulder type.....	239
Table B.6 Functional class.....	240
Table B.7 Item 17 — Functional system code.....	240
Table B.8 Outside shoulder presence.....	241
Table B.9 Left turn lane and continuous left turn lane	241
Table B.10 Item 88 — Left turn lane codes.....	242
Table B.11 ST_Surfaced, ST_Stabilized, ST_Combined, ST_Earth.....	242
Table B.12 Linking model variables to HPMS data available in TransCAD for RTL.....	244
Table B.13 Linking model variables to HPMS data available in TransCAD for RML.....	245
Table B.14 Linking model variables to HPMS data available in TransCAD for RI	246
Table B.15 Linking model variables to HPMS data available in TransCAD for UTL	247
Table B.16 Linking model variables to HPMS data available in TransCAD for UML	248
Table B.17 Linking model variables to HPMS data available in TransCAD for UI.....	249
Table B.18 Model variables with direct link to TransCAD variables	249
Table B.19 Model variables with no direct link to TransCAD variables	250
Table B.20 Surrounding land use linked to HPMS Field Manual 2000, Item 13.....	250
Table B.21 Item 13 - Rural/Urban designation.....	251
Table B.22 Right turn lane linked to HPMS Field Manual 2000, Item 89	251
Table B.23 Item 89 Right turn lane codes	252
Table B.24 Left turn lane linked to HPMS Field Manual 2000, Item 89	252
Table B.25 Item 88 Left turn lane codes.....	253
Table B.26 Median type linked to HPMS Field Manual 2000, Item 56.....	253
Table B.27 Median type linked to HPMS Field Manual 2000, Item 56.....	254
Table B.28 Paved shoulder linked to HPMS Field Manual 2000, Item 58	254
Table B.29 Item 58 - Shoulder type.....	255
Table B.30 Linking model variables to HPMS data available in TransCAD for signalized intersections (Total Accidents)	256
Table B.31 Linking model variables to HPMS data available in Transcad for signalized intersections (Injury Accidents).....	256

Table B.32 Linking model variables to HPMS data available in TransCAD for Type I - TWSC intersections (Total Accidents)	257
Table B.33 Linking model variables to HPMS data available in TransCAD for Type II - TWSC intersections (Total Accidents)	257
Table B.34 Linking model variables to HPMS data available in Transcad for Type III - TWSC intersections (Total Accidents)	258
Table B.35 Linking model variables to HPMS data available in Transcad for Type IV - TWSC intersections (Total Accidents)	258
Table B.36 Linking model variables to HPMS data available in TransCAD for Type I - TWSC intersections	259
Table B.37 Linking model variables to HPMS data available in TransCAD for Type II - TWSC intersections	259
Table B.38 Linking model variables to HPMS data available in TransCAD for Type III - TWSC intersections	259
Table B.39 Linking model variables to HPMS data available in TransCAD for Type IV - TWSC intersections	260
Table B.40 Parameter Estimates for RTL.....	261
Table B.41 Parameter Estimates for RML.....	261
Table B.42 Parameter Estimates for RI	262
Table B.43 Parameter Estimates for UTL.....	262
Table B.44 Parameter Estimates for UML	263
Table B.45 Parameter Estimates for UI	263
Table B.46 Parameter Estimates for Signalized Intersection.....	264
Table B.47 Parameter Estimates for Type 1 - TWSC Intersection.....	264
Table B.48 Parameter Estimates for Type1I - TWSC Intersection	265
Table B.49 Parameter Estimates for Type 1II - TWSC Intersection	265
Table B.50 Parameter Estimates for Type1V - TWSC Intersection.....	266

1 INTRODUCTION

Traffic safety is a vital component of highway design and transportation planning. Prior to 1991, safety was not an explicit consideration in the planning process. However, the stipulations of the Inter-modal Surface Transportation Efficiency Act (ISTEA) and the Transportation Equity Act for the 21st Century (TEA-21) have resulted in research efforts for making safety a proactive consideration in the planning process. As a part of this effort, the Indiana Department of Transportation (INDOT) has been developing a Safety Management System for Indiana.

The traditional approach to safety management has been to identify existing safety problems and remedy them. Safety conscious planning adopts a pro-active approach to identify hazardous locations and take appropriate remedial measures. The safety concept in planning has been largely ignored; for example, the evaluation of the safety aspects of the I-69 project in Indiana assigned an arbitrary value of 10 points to all corridors under consideration without any scientific basis. This approach is inadequate, especially in light of the ISTEA and TEA requirements.

The engineering tools used by INDOT are MCIBAS (Major Corridor Investment Benefit Analysis System) and HERS_IN. (Indiana Highway Economic Requirements System). MCIBAS provides a framework for corridor studies by providing a primary approach to analyzing large-scale transportation investments in Indiana. NET_BC, which is a post-processor program, is used to estimate the user benefits including the reduced crash costs. NET_BC predictions are mostly based on average crash rates estimated for several class of roads and that cannot be used to quantify or predict changes in the crash frequency in response to geometric or control features. HERS_IN predicts the reduction in the number of crashes and the associated crash costs using the crash prediction equations, but it does not include all the types of roads and intersections. Furthermore, the equations have not been validated for Indiana conditions.

The presented research project is a result of an ongoing effort to refine the prevalent practices in safety-conscious planning and assessment of safety benefits in corridor studies. The research project has developed a transportation safety planning tool which is capable of predicting frequencies of crashes at various levels of severity and for all types of roads and intersections in Indiana. A complete set of crash prediction equations developed based on Indiana data are implemented through add-on tools in a selected GIS-based planning package. The developed method offers a systematic framework for safety considerations for large transportation projects.

1.1 Objectives

The objectives of the project can be broadly summarized as follows:

- To develop a consistent and practical framework for predicting crashes useful for corridor studies in Indiana
- To integrate the methods and procedures into the current planning tools used in Indiana.

1.2 Scope

The **outcome** of the project will be methods and procedures for predicting the safety benefits of corridor-wide solutions and its implementation in the existing system. In the broadest sense, the safety benefits would be measured by (the change in) crash frequencies at various severity levels. All major types of road facilities would be considered: rural two-lane, rural multilane, rural interstate, urban two-lane, urban multilane and urban interstate. The study will consider both road sections and intersections. Modifications will be made to NET_BC in order to make crash predictions (and consequently, safety benefits) sensitive to typical geometric and control differences between the various corridor scenarios. Traffic diversion is assumed to be an input. The costs associated with crashes are already included in the current version of NET_BC.

2 SAFETY IN PLANNING IN INDIANA AND RELEVANT RESEARCH

This chapter discusses the current state of the various models and techniques adopted for safety modeling in Indiana. The tools currently used in Indiana are discussed and their shortcomings are explained. The need for a new tool/methodology for safety modeling is also presented.

Incorporation Safety in Long Range Planning Process

Most Metropolitan Planning Organizations (MPOs) and Departments of Transportation (DOTs) have safety objectives. But safety has not been an explicit consideration in planning mainly because of the lack of performance measures for appraisal. The TEA-21 requirement that both the statewide and metropolitan planning process “shall provide for the consideration of projects and strategies that will increase the safety and security of the transportation system for motorized and non-motorized users” has brought about a paradigm shift in the traditional approach to safety management and adopting safety considerations in long-range planning.

The prioritization of safety in the planning process requires its consideration as an explicit criterion. The following are some of the factors that planners must keep in mind while formulating safety policies and objectives:

- Safety goals must be **reasonable** and **achievable**. The goals have to be compatible with the resources and investments available.

The short-term objective is to integrate safety considerations into the transportation planning processes at all levels, specifically the Transportation Improvement Programs (TIPS) and the Statewide Transportation Improvement Programs (STIPS) developed by MPOs and DOTs respectively.

-Safety considerations in long-range transportation planning must be **multimodal**, in keeping with the TEA stipulation of enhanced safety for both “motorized and non-motorized users” of the transportation system. This is a very delicate issue because enhancing the safety of one particular set of users of the transportation system could probably lead to its deterioration for another set of users or modes. The organization of this long-range transportation plan document reflects the flow of activities reflected in Figure 2.1.

Figure 2.1 Long Range Plan Development Process

-Since the accomplishment of safety goals and objectives depend to a large extent on the involvement of the public, attempts must be made to mobilize **public involvement**.

-**Time frames** must be established and comparison between in-the-field progress and objectives must be done on a regular basis to monitor progress.

The selection of appropriate performance measures is crucial to the evaluation and monitoring of the progress of safety projects. This feedback could be of great help in refining and redefining goals and for the reallocation of resources and investments. Some of the safety-related performance measures include the number of fatal crashes on a certain section of the highway, proportion of fatalities in crashes, etc.

2.1 Corridor Planning Studies

In 1991, the Indiana General Assembly passed legislation that directed INDOT to establish “commercial corridors” in the state.

Commercial corridors are defined as “that part of a recognized system of highways that

- directly facilitates intrastate, interstate, or international commerce or travel
- enhances economic vitality and international competitiveness
- provides service to all parts of Indiana and the United States.”

A typical issue considered in corridor studies includes evaluation of the potential improvement of the corridor to freeway standards, which comprises estimation of travel time savings, reduced vehicle operating costs and reduced accident costs. Corridor studies also evaluate potential improvements to the economy, including the possibility of generation of new jobs.

First of all, the primary need for the project is identified. For the I-69 project in Indiana, the primary need was “to support the completion of I-69 as a national and international trade corridor.” This is evaluated by the ability to meet current Interstate Standards. Factors such as reduction of congestion, *improvement of safety*, construction costs, constructability, and environmental issues are considered in this stage. The alternatives are ranked on the basis of the extent to which they satisfy the objectives. The feasible ones are considered for further evaluation.

2.2 Safety Consideration in Existing Planning Processes in Indiana

INDOT is committed to the development of technical planning tools that will allow for system-level analysis of the state transportation system. Some of the tools developed recently include:

- TRANSCAD-based statewide travel demand model and geographic information system.
- Major Corridor Investment Benefit Analysis System
- Corridor Travel Demand Analysis
- Benefit-Cost Analysis Framework
- The Highway Economic Requirements System
- User Benefit Analysis (NET_BC)
- Economic impact modules

HERS_IN

HERS_IN (Highway Economic Requirements System for Indiana) is the software that is widely used for planning purposes. It is used to compare the relative costs and benefits of alternate options for long-range plans. The software can be used to predict the investments required for highway projects to achieve predetermined levels of service. The models used provide the optimal investment for highway improvements by considering factors such as travel time, safety, VOC, emissions cost, and highway agency costs. The benefit-cost ratio method is used for comparing the relative merits of alternate projects. The procedure adopted by HERS_IN for safety analysis is as follows. Highways are classified into six facility types:

- Rural two-lane roads, (non-intersection and intersection)
- Rural multilane roads
- Rural freeways
- Urban freeways
- Urban multilane surface streets
- Urban two-lane streets.

Separate models have been developed for each type of facility mentioned above. The expected number of crashes is estimated using empirical/statistical relationships. The number of crashes

for each level of severity (fatal, injury, and property damage only) is determined by multiplying the estimated number of crashes by the functional-class-specific ratios for that type. The number of crashes for each severity level is multiplied by the corresponding cost parameters and summed up to obtain the total estimated crash cost. The parameters used for estimating crash rates (in units of crashes per 100 million VMT) include AADT, ACR (AADT to Capacity Ratio), access control (using binary variables), and geometric features such as section length, lane width, shoulder width, road hazard rating and the degree of curvature of the section.

The crash costs are estimated as the sum of the values of lives lost and costs of injuries, property damage, and delay to other highway users.

Drawback of HERS model

1. HERS does not address the effect of congestion on safety explicitly.
2. A major drawback of HERS is that it estimates the crash rates for each level of severity by multiplying the estimate of the total number of crashes by a predetermined fixed ratio. It would be a better idea to develop separate expressions for each severity level because the impact of a safety improvement on different crash severity levels could be different.
3. Another shortcoming is that HERS views segments in isolation; therefore, the impact of a neighboring segment is not considered in the analysis. HERS does not provide confidence intervals for the outputs. Also, the economic benefits are the only criteria used by HERS. The treatment of assessment of safety benefits in HERS requires further modifications.
4. The developed equations are based on nationwide data. The confidence in crash prediction can be enhanced if the models developed are specific to Indiana.

MCIBAS

MCIBAS (Major Corridor Investment Benefit Analysis System), developed by Bernardin, Lochmueller and Associates for INDOT, is the state-of-the-practice tool used in Indiana for corridor studies. MCIBAS considers the regional economic impacts of highway projects in

addition to user benefit-cost analysis. Graphic user interfaces that analyze the databases available in GIS (TRANSCAD) for ranking the severity of need and user benefits of competing highway projects were also developed. The project elements include transportation model design, identification of major prototype corridors, highway priority setting analysis, development of future year trip tables, prototype corridor user benefit analysis, and corridor-level economic assessment.

NET_BC estimates crash frequencies by change in volume and road classification. However, it does not take into account factors such as geometry, control features, and traffic diversions, which are significant parameters that have impacts on the safety of transportation systems. Remedying these drawbacks is the principal focus of this research.

The objective of this research is to integrate the tools developed into **NET_BC**.

2.3 Safety performance functions (Past Studies)

Safety performance functions return the expected annual number of crashes for a particular facility based on the traffic volume and length of the intersection. They are usually obtained by negative binomial modeling using the traffic volume and section length. The exponent of length is restricted to 1.0 in order to avoid the effect of intersection crashes. Also, if there is a long segment of roadway without an intersection, then the length less than 1.0 mile would not be an appropriate solution.

The safety analysis may also be done by analyzing the crash frequency or crash rate of the site/project of interest. Crash frequency, a systems perspective criterion, is estimated by dividing the total number of crashes by the number of years. Crash frequencies are compared with critical crash frequencies to identify High Crash Locations (HCLs). Crash frequency does not consider exposure to risk. Crash rate, a user perspective criterion, is the number of crashes divided by the amount of exposure (usually the Vehicle Miles Traveled (VMT)). For intersections, the crash rate is defined as the number of crashes per million vehicles passing;

for roadways, it is the number of crashes per 100 million VMT. Some of the previous studies in this direction are summarized below.

Any statistical model for crashes should take into account the following factors related to crash counts:

- Crashes are relatively rare events.
- Crash counts are discrete.
- The model should therefore preferably preclude negative values for the dependent variable.

The following models have been used in the past for crash modeling:

1. Poisson;
2. Negative Binomial;
3. Zero Inflated Poisson; and
4. Zero Inflated negative binomial.

1. Poisson Models

Poisson models are suitable for modeling crash counts because they are largely used for discrete non-negative events. The crashes are assumed to have a Poisson mean λ which depends on the roadway geometric and traffic characteristics.

The Poisson probability is given by:

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad (2.1)$$

The mean is given by:

$$\lambda = \exp(\sum b_i y_i), \quad (2.2)$$

where: b_i is the parameters to be estimated and y_i represents the geometric/traffic characteristics.

Miaou et al (1993) used a Poisson model to estimate the frequency of truck crashes along highway segments:

$$\lambda_i = \exp(0.0818 + 0.1022x_1 + 0.09049x_2 + 0.0426x_3 + 0.0341x_4 - 0.0263x_5), \quad (2.3)$$

where:

X_1 = AADT in thousands,

X_2 = horizontal curve length in miles,

X_3 = shoulder width,

X_4 = deviation of stabilized shoulder width from 12 ft,

X_5 = percent of trucks in traffic stream.

Saccomento (2001) developed a Poisson model for estimating the expected crash frequency along highway segments in Italy and came up with the following model. In this model, the section length was used to model the exposure.

$$\lambda_i = \exp(-1.420 + 0.003x_1 + 0.056x_2 + 0.539x_3 - 0.018x_4), \quad (2.4)$$

where:

x_1 = length of section,

x_2 = number of private driveways in the section,

x_3 = number of intersections in the section,

x_4 = 85th percentile speed from previous section.

2. Negative Binomial Models

Poisson models have the drawback that the mean of the crash counts is being restricted to be equal to the variance. However, it is observed that the variance of crash counts is much larger than its mean; so if this is a large amount, the crash counts are over-dispersed. Negative binomial models were proposed to address this deficiency of Poisson models.

Previous work by Abbess et al. (1981) and Maher (1987) has validated the use of the Negative Binomial distribution using crash data. Brown et al (1998) used negative binomial models for computing modeling crashes on roadway segments using the negative binomial model. The dependent variables included the geometric and access control characteristics.

The following models were arrived at:

$$TOTAL = 0.494 \cdot LENG \cdot YRS \cdot AADT \cdot EXP(0.0285AC - 0.631SHLDR + 2.52PS - 0.748TWLTL - 0.604NOMEDO), \quad (2.5)$$

$$PDO = 0.374 \cdot LENG \cdot YRS \cdot AADT \cdot EXP(0.0261AC - 0.669SHLDR + 2.63PS - 0.686TWLTL - 0.684NOMEDO), \quad (2.6)$$

$$FATINJ = 0.127 \cdot LENG \cdot YRS \cdot AADT \cdot EXP(0.0325AC + 0.525SHLDR + 2.28PS - 0.865TWLTL - 0.493NOMEDO), \quad (2.7)$$

where:

LENG = length of section,

AADT = annual average daily traffic,

AC = access control,

SHLDR = shoulder,

TWTL = two-way left turn lane,

NOMEDO = no median presence.

One of the most significant and widely cited modeling approaches has been that of Vogt and Bared (1998) for rural two-lane and at-grade intersections. These models are currently used in the crash prediction module of the Interactive Highway Safety Design Model (IHSDM). The data consisted of crash data from 619 rural two-way segments in Minnesota and 712 roadway segments in Washington. An important advantage of this model is that sections with identical geometric characteristics are treated separately. Thus, the model results are scaleable, unlike in other models where sections having non-uniform geometric features are averaged out.

The model has a very good explanatory power (65% of variance explained). The over-dispersion factor was significant, suggesting that the negative binomial model was appropriate. Lane width, shoulder width, roadside hazard rating, driveway density, degree of horizontal curvature, and traffic exposure were found to be significant factors in increased crash rates.

3. Non-Linear models

Zeeger et al (1986) used non-linear least squares regression to predict the number of crashes on roadway segments. The objective of this study was to quantify the benefits of various measures such as lane widening, shoulder widening, etc. The model developed was:

$$A = 0.0109 * ADT^{0.8786} \cdot 0.9192^W \cdot 0.9316^{PA} \cdot 0.8824^{UP} \cdot 1.2365^H \cdot 0.8822^{TI} \cdot 1.3221^{T2}, \quad (2.8)$$

where:

ADT = daily traffic,

W = lane width,

PA = shoulder width (paved),

UP = unpaved shoulder width,

H = roadside hazard rating,

$TI = 1$ for flat terrain,

$T2 = 1$ for mountainous terrain.

The problem with this model is that it assumes normal distribution of residuals. This assumption has been challenged because accident counts are almost always skewed to the right. A semi-log regression method was developed by Forkenberg et al (1997) using data on 17,767 roadway segments in Iowa. The model has the limitation that it assumes a linear relationship between the number of crashes and traffic exposure, which may not hold true if there is high interaction between the vehicles in the flow stream.

4. Zero-Inflated Models

An important characteristic of crash counts is that they contain a large number of zeros. This could be one reason for the reduced predictive power of most models. This problem has been circumvented by the use of zero-inflated models which assume that a roadway has two-states: perfectly safe with probability p and not perfectly safe with probability $1-p$. Thus, these models combine other models (Poisson or Negative Binomial) with this concept. Zero-inflated models have a much better fit than "pure" Poisson or negative binomial models [Mannering et al, 1996].

The statistical models commonly applied to crash counts include binomial, Poisson, Poisson-Gamma (or Negative Binomial), Zero-Inflated Poisson and Negative Binomial Models (ZIP and ZINB), and Multinomial probability models. Given the range of possible modeling approaches and assumptions with each modeling approach, making an intelligent choice for modeling crashes is not easy [Lord et al, 2004].

The objective of this study is to provide defensible guidance on how to appropriate model crash data. It has been shown that the fundamental crash process follows a Bernoulli trial with unequal probability of independent events, also known as Poisson trials. A simulation experiment is then conducted to demonstrate how crash data give rise to the “excess” zeros frequently observed in crash data. It is shown that the Poisson and other mixed probabilistic structures are approximations assumed for modeling the motor vehicle crash process. It has been demonstrated that under certain (fairly common) circumstances, excess zeros are observed and that these circumstances arise from low exposure and/or inappropriate selection of time/space scales rather than an underlying dual-state process. Carefully selecting the time/space scales for analysis, including an improved set of explanatory variables and/or unobserved heterogeneity effects in count regression models, or applying small area statistical methods (observations with low exposure) represent the most defensible modeling approaches for datasets with a preponderance of zeros. Other models, such as the Multivariate Zero-Inflated Poisson (MZIP) models, multi-logit Poisson, multinomial Poisson, spline functions, and small area statistical methods, show good promise for crash data modeling [Lord et al, 2004].

2.4 An overview of crash reduction factors from previous studies

CRFs are defined as “mathematical and statistical tools used to estimate the effects on safety of planned improvements to streets and highways.” [Harkey, 2004] CRFs are developed using data on vehicle crashes, roadway geometric data, traffic volume data, and other types of roadway and traffic data. CRFs are used to measure the effectiveness of a safety improvement. Mathematically, a CRF is the expected percent reduction in the number of crashes due to the implementation of the safety project. The CRF for multiple improvements is obtained by multiplying the corresponding CRFs for individual improvements. CRF values are typically obtained from before-and-after studies on selected sites on which safety improvements are done. In some studies they are also obtained from regression analysis [Tarko and Eranky, 1997].

In the case of before-and-after studies, the CRF is obtained using the following relations:

$$CRF = \frac{C_b - C_a}{C_b}, \quad (2.9)$$

where:

C_b is the number of crashes before improvement and

C_a is the number of crashes after improvement.

The above equation does not take into account the effect of traffic growth. Methods that allow for growth changes are provided by Elmer [1992]. There are two major methods of developing a CRF:

- i. **Before and After Studies:** The most widely used approach, these studies are comparisons of crashes occurring before and after an improvement. One of the advantages of this method is that it conforms to the idea of a controlled experiment. On the other hand, there is a drawback in terms of cost for proper design and execution of these studies, particularly, the range of countermeasures can be beyond the means of some state DOTs [Shen and Gan, 2003].
- ii. **Cross-Sectional Approach:** This method is usually based on regression methods utilizing a large sample of roadway segments whose design attributes vary systematically. It was used by Tarko et al (2000) to develop CRFs for Indiana. The advantage of this method is that it can be used in the sensitivity analysis of an alternative highway improvement. One of the disadvantages is that it does not account for factors not included in the regression model.

There are three methods of before-and after studies:

- i. *The Naïve Method:* This simple method computes the CRF as the percentage of change in the number of crashes before and after the treatment of the site. This method has the following potential problems:

- Regression-to-mean effect
- Maturation effect
- Crash migration
- External causal factors.

ii. *Comparison Group Method:* A comparison group is a control group that is similar enough to the treated sites in traffic volume and geographic characteristics. This method takes into account factors such as maturation and external causal factors.

The CRF is computed as:

$$\text{CRF} = \text{odds ratio} - 1,$$

$$\text{Odds ratio} = (M/N)/(K/L).$$

where:

(M,N) and (K,L) are the before and after crash counts for comparison and treatment groups respectively

Girffith (1999) and Hauer (1997) used a different form of odds ratio based on the sum of the number of crashes before and after treatment.

iii. *Empirical Bayes (EB) Method:* An assumption of this method is that the number of crashes is Poisson-distributed at any site. Also, the means of populations of systems is assumed to be approximated by a gamma distribution, and changes from year to year from different factors are assumed to be similar across all sites.

The idea behind EB is to correct for the regression to mean effect by computing the expected number of crashes that would have occurred at the site if no treatment was provided. An estimate of the expected crash count is obtained by combining the expected value computed beforehand with the actual crash count history, using appropriate weights. This method is superior because it is more accurate than the other two methods and it corrects for the regression to mean effect. However, the implementation is cumbersome.

Michaels (1996) suggests consideration of the following factors in before-and-after studies:

- i. Account for variation in exposures
- ii. Traffic volumes that do not vary substantially in the before and after period
- iii. Composition of traffic should be the same
- iv. Total number of crashes must be corrected for any existing trends
- v. If crash data for many years show less than 20% variability from year to year, they may be averaged

Tarko et al, 2000 summarized the differences between the two approaches as follows: “The before and after study focuses on changes in safety over time while the cross-sectional analysis focuses on the differences in safety between locations.”

2.5 Issues related to the Estimation of CRF's

CRFs are typically overestimated due to the following reasons:

1. Biased Samples

It is not possible to conduct studies on the entire population due to issues such as data availability, lack of resources, and ethical issues. It is not possible to do controlled experiments for safety problems; rather, researchers have to rely on observational studies. This is an impediment to obtaining truly “random” samples. Different researchers conduct studies on different samples. Due to the inherent randomness of statistical samples, the CRF values obtained by each of them would be different. Researchers who obtain insignificant values of CRF do not report it. Thus, a person who looks at the literature to study the effect of a particular improvement obtains an inflated value for the CRF.

2. Regression to Mean Effect

The traditional method of identification of safety management has been remedying locations that seem to have an above-average crash rate. However, the crash rate could be due to random fluctuations. A future crash reduction observed at that site after a

safety improvement is implemented therefore is not necessarily due to the causal effect of the improvement. CRF values are obtained from before-and-after studies on such sites though, may give an incorrect assessment of the effect of the improvement.

Therefore, the crash frequency approach is based on maximizing the overall efficiency of the system, whereas the crash rate approach is based on the equity of the risk distribution. The locations for improvement are identified depending on the priorities of the agency.

The Michigan DOT SEMCOG Manual has synthesized the CRFs from a wide variety of sources, and the Kentucky Department of Transportation has developed a manual for CRF by combining the values from other states.

3. Reason for Not Using CRFs in this Study

The CRF reflect the change in traffic after an improvement and with the passage of time. For instance, if a two-lane road is increased to four lanes, that means more traffic may be attracted and diverted to this road. Therefore, more traffic could cause more crashes on this road segment; but conversely, crashes on the roads from where this traffic was diverted may decrease. In this case, the CRF do not properly replicate the existing situation. In order to use CRFs, in this research, it will be necessary to go for network modeling. Other major reasons for not using the CRF would be:

- CRFs overstate the safety impact,
- Using recalibrated CRFs does not improve the prediction accuracy, and
- Many CRFs have vague descriptions.

2.6 The effect of congestion on safety

Many researchers have attempted to identify the relationship between traffic volume and accident rates. The findings in some of the previous studies in this area are documented below.

Chang et al (2001) examined the relationship between v/c ratios and accident rates at various freeway facilities, tunnels, and toll gate facilities. The traffic volume and the number of accidents from 1992 to 1997 on the Shingal-Ansan freeway in Korea were analyzed in this study. The accident rate models, having an independent variable of v/c, were established by regression analysis and compared with each other. The relationship between the accident rates and the v/c ratios represented a U-shaped pattern for all sections. The accident rates were highest in the low hourly v/c range, decreasing with an increasing v/c ratio, and then increasing as the v/c ratio increased. The basic freeway, tunnel, and tollgate sections had minimum accident rates when the v/c was 0.78, 0.75, and 0.57, respectively.

An article by the American Automobile Association (AAA) in 1997 from the Michigan DOT and Michigan State University who looked at a 16-mile segment of interstate in the Detroit area over two years to examine relationships between different levels of congestion, measured by volume to capacity (v/c) ratios, and crash rates, measured by crashes per 100 million vehicle miles traveled (VMT). The results indicated that crash rates were very high at low levels of congestion but rapidly decreased with increasing v/c ratios, before gradually increasing again at peak levels of congestion. This U-shaped model holds true for overall weekday and weekend crashes, multi-vehicle crashes, rear-end crashes, and property-damage-only crashes. On the other hand, injury and fatal crashes tended to decrease steadily as the v/c ratios increased. The following plausible explanation was offered: low traffic levels permit higher speeds. Single-car, fixed-object, and rollover crashes are more prevalent. Also, lower v/c ratios typically occur late at night and in the early morning, when drinking and drowsy drivers are prevalent. The authors attributed the increase in general crashes at very high v/c ratios to increased traffic conflict. Interestingly, this increase appeared to limit the number of fatal and injury crashes Figure 2.2.

Figure 2.2 Minimum Accident Rate when $v/c=0.78$

[Source: AAA Foundation on Traffic Safety (1997)]

The details of the relationship between congestion and safety are not well understood (with the exception of lower crash severities, which have been documented in a general way for congested conditions, and the associated lower speeds). Based on the limited work that has been performed, a few tentative conclusions may be drawn:

- Crash potential (e.g., crashes per vehicle-mile traveled) probably increases as congestion increases.
- There is a lower proportion of single vehicle crashes (e.g., run-off-road, rollover, collision with fixed object) during congested conditions and a higher proportion of multiple vehicle crashes.
- Crash severities (extent and nature of personal injuries) are lower during congested conditions due to lower vehicle speeds at the moment of crash impact. In general, it can be assumed that any operational improvement that reduces congestion will lead to fewer crashes. The severity of crashes that occur will be higher, however, and it is likely that a greater proportion will be single-vehicle crashes. Knowing these facts can target mitigation strategies to single-vehicle crashes and higher severities—such as wider roadside recovery zones, protection of highway "furniture," and coordination

with emergency medical services. Moreover, an operations philosophy must take a systems-oriented view, where the consequences of a specific action (e.g., flow improvements) consider linked impacts such as safety [Neudorff et al, 2003].

Crash rates for injury and fatal crashes show a slight decrease with an increase in v/c . For multi-vehicle crashes, a U-shaped variation for accident rates against the v/c ratios was observed. For single-vehicle crashes, it was a steady decreasing trend [Zhou et al, 1997].

Traffic volume is a measure of exposure widely used at present in traffic accident analysis. However, the same traffic volume on road sections with different capacities creates different operating conditions and therefore different probabilities for accidents. Thus, the volume-to-capacity (v/c) ratio may offer a better measure of exposure than traffic volume per section [Frantzeskakis and Iordanis 1987].

Cerder and Liveh (1982) also showed a U-shaped relationship when accident rates were plotted as a function of hourly volume. The study conducted both a cross-sectional and time-sequence analysis of the crash data. The impact of volume on crash rates was studied by developing quantitative models (power functions) to represent the possible dependency on an aggregate level.

Gwynn (1967) examined the hourly accident rate on a 5.9 km section of a four-lane divided highway in New Jersey between 1959 and 1963 and reported that the highest accident rate occurred when the traffic volume was low and a U-shaped function would display the observed relationship.

Hall and Pendelton (1989) concluded that accident rates decreased with increasing traffic volume. Recent studies in this area (Golob et al, 2003) use disaggregate data. This study used sophisticated techniques, such as multivariate statistical analysis and non-linear canonical analysis, to arrive at the relationships between real-time flow characteristics and the likelihood of crash occurrence. The real time flow characteristics were obtained from loop detector data during the time window immediately preceding the occurrence of the crash. This methodology

is however infeasible for this project because the data and resources available for the project do not support the use of disaggregate data.

A promising attempt in predicting the safety performance of proposed and existing projects was the following work by the Federal Highway Administration (FHWA). An attempt was made to combine SPFs (or a base model) with appropriate AMFs obtained from expert judgment and the synthesis of past works in that area [FHWA-RD-99-207, 2000].

2.7 Methodologies for obtaining CRF's

2.7.1 Historical Accident Data

Historical accident data are an important indicator of the safety performance of a roadway, but they suffer from the weakness of being highly variable. Given this high variability, it is difficult to estimate the long-term expected accident rate using a relatively short-duration sample of one to three years of accident data. This is especially true for rural roadway sections and intersections where accidents are very rare events. Many locations experience no accidents, or at most one accident, over a period of several years. If a location has experienced no accidents in the past several years, it is certainly not correct to think that it will never experience an accident, yet the available data for that site alone provide an insufficient basis for estimating its long-term expected safety performance.

2.7.2 Regression Models

Regression models are very accurate tools for predicting the expected total accident experience for a location or a class of locations, but they have not proved satisfactory in isolating the effects of individual geometric or traffic control features. There is a strong temptation to interpret each coefficient in a regression model as representing the true effect of an incremental change in its associated roadway feature. This is a reasonable assumption in some cases, but not in others. A key drawback of regression models is that they are based on statistical correlations between roadway characteristics and accidents that do not necessarily represent cause-and-effect relationships. Furthermore, if the independent variables in the model

are strongly correlated to one another, it is difficult to separate their individual effects. In addition, if a variable in the model is strongly correlated to an important variable that happens not to be included in the available database, the coefficient of the variable in the model may represent the effect of the unavailable variable rather than its own effect. Thus, the value of the coefficient of a particular geometric feature may be a good estimate of the actual effect of that feature on safety, or it may be merely an artifact of, or a surrogate for, its correlation to other variables.

2.7.3 Before-and-After Studies

Before-and-after studies have been used for many years to evaluate the effectiveness of highway improvements in reducing accidents. However, most before-and-after studies reported in the literature have design flaws such that the study design cannot account for the effects of regression to the mean. Therefore, the potential user of the before-and-after study results cannot be certain whether they represent the true effectiveness of the potential improvement in reducing accidents or an over-optimistic forecast that is biased by regression to the mean.

Safety experts are generally of the opinion that, if the potential bias caused by regression to the mean can be overcome, a before-and-after study may provide the best method to quantify the safety effects of roadway geometric and traffic control features. Hauer has developed a new approach that remedies the problem of regression to the mean that has, in the past, caused before-and-after studies to provide unreliable results. However, very few of these well-designed before-and-after studies have been conducted.

2.7.4 Expert Judgment

Expert judgment, developed from many years of experience in the highway safety field, can have an important role in making reliable safety estimates. Experts may have difficulty in making quantitative estimates with no point of reference, but experts are usually very good at making comparative judgments. Experts use a frame of reference based on historical accident data, statistical models, or before-and-after study results to make useful judgments.

This study developed a new model that synthesizes the essence of all of the above approaches into one framework. The flow chart shown in Figure 2.2 illustrates framework. Expert judgment was combined with a base model developed on rural two-way segments in Washington. This in turn was combined with a calibration procedure for accident modification factors that are suited to the local needs of the respective highway agency. Accident modification factors are provided as a function of the traffic volume. The EB procedure is then applied to obtain the final predicted value. The expected number of crashes for a proposed project is computed as the sum of the expected number for crashes in its functional components.

This method is a bold attempt at addressing several of the inadequacies in the existing safety planning tools. However, this method has a few limitations, the most important of them being that the concept of expanding a base model with accident modification factors was not verified using data from any other state. The method uses the principle of moments for arriving at the mean value of crashes in other states. This means may not be scaleable when the negative binomial base model is used. The methodology relies significantly on “expert judgement,” especially when it comes to the development of accident modification factors.

Figure 2.3 Flow Diagram of the Accident Prediction Algorithm for a Single Roadway Segment or Intersection

2.8 TransCAD as a Planning Tool (short and long range planning)

The development of the transportation planning tool was initiated in the 1995-1997 Intermodal Management System Project. This project was provided for the statewide geographic information system project (GIS) which could display several modal transportation networks (e.g., highway and rail systems) plus a variety of transportation hubs and intermodal transfer services (e.g., airports, inter-city train and bus stations, rail/truck terminals, port facilities). Initially the TransCAD GIS incorporated a routing system that allows the display of highway attribute information (number of lanes, functional classification, and average daily traffic, etc.) from the INDOT highway inventory file. This connection provided for the development of a statewide travel demand model.

In order to conduct regional long-term planning using TransCAD, there is something missing that needs to be incorporated within this tool. This study focuses on using the Indiana travel demand model by incorporating crashes on segments and nodes in TransCAD for regional long-term planning. Prediction of crashes is being added in the existing networking of TransCAD. By doing so, the planners will take data from travel demand models and TransCAD in order to utilize it for long-term planning. The developed tool also predicts the changes in crash frequency and severity in response to geometry improvement and therefore can quantify the safety implications of geometry and control differences between alternative projects. This has been accomplished by extracting Crash Modification Factors from existing models and combining them. Finally, the combined model variables have been linked with TransCAD variables. The safety analysis tool has been developed for corridor studies. This task has been accomplished by developing the components of the method for corridor studies, including its manual.

2.9 Chapter summary

This chapter reviewed the various components and methodologies currently in use by Indiana and other states for highway safety in the planning process, such as crash prediction, the use of crash reduction factors, and safety software. The chapter also discussed various methods of predicting crashes and modeling techniques used for developing these models in past studies. An overview of methods for CRFs, along with their drawbacks, was also presented in this chapter.

The current models available for Indiana employed in practice need to be updated due to the following factors:

1. Better data is available through TransCAD.
2. Some data may be irrelevant; other data may not be readily obtained.
3. Current models do not contain enough variables.
4. Some are outdated.

There is a need to develop a model that can address these shortcomings in addition to performing the following tasks:

1. The models should be capable of predicting the expected number of crashes in the planning stage when more detailed actual information is not available.
2. The process of crash prediction should become more refined as more data becomes available.

The methodology adopted for safety analysis uses SPFs for various types of facilities and CMFs for various safety improvements. Several research studies have been conducted in the past for developing safety performance functions and crash reduction factors.

3 CONCEPT OF RESEARCH APPROACH

This chapter discusses the concept of the research approach to develop and implement the tool. The essence of the concept is the use of existing safety performance functions (SPF) and crash modification factors (CMFs) to predict crashes. The research approach adopted for addressing the problem objectives is presented in this chapter.

As mentioned earlier in Chapter 2, there exist lacunae in the current procedures adopted for safety analysis of corridor solutions. One of the greatest problems is the potential for a large number of safety solutions. It is not expedient to develop one comprehensive predictive model from regression analysis including an infinite number of variables. It would be impractical from both the user and the research development points of view. The transportation analyst would have to collect and prepare a large amount of data, including some that may possibly be irrelevant to the proposed improvements. As past experience indicates, in spite of having a large amount and wide scope of data, the effects of many road improvements would not be included in the predictive equation because the resources that constrain the amount of data that can be collected are always limited while the need for data grows in proportion to the number of variables included in the equation. To avoid these methodological complications, a more effective method of developing a crash predictive equation that meets the requirements of corridor studies for long-range planning is proposed.

Instead of trying to develop a set of equations with many variables, we propose to use SPFs existing in Indiana supplemented from existing models developed by other states to predict safety for segments and intersections. The results from these developed equations will be then aggregated for corridors, counties, towns, regions, etc. A safety analysis tool for corridor studies will be developed. This task will develop the components of the method for a corridor study, including its manual. It will also test the method by a comparison to real data. This will be then facilitated with TransCAD-based planning tools presently used in Indiana. Since these models have also been calibrated, we therefore need crash data as we also bring crashes to

TransCAD. These crashes will be then assigned to segments and nodes or intersections in TransCAD using the crash data. In this way the planner will be able to compute the number of crashes by dividing the corridor into functional components: road sections (segments) and nodes or intersections for planning perspective. These models then require calibration and validation by using method of calibration for the user defined set of network partitions. The models developed are based on data from different regions. Therefore the differences in regions and the changes in safety over a period of time call for calibration of the SPFs. For example, different parts of the region may experience different weather conditions (e.g., northern and southern California) or topographical conditions (e.g., northern and southern Indiana). In such cases, the planner may want to use calibration factors to consider these differences when predicting future safety in road networks.

Once the base model is processed, then the calibration process will start, as we will be adding calibration in the planning tool.

Finally, the corridor safety analysis tool and TransCAD will be applied to an example corridor project to illustrate the methodology and to further test it. The method will be applied to a recently developed project so all of the needed data are available for testing. This example will be evidence that the developed method is practical. A description of the developed method with its manual, and an example application of the method to an existing Indiana corridor will be provided.

The INDOT 2000-2025 Long-Range Plan (INDOT, 2002) defines the plan development process where a Technical Planning Analysis is the core component that feeds the Indiana long-range plan with engineering-based inputs. Development of a GIS-based method of predicting safety in transportation networks for long and short-term planning is underway in Indiana. As emphasized by Washington et al. (2004), safety should be introduced in all steps of the planning process, starting with incorporating safety into the vision, goals, and objectives, through technical analysis, development of programs, and monitoring of the system. The objective is to integrate safety consideration into the transportation planning process at all

Figure 3.1 Flow Diagram Illustrating Proposed Safety Conscious Technical Planning

levels starting with the Statewide Transportation Improvement Plans (STIP) and followed by consideration of safety objectives in long-range planning (FHWA, 2003). This long range plan of Indiana will be supplemented with the proposed safety-conscious technical planning components as shown in the flow diagram.

4 REVIEW OF EXISTING SAFETY PERFORMANCE FUNCTIONS

The roadway network plays an important role in every country's economy. This chapter presents some valuable information from the available literature for different states related to the development of crash prediction models for roadway segment and intersection. Various existing road segment and intersection models will be presented and shortly discussed. The literature contained sufficient information for different roadway and intersection classifications, which was considered to be useful for this study. The models presented in the literature are classified into six categories for roadway segments and four categories for intersection types, which are as follow;

Roadway Segment Models

- 1- Rural Two-lane (RTL)
- 2- Rural Multilane (RML)
- 3- Rural Interstate (RI)
- 4- Urban Two-lane (UTL)
- 5- Urban Multilane (UML)
- 6- Urban Interstate (UI)

Intersection Models

- 1- Signalized Intersections

2- Two-way stop-controlled (TWSC)

3- All-way stop-controlled (AWSC)

4- Roundabouts

The review of existing SPFs was based on the methodology adopted by different states and the completeness of the model on the basis of variables and standard error. The performance of the models in terms of goodness of fit was also one of the reviewing parameters. Local models were also reviewed, discussed, and also checked for their utilization in this study after improvements. Discussion of all the models is presented in the subsequent sections.

4.1 Rural Two-Lane Model 1

Source

Tarko et. al. (2000), Crash Reduction Factors for Improvement Activities in Indiana, Unpublished Report.

Model

Negative Binomial Models. Refer Table 4.1, Table 4.2 and Table 4.3 for parameter estimates of total, PDO and Fatal/Injury Crashes.

Data

Indiana Crash Data-1992-1995, 434 observations.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.1)$$

where:

β_0 = constant,

X represents various roadway and traffic variables.

The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description

L_LEN = logarithm of the section length in miles,

L_AADT = logarithm of AADT in thousands vehicle per day,

LW = lane Width in feet,

FR = skid resistance factor (value based on INDOT research division skid resistance test),

SI = pavement serviceability index,

PT_CONC = pavement type-concrete (1= Portland Cement Concrete Road, 0=otherwise),

ST_PAV = shoulder type-paved (0=not paved, 1=paved shoulder),

ALPHA = over-dispersion factor,

TOTAL = total number of crashes,

PDO = property damage only crashes,

FATINJ = fatal and injury crashes.

Table 4.1 Parameter Estimates for total crashes for Rural-Two-Lane Model 1

Variable	Coefficient	std.err	p-value
Constant	-8.436	0.7621	0
		Fixed	
L_LEN	1	parameter	
L_AADT	1.068	0.0818	0
LW	-0.169	0.0801	0.0348
FR	-0.046	0.0231	0.0453
SI	-0.026	0.0129	0.0449
PT_CONC	-0.975	0.2298	0
ALPHA	1.011	0.0952	0

Table 4.2 Estimates for PDO crashes for Rural-Two-Lane Model 1

Variable	Coefficient	std.err	p-value
Constant	-8.871	0.7624	0
L_LENG	1	Fixed parameter	
L_AADT	1.025	0.0822	0
LW	-0.124	0.0803	0.1236
FR	-0.055	0.0228	0.0168
SI	-0.027	0.0123	0.0305
PT_CONC	-0.88	0.2217	0.0001
ALPHA	0.957	0.0991	0

Table 4.3 Parameter Estimates for Fatal/Injury crashes for Rural-Two-Lane

Variable	Coefficient	std.err	p-value
Constant	-10.962	0.8793	0
L_LENG	1	Fixed parameter	
L_AADT	1.273	0.1154	0
LW	-0.227	0.2269	0.0135
SI	-0.026	0.0153	0.0946
PT_CONC	-1.003	0.4298	0.0197
ST_PAV	-0.363	0.176	0.0389
ALPHA	0.908	0.1402	0

Discussion

Model was developed using Indiana data using negative binomial regression, which is appropriate. The model will be considered for use in the final tool. The model contains some variables which are not relevant from a corridor planning perspective. These variables will be removed during recalibration.

4.2 Rural Two-Lane Model 2

Source

Andrew Vogt and Joe Bared, [Accident Models for Two-Lane Rural Roads: Segments and Intersections](#), FHWA-RD-98-133, October 1998.

Data

Minnesota (1985-1989) and Washington (1993-1995).

Model

Negative binomial models. Refer

Table 4.4A and

Table 4.4B for parameter estimates.

Table 4.4A Parameter Estimates for total crashes for Rural-Two-Lane Model 2

Variables (offset = exposure EXPO)	Minnesota 1985-89	Washington 1993-95	Combined
Intercept	1.9456 (.6992, .0054)	0.0358 (.2719, .8953)	0.6883 (.4779, .1492)
AVGM (ADT/1,000)	--	-0.0242 (.0137, .0787)	-0.0109 (.0107, .3067)
Lane Width LW	-0.1821 (.0573, .0015)	TOTWIDTH -0.0127 (.0071, .0720)	-0.0857 (.0405, .0343)
Shoulder Width SHW	-0.08 (.0158, .0001)		-0.0577 (.0106, .0001)
Roadside Hazard Rating RHR (1-7 scale)	--	0.0642 (.0254, .0116)	0.0622 (.0219, .0046)
Driveway Rate DD	0.0079 (.0042, .0630)	0.01 (.0035, .0045)	0.0091 (.0027, .0007)
Degree of Curve H	0.1421 (.0545, .0092)	0.0735 (.0154, .0001)	0.0856 (.0126, .0001)
VC (MN/COM) VMC (WA/COMV)	1.0495 (.4964, .0345)	0.0333 (.0168, .0468)	0.3748 (.2605, .1502)
Absolute Grade GR	0.199 (.0928, .0320)	0.08 (.0295, .0066)	0.0976 (.0280, .0005)
State	--	--	0.142 (.0679, .0366)
n, p $D^m/(n - p - 1)$	619, 7 1.4938	712, 8 1.4767	1331, 10 1.4993
K	0.2657 (.0385, .0001)	0.2821 (.0385, .0001)	0.3022 (.0285, .0001)

R^2_K	0.8609	0.8302	0.831
R^2	0.7251	0.6268	0.6489
R^2_{D, P^2_D}	.3720, .5607	.3455, .5300	.3518, .5464
R^2_{PD}	0.6634	0.6518	.6438

Table 4.4B: Parameter Estimates for IF crashes for Rural-Two-Lane Model 2

Variables (offset = exposure EXPO)	Minnesota 1985-89	Washington 1993-95	Combined
Intercept	1.9998 (.8205, .0148)	-0.2375 (.3511, .4988)	0.1675 (.6108, .7839)
Lane Width LW	-0.2458 (.0694, .0004)	TOTWIDTH -0.0279 (.0089, .0017)	-0.1155 (.0531, .0296)
Shoulder Width SHW	-0.1053 (.0212, .0001)		-0.074 (.0143, .0001)
Roadside Hazard Rating RHR	--	0.0506 (.0314, .1077)	0.041 (.0272, .1315)
Driveway Rate DD	--	0.0065 (.0041, .1193)	0.0054 (.0035, .1192)
Degree of Curve H	0.2158 (.0667, .0012)	0.0598 (.0194, .0020)	0.073 (.0161, .0001)
Crest VMC	--	0.0405 (.0219, .0648)	0.0399 (.0177, .0239)
Absolute Grade GR	--	0.0725 (.0377, .0543)	0.0574 (.0360, .1109)
State	--	--	0.4149 (.0879, .0001)
n, p	619, 4	712, 7	1331, 9
$D^m/(n - p - 1)$	1.0702	1.1593	1.1212
K	0.2398 (.0786, .0023)	0.2751 (.0682, .0001)	0.271 (.0518, .0001)
R^2_K	0.8934	0.8444	0.8628
R^2	0.5859	0.4824	0.5386
R^2_{D, P^2_D}	.3483, .4468	.3185, .4334	.3303, .4399
R^2_{PD}	0.7795	0.7348	0.7509

Discussion

Model developed from data from two sources: Washington and Minnesota. A combined model was also developed by pooling the data from the two sources. The methodology adopted is sound, the data source is acceptable, and the model has a large number of variables and has good performance in terms of goodness of fit.

4.3 Rural Two-Lane Model 3

Source

Lampthey, G., Development of analytical and software tools for highway safety, MS Thesis, Purdue University, May 2005.

Data

Indiana Crash Data-1997-2000, 2392 observations.

Model

Negative binomial models. Refer Table 4.5, Table 4.6 and Table 4.7 for parameter estimates of total, PDO and Fatal/Injury Crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.2)$$

where:

β_0 is constant, X represent various roadway and traffic variables.

Variable Description

L_LENG = logarithm of the section length in miles,

L_AADT = logarithm of AADT in thousands,

LW = lane width in feet,

FR = skid resistance factor (value based on INDOT research division skid resistance test),

RSW = right shoulder width in ft,

ARAD = average radius of curvature of horizontal curves in the segment,

AGRAD = average grade for vertical curves in the segment,

ALPHA = over-dispersion factor,

TOTAL = total number of crashes,

PDO = property damage only crashes,

FATINJ = fatal and injury crashes.

Table 4.5 Parameter Estimates for total crashes for Rural-Two-Lane Model 3

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-4.6915	0.177	-26.505	0
		Fixed		
L_L	1	parameter		
L_Q	0.7134	0.025	28.5746	0
LW	-0.0486	0.0144	-3.3636	0.0008
RSW	-0.017	0.0063	-2.6903	0.0071
FR	-0.0095	0.0041	-2.33	0.0198
ARAD	-0.0248	0.0069	-3.5855	0.0003
AGRAD	0.0708	0.0091	7.8279	0
ALPHA	0.2887	0.0115	25.0587	0

Table 4.6 Parameter Estimates for PDO crashes for Rural-Two-Lane Model 3

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-4.6854	0.1919	-	0
L_L	1	Fixed parameter		
L_Q	0.6669	0.0272	24.5549	0
LW	-0.0424	0.0158	-2.6838	0.0073
RSW	-0.0105	0.0069	-1.5238	0.1276
FR	-0.0089	0.0043	-2.0427	0.0411
ARAD	-0.0252	0.0075	-3.3602	0.0008
AGRAD	0.0775	0.0096	8.042	0
ALPHA	0.3172	0.0137	23.167	0

Table 4.7 Parameter Estimates for Fatal/Injury Rural-Two-Lane Model 3

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-7.0754	0.2502	-28.278	0
L_L	1	Fixed parameter		
L_Q	0.865	0.0348	24.8364	0
LW	-0.0696	0.0192	-3.6195	0.0003
RSW	-0.0342	0.0086	-3.9633	0.0001
FR	-0.0122	0.0058	-2.0915	0.0365
ARAD	-0.0215	0.0093	-2.314	0.0207
AGRAD	0.0528	0.0128	4.1274	0
ALPHA	0.2705	0.0221	12.2519	0

Discussion

Model was developed using Indiana data using negative binomial regression, which is appropriate. The model will be considered for use in the final tool. The model contains some

variables which are not relevant from a corridor planning perspective. These variables will be removed during recalibration.

4.4 Rural Two-Lane Model 4

Source

HERS-ST Technical Report, Draft, USDOT, FHWA, 2002.

Model

A modified form of the equation presented in Model 4.2 is used.

$$C = 100 \cdot (ADJST / SLEN) \cdot \exp(0.72 - 0.085LW - 0.059SHW + 0.067RHR + 0.0085DD + 0.44CCGR) \cdot (\sum CURV_i \cdot \exp(0.045 \cdot CURV_i)) / SLEN, \quad (4.3)$$

where:

C = number of segment crashes per 100 million VMT,

SLEN = section length in miles,

ADJSL = adjusted section length,

LW = lane width in feet,

SHW = shoulder width in feet,

RHR = roadside hazard rating (3.0),

DD = driveway density (3.7 per mile for rural type of development, 50 for dense development),

$CURV_i$ = average degrees of curvature in HPMS curve class i ,

$LCURV_i$ = total length of all curves in HPMS grade class i ,

GRD_i = average grade in HPMS curve class i ,

$LGRD_{ia}$ = total length in miles of all curves in HPMS grade class i ,

CCGR = crest curve grade rate in percent per hundred feet (0 for flat terrain, 0.03 for hilly and 0.4 for mountainous terrains).

Discussion

HERS uses a modified version of the model 1.2. An (ADJSL/SLEN) factor is used to estimate the safety effects of segments only. Intersection crashes are assumed to occur within a 250 feet radius from the center of an intersection. The dummy variable STATE was given a default value of 0.5 and removed from the equation. Since HERS_ST is widely used nationwide, it will be considered for use in the planning tool.

4.5 Rural Multilane Model 1

Source

Wang, Hughes and Stewart, Report: FHWA-RD-98-071.

Data

HSIS database (Data pooled from Illinois, Maine, Michigan, Minnesota and Maine).

Model

Poisson models. Refer Table 4.8 for parameter estimates of total crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.4)$$

where:

β_0 is constant, X represent various roadway and traffic variables.

Variable Description

X1 = average roadside hazard rating,

X2 = access control (1=if partial control, 0=if no control),

X3 = driveways/mile,

X4 = intersections with turn lanes/mile,

X5 = intersections without turn lanes/mile,

X6 = functional class (1=for rural principal arterial, 0=otherwise),

X7 = shoulder width (outside shoulder width in feet),

X8 = median width (median including inside shoulder in feet),

X9 = area location type (1=if rural municipal are, 0=otherwise),

DVMT = daily vehicle miles traveled.

Table 4.8 Parameter Estimates for total crashes for Rural Multi-lane Model 1

Variable	Estimate	Std.Err	Chi-Sqd	P-value
Intercept	-6.572	0.293	501.8	0.0001
X1	0.131	0.025	28.09	0.0001
X2	-0.151	0.047	10.43	0.0012
X3	0.034	0.008	19.36	0.0001
X4	0.163	0.019	70.99	0.0001
X5	0.052	0.008	40.99	0.0001
X6	-0.572	0.07	66.82	0.0001
X7	-0.094	0.011	70.15	0.0001
X8	-0.003	0.009	10.01	0.0016
X9	0.429	0.064	44.48	0.0001
log(DVMT)	1.073	0.028	1428.42	0.0001

Discussion

Data was pooled from five states, most of which are identical to Indiana. The model has a reasonably good number of variables. The model also includes a categorical variable for functional class, which is a big advantage from a planning perspective. The modeling method, Poisson, is not strictly appropriate; however, considering the merits of the data source and variables, this model will be advanced to the next stage.

4.6 Rural Multilane Model 2

Source

Tarko et. al. (2000), Crash Reduction Factors for Improvement Activities in Indiana.

Data

Indiana crash data 1992-1995, 47 observations.

Model

Negative Binomial Models. Refer Table 4.9, Table 4.10 and Table 4.11 for parameter estimates of total, PDO and Fatal/Injury crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.5)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description

L_LEN = logarithm of the section length in miles,

L_AADT = logarithm of AADT in thousands,

FR = skid resistance factor (value based on INDOT research division skid resistance test),

MW = median width (median width in feet excluding inside shoulders),

A = access control (1=no control, 2=partial control, 3=full control),

SI = pavement serviceability index,

ALPHA = over-dispersion factor,

ISW = inside shoulder width (ft).

Table 4.9 Parameter Estimates for total crashes for Rural Multi-lane Model 2

Variable	Coefficient	Std. err	p-value

Constant	-9.91	6.6179	0.1343
		Fixed	
L LENG	1	parameter	
L AADT	1.141	0.6729	0.0898
MW	-0.054	0.0145	0.0002
A	-1.131	0.4735	0.017
FR	-0.141	0.1118	0.2072
ALPHA	1.563	0.4839	0.0012

Table 4.10 Parameter Estimates for PDO crashes for Rural Multi-lane Model 2

Variable	Coefficient	Std. err	p-value
Constant	-10.562	7.9098	0.1818
		Fixed	
L LENG	1	parameter	
L AADT	1.089	0.8175	0.183
MW	-0.058	0.015	0.0001
A	-0.855	0.4823	0.0764
ALPHA	1.684	0.6163	0.0063

Table 4.11 Parameter Estimates for Fatal/Injury crashes for Rural Multi-lane Model 2

Variable	Coefficient	std.err	p-value
Constant	-9.926	5.9163	0.0934
		Fixed	
L LENG	1	parameter	
L AADT	1.077	0.6342	0.0895
MW	-0.026	0.0184	0.1537
A	-1.502	0.5407	0.0055

FR	-0.248	0.1283	0.0533
ISW	-0.304	0.1139	0.0076
ALPHA	0.544	0.6453	0.3994

Discussion

The model was developed from Indiana data using negative binomial regression. Methodology and data source are appropriate. The overdispersion factor is acceptable. The model will be considered in the subsequent stage.

4.7 Rural Multilane Model 3

Source

Lamprey, G., Development of analytical and software tools for highway safety, MS Thesis, Purdue University, May 2005.

Data

640 observations, Indiana Crash Data-1997-2000.

Model

Negative binomial models. Refer Table 4.12, Table 4.13 and Table 4.14 for parameter estimates of total, PDO and Fatal/Injury crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.6)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description:

L_L = logarithm of the section length in miles,

L_Q = logarithm of AADT in thousands,

LW = lane width in feet,
 MW = median width (median width in feet excluding inside shoulders),
 LSW = left shoulder width in feet (inside shoulder width in feet),
 AC = access control (1=no control, 2=partial control, 3=full control),
 ALPHA = over-dispersion factor,
 TOTAL = total number of crashes,
 PDO = property damage only crashes,
 FATINJ = fatal and injury crashes.

Table 4.12 Parameter Estimates for PDO crashes for Rural Multi-lane Model 3

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-3.3554	0.4533	-7.4026	0
L_L	1	Fixed parameter		
L_Q	0.6854	0.6854	16.2496	0
LW	-0.1458	0.0328	-4.4426	0
AC	-0.1046	0.058	-1.8026	0.0714
LSW	-0.0455	0.0225	-2.0276	0.0426
MW	-0.004	0.0018	-2.1577	0.031
ALPHA	0.273	0.0175	15.5989	0

Table 4.13 Parameter Estimates for Fatal/Injury crashes for Rural Multi-lane Model 3

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-5.3988	0.6115	-8.8281	0
L_L	1	Fixed parameter		
L_Q	0.801	0.0563	14.2234	0
LW	-0.1472	0.0442	-3.3301	0.0009
MW	-0.0048	0.0022	-2.1673	0.0302
AC	-0.2182	0.0731	-2.9832	0.0029
LSW	-0.0581	0.0254	-2.2845	0.0233
ALPHA	0.2915	0.0296	9.8341	0

**Table 4.14 Parameter Estimates for total crashes for Rural Multi-lane
Model 3**

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-3.4361	0.441	-7.7924	0
L_L	1	Fixed parameter		
L_Q	0.7188	0.041	17.5122	0
LW	-0.1379	0.032	-4.3101	0
AC	-0.0042	0.0018	-2.332	0.0197
LSW	-0.1329	0.0562	-2.3369	0.0179
MW	-0.0504	0.0216	-2.3278	0.0199
ALPHA	0.2637	0.0154	17.1296	0

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is acceptable. The model will be considered in the subsequent stage.

4.8 Rural Multilane Model 4

Source

HERS-ST Technical Report, Draft, USDOT, FHWA, 2002.

Model

$$\begin{aligned}
 CRASHES = & \\
 & 132.2 \cdot AADT^{0.073} \cdot \exp(0.131RHRML - 0.151AC + 0.034DDRML + 0.078INTSPM \\
 & - 0.572RPA + 0.0082(12 - LW) - 0.094SHLDW - 0.003MEDW + 0.429(DEVEL - 1),
 \end{aligned} \tag{4.7}$$

Variable Description

CRASHES = crashes per 100 million VMT,

RHRML = Roadside hazard rating for rural multilane roads (2.45),

AC = 0 for no access control, 1 otherwise (i.e. both partial and full),

DDRML = driveway density for rural multilane roads (0.41 for rural type development, 5.6 for dense development),

INTSPM = intersections per mile (10/mile),

RPA = 1 for principal arterials, 0 for lower functional class,

LW = lane width in feet,

SHLDW = shoulder width,

MEDW = median width,

DEVEL = type of development (1 for rural, 2 for dense).

- The value 2.45 used for roadside hazard rating (RHRRML) approximates the average value of the ratings used by Wang, Hughes, and Stewart.
- Values used for driveway density were obtained by multiplying the corresponding values used for rural two-lane roads by 0.112 (0.112 is the ratio of driveway density on two-lane rural roads in Minnesota used by Bared and Vogt to the driveway density on four-lane roads in Minnesota used by Wang, Hughes and Stewart).
- Lane width (which was found not to be a significant variable in the original analysis) has been assumed to have one-tenth as much influence as it has for rural two-lane roads (refer to Model 4 for rural two-lane roads).
- Maximum values of 12 feet for right shoulder width and 50 feet for median width have been assumed. Increases in shoulder and median widths beyond these values are likely to have appreciably smaller effects on crash rates than would be indicated by the original form of the equation.
- A maximum value of 10 has been assumed for the intersections per mile (INTSPM). None of the sections analyzed had more than 10 intersections per mile. The equation is relatively sensitive to INTSPM so a maximum value was prescribed.
- The HPMS and HERS variable “type of development” (rural dense or rural rural) is assumed to be a proxy for the “area location type” variable (rural municipal or rural non-municipal) used in the original analysis, though the match may be imperfect. (“Rural municipal” may include some small urban places that would be coded as “urban” by HPMS and HERS, and it

may exclude some small rural developments that would be coded as “rural dense” by HPMS and HERS).

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is slightly high; but considering the small sample size (which is usually the case for urban interstates), it is acceptable. The model will be considered in the subsequent stage.

4.9 Rural Interstate Model 1

Source

Lamptey, G., Development of analytical and software tools for highway safety, MS Thesis, Purdue University, May 2005.

Model

Negative binomial model. Refer Table 4.15,

Table 4.16 and Table 4.17 for parameter estimates of total, PDO and Fatal/Injury crashes.

Data

640 observations, Indiana crash data 1997-2000.

Variable Description

L_L = logarithm of section length in miles,

L_Q = logarithm of AADT in thousands,

LW = lane width in feet,

MW = median width (median width in feet excluding inside shoulders),

LSW = left shoulder width,

ALPHA = overdispersion factor.

Table 4.15 Parameter Estimates for PDO crashes for Rural Interstate Model 1

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	6.3275	1.9154	3.3036	0.001
		Fixed		
L_L	1	parameter		
L_Q	0.7784	0.1093	7.1236	0
LW	-0.5706	0.1178	-4.8439	0
MW	-0.0108	0.023	-4.6665	0
LSW	-1.3338	0.1839	-7.2528	0
ALPHA	0.6884	0.0313	22.0211	0

Table 4.16 Parameter Estimates for Fatal/Injury crashes for Rural Interstate Model 1

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	0.0964	1.5296	0.063	0.9497
		Fixed		
L_L	1	parameter		
L_Q	0.87	0.0908	9.5807	0
LW	-0.4179	0.0896	-4.6634	0
LSW	-1.0011	0.1856	-5.3938	0
ALPHA	0.3682	0.0314	11.7363	0

Table 4.17 Parameter Estimates for total crashes for Rural Interstate Model 1

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	6.4503	1.9802	3.3803	0.0007
		Fixed		
L_L	1	parameter		
L_Q	0.7942	0.1119	7.099	0
LW	-0.5658	0.1178	-4.8018	0
MW	-0.01	0.0023	-4.2688	0

LSW	-1.3779	0.1812	-7.6039	0
ALPHA	0.6802	0.0295	23.0237	0

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is acceptable. The model will be considered in the subsequent stage.

4.10 Rural Interstate Model 2

Source

HERS-ST Technical Report, Draft, USDOT, FHWA, 2002.

Model

Error! Objects cannot be created from editing field codes., (4.8)

Variable Description

CRASH = crashes per 100 million VMT,

LW = lane width in feet.

This equation is a modified form of the model from the following source: *Roadway Safety: A Review of the Ontario Experience and of Relevant Work Elsewhere*, prepared for the Ministry of Transportation, Ontario (Persaud 1992). The original equation was expanded to incorporate the term for lane width, (the term for lane width was obtained from the HERS model for rural multilane roads) multiplied by the section length, divided by the VMT, and then multiplied by the calibration factor.

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is slightly high; but considering the small sample size (which is usually the case for urban interstates), it is acceptable. The model will be considered in the subsequent stage.

4.11 Urban Two-Lane Model 1

Source

Tarko et. al. (2000), Crash Reduction Factors for Improvement Activities in Indiana.

Data

Indiana Crash Data-1992-1995, 182 observations.

Model

Negative Binomial Models. Refer Table 4.18, Table 4.19 and

Table 4.20 for parameter estimates of total, PDO and Fatal/Injury crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.9)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description

L_LEN = logarithm of the section length in miles,

L_AADT = logarithm of AADT in thousands,

LW = lane width in feet,

FR = skid resistance factor (value based on INDOT research division skid resistance test),

T_L = presence of turning lane (indicates strong left turning volumes: 1=left turn lane present, 0=otherwise),

T_C = presence of continuous turning lane (1=left turn lane present, 0=otherwise),

ALPHA = overdispersion factor.

Table 4.18 Parameter Estimates for total crashes for Urban Two-Lane Model 1

Variable	Coefficient	Std.err	p-value
Constant	-10.71	3.1317	0.0006
		Fixed	
L_LEN	1	parameter	
L_AADT	1.22	0.3172	0.0001
LW	-0.214	0.1329	0.1075
T_L	1.182	0.5132	0.0213
T_C	-1.544	1.1857	0.1929
ALPHA	1.525	0.2597	0

Table 4.19 Parameter Estimates for PDO crashes for Urban Two-Lane Model 1

Variable	Coefficient	Std.err	p-value
Constant	-12.946	2.93	0
		Fixed	
L_LEN	1	parameter	
L_AADT	1.154	0.3211	0.0003

FR	-0.048	0.0329	0.1467
T_L	0.995	0.4796	0.038
ALPHA	1.371	0.2691	0

Table 4.20 Parameter Estimates for Fatal/Injury crashes for Urban Two-Lane Model 1

Variable	Coefficient	Std.err	p-value
Constant	-7.442	3.0078	0.0134
L_LENG	1	Fixed parameter	
L_AADT	0.873	0.328	0.0077
LW	-0.332	0.1637	0.0426
T_L	1.716	0.4383	0.0001
ALPHA	1.326	0.3651	0.0003

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is slightly high; but considering the small sample size (which is usually the case for urban interstates), it is acceptable. The model will be considered in the subsequent stage.

4.12 Urban Two-Lane Model 2

Source

Lamprey, G., Development of analytical and software tools for highway safety, MS Thesis, Purdue University, May 2005.

Data

613 observations, Indiana Crash Data-1997-2000.

Model

Negative binomial models. Refer Table 4.21, Table 4.22 and Table 4.23 for parameter estimates of total, PDO and Fatal/Injury crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.10)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description

L_L = logarithm of the section length in miles,

L_Q = logarithm of AADT in thousands,

LW = lane width in feet,

ST = shoulder type (0=no shoulder, 1=earth shoulder, 2=stabilized shoulder, 3=paved shoulder),

TL = presence of turning lane (indicates strong left turning volumes: 1=left turn lane present, 0=otherwise),

RSW = right shoulder width (outside shoulder width in feet),

CRB = presence of curb (1=curb present, 0=otherwise),

ALPHA = overdispersion factor.

Table 4.21 Parameter Estimates for PDO crashes for Urban Two-Lane Model 2

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-6.3069	0.7605	-8.2934	0
L_L	1	Fixed parameter		
L_Q	1.1078	0.1081	10.2433	0
LW	-0.1658	0.0326	-5.0806	0
RSW	0.0524	0.0223	2.3473	0.0189
ST	-0.2849	0.078	-3.6523	0.0003
TL	0.3267	0.1387	2.3551	0.0185
CRB	-0.2686	0.1323	-2.0299	0.0424
ALPHA	0.799	0.0525	15.2088	0

Table 4.22 Parameter Estimates for Fatal/Injury crashes for Urban Two-Lane Model 2

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-	-	-	-
Constant	10.3913	0.967	10.7461	0
L_L	1	Fixed parameter		
L_Q	1.3572	0.1353	10.0319	0
LW	-0.1412	0.0414	-3.4149	0.0006
RSW	0.0754	0.024	3.1475	0.0016
ST	-0.2569	0.0933	-2.7529	0.0059
TL	0.2781	0.1584	1.7555	0.0792
CRB	-0.232	0.1676	-1.3842	0.1663
ALPHA	0.8782	0.0791	11.1028	0

Table 4.23 Parameter Estimates for Total crashes for Urban Two-Lane Model 2

Variable	Estimate	Std.Err	t-ratio	p-value

Constant	-6.2872	0.7715	-8.1498	0
		Fixed		
L_L	1	parameter		
L_Q	1.1436	0.1092	10.4765	0
LW	-0.1736	0.0324	-5.3541	0
RSW	0.0563	0.0216	2.5994	0
ST	-0.2963	0.0727	-4.0782	0.0093
TL	0.3224	0.145	2.2238	0.0262
CRB	-0.2951	0.1334	-2.2111	0.027
ALPHA	0.8523	0.0525	16.2241	0

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The overdispersion factor is acceptable. The model will be considered in the subsequent stage.

4.13 Urban Two-Lane Model 3

Source

HERS Technical Report, (2002), USDOT, FHWA.

Model

$$\text{Error! Objects cannot be created from editing field codes.,} \quad (4.11)$$

Obtained from linear regression, with R-squared value=0.99.

Discussion

The model uses linear regression, which is unacceptable. Also, there is only one variable (AADT). This model will not be used.

4.14 Urban Multilane Model 1

Source

Sawalha, Z. and Sayed, T., Evaluating the safety of urban arterial roads, Journal of Transportation Engineering, Volume 127, Issue 2, March-April 2001.

Data

Urban arterial roadways in Greater Vancouver general district, British Columbia.

Model

GLIM model. GLIM 4 statistical software package. Refer Table 4.24 for parameter estimates of total crashes.

$$\begin{aligned} \text{Crash}/3\text{years} = & 0.0228L^{0.763}V^{0.6459}\exp(0.09097USD + 0.08274CROD \\ & + 0.08515NL + 0.1553I_{UND} + 0.01683(DD)(I_{BUS})), \end{aligned} \quad (4.12)$$

Variable Description

L = segment length in kilometers,

V = AADT (vehicles/day),

USD = unsignalized intersection density (per km),

CROD = crosswalk density per km,

NL = number of lanes,

DD = driveway density (/km),

IUND = indicator variable for undivided median treatment (1 if undivided),

IBUS = indicator variable for business land use (1 if business land use).

Table 4.24 Parameter Estimates for Total crashes for Urban Multi-Lane Model 1

Variable	Coefficient	t-ratio

Constant	0.0228	5.69
L	0.7631	12.08
V	0.6459	8.79
USD	0.09097	9.83
CROD	0.08274	6.17
NL	0.08515	2.71
I _{UND}	0.1553	2.14
DD*I _{BUS}	0.01683	4.43

Degrees of freedom=410.90 (379)

K=3.422

Pearson chi-squared=391.1 (425.4)

Discussion

The model was developed from Canada data using GLIM models. Unit conversion will be necessary (Driveway Density). The model will be considered.

4.15 Urban Multilane Model 2

Source

Tarko et. al. (2000), Crash Reduction Factors for Improvement Activities in Indiana.

Data

Indiana Crash Data-1992-1995, 331 observations.

Model

Negative Binomial Models. Refer Table 4.25, Table 4.26 and Table 4.27 for parameter estimates of total, PDO and Fatal/Injury Crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.13)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description

L_LENG = logarithm of section length,

L_AADT = logarithm of AADT,

NR_LAN = number of Lanes,

A = access control (1=for no control, 2= partial control, 3=full control),

FR = friction Number,

SI = serviceability Index (scale 0-5, 5 being new or nearly new),

ST_PAV = shoulder type-paved (1=paved shoulder present, 0=otherwise),

P_NUM = number of parking lanes (0, 1 or 2),

ISW = inside shoulder width in feet,

TL = presence of left turn (1=left turn lane exists on section, 0=otherwise),

T_C = continuous turn lane on section (1=continuous left turn lane present, 0=otherwise),

ALPHA = overdispersion factor.

Table 4.25 Parameter Estimates for Total crashes for Urban Multi-Lane Model 2

Variable	Coefficient	std.err	p-value
Constant	-13.577	3.783	0.0003
L_LENG	1	Fixed parameter	
L_AADT	1.67	0.4026	0
NR_LANE	-1.026	0.2096	0
A	-1.483	0.3836	0.0001
FR	-0.132	0.0339	0.0001
SI	0.075	0.0273	0.0062
ST_PAV	-0.759	0.336	0.0239
P_NUM	-1.234	0.4831	0.0106
T_C	-1.936	0.7171	0.007
ALPHA	3.626	0.4057	0

Table 4.26 Parameter Estimates for PDO crashes for Urban Multi-Lane Model 2

Variable	Coefficient	std.err	p-value
Constant	-13.143	4.0183	0.0011
L_LENG	1	Fixed parameter	
L_AADT	1.513	0.4211	0.0003
NR_LANE	-0.949	0.2132	0
A	-1.33	0.3916	0.0007
FR	-0.137	0.0376	0.0003
SI	0.083	0.0277	0.0003
ST_PAV	-0.693	0.3336	0.0377
P_NUM	-1.08	0.4853	0.026
T_C	-1.767	0.6837	0.0097
ALPHA	3.644	0.4469	0

Table 4.27 Parameter Estimates for Fatal/Injury crashes for Urban Multi-Lane Model 2

Variable	Coefficient	Std.err	p-value
Constant	-19.331	4.3339	0
L_LENG	1	Fixed parameter	
L_AADT	2.375	0.4633	0
NR_LANE	-1.095	0.3243	0.0007
A	-2.009	0.6391	0.0017
FR	-0.065	0.0473	0.1701
ISW	-0.205	0.0675	0.0024
T_C	-2.358	0.8717	0.0068
ALPHA	3.75	0.6602	0

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The overdispersion factor is high. Since the model was developed from Indiana data, it will be considered in the subsequent stage.

4.16 Urban Multilane Model 3

Source

Brown, H. and A. Tarko, The Effects of Access Control on Safety on Urban Arterial Streets, paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington D.C., January 1999.

Functional Form

$$a = \exp(\beta_0 + \sum \beta_i X_i), \quad (4.14)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Refer Table 4.28, Table 4.29 and Table 4.30 for parameter estimates of total, PDO and Fatal/Injury crashes.

Variable Description

LNLEN = logarithm of the section length in miles,

LNYSRS = Logarithm of number of years,

LNAADT = logarithm of AADT in thousands,

LW = lane width in feet,

ACCESS = number of access points per km,

ALPHA = overdispersion factor,

SHLDR = shoulder presence (1=outside shoulder present, 0=otherwise),

TWLTL = presence of two way left turn lane (1=left turn lane exists on section, 0=otherwise),

TOTAL = total number of crashes,

PDO = property damage only crashes,

FATINJ = fatal and injury crashes,

NOMEDO = dummy variable to indicate no median opening between signalized intersections (1=if no opening, 0=otherwise),

PS = proportion of access points that are signalized.

Table 4.28 Parameter Estimates for Total crashes for Urban Multi-Lane Model 3

Variable	Coefficient	Std.Err	t-ratio	p-value
Constant	-1.182	0.298	-3.972	0.0001
LNLEN	1	Fixed Parameter		
LNYSRS	1	Fixed Parameter		
LNAADT	1	Fixed Parameter		
ACCESS	0.0285	0.00843	3.379	0.0007
SHLDR	-0.631	0.257	-2.455	0.0141
PS	2.52	0.913	2.761	0.0058
TWLTL	-0.748	0.343	-2.179	0.0293
NOMEDO	-0.604	0.228	-2.649	0.0081
ALPHA	1.147	0.131	8.78	0

Table 4.29 Parameter Estimates for PDO crashes for Urban Multi-Lane Model 3

Variable	Coefficient	Std.Err	t-ratio	p-value
Constant	-1.459	0.277	-5.272	0
LNLEN	1	Fixed Parameter		
LNYSRS	1	Fixed Parameter		
LNAADT	1	Fixed Parameter		
ACCESS	0.0261	0.00814	3.21	0.0013
SHLDR	-0.669	0.25	-2.68	0.0074
PS	2.627	0.861	3.05	0.0023
TWLTL	-0.686	0.321	-2.134	0.0329
NOMEDO	-0.684	0.212	-3.234	0.0012
ALPHA	1.105	0.134	8.26	0

Table 4.30 Parameter Estimates for Fatal/Injury crashes for Urban Multi-Lane Model 3

Variable	Coefficient	Std.Err	t-ratio	p-value
Constant	-2.54	0.312	-8.141	0
LNLEN	1	Fixed Parameter		
LNYSRS	1	Fixed Parameter		
LNAADT	1	Fixed Parameter		
ACCESS	0.0325	0.00779	4.166	0
SHLDR	-0.525	0.252	-2.081	0.0374
PS	2.28	0.87	2.62	0.0088
TWLTL	-0.865	0.356	-2.432	0.015
NOMEDO	-0.493	0.252	-1.956	0.0505
ALPHA	1.041	0.138	7.54	0

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The overdispersion factor is acceptable. The model will be considered in the subsequent stage.

4.17 Urban Multilane Model 4

Source

Lamptey, G., Development of analytical and software tools for highway safety, MS Thesis, Purdue University, May 2005.

Data

Indiana Crash Data-1997-2000, 316 observations.

Model

Negative binomial models. Refer Table 4.31, Table 4.32 and Table 4.33 for parameter estimates of total, PDO and Fatal/Injury Crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.15)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description

L_L = logarithm of the section length in miles,

L_Q = logarithm of AADT in thousands,

LW = lane width in feet,

AC = access control (1=no control, 2=partial control, 3=full control),

TL = presence of left turning lane on section (1=left turn present, 0=otherwise),

CRB = presence of curbed shoulder on section (1=outside curb present on section, 0=otherwise),

ALPHA = overdispersion factor,

TOTAL = total number of crashes,

PDO = property damage only crashes,

FATINJ = fatal and injury crashes.

Table 4.31 Parameter Estimates for Total crashes for Urban Multi-Lane Model 4

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-7.1531	1.3061	-5.4767	0
L_L	1	Fixed parameter		
L_Q	1.2402	0.1152	10.7703	0
LW	-0.1541	0.059	-2.6113	0.009
AC	-0.43	0.146	-2.9459	0.0032
CRB	-0.299	0.1265	-2.3641	0.0181
TL	-0.2195	0.1515	-1.4483	0.1475
-				
FR	0.03758	0.01968	-1.9101	0.0561
ALPHA	0.7476	0.0646	11.5691	0

Table 4.32 Parameter Estimates for PDO crashes for Urban Multi-Lane Model 4

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-7.60027	1.275	-	0
L_L	1	Fixed parameter		
L_Q	1.2525	0.1125	11.1297	0
LW	-0.1525	0.0588	-2.593	0.0095
AC	-0.4135	0.1443	-2.8657	0.0042
CRB	-0.2786	0.1246	-2.2367	0.0253
TL	-0.2252	0.1471	-1.5304	0.1259
FR	-0.0323	0.0198	-1.6306	0.103
ALPHA	0.722	0.0632	11.4267	0

Table 4.33 Parameter Estimates for Fatal/Injury crashes for Urban Multi-Lane Model 4

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-7.5405	1.5202	-4.9604	0
L_L	1	Fixed parameter		
L_Q	1.164	0.1385	8.4022	0
LW	-0.1521	0.0663	-2.2955	0.0217
AC	-0.5637	0.1596	-3.5323	0.0004
CRB	-0.3766	0.1495	-2.5192	0.0118
TL	-0.2505	0.1619	-1.5473	0.1218
FR	-0.0675	0.0217	-3.1086	0.0019
ALPHA	0.8176	0.0833	9.8187	0

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is slightly high; but considering the small sample size (which is usually the case for urban interstates), it is acceptable. The model will be considered in the subsequent stage.

4.18 Urban Multilane Model 5

Source

HERS-ST Technical Report, Draft, USDOT, FHWA, 2002.

Model

$$CRASH = A * AADT^B NSIGPM^C \quad (4.16)$$

Variable Description

CRASH = crashes per 100 million VMT,

NSIGPM = number of signals per mile (minimum=0.1, maximum=8).

A, B and C are obtained from Table 4.34.

Table 4.34 Estimates

Type of Section	A	B	C
Two-Way with Left Turn Lane	95.1	0.1498	0.4011
One -way or Two-way with a median:			
1. Wider than 4 feet OR			
2. Curbed OR	82.6	0.1749	0.2515
3. A "positive barrier"			
Otherwise	115.8	0.1749	0.2515

This equation was derived from an equation for estimating annual crashes per mile that was estimated by Margiotta using data from Bowman and Vecellio (1994) The derivation involved multiplying Margiotta's equations by section length, dividing by annual VMT, multiplying by 108, and incorporating calibration factors.

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data source are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is slightly high; but considering the small sample size (which is usually the case for urban interstates), it is acceptable. The model will be considered in the subsequent stage.

4.19 Urban Interstate Model 1

Source

Lamprey, G., Development of analytical and software tools for highway safety, MS Thesis, Purdue University, May 2005.

Data

Indiana Crash Data-1997-2000, 108 observations.

Model

Negative binomial models. Refer Table 4.35, Table 4.36 and Table 4.37 for parameter estimates of total, PDO and Fatal/Injury crashes.

Functional Form

$$a = \exp\left(\beta_0 + \sum \beta_i X_i\right), \quad (4.17)$$

where:

β_0 is constant, X represent various roadway and traffic variables. The estimates β_0 of each of the roadway/traffic characteristics is presented below:

Variable Description

L_L = logarithm of section length in miles,

L_Q = logarithm of AADT in thousands,

LW = lane width,

MW = median width,

LSW = left shoulder width,

ALPHA = overdispersion factor.

Table 4.35 Estimates

FATINJ Crashes

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-9.2789	5.9018	-1.5722	0.1159
L_L	1	Fixed parameter		

L_Q	2.0918	0.3293	6.3522	0
LW	-0.9166	0.3422	-2.6786	0.0074
LSW	-0.1783	0.0645	-2.7646	0.0057
ALPHA	1.6607	0.3077	5.3976	0

Table 4.36 Estimates

TOTAL Crashes

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	- 10.3185	6.1539	-1.6767	0.0936
L_L	1	Fixed parameter		
L_Q	2.0693	0.3241	6.3839	0
LW	-0.6507	0.3858	-1.6867	0.0917
MW	-0.0176	0.0112	-1.572	0.116
LSW	-0.2438	0.0719	-3.3878	0.0007
ALPHA	1.2556	0.2395	5.243	0

Table 4.37 Estimates

PDO Crashes

Variable	Estimate	Std.Err	t-ratio	p-value
Constant	-8.9996	6.6557	-1.3522	0.1763
L_L	1	Fixed parameter		
L_Q	2.1218	0.35	6.0628	0
LW	-0.7974	0.4087	-1.951	0.0511
MW	-0.0156	0.0113	-1.3829	0.1667
LSW	-0.2544	0.076	-3.3478	0.0008
ALPHA	1.3698	0.2596	5.2776	0

Discussion

The model was developed from Indiana data using negative binomial regression. The methodology and data sources are appropriate. The model has lane width, median width, and shoulder width at a statistically significant level. The overdispersion factor is slightly high; but considering the small sample size (which is usually the case for urban interstates), it is acceptable. The model will be considered in the subsequent stage.

4.20 Urban Interstate Model 2

Source

HERS Technical Report, Draft, (2002), USDOT, FHWA

Variable Description

CRASHES = number of crashes per 100 million VMT,

ACR = AADT divided by two-way hourly capacity,

LW = lane width in feet.

The original form of the equation can be found in “Development of a Safety Model to Assess the Impact of Implementing IVHS User Services,” Shelby A. Tedesco, et. al., Proceedings of the IVHS America 1994 Annual Meeting, April 1994. The original equation was expanded to incorporate a lane width term. The effect of lane width on safety on freeways is assumed to be the same as that on rural multilane roads.

Discussion

The model combines results from linear regression with a term from another model (for lane width). The methodology does not appear sound. The model is not likely to be advanced for further consideration.

4.21 Signalized Intersections Model 1

Source

Juan C. Pernia, Jian Jhon Lu, Xiangli Zie, Michael Weng and Debroah Snyder, Florida, Transportation Research Board, 2002.

Data

Development of Models to Quantify the Impacts of Signalization on Intersections crashes.

Model

Poisson regression was initially used but due to the problems associated with overdispersion, Negative binomial regression was applied. The model developed for calculating the expected number of crashes was as follows:

$$a = \exp(b_o + b_1X_1 + b_2X_2 + \dots + b_qX_n), \quad (4.18)$$

where:

$X_1, X_2, \dots, X_n$ = model variables,

$\beta_0, \beta_1, \dots, \beta_q$ = model parameters.

The variables and parameter estimates are presented in Table 4.38.

For the above models, parameter estimates were established for signalized intersection on two, four, and six lane-urban arterial streets.

Table 4.38 Parameter Estimates

Variable Description	Name	Level	Value	DOF	Coefficient Estimate	Standard Error
Intercept				1	0.5718	0.1142
Average AADT	AVGADT	<15,000 vpd	0	1	0.4868	0.0578
		15,000~30,000	1			
		>=30,000 vpd	2			
Surrounding Land use	URBRUR	Urban	1	1	0.0949	0.0766
		Rural	0			
Location	LOCATYP	Business	1	1	0.1728	0.0966

Type		Other	0			
Number of lanes on major road	LANE	>4	1	1	0.2654	0.0928
		<=4	0			
Posted Speed on major road	SPEED	>45 mph	1	1		
		>=45 mph	0			
Presence of Median on major road	MEDIAN	Divided	1	1	0.1845	0.0876
		Undivided	0			
Shoulder Treatment	SHOULD	Paved	1	1	-0.1102	0.0813
		Other	0			
Dispersion Parameter				1	0.3450	0.0400

Discussion

The methodology is correct and all parameters and variable are well defined. The demographic and geographic conditions of Florida are different then that of Indiana. Further, the omission of AADT on minor streets is a considerable deficiency.

4.22 Signalized Intersections Model 2

Source

Andrew Tarko, Mayank Kanodia, FHWA/IN/JTRP-2003/19 (Joint transportation research program, Purdue University), Indiana, 2004.

Data

Hazard Elimination Program—Manual on Improving Safety of Indiana Road Intersections and Sections.

Model

Negative binomial methodology was used to estimate the number of crashes. Following two models were developed keeping in view the crash severity as shown in Table 4.39.

Table 4.39 Signalized Intersection Model 2

S.no.	Model	Over-dispersion Parameter
1	$a_{IF} = 0.1954Q^{0.723}$	0.639
2	$a_{PD} = 0.1758Q^{1.0334}$	0.646

where:

a_{IF} = typical I/F crash frequency, in I/F crashes per year,

a_{PD} = typical PDO crash frequency, in PDO crashes per year,

Q = AADT entering an intersection, in thousand veh/day.

Discussion

This model was developed for Indiana using sound methodology. It does not include variables other than AADT.

4.23 Signalized Intersections Model 3

Authors

Andrew P. Tarko, Kumares C.Sinha, Shyam Eranky, Henry Brown, Elizabeth Roberts, Rodian Scinteie, Samantha Islam, Joint Transportation Research Program, Purdue University, Indiana, 2000.

Data

Crash Reduction Factors for Improvement Activities in Indiana.

Model

Negative Binomial regression was used to model the frequency of crashes. The models were developed for injury crashes, PDO crashes, and total number of crashes. The model is as follows:

$$a_j = e^K Y Q^b \exp(\sum \gamma_i X_i) \quad (4.19)$$

where:

C_j = number of crashes at j severity,

Y = number of years,

K, γ_i , β = coefficients,

Q = AADT entering the intersection, in vehicles per day,

X_i = explanatory variable representing factor i.

Parameter estimates for all the variables mentioned above for the three types of crash severity are shown in Table 4.40 and Table 4.41.

Table 4.40 Parameter Estimates

Variables	Injury Crashes		PDO Crashes		Total	
	Coeff.	Std.Error	Coeff.	Std.Error	Coeff.	Std.Error
ONE	-5.684	1.8783	-8.6517	1.9626	-6.3771	1.8562
L_Y	1.0000		1.0000		1.0000	
L_Q	0.5848	0.1947	0.9965	0.2025	0.7821	0.1913
No_LANES	0.0916	0.0280	0.0449	0.0326	0.0673	0.0314
SEP	-0.6998	0.1858	-0.5046	0.1603	-0.5499	0.1697
SEP_R	0.5277	0.1892	0.4204	0.1430	0.4627	0.1536
SEP_W			-0.0307	0.0140	-0.0257	0.0134
URB_IND	-0.2677	0.1835				
NO_1_WAY			-0.2672	0.1936		
Alpha	0.5471	0.0728	0.5631	0.0557	0.5919	0.0540

Table 4.41 Variables and units

Variables	Variable Description	Unit
NO_LANES	Total number of approach lanes including turn lanes	-
SEP	Total number of directions separations on approaches (median and divisional islands)	-
SEP_R	Number of raised separation at the intersection	-
SEP_W	Average width of separation	Feet
URB_IND	Enter 1 for urban area, otherwise enter 0	-
NO_1_WAY	Number of one-way approaches at the intersection	-

Discussion

This model was developed for Indiana using sound methodology. The model is useful for the study.

4.24 Signalized Intersections Model 4**Source**

Vogt Andrew, FHWA-RD-99-128, 1999.

Data

Crash Models for Rural Intersections: Four-Lane by Two-Lane Stop-Controlled and Two-Lane by Two-Lane Signalized ,California 1993-1995 and Michigan 1993-1995 (49 sites).

Model

Negative binomial regression methodology and maximum likelihood methods available in the software packages SAS and LIMDEP were used to fit the models to the data. The following model was developed to estimate the expected number of accidents:

$$a = Q_1^{g_1} Q_2^{g_2} \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_q X_n), \quad (4.20)$$

where:

$Q, X_1, X_2, \dots, X_n$ = model variables,

$\gamma_1, \gamma_2, \beta_1, \beta_2, \dots, \beta_q$ = model parameters.

List of variables available for the analyses, their units, and the abbreviations used in the model estimation results are summarized in Table 4.42, while the parameter estimates used for total accidents have also been summarized separately in Table 4.43.

Table 4.42 Variable description and units

Variables	Variable Description and Units
ADT1	Average daily traffic on mainline (vehicles per day)
ADT2	Average daily traffic on minor road (vehicles per day) AADT1>AADT2
PKTRUCK	Truck percentage passing through the intersection during the peak hour (%)
PKLEFT2	Left turn percentage of the minor road during the peak hour (%)
PROT_LT	Protected Left-turn (0=No, 1=Yes)
VEI1	Sum of absolute change of grade (percent per 100 ft) for each crest curve on major road any portion of which is within 800 feet of the intersection center, divided by the number of such curves
VEI2	Sum of absolute change of grade (percent per 100 ft) for each crest curve on minor road any portion of which is within 800 feet of the intersection center, divided by the number of such curves
VEICOM	(1/2) (VEI1 + VEI2)

Table 4.43 Parameter Estimates

Variables	Total Accidents (Standard Error, P-value)	Injury Accidents (Standard Error, P-value)
Intercept	-6.9636 (2.7911, 0.0132)	-3.2562 (2.9932, 0.2767)
LOG of AADT1	0.6199 (0.2504, 0.0133)	-
LOG of AADT2	0.3948 (0.1737, 0.0230)	-
LOG of (AADT1*AADT2)	-	0.2358 (0.1722, 0.1707)
PKTRUCK	0.0315 (0.0143, 0.0275)	0.0323 (0.0146, 0.0267)
PKLEFT2	-0.0142 (0.0047, 0.0023)	-0.0113 (0.0062, 0.0678)
PROT_LT	-0.6754 (0.1824, 0.0002)	-0.2943 (0.1864, 0.1144)
VEICOM	0.1299 (0.0323, 0.0039)	0.0822 (0.0662, 0.0138)
Overdispersion	0.1161	0.1630

Discussion

The methodology is sound and all the parameters estimates including the overdispersion variables were well defined. The regional weather conditions also resemble that of Indiana. Therefore, the model is acceptable from the viewpoint of methodological soundness, completeness, and regional adequacy.

4.25 Signalized Intersections Model 5

Source

Washington Simon, Persaud Bhagwant, Lyon Craig and Oh Juttaek, FHWA-RD-03-037, 2003.

Data

Validation of Accident Models for Intersections, California 1991-1998, Michigan 1993-97 and Georgia 1996-97 (Total 124 sites, 3417 total accidents).

Model

Negative binomial regression methodology. The following model was developed to estimate the expected number of accidents:

$$a = Q_1^{g_1} Q_2^{g_2} \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_q X_n), \quad (4.21)$$

where:

$Q, X_1, X_2, \dots, X_n$ = model variables,

$\gamma_1, \gamma_2, \beta_1, \beta_2, \dots, \beta_q$ = model parameters.

A list of variables available for the analyses, their units, and the abbreviations used in the model estimation results are summarized in Table 4.44, while the parameter estimates used for total accidents have also been summarized separately in Table 4.45.

Table 4.44 Variable description and units

Variables	Variable Description and Units
ADT1	Average daily traffic on mainline (vehicles per day)
ADT2	Average daily traffic on minor road (vehicles per day) AADT1 > AADT2
COMDRWY1	Number of commercial driveways on the busier road within 250 feet of the intersection center
SPD1	The average posted speed on major road in vicinity of the intersection (mph)
LIGHT	Light at intersection (0 = no, 1 = yes)
HEI1	Average degree of curve for horizontal curves on major road within 800 feet of the intersection center (degree/100ft.)
HEI2	Average degree of curve for horizontal curves on minor road within 800 feet of the intersection center (degree/100ft.)
HEICOM	(1/2) (HEI1 + HEI2)

Table 4.45 Parameter Estimates

Variables	Total Accidents (Standard Error, P-value)	Injury Accidents (Standard Error, P-value)

Intercept	-5.1527 (1.8653, 0.0057)	-9.0707 (1.9064, 0.0000)
LOG of AADT1	0.4499 (0.1968, 0.0223)	0.6697 (0.1899, 0.0004)
LOG of AADT2	0.2699 (0.0767, 0.0004)	0.2509 (0.0929, 0.0069)
COMDRWY1	0.0539 (0.0304, 0.0757)	-
SPD1	0.0177 (0.0090, 0.0482)	0.0397 (0.0093, 0.0000)
LIGHT	-0.2938 (0.1837, 0.1098)	-0.3985 (0.1702, 0.0192)
HEICOM	-0.0288 (0.0153, 0.0597)	-
HEI2	-	-0.0284 (0.0126, 0.0244)
Overdispersion	0.4019	0.2360

Discussion

The methodology is sound and all the parameters estimates including the overdispersion variables were well defined. Therefore, the model is acceptable from the viewpoint of methodological soundness and completeness adequacy.

4.26 TWSC Intersections Model 1

Source

Bhangwant N Persuad, Richard Garder, Dominique Lord, Transportation Research Board Record, 1999.

Data

Observational Before-After Study of the Safety Effect of US Roundabout Conversions Using Empirical Bayes Method, Toronto, Maryland and Florida.

Model

The Negative Binomial regression and GLIM were used. The following model was developed for the estimation of crashes for four-legged and three-legged TWSC intersections.

$$a = (\gamma) (\text{total entering AADT})^{\beta_1} (\text{minor road proportion of AADT})^{\beta_2}, \quad (4.22)$$

where:

γ, β_1 and β_2 = parameters.

For the above models, parameter estimates have been established for three-legged and four-legged intersections. The results of the parameter estimates for the aforementioned intersections with respect to crash severity are summarized in Table 4.46.

Table 4.46 Parameter Estimates

Crash Severity	$\ln(\alpha)$ (S.E)	β_1 (S.E)	β_2 (S.E)	$k=1/\alpha$
Four-legged Intersection				
All Combined	-9.886 (2.01)	1.202 (0.213)	0.376 (0.107)	3.10
Injury	-8.613 (2.31)	0.901 (0.245)	0.197 (0.122)	3.24
Three-legged Intersection				
All Combined	-8.051 (1.60)	1.053 (0.163)	0.559 (0.066)	4.31
Injury	-11.364 (2.16)	1.225 (0.270)	0.466 (0.0937)	3.76

S.E = Standard error

Discussion

The methodology used in this model is good and all the parameters and variables are well defined. The demographic condition of the regions of Florida, Maryland, and Toronto are not similar to that of Indiana. The model is acceptable from the viewpoint of methodological soundness and completeness.

4.27 TWSC Intersections Model 2

Source

Juan C.Pernia, Jian Jhon Lu, Xiangli Xie, Michael Weng and Debroah Snyder., Transportation Research Board Record, Florida, 2002.

Data

Development of Models to Quantify the Impacts of Signalization on Intersections Crashes.

Model

Poisson regression and Negative Binomial regression were used. The model developed for calculating the expected number of crashes was as follows:

$$a = \exp(b_o + b_1X_1 + b_2X_2 + ... + b_qX_n), \quad (4.23)$$

where:

$X_1, X_2, \dots, X_n$ = model variables,

$\beta_0, \beta_1, \dots, \beta_q$ = model parameters.

For the above models, the parameters estimates have been established for three-legged and four-legged intersections. The results of the parameter estimates for the aforementioned intersections with respect to all crashes are summarized in Table 4.47.

Table 4.47 Parameter Estimates

Variable Description	Name	Level / Units	Value	DOF	Coefficient Estimate	Standard Error
Intercept				1	0.6827	0.1229
Major road AADT	ADT	<15,000 vpd	0	1	0.2777	0.0641
		15,000~30,000	1			
		>=30,000 vpd	2			
Surrounding Land use	URBRUR	Urban Rural	1 0	1	0.1193	0.0866
Location	LOCATYP	Business	1	1	0.1705	0.1082

Type		Other	0			
Number of Lanes on major road	LANE	>4	1	1	0.2614	0.1214
		<=4	0			
Speed limit on major road	SPEED	>45 mph	1	1	-0.1695	0.1071
		>=45 mph	0			
Presence of Median on major road	MEDIAN	Divided	1	1	0.2752	0.0954
		Undivided	0			
Shoulder Treatment	SHOULD	Paved	1	1	-0.1679	0.0954
		Other	0			
Dispersion Parameter				1	0.4480	0.0522

Discussion

The methodology of Negative Binomial distribution used in this model is good and all the parameters and variables are well defined. The demographic condition of Florida is not similar to that of Indiana. The model is acceptable from the viewpoint of methodological soundness and completeness.

4.28 TWSC Intersections Model 3

Source

Carl Belanger, Traffic and Roadway Accident Analysis and Traffic Records Research (No.1467), Canada (Qubec), 1994.

Data

Estimation of Safety of Four-Legged Unsignalized Intersections.

Model

Negative Binomial regression and GLIM were used for the estimation of the safety of four-legged two-way stop controlled intersections. The model developed for the total intersection accidents is as follows:

$$a = b_0 (Q_1^{g_1} Q_2^{g_2}), \quad (4.24)$$

where:

Q_1 and Q_2 = major and minor flows in vpd.

The values of the parameters are defined in Table 4.48.

Table 4.48 Parameter Estimates

Functional form	γ_1	γ_2	K	Deviance (d.f.)	R^2
$a = 5.59 \times 10^{-6} Q_1^{g_1} Q_2^{g_2}$	0.42	0.51	2.95	164.64 (146)	.47, .56

Discussion

The results are incomplete with respect to error estimates. Moreover, the Quebec conditions are not similar to that of Indiana. Therefore, the model will not be used.

4.29 TWSC Intersections Model 4

Source

Andrew P. Tarko, Mayank Kanodia, FHWA/IN/JTRP-2003/19 (Joint transportation research program, Purdue University), Indiana, 2004.

Data

Hazard Elimination Program—Manual on Improving Safety of Indiana Road Intersections and Sections.

Model

Negative binomial distribution methodology was used to estimate the number of crashes. The following two models were developed keeping in mind crash severity:

$$a = kQ_1^{\beta_1} Q_2^{\beta_2} e^{\gamma F} \quad (4.25)$$

where:

a = expected number of crashes,

k, γ , β_i = regression coefficients,

Q_1 = average AADT on two legs of major road, in thousand vehicles per day,

Q_2 = average AADT on two legs of minor road, in thousand vehicles per day,

F = 1 if the intersection has a flasher, = 0 otherwise.

The parameter estimates for the model have been summarized in Table 4.49.

Table 4.49 Parameter Estimates

Variables	Injury Crashes		PDO Crashes		Total	
	Coeff.	Std.Error	Coeff.	Std.Error	Coeff.	Std.Error
LOGPROAADT	0.539	0.145	0.549	0.104	0.544	0.874
FLASH	0.593	0.282	0.512	0.224	0.539	0.208
Alpha (overdispersion)	0.455	0.139	0.127	0.536	0.192	0.571

Discussion

The model was developed for Indiana using sound methodology. Therefore, the model is useful for the study.

4.30 TWSC Intersections Model 5

Source

Vogt Andrew and Bared John, FHWA-RD-98-133, Minnesota 1985-1989, year of publication 1998.

Data

Accident Prediction Models for Two-Lane Rural Roads: Segments and Intersections.

Model

Negative binomial regression methodology and maximum likelihood methods available in the software packages SAS and LIMDEP were used to fit the models to the data. The model is applicable to four-legged rural intersections, with stop control on minor legs. The following model was developed to estimate the expected number of accidents:

$$a = Q_1^{g_1} Q_2^{g_2} \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_q X_n), \quad (4.26)$$

where:

Q, $X_1, X_2, \dots, X_n$ = model variables,

$\gamma_1, \gamma_2, \beta_1, \beta_2, \dots, \beta_q$ = model parameters.

The list of variables available for the analyses, their units, and the abbreviations used in the model estimation results are summarized in Table 4.50, while the parameter estimates used for total and injury accidents have also been summarized separately for the four types of intersections in Table 4.51, Table 4.52, Table 4.53, and Table 4.54.

Table 4.50 Variable description and units

Variable Abbreviation	Variable Description and Units
ADT1	Average daily traffic on mainline (vehicles per day)
ADT2	Average daily traffic on minor road (vehicles per day) AADT1>AADT2
HI	Sum of degree of curve (degrees per 100 ft) of each horizontal curve on major road any portion of which is within 250 feet of the intersection center divided by the number of such curves
VCI1	Sum of absolute change of grade (percent per 100 ft) for each crest curve on major road any portion of which is within 250 feet of the intersection center, divided by the number of such curves
SPD1	The average posted speed on major road in vicinity of the intersection (mph)
HAU	Intersection Angle (α in degrees) The absolute deviation from 90 degrees of the minor leg from the increasing direction of the major road
DRWY1	Number of residential and commercial driveways on the major road within 250 feet of the intersection center
HAZRAT1	Roadside Hazard rating on major road within 250 ft of the intersection center (from 1, least hazardous case, to 7, most hazardous case)
RTMAJ	1 if right-turn exists on major roads, 0 otherwise
PKLEFT1	Left turn percentage on major road during the peak hour (%)
LTLN1S	Left turn lane on major roads (0 = No, 1 = Yes)
SPD2	The average posted speed on minor road in vicinity of the intersection (mph)
MEDWDTH1	Median width on major roads (ft)

The four types of intersections addressed in this research include:

- I- Type I-Three-Legged stop controlled intersections of two-lane roads
- II- Type II-Four-Legged stop controlled intersections of two-lane roads
- III- Type III-Three-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads
- IV- Type IV-Four-Legged stop controlled intersections with two lanes on the minor and four lanes on the major roads

4.30(I) Three-Legged stop controlled intersections of two-lane roads

Table 4.51 Parameter Estimates

Variables	Total Accidents (Type I) Parameter Estimates (std. error, p-value)	Injury Accidents (Type I) Parameter Estimates (std. error, p-value)
Intercept	-12.992 (1.151, <0.001)	-13.037 (1.791, 0.0001)
LOG of AADT1	0.8052 (0.0639, 0.0001)	0.8122 (0.0973, 0.0001)
LOG of AADT2	0.5037 (0.0708, 0.0001)	0.4551 (0.0977, 0.0001)
HI	0.0339 (0.0327, 0.3004)	0.0335 (0.0327, 0.3047)
VCII	0.1726 (0.0677, 0.0108)	0.1869 (0.3657, 0.6092)
SPD1	0.0285 (0.0177)	0.0156 (0.0269, 0.5618)
HAU	0.0045 (0.0032, 0.1578)	0.0051 (0.0045, 0.2594)
DRWY1	-	-0.0120 (0.0714, 0.8671)
HAZRAT1	0.1726 (0.0677, 0.0108)	0.2065 (0.0930, 0.0263)
RTMAJ	0.2671 (0.1398, 0.0561)	0.3620 (0.1814, 0.0460)
Overdispersion	0.481	0.494

Discussion

The methodology is sound and all the parameters estimates were well defined. The regional weather conditions also resemble that of Indiana. Therefore, the model is acceptable from the viewpoint of methodological soundness, completeness, and regional adequacy.

4.30(II) Four-Legged stop controlled intersections of two-lane roads

Table 4.52 Parameter Estimates

Variables	Total Accidents (Type II) Parameter Estimates (std. error, p-value)	Injury Accidents (Type II) Parameter Estimates (std. error, p-value)
Intercept	-10.426 (1.317, <0.001)	-10.783 (1.77, 0.0001)
LOG of AADT1	0.6026 (0.0836, 0.0001)	0.6339 (0.1055, 0.0001)
LOG of AADT2	0.6091 (0.0694, 0.0001)	0.6229 (0.0870, 0.0001)
HI	0.0449 (0.0473, 0.3431)	0.0729 (0.0635, 0.2513)
VCI1	0.2885 (0.2576, 0.2628)	0.2789 (0.4623, 0.5464)
SPD1	0.0187 (0.0176, 0.2875)	0.0112 (0.0251, 0.6567)
HAU	-0.0049 (0.0033, 0.1341)	-0.0043 (0.0044, 0.3258)
DRWY1	0.1235 (0.0519, 0.0173)	0.0857 (0.0639, 0.1799)
HAZRAT1	-	-0.1255 (0.0720, 0.0889)
RTMAJ	-	0.0451 (0.1655, 0.7865)
Overdispersion	0.206	0.181

Discussion

The methodology is sound and all the parameter estimates were well defined. The regional weather conditions also resemble that of Indiana. Therefore, the model is acceptable from the viewpoint of methodological soundness, completeness, and regional adequacy.

4.30(III) Three-Legged stop controlled intersections with two lanes on the minor and four lanes on the major road

Table 4.53 Parameter Estimates

Variables	Total Accidents (Type III) Parameter Estimates (std. error, p-value)	Injury Accidents (Type III) Parameter Estimates (std. error, p-value)
Intercept	-12.2196 (2.3575, 0.0001)	-12.3246 (2.8076, 0.0001)
LOG of AADT1	1.1479 (0.2527, 0.0001)	1.1436 (0.2763, 0.0001)
LOG of AADT2	0.2624 (0.0866, 0.0024)	0.1357 (0.1029, 0.1872)
HAU	-	0.0230 (0.0131, 0.0790)
DRWY1	0.0391 (0.0239, 0.1023)	-
MEDWDTH1	-0.0546 (0.0249, 0.0285)	-
Overdispersion	0.3893	0.3787

Discussion

The methodology is sound and all the parameter estimates were well defined. The regional weather conditions also resemble that of Indiana. Therefore, the model is acceptable from the viewpoint of methodological soundness, completeness, and regional adequacy.

4.30(IV) Four-Legged stop controlled intersections with two lanes on the minor road and four lanes on the major road

Table 4.54 Parameter Estimates

Variables	Total Accidents (Type IV) Parameter Estimates (std. error, p-value)	Injury Accidents (Type IV) Parameter Estimates (std. error, p-value)
Intercept	-9.4631 (2.5991, 0.0003)	-12.5296 (2.9908, 0.0001)
LOG of AADT1	0.8503 (0.2779, 0.0022)	0.9505 (0.3284, 0.0038)
LOG of AADT2	0.3294 (0.1255, 0.0087)	0.3237 (0.1645, 0.0491)
PKLEFT1	0.1100 (0.0412, 0.0076)	0.0994 (0.0433, 0.0216)
LTLN1S	-0.4841 (0.2311, 0.0362)	-
SPD2	-	0.0339 (0.0179, 0.0577)
Overdispersion	0.4578	0.4308

Discussion

The methodology is sound and all the parameter estimates were well defined. The regional weather conditions also resemble that of Indiana. Therefore, the model is acceptable from the viewpoint of methodological soundness, completeness, and regional adequacy.

4.31 TWSC Intersections Model 6

Source

Washington Simon, Persaud Bhagwant, Lyon Craig and Oh Jutaek , FHWA-RD-03-037, 2003.

Validation of Accident Models for Intersections.

Type I: Minnesota 1985-98(270 sites), Georgia 1996-97 (116 sites)

Type II: Minnesota 1985-98(250 sites), Georgia 1996-97 (108 sites)

Type III: California 1991-1998 (218 sites), Michigan 1993-97 (24 sites) and Georgia 1996-97 (52 cases)

Type IV: California 1991-1998 (152 sites), Michigan 1993-97 (18 sites) and Georgia 1996-97 (52 cases)

Model

Negative binomial regression methodology. The model is applicable to four-legged rural intersections, with stop control on the minor legs. The following model was developed to estimate the expected number of accidents:

$$a = Q_1^{s_1} Q_2^{s_2} \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_q X_n), \quad (4.27)$$

where:

$Q, X_1, X_2, \dots, X_n$ = model variables,

$\gamma_1, \gamma_2, \beta_1, \beta_2, \dots, \beta_q$ = model parameters.

The list of variables available for the analyses, their units, and the abbreviations used in the model estimation results are summarized in Table 4.55, while the parameter estimates used for total and injury accidents have also been summarized separately for the four types of intersection in Table 4.56, Table 4.57, Table 4.58 and Table 4.59.

The four types of intersection addressed in this research include:

Type I-Three-Legged stop controlled intersections of two-lane roads,

Type II-Four-Legged stop controlled intersections of two-lane roads,

Type III-Three-Legged stop controlled intersections with two lanes on the minor and four lanes on the major roads,

Type IV- Four-Legged stop controlled intersections with two lanes on the minor and four lanes on the major roads.

Table 4.55 Variable description and units

Variable Abbreviation	Variable Description and Units
ADT1	Average daily traffic on mainline (vehicles per day)
ADT2	Average daily traffic on minor road (vehicles per day) AADT1>AADT2
HI	Sum of degree of curve (degrees per 100 ft) of each horizontal curve on major road any portion of which is within 250 feet of the intersection center divided by the number of such curves
VC11	Sum of absolute change of grade (percent per 100 ft) for each crest curve on major road any portion of which is within 250 feet of the intersection center, divided by the number of such curves
DRWY1	Number of residential and commercial driveways on the major road within 250 feet of the intersection center
SDR2	Right-side sight distance on minor road (feet)
PKTRUCK	Truck percentage passing through the intersection during the peak hour (%)
PKLEFT	Left turn percentage during the peak hour (%)
SPD2	The average posted speed on minor road in vicinity of the intersection (mph)
PKLEFT1	Left turn percentage on major road during the peak hour (%)
COMDRWY1	Commercial driveways on the major road within 250 feet of the intersection center
VE11	Sum of absolute change of grade (percent per 100 ft) for each crest curve on major road any portion of which is within 800 feet of the intersection center, divided by the number of such curves
HAU	Intersection Angle (degrees) variable defined where the angle between the major and minor road is measured from the far side of the minor road. Three-legged Intersections: Angle minus 90 if minor road is to the right of the major road in the increasing direction; 90 minus angle if minor road is to the left of the major road in the increasing direction Four-legged Intersections: (right angle - left angle) / 2
MEDWDTH1	Median width on major roads (ft)
MEDTYPE1	Median Type (0=no median, 1=panted, 2=curbed, 3=others)
HAZRAT1	Roadside Hazard rating on major road within 250 ft of the intersection center (from 1, least hazardous case, to 7, most hazardous case)
RT MAJ	1 if right-turn exists on major roads, 0 otherwise
LT MAJ	1 if left-turn exists on at least one approach of major roads, 0 otherwise

4.31(I) Three-Legged stop controlled intersections of two-lane roads

Table 4.56 Parameter Estimates

Variables	Total Accidents (Type I) Parameter Estimates (std. error, p-value)	Injury Accidents (Type I) Parameter Estimates (std. error, p-value)
Intercept	-8.825 (0.438, <0.001)	-9.727 (0.566, <0.001)
LOG of AADT1	0.7001 (0.0460, <0.001)	0.7174 (0.0596, <0.001)
LOG of AADT2	0.3785 (0.0470, <0.001)	0.3251 (0.0569, <0.001)
HI	0.0314 (0.0113, 0.005)	0.0337 (0.0139, 0.016)
RT MAJ	-0.1887 (0.0841, 0.025)	-
LT MAJ	-0.155 (0.110, 0.157)	-0.258 (0.132, 0.051)
HAZRAT1	0.1433 (0.0362, <0.001)	0.1678 (0.0457, <0.001)
VCI1	0.1204 (0.0362, <0.001)	0.0846 (0.0557, 0.129)
Overdispersion	0.263	0.278

Discussion

The methodology is sound and all the parameter estimates were well defined. The regional weather conditions also resemble that of Indiana. Therefore, the model is acceptable from the viewpoint of methodological soundness, completeness, and regional adequacy.

4.31(II) Four-Legged stop controlled intersections of two-lane roads

Table 4.57 Parameter Estimates

Variables	Total Accidents (Type II) Parameter Estimates (std. error, p-value)	Injury Accidents (Type II) Parameter Estimates (std. error, p-value)
Intercept	-9.248 (0.475, <0.001)	-10.303 (0.565, <0.001)
LOG of AADT1	0.7079 (0.0537, <0.001)	0.6392 (0.0638, <0.001)
LOG of AADT2	0.5153 (0.0529, <0.001)	0.6584 (0.0614, <0.001)
HI	-	0.0626 (0.0272, 0.021)
VCI1	0.0766 (0.0554, 0.167)	-
DRWY1	0.1375 (0.0282, <0.001)	0.0944 (0.0337, 0.005)
Overdispersion	0.278	0.278

Discussion

The methodology is sound and all the parameter estimates were well defined. The regional weather conditions also resemble that of Indiana. Therefore, the model is acceptable from the viewpoint of methodological soundness, completeness, and regional adequacy.

4.31(III) Three-Legged stop controlled intersections with two lanes on the minor road and four lanes on the major road

Table 4.58 Parameter Estimates

Variables	Total Accidents (Type III) Parameter Estimates (std. error, p-value)	Injury Accidents (Type III) Parameter Estimates (std. error, p-value)
Intercept	-10.1914 (1.5232, 0.0000)	-10.6443 (2.0474, 0.0000)
LOG of AADT1	0.8877 (0.1666, 0.0000)	0.8498 (0.2097, 0.0001)
LOG of AADT2	0.3228 (0.0585, 0.0000)	0.2188 (0.0949, 0.0212)
PKTRUCK	-	-0.0253 (0.0135, 0.0605)
PKLEFT	-	0.0333 (0.0188, 0.0758)
COMDRWY1	0.0681 (0.0281, 0.0154)	0.0627 (0.0353, 0.0756)
VEI1	0.1081 (0.0556, 0.0519)	-
HAU	0.0101 (0.0059, 0.0861)	0.0163 (0.0053, 0.0021)
MEDWDTH1	-0.0106 (0.0060, 0.0760)	-
MEDTYPE1	-0.3209 (0.1771, 0.0700)	-
HAZRAT1	-	0.1889 (0.0923, 0.0407)
Overdispersion	0.4229	0.5102

Discussion

The methodology is sound and all the parameter estimates were well defined. Therefore, the model is acceptable from the viewpoint of methodological soundness and completeness adequacy.

4.31(IV) Four-Legged stop controlled intersections with two lanes on the minor road and four lanes on the major road

Table 4.59 Parameter Estimates

Variables	Total Accidents (Type IV) Parameter Estimates (std. error, p-value)	Injury Accidents (Type IV) Parameter Estimates (std. error, p-value)
Intercept	-7.4713 (1.8930, 0.0001)	-7.3927 (2.1279, 0.0005)
LOG of AADT1	0.7350 (0.1849, 0.0001)	0.5008 (0.2186, 0.0220)
LOG of AADT2	0.2390 (0.0926, 0.0099)	0.3027 (0.1341, 0.0240)
SDR2	-0.0003 (0.0001, 0.0403)	-
PKTRUCK	-0.0479 (0.0110, 0.0000)	-0.0520 (0.0127, 0.0000)
PKTHRU2	0.249 (0.0085, 0.0034)	-
PKLEFT	0.0229 (0.0118, 0.0525)	-
SPD2	-	0.0289 (0.0145, 0.0465)
PKLEFT1	-	0.0523 (0.0128, 0.0000)
Overdispersion	0.4001	0.4671

Discussion

The methodology is sound and all the parameter estimates were well defined. Therefore, the model is acceptable from the viewpoint of methodological soundness and completeness adequacy.

4.32 AWSC Intersections Model 1

Source

Andrew P. Tarko, Mayank Kanodia, FHWA/IN/JTRP-2003/19 (Joint transportation research program, Purdue University), Indiana, 2004.

Data

Hazard Elimination Program—Manual on Improving Safety of Indiana Road Intersections and Sections.

Model

Negative binomial distribution methodology was used to estimate the number of crashes. The following two models were developed keeping in mind crash severity:

$$a = e^k (Q_1 Q_2)^b e^{gF}, \quad (4.28)$$

where:

a = expected number of crashes,

K, γ , β_i = Coefficients to be determined,

Q_1 = average AADT on two legs of major road, in thousand vehicles per day,

Q_2 = average AADT on two legs of minor road, in thousand vehicles per day,

F = whether the intersection has a flasher.

* $Q_1 > Q_2$

The parameter estimates for the model have been summarized in Table 4.60.

Table 4.60 Parameter Estimates

Variables	Injury Crashes		PDO Crashes		Total	
	Coeff.	Std.Error	Coeff.	Std.Error	Coeff.	Std.Error
LOGPROAADT	0.342	0.287	0.674	0.128	0.611	0.118

FLASH	0.608	0.794	0.437	0.360	0.469	0.342
Alpha (overdispersion)	1.992	0.627	0.284	0.786	0.443	0.916

Discussion

The model was developed for Indiana using sound methodology. Therefore, the model is useful for the study.

4.33 Roundabouts Model 1

Source

Bhangwant N Persuad, Richard Garder, Dominique Lord, Transportation Research Record, Toronto, Maryland and Florida, 1999 and also at

http://www.dot.state.ny.us/roundabouts/files/insurance_report.pdf

Data

Observational Before-After Study of the Safety Effect of US Roundabout Conversions Using Empirical Bayes Method.

Model

The model is a modified version of model 1 for TWSC intersections. It is obtained by multiplying the model 1 with the crash modification factors shown in Table 4.61.

Table 4.61 Crash Modification Factors

Severity	Single-lane Roundabout		Multi-lane Roundabout
	Urban location	Rural location	Urban location
ALL	0.82	0.42	0.65
INJURY	0.12	0.18	0.26

4.34 Roundabouts Model 2

Source

Maycock G., and Hall R.D, Transport and Road Research Laboratory (TRRL), England, 1984.

Data

Crashes at four-arm roundabouts. Crash data for periods of four to six years for 84 four-leg roundabouts of all sizes with various number of approach and entry lanes..

Model

The models are based on generalized linear regression of the exponential form, which assumes a Poisson distribution.

1. Entry-Circulating:

$$A = 0.052Q_e^{0.7}Q_c^{0.4} \exp(-40C_e + 0.14e - 0.007ev - \frac{1}{1 + \exp(4R - 7)}) + 0.2P_m - 0.01\theta, \quad (4.29)$$

where:

A = personal injury crashes (including fatalities) per year per roundabout approach,

Q_e = entering flow (1,000s of vehicles/day),

Q_c = circulating flow (1,000s of vehicles/day),

C_e = entry curvature = $1/R_e$,

e = entry width (m),

v = approach width (m),

R = ratio of inscribed circle diameter/central island diameter,

P_m = proportion of motorcycles (%),

θ = angle to next leg, measured centerline to centerline (degrees).

2. Approaching:

$$A = 0.057Q_e^{1.7} \exp(20C_e - 0.1e), \quad (4.30)$$

where:

A = personal injury crashes (including fatalities) per year at roundabout approach or leg,

Q_e = entering flow (1,000s of vehicles/day),

C_e = entry curvature = $1/R_e$,

R_e = entry path radius for the shortest vehicle path (m),

e = entry width (m).

3. Single Vehicle:

$$A = 0.064Q_e^{0.8} \exp(25C_e + 0.2V - 45C_a), \quad (4.31)$$

where:

A = personal injury crashes (including fatalities) per year at roundabout approach or leg,

Q_e = entering flow (1,000s of vehicles/day),

C_e = entry curvature = $1/R_e$,

R_e = entry path radius for the shortest vehicle path (m),

V = approach width (m),

C_a = approach curvature = $1/R_a$,

R_a = approach radius (m), defined as the radius of a curve between 50 m (164 ft) and 500 m (1,640 ft) of the yield line.

4. Other (Vehicle):

$$A = 0.064Q_e^{0.8} \exp(25C_e + 0.2v - 45C_a) , \quad (4.32)$$

where:

A = personal injury crashes (including fatalities) per year at roundabout approach or leg,

Q_{ec} = product $Q_e * Q_c$,

Q_e = entering flow (1,000s of vehicles/day),

Q_c = circulating flow (1,000s of vehicles/day),

P_m = proportion of motorcycles.

5. Pedestrian:

$$A = 0.029Q_{ep}^{0.5} , \quad (4.33)$$

where:

A = personal injury crashes (including fatalities) per year at roundabout approach or leg,

Q_{ep} = product ($Q_e + Q_{ex}$). Q_p ,

Q_e = entering flow (1,000s of vehicles/day),

Q_{ex} = exiting flow (1,000s of vehicles/day),

Q_p = pedestrian crossing flow (1,000s of pedestrians/day),

Discussion

No crash prediction models exist yet for U.S. roundabouts. As a result, crash prediction models from other countries can be useful in understanding the relative effects of various geometric characteristics on the number of accidents that might be expected. However, since crash magnitude and types vary from country to country, the application of a crash prediction model from another country may not accurately predict crash frequencies in the U.S.

All-way Stop Controlled Intersections

Only one model is presented concerning all-way stop controlled intersections. The variables for this model are the AADT on major and minor road in thousand vehicles per day and the presence of a flasher at the intersection.

Roundabouts

Model 1 is a modified version of model 1 for TWSC intersections. It is obtained by multiplying the model 1 with the crash modification factors shown in Table 4.61. Model 2 was developed in England and thus the application of this model to the U.S. may not accurately predict crash frequencies.

4.35 Models Selected for Implementation

Based on the literature review for all types of segment and intersection, the following models have been selected for implementation and are presented in Table 4.62.

Table 4.62 Models with Selected Variables for Implementation

Facility Type	Model Selected
Roadway Segment	
Rural Two Lane (RTL)	1, 2, 3
Rural Multilane (RML)	1, 2, 3
Rural Interstate (RI)	1
Urban Two Lane (UTL)	1, 2
Urban Multilane	1, 2, 3, 4
Urban Interstate	1, 2
Intersection	
Signalized	1, 3, 4, 5
Two-way Stop-Controlled (TWSC)	2, 6

All-way Stop-Controlled (AWSC)	1
Roundabout	1

The above mentioned models have been selected on the basis of the following criteria:

- Quality of the variable
- Region
- Variable Indiana
- Data available in Indiana

These adopted models for Indiana with exception to roundabouts have been adjusted by redefining model variables and combining slopes. The details are shown in Appendix A.

5 PRELIMINARY INDIANA MODELS

5.1 Basic Safety Performance Functions

5.1.1 Basic safety performance functions (BSPF) overview

Safety Performance Functions return the expected annual number of crashes for a particular facility based on the traffic volume and length of the segment. They are usually obtained by negative binomial modeling using traffic volume and section length. The exponent of the length is restricted to 1.0.

The general form of the BSPF is:

$$\begin{aligned}
 A &= k_1 \cdot L \cdot Q^{\beta_1} && , \text{ for segments} \\
 A &= k_2 \cdot Q^{\beta_1} \cdot Q^{\beta_2} && , \text{ for Signalized and TWSC intersections} \\
 A &= k_3 \cdot (Q \cdot Q)^{\beta_3} && , \text{ for AWSC intersections}
 \end{aligned}$$

where:

A = predicted number of crashes on a segment,

L = length of the segment as defined in section X.1.2,

Q = annual average daily traffic as defined in section X.1.2,

k_i, β_i = constants for specific severity level and facility type.

5.1.2 Basic safety performance functions used

The BSPFs used for this research were selected from the most recent and adequate models available for Indiana. Only Indiana models were used for basic safety performance function that returns an initial crash prediction value which later is modified by crash modification factors dependent on the roadway characteristics of a segment. Crash modification factors were developed by combining models from Indiana and different states as explained in the next section. The BSPFs used in this research are provided below in Table 5.1.

Table 5.1 Basic Safety Performance Functions: segment and intersections

Facility	Basic Safety Performance Functions
Rural two-lane road segment	$a_{IF} = 0.208 \times L \times Q^{0.604}$
	$a_{PDO} = 0.712 \times L \times Q^{0.592}$
Rural multilane road segment	$a_{IF} = 0.107 \times L \times Q^{0.814}$
	$a_{PDO} = 0.634 \times L \times Q^{0.615}$
Urban two-lane road segment	$a_{IF} = 0.105 \times L \times Q^{1.080}$
	$a_{PDO} = 0.603 \times L \times Q^{0.896}$
Urban multilane road segment	$a_{IF} = 0.674 \times L \times Q^{0.435}$
	$a_{PDO} = 2.028 \times L \times Q^{0.460}$
Rural interstate	$a_{IF} = 0.044 \times L \times Q^{0.917}$
	$a_{PDO} = 0.169 \times L \times Q^{0.943}$
Urban interstate	$a_{IF} = 0.00048 \times L \times Q^{2.238}$
	$a_{PDO} = 0.0057 \times L \times Q^{1.954}$
Signalized intersection	$a_{IF} = Q_1^{0.6697} \times Q_2^{0.2509}$
	$a_{TOTAL} = Q_1^{0.4499} \times Q_2^{0.2699}$
Two-way stop-controlled intersection	$a_{IF} = Q_1^{0.7174} \cdot Q_2^{0.3251}$

Type I	$a_{TOTAL} = Q_1^{0.7001} \cdot Q_2^{0.3785}$
Two-way stop-controlled intersection Type II	$a_{IF} = Q_1^{0.6392} \cdot Q_2^{0.6584}$
	$a_{TOTAL} = Q_1^{0.7079} \cdot Q_2^{0.5153}$
Two-way stop-controlled intersection Type III	$a_{IF} = Q_1^{0.8498} \cdot Q_2^{0.2188}$
	$a_{TOTAL} = Q_1^{0.8877} \cdot Q_2^{0.3228}$
Two-way stop-controlled intersection Type IV	$a_{IF} = Q_1^{0.5008} \cdot Q_2^{0.3027}$
	$a_{TOTAL} = Q_1^{0.7350} \cdot Q_2^{0.2390}$
All-way stop-controlled intersection	
a_{IF} = injury/fatal crash frequency in Indiana, in crash per year, a_{PDO} = pedestrian damage only (PDO) crash frequency in Indiana, in crash per year, Q = AADT entering the intersection or moving along the road segment, in thousand vehicles per day, L = road segment length, in miles.	

Table 5.2 Basic Safety Performance Functions roundabouts

A		Single Lane Roundabout		Multi Lane Roundabout
		Urban Location	Rural Location	Urban Location
Four-legged Intersection	All	$0.82 \cdot \ln(-9.886) \cdot Q^{1.202} \cdot P^{0.376}$	$0.42 \cdot \ln(-9.886) \cdot Q^{1.202} \cdot P^{0.376}$	$0.65 \cdot \ln(-9.886) \cdot Q^{1.202} \cdot P^{0.376}$
	Injury	$0.12 \cdot \ln(-8.613) \cdot Q^{0.901} \cdot P^{0.197}$	$0.18 \cdot \ln(-8.613) \cdot Q^{0.901} \cdot P^{0.197}$	$0.26 \cdot \ln(-8.613) \cdot Q^{0.901} \cdot P^{0.197}$
Three-legged Intersection	All	$0.82 \cdot \ln(-8.051) \cdot Q^{1.053} \cdot P^{0.559}$	$0.42 \cdot \ln(-8.051) \cdot Q^{1.053} \cdot P^{0.559}$	$0.65 \cdot \ln(-8.051) \cdot Q^{1.053} \cdot P^{0.559}$
	Injury	$0.12 \cdot \ln(-11.364) \cdot Q^{1.225} \cdot P^{0.466}$	$0.18 \cdot \ln(-11.364) \cdot Q^{1.225} \cdot P^{0.466}$	$0.26 \cdot \ln(-11.364) \cdot Q^{1.225} \cdot P^{0.466}$
a – Number of crashes. Q – Total entering AADT, in thousand vehicles per day. P – Minor road proportion of AADT.				

5.2 Crash Modification Factors

This research has attempted to use as much of the existing safety performance function as possible by extracting crash modification factors from existing models and combining them. In some cases, different models represented the same roadway characteristic differently, which required redefining some of the original variables by changing units or scale. The overall process of combining various crash modification factors included four steps:

- (1) Redefining the original model variables and recalculating corresponding slopes
- (2) Combining the slopes for the same model variables,
- (3) Linking the model variables and TransCAD variables, and
- (4). Calculating the average values of model variables.

Steps 1, 2 and 3 are described in detail in Appendix A.

5.2.1 Crash modification factors overview

A crash modification factor (CMF) is defined as the increment effects of individual geometric design and traffic control elements. CMFs are used in accident prediction methodology to represent the effects on safety of specific geometric design and traffic control features. In order to develop CMFs, a critical review of the safety literature discussed in Chapter 2 was undertaken. By using the findings of the literature review as the basis for (1) selecting the final set of geometric elements/variables for which CMFs could be developed and (2) quantifying those CMFs. These CMFs are based on the coefficients or parameter values and standard error of the coefficients. For roadway segments, the final CMFs included all of the variables available in the roadway segment-based models from different states.

The general form of the crash modification factors is:

$$CMF = \prod_i CMF_i = \prod_i \exp(a_i \cdot (X_i - \bar{X}_i)), \quad (5.5)$$

where:

CMF = overall crash modification factor for a segment,

CMF_i = crash modification factor for variable i,

a_i = coefficient of the variables i,

$\bar{X}_i$ = average value for variable i for a specific facility type.

The preliminary CMF for segments are presented in **Error! Reference source not found..**

Table 5.3 Crash modification factors (CMF)

Facility	Severity Type	CMF=EXP(SLOPE(X-X_MEAN))											
		LW		SW		ISW		MW		AC		NTL	
		SLOPE	X_MEAN	SLOPE	X_MEAN	SLOPE	X_MEAN	SLOPE	X_MEAN	SLOPE	X_MEAN	SLOPE	X_MEAN
Rural Two-Lane	TOTAL	-0.0598	11.34	-0.0256	3.69								
	IF	-0.0831	11.34	-0.0442	3.69								
	PDO	-0.0524	11.34	-0.0105	3.69								
Rural Multilane	TOTAL	-0.1379	11.89	-0.094	8.39	-0.1329	3.29	-0.0172	42.38	0.0044	2.60		
	IF	-0.1472	11.89	-		-0.0697	3.29	-0.0051	42.38	0.2412	2.60		
	PDO	-		-		-0.0455	3.29	-0.0048	42.38	0.1153	2.60		
Rural Interstate	TOTAL	-0.5658	12.00			-1.3799	4.70	-0.0100	63.27				
	IF	-0.4179	12.00			-1.0011	4.70	-					
	PDO	-0.5706	12.00			-1.3338	4.70	-0.0108	63.27				
Urban Two-Lane	TOTAL	-0.1759	11.85	0.0563	2.89								
	IF	-0.1527	11.85	0.0754	2.89								
	PDO	-0.1658	11.85	0.0524	2.89								
Urban Multilane	TOTAL	-0.1541	11.88			-				0.5632	2.80	-1.026	4.1
	IF	-0.1521	11.88			-0.205	1.75			0.6485	2.80	-1.095	4.1
	PDO	-0.1525	11.88			-				0.5231	2.80	-0.949	4.1
Urban Interstate	TOTAL	-0.6507	12.00			-0.2438	5.57	-0.0176	52.05				
	IF	-0.9166	12.00			-0.1783	5.57	-					
	PDO	-0.7974	12.00			-0.2544	5.57	-0.0156	52.05				

Table 5.3 continued

[illegible]

5.3 Preliminary Models

In this section, a preliminary set of models for all facility types, including roadway segment and intersections, will be presented along with the definitions of all the variables used. These models have been selected from the models discussed in the previous section of this chapter, after redefining the original model variables and recalculating the corresponding slopes. The selected models per facility are shown below in Table 5.4 to Table 5.9. The subsequent section contains the definitions of the model variables.

5.3.1 Segments

Model for RTL

Table 5.4 Parameter Estimates for RTL

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.0831	-0.0524	-0.0598
SW	-0.0442	-0.0105	-0.0256
ADCURV	0.2158	-	0.1421
AGRAD	0.0528	0.0775	0.0720

Model for RML

Table 5.5 Parameter Estimates for RML

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.1472	-	-0.1379
SW	-	-	-0.0940
ISW	-0.0697	-0.0455	-0.1329
MW	-0.0051	-0.0048	-0.0172
AC	0.2412	0.1153	0.0044
FC	-	-	-0.5720

Model for RI**Table 5.6 Parameter Estimates for RI**

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.4179	-0.5706	-0.5658
ISW	-1.0011	-1.3338	-1.3799
MW	-	-0.0108	-0.0100

Model for UTL**Table 5.7 Parameter Estimates for UTL**

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.1527	-0.1658	-0.1759
LTL	0.4522	0.3783	0.3859
CRB	-0.2320	-0.2686	-0.2951
ST Surfaced	-0.7707	-0.8547	-0.8889
ST Stabilized	-0.5138	-0.5698	-0.5926
ST combination	-0.6423	-0.7123	-0.7408
ST Earth	-0.2569	-0.2849	-0.2963
CELT	-	-	-1.5440
SW	0.0754	0.0524	0.0563

Model for UML**Table 5.8 Parameter Estimates for UML**

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.1521	-0.1525	-0.1541
NTL	-1.0950	-0.9490	-1.0260
AC	0.6485	0.5231	0.5632
CRB	-0.3766	-0.2786	-0.2990
OSP	-0.5250	-0.6690	-0.6310
CLTL	-2.3580	-1.7670	-1.9360
LTL	-0.3558	-0.3052	-0.3058
ISW	-0.2050	-	-

Model for UI

Table 5.9 Parameter Estimates for UI

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.9166	-0.7974	-0.6507
ISW	-0.1783	-0.2544	-0.2438
MW	-	-0.0156	-0.0176

5.3.2 Intersection models

➤ Type 1- Three-Legged stop controlled intersections of two-lane roads

Table 5.10 Parameter Estimates for Type 1 - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.7174	0.7001
LOG AADT2	0.3251	0.3785
URBRUR	-	0.1193
SHOULD	-	-0.1679
RTMAJ	-	-0.1887
LTMAJ	-0.258	-0.155

➤ **Type II- Four-Legged stop controlled intersections of two-lane roads**

Table 5.11 Parameter Estimates for Type II - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.6392	0.7079
LOG AADT2	0.6584	0.5153
URBRUR	-	0.1193
SHOULD	-	-0.1679

➤ **Type III- Three-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads**

Table 5.12 Parameter Estimates for Type III - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.8498	0.8877
LOG AADT2	0.2188	0.3228
URBRUR	-	0.1193
SHOULD	-	-0.1679
MEDWDTH1	-	-0.0106
MEDTYPE	-	0.6418
PKTRUCK	-0.0253	-

- **Type IV- Four-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads**

Table 5.13 Parameter Estimates for Type IV - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.5008	0.7350
LOG AADT2	0.3027	0.2390
URBRUR	-	0.1193
SHOULD	-	-0.1679
PKTRUCK	-0.0520	-0.0479

Table 5.14 Parameter Estimates for Type V - All-Way Stop Controlled Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.3423	0.611
LOG AADT2	0.3423	0.611

Table 5.15 Parameter Estimates for Type VI – Signalized Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.6687	0.4499
LOG AADT2	0.2509	0.2699
URBRUR	-0.2677	0.0949
SHOULD	-	-0.1102
LANE	0.0916	0.0673
PKTRUCK	0.0323	0.0315

Table 5.16 Preliminary equations for segments

Facility	Safety Performance Functions
Rural two-lane road segment	$a_{IF} = 0.208 \times L \times Q^{0.604} \times e^{-0.0831(LW-11.34)-0.0442(SW-3.69)+0.2158(ADCURVE-1.10)+0.0528(AGRAD-1.57)}$
	$a_{PDO} = 0.712 \times L \times Q^{0.592} \times e^{-0.0524(LW-11.34)-0.1105(SW-3.69)+0.0775(AGRAD-1.57)}$
Rural multilane road segment	$a_{IF} = 0.107 \times L \times Q^{0.814} \times e^{-0.1472(LW-11.89)-0.0697(ISW-3.29)-0.0051(MW-42.38)+0.2412(AC-2.60)}$
	$a_{PDO} = 0.634 \times L \times Q^{0.615} \times e^{-0.0455(ISW-3.29)-0.0048(MW-42.38)+0.1153(AC-2.60)}$
Urban two-lane road segment	$a_{IF} = 0.105 \times L \times Q^{1.080} \times e^{\frac{-0.1527(LW-11.85)+0.4522(LTL-0.136)-0.2320(CRB-0.49)+0.0754(SW-2.89)-0.7707(ST\_Surfaced-0.15)-0.5138(ST\_Stabilized-0.21)-0.6423(ST\_Combination-0.19)-0.2569(ST\_Earth-0.01)}{0.2569(ST\_Earth-0.01)}}$
	$a_{PDO} = 0.603 \times L \times Q^{0.896} \times e^{\frac{-0.1658(LW-11.85)+0.3783(LTL-0.136)-0.2686(CRB-0.49)-0.8547(ST\_Surfaced-0.15)-0.5689(ST\_Stabilized-0.21)-0.7123(ST\_Combination-0.19)-0.2849(ST\_Earth-0.01)+0.0524(SW-2.89)}{0.0524(SW-2.89)}}$
Urban multilane road segment	$a_{IF} = 0.674 \times L \times Q^{0.435} \times e^{\frac{-0.1521(LW-11.88)-1.095(NTL-4.1)+0.6485(AC-2.80)-0.3766(CRB-0.0953)-0.5250(OSP-0.55)-2.358(CELT-0.125)-0.3558(LTL-0.191)-0.205(ISW-1.75)}{0.5250(OSP-0.55)-2.358(CELT-0.125)-0.3558(LTL-0.191)-0.205(ISW-1.75)}}$
	$a_{PDO} = 2.028 \times L \times Q^{0.460} \times e^{\frac{-0.1525(LW-11.88)-0.9490(NTL-4.1)+0.5231(AC-2.80)-0.2786(CRB-0.0953)-0.669(OSP-0.55)-1.767(CELT-0.12)-0.3052(LTL-0.191)}{0.669(OSP-0.55)-1.767(CELT-0.12)-0.3052(LTL-0.191)}}$
Rural interstate	$a_{IF} = 0.044 \times L \times Q^{0.917} \times e^{-0.4179(LW-12)-1.0011(ISW-4.70)}$
	$a_{PDO} = 0.169 \times L \times Q^{0.943} \times e^{-0.5706(LW-12)-1.3338(ISW-4.70)-0.0108(MW-63.27)}$
Urban interstate	$a_{IF} = 0.00048 \times L \times Q^{2.238} \times e^{-0.9166(LW-12)-0.1783(ISW-5.57)}$
	$a_{PDO} = 0.0057 \times L \times Q^{1.954} \times e^{-0.7974(LW-12)-0.2544(ISW-5.57)-0.0156(MW-52.05)}$
a_{IF} = injury/fatal crash frequency in Indiana, in crash per year, a_{PDO} = PDO crash frequency in Indiana, in crash per year, Q = AADT entering the intersection or along the road segment, in thousand vehicles per day, and L = road segment length, in miles.	

Table 5.5.17 Preliminary equations for intersections (Total crashes)

Intersection type	INT_TYPE	Equation
TWSC Type-I	1	$0.000147 \cdot Q_1^{0.7001} \cdot Q_2^{0.3785} \cdot e^{\{0.1193 \cdot (URBRUR - 1) - 0.1679 \cdot (SHOULD - 1) - 0.1887 \cdot (RTMAJ - 1) - 0.1550 \cdot (LTMAJ - 1)\}}$
TWSC Type II	2	$0.000096 \cdot Q_1^{0.7079} \cdot Q_2^{0.5153} \cdot e^{\{0.1193 \cdot (URBRUR - 1) - 0.1679 \cdot (SHOULD - 1)\}}$
TWSC Type III	3	$0.000037 \cdot Q_1^{0.8877} \cdot Q_2^{0.3228} \cdot e^{\{0.1193 \cdot (URBRUR - 1) - 0.1679 \cdot (SHOULD - 1) - 0.0106 \cdot (MEDWIDTH - 1) + 0.6418 \cdot (MEDTYPE - 1)\}}$
TWSC Type IV	4	$0.000569 \cdot Q_1^{0.7350} \cdot Q_2^{0.2390} \cdot e^{\{0.1193 \cdot (URBRUR - 1) - 0.1679 \cdot (SHOULD - 1) - 0.0479 \cdot (PKTRUCK - 1)\}}$
AWSC	5	$0.000210 \cdot (Q_1 \cdot Q_2)^{0.6111}$
Signalized Intersection	6	$0.005784 \cdot Q_1^{0.4499} \cdot Q_2^{0.2699} \cdot e^{\{0.0949 \cdot (URBRUR - 1) - 0.1102 \cdot (SHOULD - 1) + 0.0673 \cdot (LANE - 1) + 0.0315 \cdot (PKTRUCK - 1)\}}$

Table 5.18 Preliminary equations for intersections (Injury crashes)

Intersection type	INT_TYPE	Equation
TWSC Type-I	1	$0.000060 \cdot Q_1^{0.7174} \cdot Q_2^{0.3251} e^{\{-0.2580(LTMAJ-1)\}}$
TWSC Type II	2	$0.000034 \cdot Q_1^{0.6392} \cdot Q_2^{0.6584}$
TWSC Type III	3	$0.000024 \cdot Q_1^{0.8498} \cdot Q_2^{0.2188} e^{\{-0.0253(PKTRUCK-1)\}}$
TWSC Type IV	4	$0.000616 \cdot Q_1^{0.5008} \cdot Q_2^{0.3027} e^{\{-0.0520(PKTRUCK-1)\}}$
AWSC	5	$0.003675 \cdot (Q_1 \cdot Q_2)^{0.3423}$
Signalized Intersection	6	$0.000115 \cdot Q_1^{0.6697} \cdot Q_2^{0.2509} e^{\{-0.2677(URBRUR-1)+0.0916(LANE-1)+0.0323(PKTRUCK-1)\}}$
Q ₁ - major road AADT (veh/day), Q ₂ -minor road AADT(veh/day)		

5.4 Model Variables

This section defines *model variables*, presents the types of model variables, and how they were derived and how they are used. In the second part of this section, a complete set of precise definition for all model variables is provided.

5.4.1 Model variables overview

Model variables are variables used either in crash modification factors (CMF) or in a basic safety performance function (BSPF), both of which are explained in the following sections of this chapter. Model variables express roadway characteristics that have an effect on the predicted number of crashes. Variable input into the BSPFs and CMFs from this chapter should be collected according to the definitions described in this section. There are two types of model variables used in CMFs: continuous and binary. Continuous variables measure specific roadway geometric characteristic:

- Lane width (LW)
- Shoulder width (SW)
- Inside shoulder width (ISW)
- Median width (MW)
- Access control (AC)
- Number of through lanes (NTL)
- Average degree of curve (ADCURVE)
- Average grade on section (AGRAD)
- Surrounding Land Use (URBRUR)
- Number of Lanes (LANE)
- Truck Percentage (PKTRUCK)
- Right-Turn Lane (RT MAJ)
- Left-Turn Lane (LT MAJ)
- Median Width (MEDWDTH1)

- Median Type (MEDTYPE1)

Binary variables express the presence of a specific roadway feature: 1- if the feature is present; 0 - if the feature is not present.

- Curb presence (CRB)
- Functional class (FC)
- Outside shoulder presence (OSP)
- Exclusive left turn lane (LT)
- Continuous left turn lane (CLTL)
- Surfaced shoulder (ST_Surfaced)
- Stabilized shoulder (ST_Stabilized)
- Combination shoulder (ST_Combined)
- Earth shoulder (ST_Earth)
- Paved Shoulder (SHOULD)

5.4.2 Model variables definitions

Segments model variables

Q (Annual average daily traffic, AADT) - An average daily value that represents all days of the reporting year. For two-way facilities, it is the AADT for both directions; for one-way streets, it is the directional AADT. Growth factors must be applied if the AADT is not derived from current year counts.

L (Length of section) - Length of a section measured along the centerline of a roadway. For independently aligned, divided highways, use centerline length as the average of the lengths of the directional roadways, measured along their centerlines.

Lane width (LW) - For multilane roads, it is the prevailing width of the travel lane on a section rounded to the nearest foot. The width of the travel lane is measured between the lane striping. For two-lane roads, lane width is the distance from the centerline or delineation to where the pavement/shoulder surface changes or to the pavement lane striping if the shoulder and pavement surface are the same. A parking lane is excluded from lane width determination for all facility types. Where there is no delineation

between the through traffic lane and the shoulder or parking lane, or where there is no centerline, estimate a reasonable split between the actual width used by the traffic and the shoulder or parking lane based on state/local design guidelines.

Shoulder width (SW) - Shoulder width is the width of the shoulder, which includes rumble strips and gutter pans and does not include bike lanes and parking lanes. For two lane roads, it is the predominant width of the shoulder on a section; if there is no predominant width, enter the average width of both shoulders. For multilane roads, it is the predominant width of the outside shoulders; if there is no predominant width, enter the average width of the outside shoulders. Shoulder width should be recorded to the nearest foot.

Inside shoulder width (ISW) - Inside shoulder width is the predominant width of the inside shoulder, which includes rumble strips and gutter pans and does not include bike lanes and parking lanes. If there is no predominant width, use the average of both inside shoulders on a given section. Shoulder width should be recorded to the nearest foot.

Median width (MW) - Predominant width of the median, which includes the inside shoulders, if any, measured between the inside edges of the through lanes to the nearest foot. If the median width is greater than 100 feet, enter 100. Median width does not include turning bays cut into the median.

Access Control (AC) - Measures the degree of access control on a sample roadway sections. Enter 1 for full access control (preference given to through traffic movements by providing interchanges with selected public roads and by prohibiting crossing at grade and direct driveway connections); 2 for partial access control (preference given to through traffic movement; in addition to interchanges, there may be some crossings at-grade with public roads, but direct private driveway connections have been minimized through the use of frontage roads or other local access restrictions); 3 for no access control (include all sections that do not meet the criteria above); control of curb cuts is not access control. If access control is not known, enter 1 for interstates, 2 for arterials, and 3 for collector and local roads.

Number of through lanes (NTL) – Determined according to the striping, if present, on multilane facilities, or according to traffic use or state/local design guidelines if no

striping or only centerline striping is present. NTL is the prevailing number of through lanes in both directions carrying through traffic in the off-peak period. Exclude auxiliary lanes, such as collector-distributor lanes, weaving lanes, frontage road lanes, parking and turning lanes, acceleration/deceleration lanes, toll collection lanes and truck climbing lanes as defined in *AASHTO Design Guide*.

Average degree of curve (ADCURVE) - expresses the average degree of curve weighted with the length of each curve on that section. It is computed with the following formula:

$$ADCURVE = \frac{\sum_i S_i \cdot D(a)_i}{\sum_i S_i} , \quad (5.2)$$

where:

S_i = length of horizontal curve i on section,

$D(a)_i$ = degree of curve i (per 100 ft arc definition).

Average grade on section (AGRAD) – is the average absolute grade on section weighted with the length of each curve on that section. Each grade and flat segment is coded as a separate curve. The sum of all curves equals the section length. AGRAD is computed to the following formula:

$$AGRAD = \frac{\sum_i S_i \cdot G_i}{\sum_i S_i} , \quad (5.3)$$

where:

S_i = length of curve i on section,

G_i = absolute grade of curve i on section.

Curb presence (CRB) – Binary variable indicating the presence of a curb in the vicinity of travel lane for two-lane roadways. For multilane roads, this variable indicates the

presence of an outside curb in the vicinity of the travel lane. Enter 1 if a curb is present in the direct vicinity of the travel lane, otherwise enter 0. If a parking lane exists and a curb is present, enter 0 for this variable.

Functional class (FC) – This binary variable indicates if a given segment is classified as a rural principal arterial. Enter 1 if the section is classified as rural principal arterial, enter 0 otherwise. For any urban facility, enter 0.

Outside shoulder presence (OSP) – Binary variable indicating the presence of outside shoulder. Enter 1 if section contains an outside shoulder, otherwise enter 0. This variable is used for a UML facility.

Left turn lane (LTL) – Binary variable indicating the presence of single exclusive left turn lane or multiple turning lanes that allow for simultaneous turns from all turning lanes; through movements are prohibited in these lanes. Enter 1 if the section contains a LTL, otherwise enter 0.

Continuous left turn lane (CLTL) - Binary variable indicating the presence of a continuous exclusive left turning lane, from intersection to intersection on a given section; through movements are prohibited in this lane. Enter 1 if the section contains a CLTL, otherwise enter 0.

Surfaced shoulder (ST_Surfaced) – Binary variable indicating the presence of a paved shoulder on a section. Enter 1 if the section contains a paved shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Stabilized shoulder (ST_Stabilized) – Binary variable indicating the presence of a stabilized shoulder on a section. Enter 1 if the section contains a stabilized shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Combination shoulder (ST_Combined) – Binary variable indicating the presence of a combination shoulder on a section. Enter 1 if the section contains a combined shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Earth shoulder (ST_Earth) - Binary variable indicating the presence of an earth shoulder. Enter 1 if the section contains an earth shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Intersections model variables

Major Road Approach Volume (Q1) – Approach volumes on the major links (maximum of two) of the intersection, average daily value that represents all days of the reporting year.

Minor Road Approach Volume (Q2) – Approach volumes on the minor links of the intersection, average daily value that represents all days of the reporting year.

Surrounding Land Use (URBRUR) - Binary variable indicating whether the intersection is located in a rural or an urbanized area. Enter 1 if the intersection is located in an urban area; enter 0 if the intersection is located in a rural area.

Paved Shoulder (SHOULD) - Binary variable indicating the presence of a paved shoulder at the side of the major road of the intersection. Enter 1 if a paved shoulder exists, enter 0 otherwise.

Number of Lanes (LANE) - The number of lanes is defined as the total number of approach lanes at an intersection including through, shared, and exclusive turning lanes. The number of lanes is determined according to the striping, if present, on multilane facilities, or according to traffic use or state/local design guidelines if no striping is present.

Truck Percentage (PKTRUCK) - Variable % indicating the percentage of trucks during the intersection's peak hour traffic.

Right-Turn Lane (RT MAJ) - Binary variable indicating the presence of right-turn lanes on major roads. Enter 1 if there is a right-turn lane on the major road, otherwise enter 0.

Left-Turn Lane (LT MAJ) - Binary variable indicating the presence of left-turn lanes on at least one approach of major roads. Enter 1 if there is a left-turn lane on at least one approach of major roads, otherwise enter 0.

Median Width (MEDWDTH1) - Variable indicating the median width on major roads measured in feet. It is the predominant width of the median, which includes the inside shoulders, if any, measured between the inside edges of the through lanes to the nearest foot.

Median Type (MEDTYPE) - Variable indicating the type of the median on the major approach of the intersection. Enter 0 if no median exists, enter 1 for a painted median, enter 2 for a curbed median, or enter 3 for other types of medians.

5.5 Evaluating preliminary models for segments

This section evaluates the usefulness of the CMFs by comparing the BSPFs with the full-form SPFs as presented in the previous section. The theory is that the full-form models perform better than the BSPFs if the CMFs improve the crash predictions and therefore can be transferred from other states to Indiana conditions.

5.5.1

5.5.2 Overall Framework

The evaluation began with an application of the obtained BSPFs and SPFs to the geometric and traffic data available for Indiana for various types of road segments. The CMFs were evaluated as follows:

1. Predict the number of crashes at two levels of severity: PDO and injury/fatal for sample road segments with the BSPFs and SPFs.
2. Calibrate the BSPFs with calibration factors k . Repeat the same for the SPFs.
3. Compare the performance of the BSPFs and SPFs.

5.5.3 Criteria for evaluation

Three criteria were used for evaluating the BSPFs and SPFs:

- Difference between the calibration factors and the value one
- Standard error of prediction
- Difference in performance between BSPFs and full-form SPFs

5.5.4 Crash Predictions with BSPFs and SPFs

Concept evaluation will be based on the results of the BSPFs and the full form of the equation by adding the CMFs to the same BSPF. A comparison between the two will better illustrate the results and evaluating the concept.

BSPFs

The basic safety performance function includes the traffic volume and the link length as previously discussed in an earlier section (see Table 5.1). The BSPFs for various facility types and severity levels were applied to the Indiana state road network of 19,151 links with crashes recorded in 2004. The BSPFs were implemented in TransCAD without the CMFs. The results were tested in Microsoft Excel. The evaluation results are shown in Tables 5.15 and 5.16.

Table 5.19 Results for Injury/Fatal Predicted Crashes Using BSPF

Facility Type	Actual Crashes	Predicted Crashes	Calibration Factor	Standard Error ¹
Rural two-lane roads	3525	3271	1.08	0.76
Rural multilane roads	770	682	1.13	1.25
Rural interstate	534	685	0.78	1.78
Urban two-lane roads	1595	876	1.82	1.05
Urban multilane roads	2463	1666	1.48	1.93
Urban interstate	1017	2289	0.44	2.54

Note ¹ The standard error calculated after adjustment with the calibration factor

Table 5.20 Results for PDO Predicted Crashes Using BSPF

Facility Type	Actual Crashes	Predicted Crashes	Calibration Factor	Standard Error ¹
Rural two-lane roads	11386	11013	1.03	1.78
Rural multilane roads	2301	2423	0.95	3.24
Rural interstate	2546	2871	0.89	6.77
Urban two-lane roads	4366	3244	1.35	2.46
Urban multilane roads	6867	5414	1.27	4.83
Urban interstate	4669	7673	0.61	11.09

Note ¹ The standard error calculated after adjustment with the calibration factor

The predictions of injury/fatal crashes on non-freeway facilities are consistently lower than for interstate rural and urban segments. The results for PDO crashes are less consistent, although the trend is similar. The predictions for the rural roads are closer to the reported crashes than for the urban roads. The standard errors of crash prediction, even after applying the calibration factors, are considerable.

SPFs

Tables 5.17 and 5.18 show the results for injury/fatal and PDO crashes. The predictions of injury/fatal crashes are lower than the reported crashes for almost all of the facility types, except rural and urban interstate segments. The standard errors of crash prediction even after applying the calibration factors are considerable. These trends are similar to the ones observed for the BSPFs.

Table 5.21 Results for Injury/Fatal Predicted Crashes Using Safety Performance

Functions

Facility Type	Actual Crashes	Predicted Crashes	Calibration Factor	Standard Error ¹
	3525	3072	1.15	0.78
Rural multilane roads	770	630	1.22	1.22
Rural interstate	534	1206	0.44	2.36
Urban two-lane roads	1595	652	2.44	1.16
Urban multilane roads	2463	1132	2.18	2.24
Urban interstate	1017	2071	0.49	3.26

Note ¹ The standard error calculated after adjustment with the calibration factor

Table 5.22 Results for PDO Predicted Crashes Using Safety Performance Functions

Facility Type	Actual Crashes	Predicted Crashes	Calibration Factor	Standard Error ¹
Rural two-lane roads	11386	9764	1.17	1.80
Rural multilane roads	2301	2271	1.01	3.24
Rural interstate	2546	5802	0.44	11.94
Urban two-lane roads	4366	3598	1.21	3.12
Urban multilane roads	6867	2748	2.50	5.67
Urban interstate	4669	8761	0.53	20.66

Note ¹ The standard error calculated after adjustment with the calibration factor

5.5.5 Comparison of BSPFs and SPFs

The performance of the BSPFs and SPFs represented with the calibration factors and standard errors of prediction are compared in Tables 5.22 and 5.23. The calibration factors represent the systematic error of prediction. The larger the difference between the calibration factor and one, the larger is the systematic error of prediction. It is obvious that the full-form SPFs perform worse than the BSFs in almost all cases. The standard error of prediction represents the variability of the prediction adjusted with the calibration factor around the true number of crashes (number of reported crashes). It is calculated with the following equation:

$$S_{est} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n - 2}} = \sqrt{\frac{SSE}{df_{errors}}} = \sqrt{MSE} \quad (5.1)$$

where:

S_{est} = standard error of the estimate,

Y = observed value of the sample,

$\hat{Y}$ = predicted value of the sample,

df = degree of freedom of the sample,

SSE = sample standard error,

MSE = mean square root of the error.

Again, the performance of the full-form SPFs was worse than the BSPFs in almost all cases. The conclusion is that the CMFs derived from the models developed for other states should not be used in Indiana.

Table 5.23 Calibration Factors for BSPFs and SPFs

Facility Type	Calibration Factors			
	Injury/Fatal		PDO	
	BSPF	SPF	BSPF	SPF
Rural two-lane roads	1.08	1.15	1.03	1.17
Rural multilane roads	1.13	1.22	0.95	1.01
Rural interstate	0.78	0.44	0.89	0.44
Urban two-lane roads	1.82	2.44	1.35	1.21
Urban multilane roads	1.48	2.18	1.27	2.50
Urban interstate	0.44	0.49	0.61	0.53

Table 5.23 Standard Error of Prediction for BSPFs and SPFs

Facility Type	Standard Error of Prediction			
	Injury/Fatal		PDO	
	BSPF	SPF	BSPF	SPF
Rural two-lane roads	0.76	0.78	1.78	1.80
Rural multilane roads	1.25	1.22	3.24	3.24
Rural interstate	1.78	2.36	6.77	11.94
Urban two-lane roads	1.05	1.16	2.46	3.12
Urban multilane roads	1.93	2.24	4.83	5.67
Urban interstate	2.54	3.26	11.09	20.66

5.5.6 Summary

The results show that the CMFs derived from the existing SPFs developed for other states cannot be used to adjust Indiana BSPFs. The standard error of prediction was higher for the full-form SPFs than for the BSPFs. This conclusion is equivalent with the statement that it is difficult to transfer SPFs from other states when recalibration is limited to a single calibration factor while the regression coefficient in the front of all model variables remains unchanged. This result calls for full recalibration of the safety performance functions, which includes adjusting all the regression coefficients in the models.

5.6 Evaluating preliminary models for intersections

5.6.1 Overall Framework

Evaluation of the preliminary models for intersections was not possible due to missing fields in the segment input table from which the generated intersection table was obtained. The first step in evaluating the intersection preliminary models was to obtain the correct classification of intersections using video logs for Indiana and reclassifying all two-way stop controlled intersections. All signalized intersections and all-way stop controlled intersections were correctly classified by the Tool.

The checking of the two-way stop controlled intersections was processed at the INDOT Office of Research and Development using the video log data and friction query software. The process of checking involved the following steps.

- Identification of reference posts (start and end) for the two intersecting links by using friction query data software.
- Running the video log data (Pathview-II only).

- In order to find the county, there are two files in the video log data that needed to be uploaded (i.e., Indot04 and Indot05).
- These Indot04 and Indot05 files also contained district data.
- Once the Indot04 or Indot05 files were run, two windows were uploaded. The two windows were image and data windows. The image window contained the road view along with the road posts and the data file containing information for road posts along with junctions.
- Finding reference posts (start and end) for the links in video log data to identify the intersection.

The following two-way stop-controlled intersections were under consideration;

- I- Type I-Three-Legged stop controlled intersections of two-lane roads
- II- Type II-Four-Legged stop controlled intersections of two-lane roads
- III- Type III-Three-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads
- IV- Type IV- Four-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads

5.6.2 Intersection reclassification results

The checking process began with the Type-II intersection and it was found that the data provided contained reporting or feeding errors. The majority of the intersections were found to be either all-way stop controlled (**45%**) or signalized intersections (**21%**), and the remaining were two-way stop controlled intersections (**34%**).

For Type-III intersections, **37%** were found to be signalized, **43%** were two-way stop controlled intersections, and the remaining were either interstates or entry/exit ramps. Similarly, for Type-IV intersections, **30%** were found to be signalized, **59%** were two-way stop controlled intersections, and the remaining were either interstate roads or mile posts not available roads. Only **one** intersection was found to be an all-way stop

controlled intersection. There was no road post (RP) data available for ten intersections. For Type-I, **76%** were two-way stop controlled intersections, **18%** were signalized, **4%** were all-way stop controlled intersections, and the remaining intersections were unidentified. The checking process for all four types of intersections was completed (refer to Table 1).

Initially, it was targeted to check a maximum of 50 intersections per day, but it was found to be difficult to reach this goal once the checking process began because sometimes an intersection could not be easily found in the many video log files for each district. It became reasonable once the process began that the maximum number of intersections that can be checked in one day was approximately 30-35.

Table 1 shows the results of the work along with the total number of intersections for all four types of intersections.

5.6.3 Conclusion

The TWSC intersection classification by the developed tool is not consistent with the actual classification in real life. Errors in classification by the tool resulted from a lack of data in the segment table from which the intersection table was generated. This research developed correct classifications of state-state intersections which can be appended to the intersection table. If the intersection information in the segment table (HERS HPMS Input) is complete, then the developed tool will give correct classification.

Table 5.24 Summary of TWSC intersection classification checking.

Intersection Type	Total no. of Intersections	Findings (out of total)			Status
		Two-way	All-way	Signalized	
Type-I	329	251	13	58	Completed
Type-II	157	53	71	33	Completed
Type-III	158	67	2	60	Completed
Type-IV	152	89	1	45	Completed

6 CALIBRATION OF PRELIMINARY MODELS

This chapter focuses on the full calibration of the segment SPFs evaluated in Chapter 5. Guided by the results in Chapter 5, the SPFs are recalibrated for each of the six road classes; namely, Rural Two-lane, Rural Multi-lane, Urban Two-lane, Urban Multi-lane, Rural Interstates, and Urban Interstates. This chapter also discusses the differences between the preliminary and final recalibrated models.

The form of the recalibrated model remains the same as of the original and it is as follows:

$$A = \exp(k)LQ^\beta \exp(\sum \gamma_i X_i) \quad (6.1)$$

where,

A = number of crashes in a year,

L = length of the section in miles,

Q = AADT of the section,

X = explanatory variables,

k, β , γ = constants.

For calibration we used all the variables present in the preliminary safety performance function. The models were then incorporated in LIMDEP v.7.0 for the calibration process using the Indiana HPMS data available in TransCAD.

6.1 Data Used in the Calibration Task

The following databases were used for recalibration:

1. Statistical Samples Correlating Segments and Crashes:

This database contains information pertaining to *road sections* and is available in TransCAD. The fields in the database include the type of facility, the length of the section, the lane width, the number of lanes, the access control, the average grade and the average curve values in the section, the presence of turn lanes, the type of curb, the median width, the median type, the shoulder width, and the shoulder type. The number of crashes for all severity levels (fatal/injury and property damage only) for three years (2003, 2004, and 2005) are available. This is the most updated form of the available data for Indiana in comparison to data used in a previous study conducted (Lampthey et al, 2004). Table 6.1 shows the sample sizes used for recalibration of various facilities.

Table 6.1 Sample Size Per Facility and Crash History

Facility	Sample Size	Total Length	Average Length	Total Crashes (No.)	
		(mi)	(mi)	Fatal/Injury	PDO
Rural Two-Lane	11132	7609.37	0.684	7518	15442
Rural Multilane	1313	851.35	0.648	1969	4988
Rural Interstates	307	769.76	2.51	1769	8035
Urban Two-Lane	3221	731.35	0.227	3492	7884
Urban Multilane	2727	666.44	0.244	6276	16216
Urban Interstates	473	401.68	0.849	3532	15313

2. Road Inventory Database:

This database, maintained by INDOT, contains the geometric features of road segments. The database consists of the geometric and traffic flow characteristics of various road segments in Indiana.

3. Crash Database for 2004:

The most up-to-date and recent crash data available in TransCAD is for three years (2003, 2004, and 2005) and for severity levels of fatal/injury and property damage only (PDO). This data has been assigned to each link in TransCAD.

6.2 Calibration Process

The number of crashes on a section depends on its length and the volume of flow. This problem also involves the modeling of count data; therefore, a Poisson or negative binomial model would be appropriate. However, for crash count modeling, a negative binomial model is the standard practice because the variance of crash data is much higher than the mean. Therefore, negative binomial theory was adopted in order to carry out regression analysis for the calibration of models. The significance level of the overdispersion factor confirmed this assumption. LIMDEP v 7.0 was used for the analysis of models.

The recalibration applied in LIMDEP is explained as follows.

1. For the purpose of calibration, it was decided to have a non-linear relationship with the AADT. The section length was used as an offset variable (i.e., its coefficient was restricted to be one ($b_1 = 1$)). The section length had to be adjusted to account for the presence of intersections. Therefore, from the total section length, 250 feet was subtracted at each end.
2. At first, a full model was assumed by including all the variables in the model.
3. The model obtained in Step 2 was then evaluated based on the overdispersion parameter and the significance level of all the variables in the original model. The significance was tested using the probability of obtaining a greater t-statistic than that observed if the null hypothesis was true. The null hypothesis in this case is that the coefficient of the independent variable is zero. A variable was considered significant if this probability was less than 0.1. See Appendix B for the results obtained in step 2.

4. The variables that did not prove to be significant were removed from the recalibrated models.

The calibrated models for all six categories of road sections are attached as Appendix B.

6.3 Comparison of the Preliminary and Final Recalibrated Models

In the present study, the safety performance functions developed earlier were recalibrated and the present model results were compared with the past model results. The significance of each explanatory variable in the present model was tested using the associated standard error and sign of the parameter estimates against the past model. In addition, it was also important that the magnitude of the estimated coefficient for the present model should be practical and justifiable from an engineering viewpoint. For each of the six road classes, the past models were recalibrated and compared with the presently developed crash prediction models for each of the following categories of crash severities:

1. Injury/fatal crashes,
2. PDO crashes.

The results are presented and discussed in the subsequent sections.

6.4 Crash Prediction Models for Rural Two-Lane

The present recalibrated crash prediction models developed for Rural Two-Lane roads along with past models are shown in Table 6.2 and

Table 6.3 for the two categories of crash severities (i.e., injury/fatal and PDO).

Table 6.2 Injury and Fatal Crash Model for Rural Two-Lane Roads

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	0.0002	0.1759	0.2080
L_L	1.0000	0.0000	1.0000
L_Q	1.0582	0.0176	0.6040
LW	-0.1017	0.0170	-0.0831
RSW	-0.0126	0.0045	-0.0442
ST_SURF	0.3076	0.0302	-
ST_COMB	-0.6518	0.0231	-
ACURV	0.1832	0.0071	0.2158
AGRAD	-	-	0.0528
Alpha	0.8519	0.2140	-

Table 6.3 PDO Crash Model for Rural Two-Lane Roads

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	4.26×10^{-6}	0.0802	0.7120
L_L	1.0000	0.0000	1.0000
L_Q	0.8284	0.0081	0.5920
LW	-0.1769	0.0079	-0.0524
RSW	-	-	-0.0105
ST_STAB	-0.5831	0.0106	
AGRAD	0.0623	0.0036	0.0775
Alpha	0.6381	0.0080	-

Discussion of Results

For all severity types (injury/fatal and PDO), it was determined that the variance of the respective crash frequency was much greater than the mean during the recalibration process of the present models, and the negative binomial model form consequently was

used. Upon comparison the present and past models coefficient (parameter estimate) values, it was very evident that both models have same signs for the coefficient with respect to the variables, which instilled higher confidence in the present model. For both of the crash models, the section length was found to be positive and was considered to be significant. This relationship is rather obvious as longer road sections are expected to have more crashes compared to shorter ones. The coefficient of the section length for both models was restricted to 1.

Traffic volume (AADT) turned out to be statistically significant for the injury/fatal and PDO present models but lower than one for the present and past models. When compared to the past model, for PDO, the regression coefficient associated with the AADT was higher for the present model. Thus, the same increase in the number of vehicles caused, on average, more PDO crashes and less injury/fatal on rural two-lane roads for the present model.

The negative coefficient for the lane width in injury/fatal and PDO in both the present and past models indicated that wider lane widths are associated with lower crashes. The magnitude of the estimated coefficients for injury/fatal and PDO was greater for the present model compared to the past model, suggesting a higher percentage reduction in crashes for every unit of increase in lane width.

The right shoulder width variable was found to be significant for the present model. The coefficient of this variable was found to be negative for the present model, which is also similar to the sign of the past model coefficient. On comparing the results, the magnitude of the estimated coefficient for the present model was close enough to that of the past model. In the case of PDO crashes, the coefficient was found to be on the higher side compared to the past model coefficient. These results suggest that a decrease in the number of crashes is associated with increasing the right shoulder width.

For the injury/fatal and PDO crash models, the average grade variable was found to be significant and also had the same sign (positive) as the past model variable. However, the results showed lower values for the present models compared to the past models. These results suggest that the present model results are more robust than the past models. The positive sign of the estimated variables also suggests that the higher the grade, the higher will be the number of crashes.

The overdispersion factors obtained for the fatal/injury and PDO crashes were also found to be highly significant, and the parameter estimate was also on the lower side. However, for the past model, more than one overdispersion factor was available as these models were combined using different states' models as mentioned earlier in this chapter. The lower estimated value for the parameter estimate of the overdispersion factor shows the high prediction ability of this model. The lower the value of the overdispersion, the higher will be the prediction ability of the model.

6.5 Crash Prediction Models for Rural Multi-Lane

The present recalibrated crash prediction models developed for Rural Multi-Lane roads along with the past models are shown in

Table 6.4 and Table 6.5 for the two categories of crash severities (i.e., injury/fatal and PDO).

Table 6.4 Injury and Fatal Crash Model for Rural Multi-Lane Roads

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	0.0904	2.387	0.1070
L_L	1.0000	0.0000	1.0000
L_Q	1.3949	0.0086	0.8140
LW	-0.8502	0.0206	-0.1472
LSW	-0.1254	0.0027	-0.0697
MW	-0.0114	0.0003	-0.0051
AC	0.1418	0.0067	0.2412
Alpha	0.1412	0.0023	-

Table 6.5 PDO Crash Model for Rural Multi-Lane Roads

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	0.0002	2.2839	0.6340
L_L	1.0000	0.0000	1.0000
L_Q	1.4679	0.1190	0.6150
LW	-0.4106	0.1911	-
LSW	-0.1413	0.0469	-0.0455
MW	-0.0140	0.0047	-0.0048
AC	0.1085	0.0779	0.1153
RSW	-0.0542	0.0311	-
Alpha	0.6248	0.0314	-

Discussion of Results

The present estimated coefficient results for the injury/fatal and PDO crash models for rural multi-lane roads show similar trends for signs of all the variables as those of past model coefficient estimate results, except for median width in the PDO crash model.

Traffic volume (AADT) turned out to be statistically significant for the injury/fatal and PDO present models, with higher values for the estimated coefficient, almost greater than one. When compared to the past model for injury/fatal and PDO, the regression coefficient associated with the AADT was much higher for the present model. Thus, the results suggest that the same increase in the number of vehicles causes, on average, additional crashes on rural multi-lane roads for the present model. One of the reasons for this high estimated coefficient would be the higher traffic volumes on multi-lane roads as compared to two-lane roads.

The negative coefficient for the lane width for injury/fatal in both the present and past models indicates that wider lane widths are associated with lower crashes. The magnitude

of the estimated coefficient was higher for the present model compared to the past model, suggesting a higher percentage reduction in crashes for every unit increase in lane width. This is normal in the case of rural multi-lane roads as they have wider lane widths. Also, wider lane widths serve as buffer zones and thus offer more opportunity for errant vehicles to recover or for temporary refuge to avoid an errant oncoming vehicle [Lampthey et al, 2005].

The left shoulder width variable was found to be significant for the present model. The coefficient of this variable for the present model also showed a similar trend with a negative sign as that of the past model. On comparing the results, the magnitude of the estimated coefficient for the present models was lower for the PDO present crash model but slightly higher for the injury/fatal present crash model compared to that of the past model. These results suggest that a decrease in the number of crashes is associated with increasing the left shoulder width. The present model results with high values for the estimated coefficient provide high confidence compared to the past model results.

For the injury/fatal and PDO crash models, the median width variable was found to be significant. The negative coefficient is consistent for the present estimated coefficient for the injury/fatal crash model, with a much lower value than the past model coefficient. In the case of the present PDO crash model, the estimated coefficient was found to be positive and not consistent with the past model, which had negative sign. The present injury/fatal crash model results also showed that the median width is a statistically significant variable affecting safety at rural multi-lane road sections. The negative coefficient is also consistent with the expectation that wider median widths enhance safety.

The access control variable was found to be significant for both severity levels in the case of the present model. The variable sign of the present model results was also positive, which is consistent with the past model results. The present model showed higher values for the estimated coefficients compared to only the PDO past crash model. These positive

signs with higher values for the coefficient were expected because the higher the level of access control, the more crashes that occur due to more conflict points.

The overdispersion factors obtained for the fatal/injury and PDO crashes were also found to be highly significant. The estimated coefficient values of the overdispersion factor for the present crash models were found to be on the lower side, which suggests that this model has high prediction ability because of the lower value of the overdispersion factor.

6.6 Crash Prediction Models for Rural Interstate

The present recalibrated crash prediction models developed for Rural Interstates along with the past models are shown in Table 6.6 and Table 6.7 for the two categories of crash severities (i.e., injury/fatal and PDO).

Table 6.6 Injury and Fatal Crash Model for Rural Interstates

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	2.5×10^{-7}	2.4511	0.0440
L_L	1.0000	0.0000	1.0000
L_Q	1.5884	0.2442	0.9170
LW	-	-	-0.4179
LSW	-0.0101	0.0425	-1.0011
Alpha	1.8257	0.1648	-

Table 6.7 PDO Crash Model for Rural Interstates

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	1.9×10^{-7}	1.3686	0.1690
L_L	1.0000	0.0000	1.0000
L_Q	1.4687	0.1327	0.9430
LW	-	-	-0.5706
MW	0.0068	0.0030	-0.0108
LSW	-0.0523	0.0304	-1.3338
Alpha	0.4833	0.0561	-

Discussion of Results

The present estimated coefficient results for the injury/fatal and PDO crash models for rural interstates showed a mixed trend for the signs of all the variables compared to the past model coefficient estimate results. For both crash models, the section length was found to be positive and was considered to be significant.

Traffic volume (AADT) was statistically significant for the injury/fatal and PDO present models with higher values for the estimated coefficient, almost greater than one for the injury/fatal crash model. When compared to the past model, for the PDO the regression coefficient associated with the AADT was also found to be higher for the present model. Thus, the results suggest that the same increase in the number of vehicles causes, on average, additional crashes on rural interstates for the present model. Higher values for the estimated coefficient for the present models also reflect the true condition for rural interstates as they carry higher traffic volumes.

For the PDO crash model, the median width variable was found to be significant. Conversely, the present estimated coefficient for the PDO crash model was found to be positive and not consistent with the past model, which had a negative sign. The present and past injury/fatal crash models do not contain this variable. The value for the present model is also on the lower side compared to the past model.

The coefficient of the left shoulder width variable for the present model showed a similar trend with a negative sign as for the injury/fatal and PDO past crash models. Upon comparing the results, the magnitude of the estimated coefficient for the present models was lower for the injury/fatal and PDO present crash models compared to that of the past models.

6.7 Crash Prediction Models for Urban Interstate

The present recalibrated crash prediction models developed for Urban Two-Lane roads, along with the past models, are shown in Table 6.8 and Table 6.8 Injury and Fatal Crash Model for Urban Two-Lane Roads

for the two categories of crash severities (i.e., injury/fatal and PDO).

Table 6.8 Injury and Fatal Crash Model for Urban Two-Lane Roads

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	1.03×10^{-6}	1.2206	0.1050
L_L	1.0000	0.0000	1.0000
L_Q	1.7534	0.0092	1.0800
LW	-0.0847	0.0111	-0.1527
RSW	0.0084	0.0018	0.0754
LTL	-0.0517	0.0152	0.4522
CRB	-	-	-0.2320
ST_SURF	0.1103	0.0219	-0.7707
ST_STAB	-	-	-0.5138
ST_COMB	-0.6092	0.0151	-0.6423
ST_EARTH	-	-	-0.2569
Alpha	0.3585	0.0052	-

Table 6.9 PDO Crash Model for Urban Two-Lane Roads

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	2.12×10^{-5}	0.3242	0.6030
L_L	1.0000	0.0000	1.0000
L_Q	1.2307	0.0226	0.8960
LW	-0.0085	0.0244	-0.1658
RSW	0.0072	0.0038	0.0524
LTL	-0.0507	0.0287	0.3783
CRB	-	-	-0.2686
ST_SURF	0.3589	0.0471	-0.8547
ST_STAB	-	-	-0.5698
ST_COMB	-0.5316	0.0319	-0.7123
ST_EARTH	-	-	-0.2849
Alpha	0.5919	0.0199	-

Discussion of Results

The present estimated coefficient results for the injury/fatal and PDO crash models for urban two-lane roads showed a similar trend for only for the shoulder type variables as compared to past model coefficient estimate results. For both of the crash models, the section length was found to be positive and was considered to be significant.

Traffic volume (AADT) turned out to be statistically significant for the injury/fatal and PDO present models but with lower values for the estimated coefficient lower than one for both types of crash models in comparison to the past model values. Thus, the results for the present model suggest that urban two-lane roads have lower traffic but that an increase in traffic will cause more crashes because of its positive coefficient. The estimated coefficient for the PDO present crash model is nearly equal to that of the past model.

The positive coefficients for lane width for the present model indicate that narrow lane widths are associated with more crashes. The magnitude and positive sign of the estimated coefficient was lower for the present model compared to that of the past model, suggesting a percentage increase in crashes for every unit of decrease in lane width for the present model.

The right shoulder width variable was found to be significant for the present model. The coefficient of this variable was also found to be positive for the present model, however, which is also similar to the sign of the past model coefficient but on the lower side. These results suggest that an increase in the number of crashes is associated with a decreasing right shoulder width for the present model results. The present results truly represent the situation for two-lane urban roads where the lane width is not sufficient. Therefore, wider shoulders definitely provide confidence and comfort to driver safely.

The type of shoulder (i.e., combined) and a left-turn lane were found to have the same negative sign as the past model, except for “surface,” which was positive. The negative

sign associated with the present model and also the lower values of the coefficients are consistent with the expectation that road shoulders (i.e., combined) reduce the number of crashes as compared to a surface shoulder. One of the reasons that a combined shoulder performs better would be better visibility due to the combination of earth and gravel, as compared to the surface shoulder. In the case of a curb, stabilized and earth shoulders for the present model were found to be insignificant.

The overdispersion factors obtained for the fatal/injury and PDO present crash models were also found to be highly significant. The estimated coefficient of the overdispersion factor was very low, which gives a good indication of the prediction ability of the model.

6.8 Crash Prediction Models for Urban Multi-Lane

The present recalibrated crash prediction models developed for Urban Multi-Lane roads along with the past models are shown in

Table 6.10 and

Table 6.11 for the two categories of crash severities (i.e., injury/fatal and PDO).

Table 6.10 Injury and Fatal Crash Model for Urban Multi-Lane

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	0.0002	0.3659	0.6740
L_L	1.0000	0.0000	1.0000
L_Q	2.0994	0.0101	0.4350
LW	-1.2461	0.0304	-0.1521
NTL	-0.1458	0.0066	-1.0950
AC	1.9532	0.0106	0.6485
CRB	-	-	-0.3766
OSP	-1.8004	0.0239	-0.5250
CLTL	-0.7831	0.0164	-2.3580
LTL	-0.1383	0.0143	-0.3558
LSW	-0.1656	0.0023	-0.2050
Alpha	0.5013	0.0046	-

Table 6.11 PDO Crash Model for Urban Multi-Lane Roads

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	0.0004	0.5316	2.0280
L_L	1.0000	0.0000	1.0000
L_Q	1.6029	0.0196	0.4600
LW	-0.7372	0.0434	-0.1525
NTL	-0.5752	0.0153	-0.9490
AC	1.4734	0.0216	0.5231
CRB	-	-	-0.2786
OSP	-0.7133	0.0549	-0.6690
CLTL	-0.1360	0.0307	-1.7670
LTL	-0.1955	0.0222	-0.3052
Alpha	0.5886	0.0124	-

Discussion of Results

The present estimated coefficient results for the injury/fatal and PDO crash models for urban multi-lane roads showed a mixed trend for all the present model variables as did the past model coefficient estimate results, except for lane width and curb. For both of the crash models, the section length was found to be positive and considered to be significant.

Traffic volume (AADT) was, as expected, statistically significant for the injury/fatal and PDO present models but with higher values of the estimated coefficient lower than one for both types of crash models. When compared to the past model, the regression coefficient associated with the AADT was higher for the present model and truly represented the situation of urban multi-lane roads, which carry higher traffic. Thus, the present model results suggest that the same increase in the number of vehicles causes, on average, additional crashes on urban multi-lane roads.

The positive coefficient of lane width for the present models indicates that narrow lane widths are associated with more crashes. The magnitude and positive sign of the estimated coefficient were higher for the present model compared to the past model, suggesting a percentage increase in crashes for every unit of decrease in lane width on urban multi-lane roads with higher volumes.

The type of shoulder (i.e., outside shoulder presence) and the number of through-lanes (NTL) was found to have the same negative sign as the past model, except for curb which was positive. The negative sign associated with the present models and also the lower values of the coefficients are consistent with the expectation that road shoulders reduce the number of crashes. In the case of curb for the present model, there was an opposite situation, with a positive coefficient.

The estimated coefficient for the continuous left-turn lane (CLTL) variable was found to be negative for the present model as compared to the past model, which also had a negative value. The results suggest that the presence of a continuous left-turn lane reduces the chances of crashes because of lower conflict opportunities between vehicles. The result for the present model in this case was appropriate when compared to the past model.

The access control variable was found to be significant for both severity levels in the case of the present model. The variable sign of the present model results was also positive, which is consistent with the past model results. The present model showed almost the same value for the estimated coefficients when compared to the past model. These results suggest that higher levels of access control can be associated with more crashes as there are more conflict points. The curb variable was found to be insignificant for the present model. However, the overdispersion factor was highly significant, and the lower value of the coefficient suggests good predictive ability for the present model.

6.9 Crash Prediction Models for Urban Interstate

The present recalibrated crash prediction models developed for Urban Interstates along with past models are shown in Table 6.12 and

Table 6.13 for the two categories of crash severities (i.e., injury/fatal and PDO).

Table 6.12 Injury and Fatal Crash Model for Urban Interstates

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	3.0×10^{-7}	1.3284	0.00048
L_L	1.0000	0.0000	1.0000
L_Q	1.5943	0.1266	2.2380
LW	-	-	-0.9166
LSW	0.0123	0.2031	-0.1783
Alpha	2.352	0.1513	-

Table 6.13 PDO Crash Model for Urban Interstates

	Present Model		Past Model
	Coefficient	Std.Error.	Coefficient
Constant	1.181×10^{-6}	1.3533	0.0057
L_L	1.0000	0.0000	1.0000
L_Q	1.1639	0.1134	1.9540
LW	-	-	-0.7974
MW	-0.0041	0.0034	-0.0156
LSW	0.0695	0.0167	-0.1783
Alpha	0.9990	0.0731	-

Discussion of Results

The present estimated coefficient results for the injury/fatal and PDO crash models for urban interstates show different/opposite trend for nearly all the variables compared to the past model coefficient estimate results, except for median width in the PDO crash model.

Traffic volume (AADT) turned out to be statistically significant for the injury/fatal and PDO present models with higher values for the estimated coefficient for both types of crash models. When compared to the past model, the regression coefficient associated with the AADT for both of the crash models was near or greater than one and truly represented the situation of urban interstates, which carry higher volumes of traffic. Thus, the present model results suggest that the same increase in the number of vehicles causes, on average, additional crashes on urban interstates.

The lane width variable for the present model was initially considered during recalibration, but due to regression, co-linearity was removed from the analysis. This co-linearity occurred as the entire data had the same lane width.

The coefficient of the left shoulder width variable for the present injury/fatal and PDO crash models showed an opposite trend with a negative sign compared to the past model. Upon comparing the results, the magnitude of the estimated coefficient for the present models was lower for the injury/fatal crash models than that of the past models. This shows that a decrease in left shoulder width will cause more PDO crashes in the case of the present model.

The estimated coefficient for the median width variable was found to be negative in the present PDO crash model. The coefficient contains the same negative sign as the past model, but the value is higher than the present model.

The overdispersion factor was found to be significant. In this case, the parameter estimate for the overdispersion factor was found to be satisfactory as the value was on the lower side (less than one). This suggests that the model has reasonable prediction ability.

6.10 Summary

This section discussed the recalibration of safety performance functions (SPFs) for the six road classes of the Indiana State Highway Network using Poisson and Negative Binomial regression analysis. A discussion of a comparison of the present and past crash model results was provided. The value of most of the estimated coefficients for the present injury/fatal and PDO crash models were found to be significant and consistent in terms of the sign associated with the coefficients compared to the past crash models. The present models should have a reasonable predictive and explanatory ability as indicated by goodness-of-fit measures and statistics extracted using LIMDEP. However, availability of more years of crash data and the addition of variables will definitely enhance the present model crash results and will also give more stable crash prediction models.

The next chapter presents new models for Indiana developed after obtaining new crash data for years 2003-2005.

7 INDIANA NEW MODELS

Final Indiana models for segments and intersections have been developed based on 2003, 2004 and 2005 crash data. Intersection models predict crashes for state intersections. A crash was considered an intersection crash if it occurred within 250 feet from an intersection and within 500 feet from an interchange. Segment models predict crashes for Indiana roadway inventory segments. Crashes that occurred within 65 feet from the outside shoulder edge were considered segment crashes. Crashes that occurred at intersections of local and state roads were classified as segment crashes because this type of intersections has not been represented in the road network available for this research.

There are six models developed for segments and four models developed for intersections. Each segment type has a corresponding crash prediction model and some of the intersection types share the same models. The two-way stop-controlled intersections include four types determined by the number of intersection legs (3 or 4) and the number of lanes (2 or 4 lanes). These four types of intersections share a common model and are represented in the model with a binary variable. Although the new data do not support different types of intersections based on the number of lanes, the four types are kept to allow future easy refinement of these models when additional data are available. This decision is justified by the existing publications, which indicate a significant safety impact of the number of lanes (see Chapter 4). The interchanges include two types: signalized and unsignalized. These two types of intersections (type 7 and 8) also share a common model and the type of intersection is represented by a single binary variable indicating the presence of signalization. Facility types are provided below:

SEGMENTS

1. Type 1: Rural two-lane
2. Type 2: Rural multilane
3. Type 3: Urban two-lane
4. Type 4: Urban multi-lane

5. Type 5: Rural interstate
6. Type 6: Urban interstate

New Model Variables in the Expanded Segment Table

All variables used in the developed models with column names as they appear in the expanded segment table and their definition are shown below.

Annual average daily traffic (H33AADT) - An average daily value that represents all days of the reporting year. For two-way facilities, is the AADT for both directions, for one-way streets is the directional AADT. Growth factors must be applied if the AADT is not derived from current year counts. This variable has been coded in units of veh/day.

Length of section (LENGTH) - Length of a section measured along the centerline of a roadway. For independently aligned, divided highways, use centerline length as the average of the lengths of the directional roadways, measured along their center lines. This variable has been coded in miles.

Lane width (H54LANEW) - For multilane roads is the prevailing width of travel lane on a section rounded to the nearest foot. The width of the travel lane is measured between lane striping. For two-lane roads lane width is the distance from the centerline or delineation to where the pavement/shoulder surface changes, or to the pavement lane striping if the shoulder and pavement surface are the same. Parking lane is excluded from lane width determination for all facility types. Where there is no delineation between the through traffic lane and the shoulder or parking lane, or where there is no centerline, estimate a reasonable split between the actual width used by traffic and the shoulder or parking lane based on State/local design guidelines. This variable has been coded in feet.

Shoulder width (H59RSHLDW) - Shoulder width is the width of the shoulder which includes rumble strips and gutter pans and does not include bike lanes and parking lanes. For two lane roads is the predominant width of the shoulder on a section if there is no predominant width enter the average width of both shoulders. For multilane roads is the predominant width of the outside shoulders if there is no predominant width enter the

average width of the outside shoulders. Shoulder width should be recorded to the nearest foot. This variable has been coded in feet.

Inside shoulder width (H60LSHLDW) - Inside shoulder width is the predominant width of the inside shoulder which includes rumble strips and gutter pans and does not include bike lanes and parking lanes. If there is no predominant width use the average of both inside shoulders on a given section. Shoulder width should be recorded to the nearest foot. This variable has been coded in feet.

Median width (H56MEDW) - predominant width of the median which includes inside shoulders, if any, measured between the inside edges of the through lanes to the nearest foot. If the median width is greater than 100 feet enter 100. Median width does not include turning bays cut into the median. This variable has been coded in feet.

Access Control (H55ACCESS) - Measures the degree of access control on sample roadway sections. Enter 1 for full access control (Preference given to through traffic movements by providing interchanges with selected public roads and by prohibiting crossing at grade and direct driveway connections), 2 for partial access control (preference given to through traffic movement. In addition to interchanges, there may be some crossings at-grade with public roads, but direct private driveway connections have been minimized through the use of frontage roads or other local access restrictions), 3 for no access control (include all sections that do not meet the criteria above). Control of curb cuts is not access control. If access control is not known enter 1 for interstates, 2 for arterials and 3 for collector and local roads.

Average degree of curve (ADCURVE) - expresses the average degree of curve weighted with the length of each curve on that section. It is computed with the following formula:

$$ADCURVE = \frac{\sum_i S_i \cdot D(a)_i}{\sum_i S_i} , \quad (7.1)$$

where:

S_i = length of horizontal curve i on section,

$D(a)_i$ = degree of curve i (per 100 ft arc definition).

Average grade on section (AGRADE) – is the average absolute grade specified in % on section weighted with the length of each curve on that section. Each grade and flat segment is coded as a separate curve. The sum of all curves equals section length. AGRAD is computed to the following formula:

$$AGRAD = \frac{\sum_i S_i \cdot G_i}{\sum_i S_i} , \quad (7.2)$$

where:

S_i = length of curve i on section specified in miles,

G_i = absolute grade of curve i on section specified in %.

Curb presence (CRB) – Binary variable indicating the presence of a curb in the vicinity of travel lane for two-lane roadways. For multilane roads this variable indicates presence of outside curb in the vicinity of the travel lane. Enter 1 if curb is present in the direct vicinity of the travel lane, otherwise enter 0. If a parking lane exists and a curb is present enter 0 for this variable.

Functional class (FC) – This binary variable indicates if a given segment is classified as a rural principal arterial. Enter 1 if the section is classified as rural principal arterial, enter 0 otherwise. For any urban facility enter 0.

Intersection with left turning bay (INT_L) – Binary variable indicating the presence of intersection with single exclusive left turn lane or multiple turning lanes allowing for simultaneous turns from all turning lanes, through movements are prohibited in these lanes. Enter 1 if section contains intersection with left turning bay(s) only, enter 0 otherwise.

Continuous left turn lane (CELT) - Binary variable indicating the presence of a continuous exclusive left turning lane, from intersection to intersection on a given

section, through movements are prohibited in this lane. Enter 1 if section contains CLTL, otherwise enter 0.

Improved shoulder (ST_IMPROVE) – Binary variable indicating the presence of an improved shoulder on a section. Enter 1 if the section contains an improved shoulder, enter 0 otherwise. Earth shoulder is not considered an improved shoulder but combined shoulder consisting of earth shoulder and paved shoulder is.

Signal presence (SIG) - Binary variable indicating the presence of a signalized intersection on a section. Enter 1 if the section contains a signalized intersection, otherwise enter 0.

INTERSECTIONS

- Type 1: Three-Legged stop controlled intersections of two-lane roads
- Type 2: Four-Legged stop controlled intersections of two-lane roads
- Type 3: Three-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads
- Type 4: Four-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads
- Type 5: All-way stop controlled intersection
- Type 6: Signalized intersection
- Type 7: Unsignalized interchange surface area
- Type 8: Signalized interchange surface area

New Model Variables in the Intersection Table

Major Road Approach Volume (H1AADT) – Approach volumes on the major links (maximum of two) of the intersection, average daily value that represents all days of the reporting year. This variable should be coded in units of (veh/day).

Minor Road Approach Volume (H2AADT) – Approach volumes on the minor links of the intersection, average daily value that represents all days of the reporting year. This variable should be coded in units of (veh/day).

Surrounding Land Use (H3URBRUR) - Binary variable indicating whether the intersection is located in a rural or an urbanized area. Enter 1 if the intersection is located in an urban area; enter 0 if the intersection is located in a rural area.

Paved Shoulder (H4SHOULD) - Binary variable indicating the presence of a paved shoulder at the side of the major road of the intersection. Enter 1 if a paved shoulder exists, enter 0 otherwise.

Right-Turn Lane (H5RTMAJ) - Binary variable indicating the presence of right-turn bay(s) on major roads. Enter 1 if there is a right-turn bay(s) on the major road, otherwise enter 0.

Left-Turn Lane (H6LTMAJ) - Binary variable indicating the presence of left-turn bay(s) on at least one approach of major roads. Enter 1 if there is a left-turn bay(s) on at least one approach of major roads, otherwise enter 0.

Number of Lanes (H7LANE) - The number of lanes is defined as the total number of approach lanes at an intersection including through, shared and exclusive turning lanes. The number of lanes is determined according to the striping, if present, on multilane facilities, or according to traffic use or State/local design guidelines if no striping is present.

Median Width (H9MEDWDTH1) - Variable indicating the median width on major roads measured in feet. It is the predominant width of the median which includes inside shoulders, if any measured between the inside edges of the through lanes to the nearest foot. This variable should be coded in units of feet.

Three Leg Intersection (H11LEG3) – Binary variable indicating if a given intersection is a T-intersection. Enter 1 if intersection consists of three legs, otherwise enter 0.

Signalized Ramp (H12SIG) – Binary variable indicating presence of signals on a ramp. Enter 1 if the ramp has a signal, otherwise enter 0.

Final models are presented in appendix C. The PASS tool used the best model as the default model to estimate annual crash frequencies for segments and intersections.

8 TRANSCAD TOOL

This chapter provides a brief introduction to the developed tools' capabilities and their use in the planning process. Detailed information can be found in the companion user's manual.

8.1 Developed tools in long term planning – concept

8.1.1 Crash prediction procedure for planning

Both tools (PASS and INPASS) have been developed to implement safety into the planning process. INPASS serves as a preprocessing tool for Indiana's transportation network so it can be used as input in the PASS tool for safety estimation of a current or planned network. The full safety prediction process involves a calibration step, which consists of pre-processing the calibration input data, updating the crash prediction models, and calibrating the crash prediction models. Once the calibration step is finished, the user can run multiple alternatives for future networks by pre-processing the prediction input data and predicting crash frequencies with the applied calibration factors.

Figure 8.1 Tools in planning.

8.1.2 Pre-Processing calibration input data

Pre-processing the calibration input data involves expanding the current network table so that it contains all the necessary variables used in segment models, creating an intersection table with all variables used in intersection models, and reading original crash data and sorting it into 15 crash tables. Only table type 10, which is the collision record, is used in calibration pre-processing. The last step of pre-processing the calibration input data is to assign crashes to the segments and intersections where they occurred. The outputs of the pre-processing calibration step are two tables: segments and intersections, with appended recorded crashes.

8.1.3 Updating crash predictions

Figure 8.2 Updating crash predictions.

In order to update the crash predication models, the user modifies existing models or creates new models which he thinks will better replicate the connection between network safety and its characteristics.

8.1.4 Calibrating crash prediction models

Figure 8.3 Calibrating crash prediction models.

In order to calibrate the crash prediction models, the user develops a calibration factor file which contains appropriate calibration factors for each facility type and one calibration factor for each calibration sub-network specified by the user. If the user does not wish to use calibration sub-networks, only the calibration factors are obtained for each facility type (standard calibration). When calibrating the crash predictions, the software compares the initial predictions with the recorded crashes for a given segment or intersection and adjusts the predictions so they better replicate the actual recorded crashes.

8.1.5 Pre-processing prediction input data

Figure 8.4 Pre-processing prediction input data.

Once the user is finished developing the calibration factor file and updating the models, the next step will be to preprocess the future network for crash prediction. When preprocessing prediction input data for the future network, the user can recreate expanded segment tables and generate intersection tables for a future network or modify within the TransCad environment a present network's expanded segment table and intersection table by adding nodes and links specific to each alternative.

8.1.6 Predicting crash frequencies for future network

Figure 8.5 Predicting crash frequencies for future network.

Having future network representation obtained by modifying present network or by pre-processing prediction input data using INPASS, the user can predict crash frequencies for a number of alternatives by applying the calibration factor file and updated models.

8.2 Expanding segment table

8.2.1 Overview

Equations developed from previous research and recalibrated for Indiana's current network have more variables inside the models than could be taken directly from the current HERS HPMS network segment table. PASS applies the appropriate equation and assigns the value for any variable to the equation directly from the input network table as described in Chapter 2. If the value for any variable in the equation cannot be taken directly from the input network table, the user should create a new column within the TransCad environment in the same input network table with the required conversion so

the value in the converted column can be applied directly to the equation. An expanded segment table thus has been created by appending additional variables (columns) to the original HERS HPMS segment network table so the values for all variables can be taken directly from the input network table.

8.2.2 Obtaining additional variables

Additional variables have been obtained by using existing columns within the HERS HPMS input network table based on the conditions shown below:

CRB=1 if H58ShldT=6

CRB=0 otherwise

FC=1 if H17FC=2

FC=0 otherwise

OSP=1 if H58ShldT>1 and H58ShldT<6

OSP =0 otherwise

LTL=1 if H88LTurn>0 and H88LTurn<4

LTL=0 otherwise

CELT=1 if H88LTurn=2

CELT=0 otherwise

ST_Surfaced=1 if H58ShldT=2

ST_Surfaced=0 otherwise

ST_Stabilized=1 if H58ShldT=3

ST_Stabilized =0 otherwise

ST_Combined=1 if H58ShldT=4

ST_Combined =0 otherwise

ST_Earth=1 if H58ShldT=5

ST_Earth =0 otherwise

$$\begin{aligned}
 ADCURVE &= \frac{\frac{1}{3} \cdot 1.7 \cdot H63LCurveA + 4.45 \cdot H64LCurveB + 6.95 \cdot H65LCurveC +}{H63LCurveA + H64LCurveB + H65LCurveC +} \\
 &\quad \frac{+ 11.2 \cdot H66LCurveD + 20.95 \cdot H67LCurveE + 28 \cdot H68LCurveF}{+ H66LCurveD + H67LCurveE + H68LCurveF}, \\
 AGRADE &= \left(\frac{0.2 \cdot H72LGradeA + 1.45 \cdot H73LGradeB + 3.45 \cdot H74LGradeC +}{H72LGradeA + H73LGradeB + H74LGradeC +} \right. \\
 &\quad \left. \frac{+ 5.45 \cdot H75LGradeD + 7.45 \cdot H76LGradeE + 8.5 \cdot H77LGradeF}{+ H75LGradeD + H76LGradeE + H77LGradeF} \right),
 \end{aligned}$$

8.3 Creating an intersection table

8.3.1 Overview

Currently INDOT does not retain any data for state-state road intersections. To develop models which could predict the annual number of crashes at specific intersections based on their characteristics, such a data table was generated from the segment network table (HPMS). The original HPMS segment network table developed in TransCad as a geographic file had two layers associated with it: a segment layer and a node layer, which contained only the geographic location and the node ID. Nodes in the node layer are composed of actual state-state road intersections and dummy nodes used to code inventory segments for the segment layer. To create the actual intersection table, only state-state road intersection nodes were chosen and all dummy nodes were removed.

8.3.2 Generated intersection table structure

The generated intersection network table contains 17 columns and has been saved as a geographic file within the TranCad environment. The first column is the ID of the node carried over from the network segment table. Keeping the original node ID allows the user to link the two network tables (segments and generated intersections) with each other using TranCad standard features. Columns two through ten contain the variables

used in crash prediction extracted from the segment table based on the conditions shown in Table 6.1. Columns 11 through 16 contain the county and route numbers for all links connecting to the particular intersection.

Table 8.1 Linking between segment network table and generated intersection network table

Column name	Linking column in segment table and calculation comment
ID	ID. Values taken directly from the HERS HPMS segment table.
H1AADT	H33AADT. Summation of volumes on major approaches.
H2AADT	H33AADT. Summation of volumes on minor approaches.
H3URBRUR	Item 13. If $\sum(\text{Item 13}) \geq 2$ then H3URBRUR=1 otherwise H3URBRUR=0. Summation performed on all links.
H4SHOULD	Item 59. If $\sum(\text{Item 58}) = 2$ or $\sum(\text{Item 58}) = 4$ place then H4SHOULD=1 otherwise H4SHOULD=0. Summation performed for major approaches.
H5RTMAJ	Item 89. If Item89 ≥ 1 and Item89 ≤ 3 for any major link then H5RTMAJ=1 otherwise H5RTMAJ=0
H6LTMAJ	Item 88. If Item88 ≥ 1 and Item88 ≤ 3 for any major link then H6LTMAJ=1 otherwise H6LTMAJ=0
H7LANE	Item34. $\sum(\text{Item 34})/2 + \sum n$, where n is determined in A. Conditions for n checked separately for each link and total value for n is obtained.
H8PKTRUCK	Item 83. $\sum[\text{Item 83}(i) * \text{Item 33}(i)] / \sum[\text{Item 33}(i)]$ for $i=1..n$, where n is the number of links connecting to particular node. Operation performed on all links.
H9MEDWDTH1	Item 57. $\sum(\text{Item 57})/2$. This operation is performed on major links.
H10MEDTYPE	Item 56. If Item 56 < 4 for either or both major links then H10MEDTYPE =1 otherwise H10MEDTYPE=0.
H11LINK	Item 10. Values taken directly from the HERS HPMS segment table.
H12LINK	Item 10. Values taken directly from the HERS HPMS segment table.
H13LINK	Item 10. Values taken directly from the HERS HPMS segment table.
H14LINK	Item 10. Values taken directly from the HERS HPMS segment table.
H15LINK	Item 10. Values taken directly from the HERS HPMS segment table.
H16LINK	Item 10. Values taken directly from the HERS HPMS segment table.

8.4 Default Values of Variables

The average values for the continuous variables were calculated based on the HPMS data available in TransCAD. The average values for the binary variables were calculated as the proportion of that variable in each facility type. The average value for a model variable is the same for a specific facility type and changes across facility types. The

average values for the model variables are shown below in order and are only reported for those facility types where they were used as input to the crash modification factors.

8.4.1 Segments default values

- Lane width (LW)

Facility type	LW_mean (ft)
RTL	11.34
RML	11.89
RI	12.00
UTL	11.85
UML	11.88
UI	12.00

- Shoulder width (SW)

Facility type	SW_mean
RTL	3.69
RML	8.39
UTL	2.89

- Inside shoulder width (ISW)

Facility type	ISW_mean
RML	3.29
RI	4.7
UML	1.75
UI	5.57

- Median width (MW)

Facility type	MW_mean
RML	42.38
RI	63.27
UI	52.05

- Access control (AC)

Facility type	AC_mean
RML	2.6
UML	2.8

- Number of through lanes (NTL)

Facility type	NTL
UML	4.1

- Average degree of curve (ADCURVE)

Facility type	ADCURVE_mean
RTL	1.1

- Average grade on section (AGRAD)

Facility type	AGRAD_mean
RTL	1.57

Binary variables means are shown below and express proportion of

- Curb presence (CRB)

Facility type	# of segments	Total # of segments in sample	Curb proportion
UTL	141	2870	0.049
UML	237	2701	0.095

- Functional class (FC)

Facility type	# of segments	Total # of segments in sample	Rural principal arterial proportion
RML	664	1294	0.51

- Outside shoulder presence (OSP)

Type of shoulder	# of segments
ST_Surfaced	1026
ST_Stabilized	226
ST_Combination	241
ST_Earth	2

Facility type	# of segments	Total # of segments in sample	Outside shoulder proportion
UML	1495	2701	0.55

- Exclusive left turn lane (LT)

Facility type	# of segments	Total # of segments in sample	Exclusive left proportion
UTL	439	3220	0.136
UML	517	2701	0.191

- Continuous left turn lane (CLTL)

Facility type	# of segments	Total # of segments in sample	Continuous exclusive left proportion
UTL	391	3220	0.121
UML	338	2701	0.125

- Surfaced shoulder (ST_Surfaced)
- Stabilized shoulder (ST_Stabilized)
- Combination shoulder (ST_Combined)
- Earth shoulder (ST_Earth)

Table 8.2 provides mean values for four binary variables created from original variable shoulder type.

Table 8.2 Mean values

Facility type	Variable	# of segments	Total # of segments in sample	Proportion
UTL	ST_Surfaced	472	3220	0.15
	ST_Stabilized	680	3220	0.21
	ST_Combination	621	3220	0.19
	ST_Earth	40	3220	0.01
	curb	172	3220	0.05
	no shoulder	1235	3220	0.38

8.4.2 Intersections default values

- **Surrounding Land Use (URBRUR)**

Facility type	H3URBRUR
Type I	0.994
Type II	1.000
Type III	1.000
Type IV	1.000
Type VI	1.000

- **Paved Shoulder (SHOULD)**

Facility type	SHOULD
Type I	0.167
Type II	0.208
Type III	0.745
Type IV	0.935
Type VI	0.606

- **Number of Lanes (LANE)**

Facility type	LANE
Type VI	5.015

- **Truck Percentage (PKTRUCK)**

Facility type	PKTRUCK
Type III	6.875
Type IV	11.483
Type VI	3.448

- **Right-Turn Lane (RT MAJ)**

Facility type	RT MAJ
Type I	0.003

- **Left-Turn Lane (LT MAJ)**

Facility type	LT MAJ
Type I	0.006

- **Median Width (MEDWDTH1)**

Facility type	MEDWDTH1
Type III	45.612

- **Median Type (MEDTYPE)**

Facility type	MEDTYPE
Type III	0.877

8.5 Procedure for Calibrating Safety Prediction Models used in Planning Road Networks

While traditional safety management identifies and remedies existing safety problems, preventing safety problems requires considering safety in the transportation planning and design phases. Prevailing planning practice addresses safety implicitly as a byproduct of adding capacity and operational efficiency to the transportation system. Safety impacts are sometimes assessed using crude methods supplemented with expert judgment.

To encourage more explicit consideration of safety in transportation planning, FHWA has been promoting *safety-conscious planning*. Safety-conscious planning is a proactive approach to the prevention of crashes by establishing inherently safe transportation networks. The objective is to integrate safety considerations into the transportation planning process at all levels, starting with the Statewide Transportation Improvement Plans (STIP) and followed by consideration of safety objectives in long-range planning (FHWA, 2003).

INDOT, assisted by Purdue University research teams, has been working on a Safety Management System for the last ten years. An advanced method of screening networks for roadway hazards, updated crash reduction factors, and crash-predictive equations have been developed and implemented through *Guidelines for Highway Safety Improvements in Indiana* (Tarko and Kanodia, 2003) and a tool for a system-wide

analysis of Indiana highways for targeted improvements (Lamptey, Labi, and Sinha, 2003).

Development of a GIS-based method of predicting safety in transportation networks for long and short-term planning is underway in Indiana. As emphasized by Washington et al. (2004), safety should be introduced in all steps of the planning process, starting with incorporating safety into the vision, goals, and objectives, through technical analysis, development of programs, and monitoring of the system. The current research project funded by INDOT through their Joint Transportation Research Program with Purdue University is aimed to implement the link and to further link safety prediction models to the long-term planning process of the state network through the GIS-based planning tools.

One of the major concerns in predicting crashes in transportation networks is the applicability and accuracy of crash prediction models for network links and nodes (Lord and Persaud, 2003). Models developed in the past were frequently based on an insufficient number of observations or for regions other than the studied one. These models require calibration and validation, and this research evaluates two methods of calibrating safety prediction models for user-defined set of network partitions. A formulation of the calibration problems consistent with the maximum likelihood approach has also been presented. The methods have been tested and evaluated for the Indiana state road network.

8.5.1 Calibration Problem

A typical network representation includes various types of nodes and links supplemented with basic traffic, control, and roadway characteristics. Although many details of geometric and control design are not known in the planning stage of a transportation network, the main features of intersections and traffic control, classes of roads with their typical cross-sections, travel speeds, and daily and rush hour traffic volumes are known.

These characteristics should be utilized to improve anticipation of future safety in the planned network.

The concept of Safety Performance Functions (SPF) seems to be the most reliable method of predicting future safety at individual road intersections and links. SPFs connect various roadway and traffic characteristics with crash frequency and various levels of severity. SPFs are expected to be used in the future Highway Safety Manual to predict and evaluate roadway safety. The initial tests of SPFs for predicting safety in road networks and traffic analysis zones indicate the usefulness of the functions (Lord and Persaud, 2003). For convenience, a safety performance function should consist of a base safety performance function (BSPF) and a set of Accident Modification Factors (AMFs).

$$a_o = BSFP \cdot AMF_1 \cdot \dots \cdot AMF_n \quad (6.3)$$

where:

a_o = crash frequency at a specific severity level;

$BSFP$ = basic safety performance function;

AMF_i = crash modification function i .

The BSPF includes traffic volume and sometimes link length while the AMFs incorporate the impact of known roadway and control characteristics. In many cases, the AMFs are adapted from research done for regions other than the studied one. The differences between regions and the changes in safety over a period of years call for calibration of SPFs. Planners may also want to reflect sub-regional and local differences in large networks. For example, different parts of the region may experience different weather conditions (e.g., northern and southern California) or topographical conditions (e.g., northern and southern Indiana). Also, some parts of the region may be more developed than the rest and their safety may somewhat differ, such as the metropolitan city of Indianapolis versus smaller urban communities in Indiana. In such cases, the planner may want to use calibration factors to consider these differences when predicting future safety in road networks.

Two types of calibration are proposed:

- (1) *Standard calibration* where one calibration factor applies to each SPF,
- (2) *User-defined calibration* where standard calibration factors are supplemented with factors for sub-networks defined by the user.

In a general case, each link or node (element) of a road network has a corresponding set of calibration factors. In the standard calibration, only one calibration factor applies to each network element and it corresponds to the link or node type. In the user-defined calibration, additional calibration factors may apply if the network element belongs to sub-networks defined by the user. A resulting set of calibration coefficients CF_j apply to crash predictions a_o made with original SPFs:

$$a = a_o \cdot \prod_j CF_j . \quad (6.4)$$

The assignment of calibration factors to network elements is defined with a *coincidence matrix* Δ with elements $\delta_{ij} = 1$ if the calibration factor applies to the network elements i and 0 otherwise. A *calibration subset* s is a set of network elements that have identical vector of coincidence elements δ_{ij} . In other words, the same set of calibration parameters are used within subset s .

To better explain the concept, let us consider a simple network where all nodes are of one type (one SPF applies to all the nodes) and also all links are of one type (another single SPF applies to all the links). The standard calibration includes two calibration factors: CF_{Node} and CF_{Link} . In addition, the user divides the region into two geographical areas (CF_{South} , and CF_{North}) and focuses the calibration on a specific corridor (CF_{Corridor}) in the southern part of the network. Figure 8.6 shows the breakdown of the network and Table 8.3 assigns calibration factors to network elements through coincidence elements δ and identifies calibration subsets s .

Figure 8.6 An example road network.

The calibration task is to find the optimal values of calibration factors CF_j , $j=1..n$, such that the adjusted crash predictions $a_0 \cdot \prod_j CF_j$ best fit the crash counts recorded in the network in the recent period.

Table 8.3 Coincidence matrix Δ for example network elements and five calibration factors (resulted subsets s are shown in the last column).

Network element	Calibration factors					Subset s
	CF_{Node}	CF_{Link}	CF_{North}	CF_{South}	CF_{Corridor}	
01	1	0	1	0	0	1
02	1	0	1	0	0	1
...	...	...	...	...	...	...
07	1	0	1	0	0	1
08	1	0	1	0	0	1
09	1	0	0	1	0	2
10	1	0	0	1	0	2
11	1	0	0	1	0	2
12	1	0	0	1	0	2
16	1	0	0	1	0	2
20	1	0	0	1	0	2
13	1	0	0	1	1	3
14	1	0	0	1	1	3
15	1	0	0	1	1	3
17	1	0	0	1	1	3
18	1	0	0	1	1	3
19	1	0	0	1	1	3
0102	0	1	1	0	0	4
0105	0	1	1	0	0	4
...	...	...	...	...	...	...
0711	0	1	1	0	0	4
0812	0	1	1	0	0	4
0910	0	1	0	1	0	5
0913	0	1	0	1	0	5
...	...	...	...	...	...	...
1620	0	1	0	1	0	5
1920	0	1	0	1	0	5
1314	0	1	0	1	1	6
1317	0	1	0	1	1	6
1415	0	1	0	1	1	6
1418	0	1	0	1	1	6
1519	0	1	0	1	1	6
1718	0	1	0	1	1	6
1819	0	1	0	1	1	6

8.5.2 ML Method for Disaggregate Crash Counts

In the example presented in Figure 8.7 each subset s includes at least one calibration factor which is included in at least one other subset s . It is not possible to calibrate individual factors separately from other factors. A single calibration problem defined for the entire road network is needed. For computational convenience, let calibration factor CF_j be an exponential function $CF_j = \exp(b_j)$. Utilizing the coincidence matrix Δ and the original crash frequency predictions, the adjusted crash frequency for network element i is:

$$a_i = a_{0i} \prod_{j=1..n} CF_j = a_{0i} \prod_{j=1..n} \exp(\delta_{ij} \cdot b_j) \quad (6.5)$$

where:

a_i = adjusted crash frequency for network element i ;

a_{0i} = original crash frequency calculated for network element i with the original SPF;

$\delta_{ij} = 1$ if calibration factor j applies to element i and 0 otherwise; and

b_j = calibration parameter in calibration factor CF_j ,

n = number of calibration factors for the studied network.

A maximum likelihood (ML) method is our initial choice. Following the current practice in modeling crashes, a Poisson model with random effects or Negative Binomial model, is proposed. The Negative Binomial distribution $NB(c_i, a_i, \alpha_s)$ estimates the likelihood of crash count c_i on link i or at node i when the expected value is a_i . An over-dispersion parameter α_s represents unknown Gamma-distributed random effects in the calibration group s to which the network element i belongs. The corresponding optimization problem includes log-likelihood for the entire transportation network. The formal optimization problem is:

$$\max_{\{b\}, \{\alpha\}} LL = \max_{\{b\}, \{\alpha\}} \sum_{s \in S} \left(\sum_{i \in I_s} \ln NB(c_i, a_{0i} \prod_j \exp(\delta_{ij} \cdot b_j), \alpha_s) \right), \quad (6.6)$$

s.t.

$\alpha_s > 0$ for all $s \in S$.

where:

$\{b\}$ = set of calibration parameters corresponding to factors CF;

$\{\alpha\}$ = set of over-dispersion parameters;

S = set of indices of subsets s ;

I_s = set of indices of the network elements that belong to subset s ;

$NB(\ )$ = Negative Binomial probability distribution function;

c_i = number of crashes associated with network element i ;

a_{0i} = crash frequency calculated with non-calibrated SPF for network element i ;

$\delta_{ij} = 1$ if calibration factor j applies to element i and 0 otherwise;

b_j = calibration parameter in calibration factor function $CF_j = \exp(b_j)$; and

α_s = over-dispersion parameter for calibration subset s .

For the robustness of the calculations, the Negative Binomial function in Equation 8.4 should include the Gamma function instead of factorials to allow searching for any positive value of α :

$$NB(c, a, \alpha) = \frac{\Gamma(c + 1/\alpha)}{\Gamma(1/\alpha) \cdot \Gamma(c + 1)} \cdot \frac{(\alpha \cdot a)^c}{(1 + \alpha \cdot a)^{1/\alpha + c}}. \quad (6.7)$$

For evaluation purposes, the proposed calibration problem was solved for the Indiana state road network consisting of 19,151 links with counts of crashes recorded in 2004. The state network created in TransCAD is shown in Figure 8.7. The safety performance functions calibrated in this test have been published (Tarko and Kanodia, 2005), and they are shown in Table 8.4. We are currently assembling more elaborate SPFs using results from the past research for Indiana and other regions. Since this work is not finished yet, the calibration procedures are tested with the base SPFs.

To test the method for user-defined calibration, the Indianapolis area has been specified as a user-defined sub-network. This addition together with six types of links has produced seven calibration factors for:

1. Rural two-lane roads,
2. Rural multilane roads,
3. Rural interstates,
4. Urban two-lane roads,
5. Urban multilane roads,
6. Urban interstates, and
7. Indianapolis area.

and nine calibration subsets:

1. Rural two-lane roads,
2. Rural multilane roads,
3. Rural interstates,
4. Non-Indianapolis urban two-lane roads,
5. Non-Indianapolis urban multilane roads,
6. Non-Indianapolis urban interstates,
7. Indianapolis urban two-lane roads,
8. Indianapolis urban multilane roads, and
9. Indianapolis urban interstates.

The breakdown of the Indiana state network between the calibration subsets by the number of links and basic link-specific data are presented in Table 8.5. The traffic volumes and link lengths have been obtained from the Indiana travel demand model implemented in TransCAD. The crash counts for 2004 were obtained from the Indiana Crash Database and assigned to the link using GIS coordinates and the TransCAD standard features. The trends in average AADT and the link lengths are as expected. The

urban links are shorter than the rural but carry heavier traffic. Traffic increases with the number of lanes and the functional importance of the road.

When dividing the road network into calibration subsets, the total number of crashes per subset should be sufficiently large to allow confident calibration. The percent standard deviation of the total **annual** number of crashes in the subset can be used as guidelines. The standard deviation can be estimated with the assumption of Poisson variability of crashes over years:

$$\sigma = \frac{100}{Y \cdot \sqrt{C}} . \quad (6.8)$$

Standard deviation exceeding the maximum value set by the user, let us say 5 %, indicates that the subset with such a large standard deviation of total crashes is too small. A solution is to use more years of data or to abandon some of the network divisions. In our case, three calibration subsets seem to be rather small (marked in Table 8.5 with bold italicized font). One year of data does not allow for fine partitioning of the network. We keep these subsets as it does not defeat the purpose of this presentation.

Figure 8.7 Indiana state road network

Table 8.4 Basic safety performance functions for Indiana
(Tarko and Kanodia, 2005)

Facility	Safety Performance Functions	Over-dispersion parameter
Signalized intersection	$a_{IF} = 0.1954 \times Q^{0.723}$	0.639
	$a_{PD} = 0.1758 \times Q^{1.0334}$	0.646
Two-way stop-controlled intersection	$a_{IF} = 0.234 \times Q^{1.099}$	0.649
	$a_{PD} = 0.307 \times Q^{1.034}$	0.292
All-way stop-controlled intersection	$a_{IF} = 0.115 \times Q^{0.835}$	2.06
	$a_{PD} = 0.182 \times Q^{1.434}$	0.265
Rural two-lane segment	$a_{IF} = 0.208 \times L \times Q^{0.600}$	0.420
	$a_{PD} = 0.712 \times L \times Q^{0.592}$	0.430
Rural multilane segment	$a_{IF} = 0.107 \times L \times Q^{0.814}$	0.451
	$a_{PD} = 0.634 \times L \times Q^{0.815}$	0.484
Urban two-lane segment	$a_{IF} = 0.105 \times L \times Q^{1.080}$	1.253
	$a_{PD} = 0.603 \times L \times Q^{0.896}$	1.349
Urban multilane segment	$a_{IF} = 0.674 \times L \times Q^{0.435}$	1.588
	$a_{PD} = 2.028 \times L \times Q^{0.460}$	1.946
Rural interstate	$a_{IF} = 0.044 \times L \times Q^{0.917}$	1.053
	$a_{PD} = 0.169 \times L \times Q^{0.943}$	1.604
Urban interstate	$a_{IF} = 0.00048 \times L \times Q^{2.238}$	2.383
	$a_{PD} = 0.0057 \times L \times Q^{1.954}$	2.704
a_{PD} = typical PDO crash frequency, in PDO crashes per year, a_{IF} = typical I/F crash frequency, in I/F crashes per year, Q = AADT entering an intersection or along the road segment, in thousand veh/day, L = road segment length, in miles, and D = over-dispersion parameter.		

Table 8.5 Calibration subsets

Calibration subset	No. of Links	Av. Length (mi)	AADT (veh/day)	Fatal/Injury		PDO	
				Total	Std Dev (%)	Total	Std Dev (%)
Rural two-lane roads	11,112	0.685	4,128	3,525	1.7	11,386	0.9
Rural multilane roads	1,313	0.648	12,682	770	3.6	2,301	2.1
Rural interstates	307	2.507	26,769	534	4.3	2,546	2.0
Non-Indianapolis urban two-lane roads	3,057	0.225	9,822	1,491	2.6	4,081	1.6
Non-Indianapolis urban multilane roads	2,068	0.248	21,215	1,614	2.5	4,538	1.5
Non-Indianapolis urban interstates	235	0.872	39,830	294	5.8	1,225	2.9
Indianapolis urban two-lane roads	163	0.272	13,969	104	9.8	284	5.9
Indianapolis urban multilane roads	659	0.233	27,942	849	3.4	2,329	2.1
Indianapolis urban interstates	237	0.820	83,208	698	3.8	3,283	1.7

Calibration of the safety performance functions was done for fatal/injury and PDO crashes separately. This practice allows incorporating different costs of crashes at different levels of severity.

Columns titled ML for Links in Table 8.6 and

Table 8.7 show the calibration results. The first thought is that the values of some of the calibration factors considerably differ from one. The explanation may be in inaccurate estimation of the original SPRs based on limited samples. The limited size of the calibration subsets also contributes to this result.

As expected, the overdispersion parameters estimated in the calibration (not shown) were typically higher than the ones reported in the original publication of the SPFs. The random effects of crash frequencies unexplained with a SPF are stronger when the model is applied to roads outside of the sample used to estimate the model.

The calibration factors for interstate rural and urban links are consistently lower than for those which indicate that the original SPFs for fatal/injury and PDO crashes overstate the crash frequencies. One possible explanation is that the SPFs were developed for links that included interchanges as part of the links. The TransCAD crash counts for interstate links exclude crashes that happened in the 250-foot vicinity of an interchange. It is symptomatic that the overestimation of crashes on urban interstate links is much higher than on rural interstate links. Denser traffic in urban areas exhibits much more interaction between vehicles (and resulted crashes) in the direct vicinity of ramp exit and entrance points than in rural areas.

There is a tendency toward calibration factors higher than one for links other than interstate roads. This tendency may reflect the improvements in crash reporting achieved in the last several years. Indiana modified its crash reports and started using GIS coordinates in 2001. These changes were accompanied by an intense training effort of

police officers involved in reporting crashes. Years 2003 and 2004 witnessed a considerable increase in the number of crashes that can be precisely assigned to links and nodes.

Although many of the obtained results can be explained, the troubling observation is that the crash predictions aggregated in calibration subsets do not match the data well. It can be seen in the second parts of Table 8.6 and

Table 8.7. Some of the adjusted estimates differ by nearly 50% of the recorded number of crashes. Transportation planning utilizes aggregated measures of performance when evaluating and comparing alternative solutions. Although it is important that the safety performance functions properly predict crashes at the level of network elements, it is even more important that the prediction errors do not accumulate beyond an acceptable level when aggregating the predictions. It seems that the maximum likelihood method of calibration applied to individual links and nodes generates results that may be questioned by planners. To address this issue, another approach to SPF calibration based on aggregated crash counts is proposed in the following section.

8.5.3 LS Method for counts aggregated in calibration subsets

The ML method applied to individual links (in the general case, to network elements including nodes) has not produced the most desirable results. To directly address the need of accurate safety estimation at the sub-network level, the maximum likelihood method will be applied to the total number of crashes in user-defined calibration subsets. Let C_s denote the total number of crashes recorded in calibration subset s . If we accept the Poisson assumption for these counts, the variance of the total crash count is also C_s . The Poisson distribution for variables with the mean sufficiently high can be approximated with the normal distribution with a standard deviation equal to $\sqrt{C_s}$. A relevant ML estimation problem for total crash counts in calibration subsets is:

$$\max_{\{b\}} LL = \max_{\{b\}} \sum_{s \in S} N(C_s, A_{0s} \cdot \prod_{j=1..n} \exp(\delta_{sj} \cdot b_j), \sqrt{C_s}). \quad (6.9)$$

where:

$\{b\}$ = set of calibration parameters corresponding to factors CF;

S = set of indices of subsets s ;

$N()$ = Normal probability density function;

C_s = total number of crashes in calibration subset s ;

A_{0s} = total number of crashes predicted in the calibration subset s with original SPF;

$\delta_{sj} = 1$ if calibration factor j applies to subset s and 0 otherwise; and

b_j = calibration parameter in calibration factor function $CF_j = \exp(b_j)$.

Although the optimization problem in Equation 8.7 can be solved in its original form, an equivalent optimization problem of minimizing the sum of squares weighted with $1/\text{variance}$ is preferred for its computational convenience:

$$\min_{\{b\}} \sum_{s \in S} (C_s - A_{0s} \cdot \prod_{j=1..n} \exp(\delta_{sj} \cdot b_j))^2 / C_s \quad (6.10)$$

where:

$\{b\}$ = set of calibration parameters corresponding to factors CF;

S = set of indices of subsets s ;

C_s = total number of crashes in calibration subset s ;

A_{0s} = total number of crashes predicted in the subset s with non-calibrated SPF;

$\delta_{sj} = 1$ if calibration factor j applies to subset s and 0 otherwise; and

b_j = calibration parameter in calibration factor function $CF_j = \exp(b_j)$.

The ordinary least square technique favors large observations. The $1/C_s$ weight reduces this tendency by considering the increase in variability of large crash counts. The solutions obtained for the optimization problems in Equations 8.7 and 8.8 are identical. The calibration solutions are presented in Table 8.6 and

Table 8.7 in the columns titled SL for Subsets. It can be seen that the number of crashes predicted with the new method better match the sums of recorded crashes.

The match between the recorded and predicted crashes is perfect for subsets with only a single calibration factor not shared with other subsets. This situation occurs for three subsets: rural two-lane roads, rural multilane roads, and rural interstates. The calibration parameters can be calculated independently from other factors as

$$CF = C_s / A_{0s}, \quad (6.11)$$

where:

C_s = total crash counts for subset s , and

A_{0s} = total original crash prediction for subset s .

Table 8.6 Calibration results for injury/fatal safety performance functions

Calibration Factor	ML for Links		LS for Subsets	
	Initial	Optimal	Initial	Optimal
Rural two-lane roads	1	1.15	1	1.08
Rural multilane roads	1	1.29	1	1.12
Rural interstates	1	0.77	1	0.78
Urban two-lane roads	1	2.14	1	1.77
Urban multilane roads	1	1.93	1	1.40
Urban interstates	1	0.48	1	0.37
Indianapolis area	1	1.01	1	1.12
Log-Likelihood LL	-16,429	-16,035	< -600	-123
Total Weighted Square Residuals	41,361	2,846	2,643	167
Calibration subsets	Adjusted Prediction			
	Recorded Crashes	Original Prediction	ML for Links	LS for Subsets
Rural two-lane roads	3,525	3,271	3,755	3,525
Rural multilane roads	770	681	878	770
Rural interstates	534	685	525	534

Non-Indianapolis urban two-lane roads	1,491	798	1,706	1,411
Non-Indianapolis urban multilane roads	1,614	1,247	2,400	1,749
Non-Indianapolis urban interstates	294	379	183	139
Indianapolis urban two-lane roads	104	78	167	155
Indianapolis urban multilane roads	849	419	812	661
Indianapolis urban interstates	698	1,855	900	764

Table 8.7 Calibration results for PDO safety performance functions

Calibration Factor	ML for Links		LS for Subsets	
	Initial	Optimal	Initial	Optimal
Rural two-lane roads	1	1.15	1	1.03
Rural multilane roads	1	1.23	1	0.95
Rural interstates	1	0.87	1	0.89
Urban two-lane roads	1	1.75	1	1.32
Urban multilane roads	1	1.81	1	1.22
Urban interstates	1	0.75	1	0.55
Indianapolis area	1	1.08	1	1.07
Log-Likelihood LL	-29058	-28676	< -700	-218
Total Weighted Square Residuals	280598	257411	3118	347
Calibration subsets	Recorded Crashes	Original Prediction	Adjusted Prediction	
			ML for Links	LS for Subsets
Rural two-lane roads	11,386	11,013	12,660	11,386
Rural multilane roads	2,301	2,423	2,979	2,301
Rural interstates	2,546	2,871	2,506	2,546
Non-Indianapolis urban two-lane roads	4,081	2,969	5,184	3,930
Non-Indianapolis urban multilane roads	4,538	4,044	7,335	4,920
Non-Indianapolis urban interstates	1,225	1,512	1,129	829
Indianapolis urban two-lane roads	284	274	517	390
Indianapolis urban multilane roads	2,329	1,370	2,684	1,791
Indianapolis urban interstates	3,283	5,994	4,834	3,532

A simple method (Eq. 8.9) was proposed for calibrating the SPF for two-lane rural roads and intersections (Harwood et al., 2000). All the calibration parameters can be calculated this way in the standard calibration case. Adding user-defined sub-networks entangles some of the calibration factors together, which requires joint calibration of the calibration factors. The ML or LS formulation of the calibration problem should be used.

The re-defined LS calibration problem is more robust. It does not use probability functions that may sometimes cause computational troubles if the original model poorly fits the data (very small likelihood values). Finding a solution to the re-defined problem is much faster because the computational burden is smaller. The calculations are performed on a limited number of calibration subsets and not on thousands of network elements.

Discussion

We presented two methods of calibrating models predicting safety for the individual links and nodes of a transportation network. The calibration has been defined to address the specifics of network modeling in transportation planning. In both methods, a planner has the freedom to partition a road network in a way that addresses expected local and sub-regional safety differences. Furthermore, a planner may identify routes, corridors, and areas to focus calibration on these locations if the planning focuses on them.

The maximum likelihood method was first applied to individual network elements. The results, although plausible and possible to explain, indicate considerable differences between the recorded and predicted total number of crashes at the sub-system level. The calibration task was redefined to focus on the prediction of the cumulative number of crashes in the user-defined sub-networks. This time, the results are much more reasonable.

We have proven the feasibility of the proposed approach. A generalization of a simple calibration method proposed for individual safety performance functions, it can be

classified as a moment-based method although in the general version it involves maximization of the likelihood of moment estimates. A robust version based on weighted least squares was proposed.

The presented calibration method is focused on safety models developed for individual roadway facilities. These models typically involve neither socioeconomic variables nor area-wide infrastructure characteristics. Models for Traffic Analysis Zones have been developed to evaluate transportation policies and transportation networks at the TAZ level. The author thinks that the calibration approach proposed here can be applied to that class of models and also to networks and models with various transportation modes.

Time-related calibration of the safety prediction methods has not been covered in this report. It requires time-series analysis or at least analyzing long-term trends in safety. This consideration is out of the scope of this research but should be included in future research to develop a complete methodology of predicting safety in transportation networks.

9 CONCLUSIONS AND RECOMMENDATIONS

9.1 Conclusions

This study investigated if existing Safety Performance Functions (SPFs) and corresponding Crash Modification Factors (CMFs) developed for states other than Indiana could be applied to Indiana conditions. This task was conducted by evaluating the Basis Safety Performance Functions (BSPFs) and the full-form Safety Performance Functions (SPFs) for six types of road segments. The BSPF includes only the length of the segments and the AADT. The full-form SPF was obtained after adding Crash Modification Factors (CMFs) to the corresponding BSPF. The systematic bias and standard error of prediction were analyzed for the two forms of SPFs. The BSPFs exhibited better performance, which indicates that the CMFs added additional noise to predictions that reduced the prediction quality. These results indicate that simple calibration by adjusting the predictions with a single calibration factor is insufficient and full calibration of all model parameters is needed.

The calibration of the SPF was conducted for the six road classes on the Indiana State Highway Network using Poisson and Negative Binomial regression analysis. A comparison was made for the present and past crash model results and the results were discussed. The recalibrated values of most of the regression coefficients were significant and their signs were consistent with the signs of the original values before calibration. The recalibrated models performed satisfactorily as they had good predictive and explanatory ability as indicated by the overdispersion factor and the statistics reported by LIMDEP. However, the recalibration was performed for three years of crash data.

Although the full re-calibration yielded satisfactory results, some of the variables in the original models became insignificant in the re-calibrated models. Removing these variables from the models would improve the predictive power of the models and might justify adding other variables missing in the original models. This conclusion justifies development of new models. Our team has developed Indiana new SPFs based on three-year crash data.

The developed new SPFs can be used in the transportation planning through developed in this research set of add-in tools for TransCAD. The two tools are called INdiana Planning Analysis Software for Safety (INPASS) and Planning Analysis Software for Safety (PASS). INPASS allows reformatting the data available in Indiana in the format that is required by the developed SPFs. It creates intersection data tables that allow predicting crashes separately for segments and separately for intersections. In addition, the INPASS reformats the crash data to the format acceptable by the TransCAD and assigns crashes to segments and intersections.

The second component – INPASS calculates crashes for segments and intersections in studied road networks. It includes a convenient SPF editor that allows modifying the existing and entering new SPFS. Also, the INPASS facilitates calibration of the SPFs based on the user-defined partitioning of the road network. A separate volume includes the user manual, which explains details of using the INPASS and PASS.

9.2 Recommendations

The presented study utilized data found in the Indiana roadway inventory and crash databases available through TransCAD. One of the major aspects of this Indiana inventory data is that it contains many important safety variables that can be added to the developed safety performance model. In addition, the horizontal and vertical alignment data is available for all the roadway classes in the database. This safety variable was only used for rural two-lane roads in this study and can be added to other facilities. An

upgrade of this database on a regular basis and adding more data is therefore needed to ensure more stability of the SPFs and also for future safety analysis. There is a need for software or code to be developed in order to combine the crash data from the crash database with the roadway inventory database coded by law enforcement.

In order to implement the developed SPFs in Indiana long-term planning, a set of TransCAD-based tools has been developed to pre-process data and predicts crashes for segments and intersections. Implementation of the developed tools is needed to make necessary adjustments that are vital for use in INDOT planning process.

LIST OF REFERENCES

AAA Foundation for Traffic Safety, Progress Report on- *“Did you know?-Congestion and Crashes”*, September/October 1997.

Abbess, C., Jarrett, D. and Wright, C.C. (1981) *‘Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment with Special Reference to the ‘Regression-to-Mean’ Effect’*, Traffic Engineering and Control, Vol. 22, No.10, October, 1981, pp 535-542.

Bowman, B.L. and Vecellio, R.L., *“Effect of Urban and Suburban Median Types on both Vehicular and Pedestrian Safety,”* Transportation Research Record 1445, 1994, pp. 169-179.)

Brown, H. and A. Tarko, *The Effects of Access Control on Safety on Urban Arterial Streets*, paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington D.C., January 1999.

Ceder, A., and M. Livneh. (1982). *Relationship between Road Accident and Hourly Traffic Flow-I and II, Accident Analysis and Prevention*, Vol.14, No.1, 1982, pp. 19–44.

Dominique Lord, Simon Washington, John Ivan, *Statistical challenges with modeling motor vehicle crashes: Understanding the implications of alternate approaches*, paper presented at 83rd TRB annual meeting March 22nd 2004.

Frantzeskakis, J.M. and D.I. Iordanis. (1987). *Volume-to-Capacity Ratio and Traffic Accident on Interurban Four-Lane Highways in Greece*, Transportation Research Record 1112, TRB, National Research Council, Washington, D.C., pp. 29–38.

Griffith, M. S. *Safety Evaluation of Rolled-in Continuous Shoulder Rumble Strips Installed on Freeways*. In Transportation Research Record: Journal of the Transportation Research Board, No. 1665, Transportation Research Board, National Research Council, Washington, D.C., 1999, pp. 28–34.

Golob, T., Recker, W., and Alvarez, V. (2003c) *Freeway Safety as a Function of Traffic Flow*. Working paper UCI-ITSWP-03-2.

Gwynn, D.W. (1967). *Relationship of Accident rate and Accident Involvements with Hourly Volumes*, *Traffic Quarterly*, July 1967, pp. 407–418.

Hall, J.W. and O.J. Pendelton. (1989). *Relationship Between V/C Ratios and Accident Rates*, Report FHWA-HPR-NM-88-02. FHWA, U.S. DOT, June 1989.

Harkey, D.L, “*Crash Reduction Factors for Traffic Engineering and Intelligent Transportation System Improvements*”. Highway Safety Research Center, University of North Carolina, 2004.

Hauer, E. (1997). ‘*Observational before-after studies in road safety*’ Elsevier Science Ltd, Oxford.

HERS-ST *Technical Report*, Draft, USDOT, FHWA, 2002.

Iqbal M.A, Panikar B and Tarko A.P, “*Safety Conscious Planning for Indiana - Part I*”, Intermediate Research Report, Joint Transportation Research Program, December 2004, unpublished.

Joan Shen and Albert Gan, (2003) *Development of Crash Reduction Factors-Methods, Problems and Research Needs*, Transportation Research record 1840, pp 50-56

Lampthey, G., *Development of analytical and software tools for highway safety*, MS Thesis, Purdue University, May 2005.

Maher, M.J., (1987) ‘*Accident Migration - a Statistical Explanation?*’ Traffic

Mannering, F.L. Poch, M., 1996. Negative Binomial Analysis of Intersection-Accident Frequency. *Journal of Transportation Engineering*, Vol. 122, No. 2, 105-113. Engineering and Control, September, 1987, pp480-483.

Miaou, S.-P., and H. Lum. *Statistical Evaluation of the Effects of Highway Geometric Design on Truck Accident Involvements*. Transportation Research Record, No. 1407. 1993.

Miaou, S.-P., P. S. Hu, T. Wright, S. C. Davis, and A. K. Rathi. *Development of Relationship Between Truck Accidents and Geometric Design: Phase I*. Report FHWA-91-124. FHWA, U. S. Department of Transportation, 1993.

Michaels, M. Richard, “*Two Simple Techniques for Determining the Significance of Crash-Reducing Measures.*” *Traffic Engineering*, September 1966, pp. 448.

Neudorff, LG; Randall, JE; Reiss, R; Gordon, R (2003) Report No: FHWA-OP-04-003, *Freeway Management and Operations Handbook*, pp 1-15 to 1-21.

Prediction of the Expected Safety Performance of Rural Two-Lane Highways, Publication No. FHWA-RD-99-207, December 2000.

Saccomanno, F.F., R. Grosst, D. Green, and A. Mehmood. *Identifying Black Spots Along Highway SS107 in Southern Italy Using Two Models*. *Journal of Transportation Engineering*, Vol. 127, No. 6, November/December 2001, pp. 515-522.

Sawalha, Z. and Sayed, T., *Evaluating the safety of urban arterial roads*, *Journal of Transportation Engineering*, Volume 127, Issue 2, March-April 2001.

Tarko, A.P., K. C. Sinha, S. Eranky, H. Brown, E. Roberts, R. Scinteie, and S. Islam, “*Crash Reduction Factors for Improvement Activities in Indiana*”, JTRP Final Report May 2000.

Tarko, A.P., S.Eranky, and K.Sinha, *Methodological Considerations in the Development and Use of Crash Reduction Factors*, paper presented at the 77th Annual Meeting of the Transportation Research Board, Washington D.C., 1998.

Tarko, A.P. and Kanodia, M., “*Hazard Elimination Program-Manual on Improving Safety on Indiana Road Intersections and Sections*”, Volume 2 - Guidelines for Highway Safety Improvements in Indiana, JTRP Final Report February 2004.

Vogt, A., and J. Bared. *Accident Models for Two-Lane Rural Segments and Intersections*, Transportation Research Record, No. 1635. 1998.

APPENDICES

Appendix A Redefining Variables and Combining Model Parameters

This appendix discusses the adopted models for Indiana and their variables for each facility type. In addition this also includes combining these variables in for accurate prediction and finally using just one model by redefining the slopes of these variables.

To obtain accurate predictions of crashes, the most adequate models that could be adopted for Indiana were selected. From these models for each facility type, variables from different models were combined. Combining variables provides more accurate prediction then using only one model because it removes bias that could exist from sampling. To obtain the combined parameter (final slope) used in the CMF, the weighted average was taken for all of the adjusted parameters for a given variable for a specific facility type and severity level with the highest weight given to the adjusted parameter with the smallest variance. The formula is given below.

$$Combined\_parameter = \frac{\sum_i a_i \cdot \frac{1}{SE_i^2}}{\sum_i \frac{1}{SE_i^2}} \quad (A.1)$$

where:

a_i = slope i from model i for given variable,

SE_i = standard deviation i for slope i .

The model variable column provides information on the final variable as described in the first section of this chapter. The model column provides information on the source of the parameters. The original model variable column provides information on the variables as defined in the original models. The original parameter column provides information on the slopes of the variables as defined in the original models. The adjusted parameters column shows the adjusted parameters, which are consistent with the model variable definitions and TransCAD scale and are grouped to obtain the combined parameter. Because some models have opposite scales or slightly different variable definitions, some adjustments were required in the combining process. These adjustments were performed on the model parameters to obtain the adjusted parameters. For model variables where only one original model was used, the adjusted slope parameter is equal to the combined slope parameter. For model variables where more than one original model was used, the adjusted parameters were combined using Equation A.1 to obtain the combined parameter. The combined parameters are the final parameters used in the crash modification factors. These adjustments are documented in the comments column, which also provides information on the scale of the binary variables in the original models. The combining process is summarized in Table A.2, Table A.3, Table A.4, Table A.5B,

Table A.6 and Table A.7

Redefining median width

Some original models defined the median width excluding the left shoulder. To be able to obtain the combined parameter, the adjusted parameters had to be redefined. The median width variable was redefined to include the left shoulder as coded in TransCAD.

Redefining has no effect on the slope taken from the original model, but the standard error was recomputed.

$SE(\text{redefined}) = SE(\text{original}) * [\text{Median}(\text{redefined}) / \text{Median}(\text{original})]$, original median was computed as $\text{Median}(\text{redefined}) - 2 * (\text{Left Shoulder Width})$ based on the HPMS data. The

redefined median was computed as the average median width for a specific facility type using the HPMS data available in TransCAD. This adjustment where required for RML, RI, and UI facility types as seen in the comment column of Table A.3, Table A.4 and Table A.7.

Redefining shoulder type

For a UTL facility, the original model variable shoulder type was redefined to four new binary variables: ST_Surfaced, ST_Stabilized, ST_Combined, and ST_Earth. The slopes for each variable were computed based on the scale of the original model and original slope and are shown in Table A.5A and Table A.5B.

Table A.1 Redefined shoulder variables and their slopes for UTL

Redefined model variable	Slope a_i		
	Total	IF	PDO
Reference (no shoulder)	0	0	0
ST_Surfaced	-0.889	-0.771	-0.855
ST_Stabilized	-0.593	-0.514	-0.570
ST_Combination	-0.741	-0.642	-0.712
ST_Earth	-0.296	-0.257	-0.285

The slope of ST_Earth for each facility type is identical to original model 3 slope for shoulder type due to the fact that the slope for this binary variable is a multiple of one of the original model slopes.

In Table B.5 the original variables taken from model one for a UML facility provides the slope value for reference only and were not used in the combination step due to the lack of standard error data.

Combining parameters of the selected models

To obtain accurate predictions of crashes, the most adequate models that could be adopted for Indiana were selected. From these models for each facility type, the variables from different models were combined. Combining variables gives more accurate prediction than using only one model because it removes bias that could exist from the sampling. To obtain the combined parameter (final slope) used in the CMF, the weighted average was taken of all the adjusted parameters for a given variable for a specific facility type and severity level with the highest weight given to the adjusted parameter with the smallest variance. The formula is given below.

$$Combined\_parameter = \frac{\sum_i a_i \cdot \frac{1}{SE_i^2}}{\sum_i \frac{1}{SE_i^2}} \quad (A.2)$$

where:

a_i = slope i from model i for given variable,

SE_i = standard deviation i for slope i .

The selected models are the chosen existing SPF whose original model variables have definition that are compatible with the model variables definition (see next section) or could be transformed to model the variable definition with alterations to the slope value and standard deviations. The process of combining selected models is to derive one aggregate slope parameter for each single model variable which considers the effect of this variable on safety prediction derived from past research.

The model variable column provides information on the final variable. The model column provides information on the source of the parameters. The original model variable column provides information on the variables as defined in the original models. The original

parameter column provides information on the slopes of variables as defined in the original models. The adjusted parameters column shows the adjusted parameters which are consistent with the model variable definitions and the TransCAD scale and are the group to obtain the combined parameter. Because some of the models have opposite scales or slightly different variable definitions, some adjustments were required in the combining process. These adjustments were performed on the model parameters to obtain the adjusted parameters. For model variables where only one original model was used, the adjusted slope parameter is equal to the combined slope parameter. For model variables where more than one original model was used, the adjusted parameters were combined using equation 5.1 to obtain the combined parameter. The combined parameters are the final parameters used in the crash modification factors. These adjustments are documented in the comments column, which also provides information on the scale of the binary variables in the original models. The combining process is summarized in Table A.2, Table A.3, Table A.4, Table A.5B,

Table A.6 and Table A.7 in the following order:

- RTL
- RML
- RI
- UTL
- UML
- UI

Combining segments original model variables

Table A.2 Combining and converting variables for RTL facility

Severity Level	Model Variable	Model	Original Model Variable	Original Parameter		Adjusted Parameter		Combined Parameter	Comments
				SLOPE	SE	SLOPE	STD.ERR		
TOTAL	LW	1	LW	-0.1690	0.0801	-0.169	0.0801	-0.0598	-
		2	Lane Width LW	-0.1821	0.0573	-0.1821	0.0573		
		3	LW	-0.0486	0.0144	-0.0486	0.0144		
	SW	2	SHW	-0.0800	0.0158	-0.08	0.0158	-0.0256	-
		3	RSW	-0.0170	0.0063	-0.017	0.0063		
	ADCURVE	2	H	0.1421	0.0545	0.1421	0.0545	0.1421	-
	AGRAD	2	GR	0.1990	0.0928	0.199	0.0928	0.0720	-
		3	AGRAD	0.0708	0.0091	0.0708	0.0091		
IF	LW	1	LW	-0.2270	0.2269	-0.227	0.2269	-0.0831	-
		2	Lane Width LW	-0.2458	0.0694	-0.2458	0.0694		
		3	LW	-0.0696	0.0192	-0.0696	0.0192		
	SW	2	SHW	-0.1053	0.0212	-0.1053	0.0212	-0.0442	-
		3	RSW	-0.0342	0.0086	-0.0342	0.0086		
	ADCURVE	2	H	0.2158	0.0667	0.2158	0.0667	0.2158	-
	AGRAD	3	AGRAD	0.0528	0.0128	0.0528	0.0128	0.0528	-
PDO	LW	1	LW	-0.1240	0.0803	-0.169	0.0801	-0.0524	-
		3	LW	-0.0424	0.0158	-0.0486	0.0144		
	SW	3	RSW	-0.0105	0.0069	-0.0105	0.0069	-0.0105	-
	AGRAD	3	AGRAD	0.0775	0.0096	0.0775	0.0096	0.0775	-

Table A.3 Combining and converting variables for RML facility

Severity Level	Model Variable	Model	Original Model Variable	Original Parameters		Adjusted Parameter		Combined Parameter	Comments
				SLOPE	SE	SLOPE	SE		
TOTAL	LW	3	LW	-0.1379	0.032	-0.1379	0.0320	-0.1379	-
	SW	1	X7	-0.094	0.011	-0.0940	0.0110	-0.0940	-
	ISW	3	LSW	-0.1329	0.0562	-0.1329	0.0562	-0.1329	-
	MW	1	X8	-0.003	0.009	-0.0030	0.0090	-0.0172	-
		2	MW	-0.054	0.0145	-0.0540	0.0172		Median width variable was redefined. 1)
		3	MW	-0.0504	0.0216	-0.0504	0.0256		
	AC	1	X2	-0.151	0.047	0.1510	0.0470	0.0044	0=no control, 1- partial control no measure for full control. TransCAD scale is reversed, slope adjusted by taking opposite sign.
		2	A	-1.131	0.4735	1.1310	0.4735		1=no control, 2-partial control, 3-full control. TransCAD scale is reversed, slope adjusted by taking opposite sign.
		3	AC	-0.0042	0.0018	0.0042	0.0018		
	ARHR	1	X1	0.131	0.025	-	-	-	not present in current data (HPMS FIELD MANUAL DECEMBER 2000, USDOT, FHWA)
	FC	1	X6	-0.572	0.007	-0.5720	0.0070	-0.5720	1=rural principal arterial, 0=otherwise. Average computed as proportion of rural principal arterials in the RML sample.
IF	LW	3	LW	-0.1472	0.0442	-0.1472	0.0442	-0.1472	-
	ISW	2	ISW	-0.304	0.1139	-0.3040	0.1139	-0.0697	-
		3	LSW	-0.0581	0.0254	-0.0581	0.0254		-
	MW	2	MW	-0.026	0.0184	-0.0260	0.0218	-0.0051	Median width variable was redefined. 1)
		3	MW	-0.0048	0.0022	-0.0048	0.0026		
	AC	2	A	-1.502	0.5407	1.5020	0.5407	0.2412	1-no control, 2-partial control, 3-full control. TransCAD scale is reversed, slope adjusted by taking opposite sign.
		3	AC	-0.2182	0.0731	0.2182	0.0731		
PDO	ISW	3	LSW	-0.0455	0.0225	-0.0455	0.0225	-0.0455	-
	MW	2	MW	-0.058	0.015	-0.0580	0.0178	-0.0048	Median width variable was redefined. 1)
		3	MW	-0.004	0.0018	-0.0040	0.0021		
	AC	2	A	-0.855	0.4823	0.8550	0.4823	0.1153	1-no control, 2-partial control, 3-full control. TransCAD scale is reversed, slope adjusted by taking opposite sign.
		3	AC	-0.1046	0.058	0.1046	0.0580		

Table A.4 Combining and converting variables for RI facility

Severity Level	Model Variable	Model	Original Model Variable	Original Parameter		Adjusted Parameter		Combined Parameter	Comments
				SLOPE	SE	SLOPE	SE		
TOTAL	LW	1	LW	-0.5658	0.1178	-0.5658	0.1178	-0.5658	-
	ISW	1	LSW	-1.3799	0.1812	-1.3799	0.1812	-1.3799	-
	MW	1	MW	-0.0100	0.0023	-0.0100	0.00270	-0.0100	Median width variable was redefined. 1)
IF	LW	1	LW	-0.4179	0.0896	-0.4179	0.0896	-0.4179	-
	ISW	1	LSW	-1.0011	0.1856	-1.0011	0.1856	-1.0011	-
PDO	LW	1	LW	-0.5706	0.1178	-0.5706	0.1178	-0.5706	-
	ISW	1	LSW	-1.3338	0.1839	-1.3338	0.1839	-1.3338	-
	MW	1	MW	-0.0108	0.023	-0.0108	0.0270	-0.0108	Median width variable was redefined. 1)

Table A.5 Combining and converting variables for UTL facility

Severity Level	Model Variable	Model	Original Model Variable	Original Parameter		Adjusted Parameter		Combined Parameter	Comments
				SLOPE	SE	SLOPE	SE		
TOTAL	LW	1	LW	-0.214	0.1329	-0.214	0.1329	-0.1759	-
		2	LW	-0.1736	0.0324	-0.1736	0.0324		
	LTL	1	T_L	1.182	0.5132	1.182	0.5132	0.3859	0-no exclusive left turn, 1-exclusive left turn lane exists
		2	TL	0.3224	0.145	0.3224	0.145		
	CRB	2	CRB	-0.2951	0.1334	-0.2951	0.1334	-0.2951	0-no shoulder curb exist, 1- shoulder curb exists
	ST_Surfaced	2	ST	-0.8889		-0.8889		-0.8889	0-no surfaced shoulder exist, 1- surfaced shoulder exists
	ST Stabilized	2		-0.5926		-0.5926		-0.5926	0-no combined shoulder exist, 1- combined shoulder exists
	ST Combination	2		-0.7408		-0.7408		-0.7408	0-no earth shoulder exist, 1- earth shoulder exists
	ST Earth	2		-0.2963		-0.2963		-0.2963	0-no stabilized shoulder exist, 1- stabilized shoulder exists
	CLTL	1	T_C	-1.544	0.2597	-1.544	0.2597	-1.5440	0-no continuous exclusive left turn, 1-continuous exclusive left turn lane exists
IF	LW	1	LW	-0.332	0.1637	-0.332	0.1637	-0.1527	-
		2	LW	-0.1412	0.0414	-0.1412	0.0414		
	LTL	1	T_L	1.716	0.4383	1.716	0.4383	0.4522	0-no exclusive left turn, 1-exclusive left turn lane exists
		2	TL	0.2871	0.1584	0.2871	0.1584		
	CRB	2	CRB	-0.232	0.1676	-0.232	0.1676	-0.2320	0-no shoulder curb exist, 1- shoulder curb exists
	ST_Surfaced	2	ST	-0.7707		-0.7707		-0.7707	0-no surfaced shoulder exist, 1- surfaced shoulder exists
	ST Stabilized	2		-0.5138		-0.5138		-0.5138	0-no combined shoulder exist, 1- combined shoulder exists
	ST Combination	2		-0.6423		-0.64225		-0.6423	0-no earth shoulder exist, 1- earth shoulder exists
	ST Earth	2		-0.2569		-0.2569		-0.2569	0-no stabilized shoulder exist, 1- stabilized shoulder exists
	SW	2	RSW	0.0754	0.024	0.0754	0.024	0.0754	-
PDO	LW	2	LW	-0.1658	0.0326	-0.1658	0.0326	-0.1658	-
	LTL	1	T_L	0.995	0.4796	0.995	0.4796	0.3783	0-no exclusive left turn, 1-exclusive left turn lane exists
		2	TL	0.3267	0.1387	0.3267	0.1387		
	CRB	2	CRB	-0.2686	0.1323	-0.2686	0.1323	-0.2686	0-no shoulder curb exist, 1- shoulder curb exists
	ST_Surfaced	2	ST	-0.8547		-0.8547		-0.8547	0-no surfaced shoulder exist, 1- surfaced shoulder exists
	ST Stabilized	2		-0.5698		-0.5698		-0.5698	0-no combined shoulder exist, 1- combined shoulder exists
	ST Combination	2		-0.7123		-0.71225		-0.7123	0-no earth shoulder exist, 1- earth shoulder exists
	ST Earth	2		-0.2849		-0.2849		-0.2849	0-no stabilized shoulder exist, 1- stabilized shoulder exists
	SW	2	RSW	0.0524	0.0223	0.0524	0.0223	0.0524	-

Table A.6 Combining and converting variables for UML facility

Severity Level	Model Variable	Model	Original Model Variable	Original Parameter		Adjusted Parameter		Combined Parameter	Comments
				SLOPE	SE	SLOPE	SE		
TOTAL	LW	4	LW	-0.1541	0.0590	-0.1541	0.0590	-0.1541	-
	NTL	1	NL	0.0852	-	0.0852	-	-1.0260	-
		2	NR_LANE	-1.0260	0.2096	-1.0260	0.2096		
	AC	2	A	-1.4830	0.3836	1.4830	0.3836	0.5632	1-no control, 2-partial control, 3-full control. TransCAD scale is reversed, slope adjusted by taking opposite sign.
		4	AC	-0.4300	0.1460	0.4300	0.1460		
	-	1	DD	0.0168	-	0.0105	-	-	Data not provided in current HPMS data. Slope adjusted from per km to per mile
	-	1	USD	0.0910	-	0.0565	-	-	Data not provided in current HPMS data. Slope adjusted from per km to per mile
	CRB	4	CRB	-0.2990	0.1265	-0.2990	0.1265	-0.2990	0-no shoulder curb exist, 1-shoulder curb exists
	OSP	3	SHLDR	-0.6310	0.2570	-0.6310	0.2570	-0.6310	0-no outside shoulder exists, 1- outside shoulder exists
	CLTL	2	T_C	-1.9360	0.7171	-1.9360	0.7171	-1.9360	0-no continuous exclusive left turn, 1-continious exclusive left turn lane exists
IF	LTL	3	TWLTL	-0.7480	0.3430	-0.7480	0.3430	-0.3058	0-no exclusive left turn, 1-exclusive left turn lane exists
		4	TL	-0.2195	0.1515	-0.2195	0.1515		
	LW	4	LW	-0.1521	0.0663	-0.1521	0.0663	-0.1521	-
	NTL	2	NR_LANE	-1.0950	0.3243	-1.0950	0.3243	-1.0950	-
	AC	2	A	-2.0090	0.6391	2.0090	0.6391	0.6485	1-no control, 2-partial control, 3-full control. TransCAD scale is reversed, slope adjusted by taking opposite sign.
		4	AC	-0.5637	0.1596	0.5637	0.1596		
	CRB	4	CRB	-0.3766	0.1495	-0.3766	0.1495	-0.3766	0-no shoulder curb exist, 1- shoulder curb exists
	OSP	3	SHLDR	-0.5250	0.2520	-0.5250	0.2520	-0.5250	0-no outside shoulder exists, 1- outside shoulder exists
	CLTL	2	T_C	-2.3580	0.8717	-2.3580	0.8717	-2.3580	0-no continuous exclusive left turn, 1-continious exclusive left turn lane exists
	LTL	3	TWLTL	-0.8650	0.3560	-0.8650	0.3560	-0.3558	0-no exclusive left turn, 1-exclusive left turn lane exists
		4	TL	-0.2505	0.1619	-0.2505	0.1619		
PDO	SW	2	ISW	-0.2050	0.0675	-0.2050	0.0675	-0.2050	-
	LW	4	LW	-0.1525	0.0588	-0.1525	0.0588	-0.1525	-
	NTL	2	NR_LANE	-0.9490	0.2132	-0.9490	0.2132	-0.9490	-
	AC	2	A	-1.3300	0.3916	1.3300	0.3916	0.5231	1-no control, 2-partial control, 3-full control. TransCAD scale is reversed, slope adjusted by taking opposite sign.
		4	AC	-0.4135	0.1443	0.4135	0.1443		
	CRB	4	CRB	-0.2786	0.1246	-0.2786	0.1246	-0.2786	0-no shoulder curb exist, 1- shoulder curb exists
	OSP	3	SHLDR	-0.6690	0.2500	-0.6690	0.2500	-0.6690	0-no outside shoulder exists, 1- outside shoulder exists
	CLTL	2	T_C	-1.7670	0.6837	-1.7670	0.6837	-1.7670	0-no continuous exclusive left turn, 1-continious exclusive left turn lane exists
	LTL	3	TWLTL	-0.6860	0.3210	-0.6860	0.3210	-0.3052	0-no exclusive left turn, 1-exclusive left turn lane exists
		4	TL	-0.2252	0.1471	-0.2252	0.1471		

Table A.7 Combining and converting variables for UI facility

Severity Level	Model Variable	Model	Original Model Variable	Original Parameter		Adjusted Parameter		Combined Parameter	Comments
				SLOPE	SE	SLOPE	SE		
TOTAL	LW	1	LW	-0.6507	0.3858	-0.6507	0.3858	-0.6507	-
	ISW	1	LSW	-0.2438	0.0719	-0.2438	0.0719	-0.2438	-
	MW	2	MW	-0.0176	0.0112	-0.0176	0.0142	-0.0176	Median width variable was redefined. 1)
IF	LW	1	LW	-0.9166	0.3422	-0.9166	0.3422	-0.9166	-
	ISW	1	LSW	-0.1783	0.0645	-0.1783	0.0645	-0.1783	-
PDO	LW	1	LW	-0.7974	0.4087	-0.7974	0.4087	-0.7974	-
	ISW	1	LSW	-0.2544	0.076	-0.2544	0.076	-0.2544	-
	MW	2	MW	-0.0156	0.0113	-0.0156	0.0144	-0.0156	Median width variable was redefined. 1)

Combining intersection original model variables

Table A.8 Variables for Signalized Intersections (Total Accidents)

Severity Level	Model Variable	Model	Original Model Variable	Parameter		Comments
				Slope	SE	
TOTAL	ADT	1	AVGADT	0.5718	0.1142	Enter 0: <15000 vpd Enter 1: 15000-30000vpd Enter 2: >30000vpd
		2	Q	PD/IF only	PD/IF only	
		3	L_Q	0.7821	0.1913	
		4	LOG AADT1	0.6199	0.2504	AADT major road
			LOG AADT2	0.3948	0.1737	AADT minor road
		5	LOG AADT1	0.4499	0.1968	AADT major road
			LOG AADT2	0.2699	0.0767	AADT minor road
	URBRUR	1	URBRUR	0.0949	0.0766	Enter 1: Urban Enter 0: Rural
	SHOULD	1	SHOULD	-0.1102	0.0813	Enter 1: Paved Shoulder Enter 0: Otherwise
	LANE	1	LANE	0.2654	0.0928	Enter 1: >4 Enter 0: <=4
		3	NO_LANES	0.0673	0.0314	
	PKTRUCK	4	PKTRUCK	0.0315	0.0143	

Table A.9 Variables for Signalized Intersections (Injury Accidents)

Severity Level	Model Variable	Model	Original Model Var.	Parameter		Comments
				Slope	SE	
INJURY	ADT	2	Q	0.723	-	
		3	L_Q	0.5848	0.1947	
		5	LOG AADT1	0.6697	0.1899	AADT major road
			LOG AADT2	0.2509	0.0929	AADT minor road
		4	LOG (AADT1 *AADT2)	0.2358	0.1722	
	URBRUR	3	URB_IND	-0.2677	0.1835	Enter 1: Urban Enter 0: Rural
	LANE	3	NO_LANES	0.0916	0.0280	
	PKTRUCK	4	PKTRUCK	0.0323	0.0146	

Table A.10 Variables for TWSC Intersections (Total Accidents)

Severity Level	Model Variable	Model	Original Model Var.	Parameter		Comments
				Slope	SE	
TOTAL	ADT	2	ADT	0.2777	0.0641	Enter 0: <15000 vpd Enter 1: 15000-30000vpd Enter 2: >30000vpd
		4	L Q	0.544	0.874	
		5-Type I	LOG AADT1	0.8052	0.0639	AADT major road
			LOG AADT2	0.5037	0.0708	AADT minor road
		5-Type II	LOG AADT1	0.6026	0.0836	AADT major road
			LOG AADT2	0.6091	0.0694	AADT minor road
		5-Type III	LOG AADT1	1.1479	0.2527	AADT major road
			LOG AADT2	0.2624	0.0866	AADT minor road
		5-Type IV	LOG AADT1	0.8503	0.2779	AADT major road
			LOG AADT2	0.3294	0.1255	AADT minor road
		6-Type I	LOG AADT1	0.7001	0.0460	AADT major road
			LOG AADT2	0.3785	0.0470	AADT minor road
		6-Type II	LOG AADT1	0.7079	0.0537	AADT major road
			LOG AADT2	0.5153	0.0529	AADT minor road
		6-Type III	LOG AADT1	0.8877	0.1666	AADT major road
			LOG AADT2	0.3228	0.0585	AADT minor road
		6-Type IV	LOG AADT1	0.7350	0.1849	AADT major road
			LOG AADT2	0.2390	0.0926	AADT minor road
	URBRUR	2	URBRUR	0.1193	0.0866	Enter 1: Urban Enter 0: Rural
	SHOULD	2	SHOULD	-0.1679	0.0954	Enter 1: Paved Shoulder Enter 0: Otherwise
	PKTRUCK	6-Type IV	PKTRUCK	-0.0479	0.0110	
	RTMAJ	5-Type I	RTMAJ	0.2671	0.1398	Enter 1: Yes, Enter 0: No
		6-Type I	RTMAJ	-0.1887	0.0841	Enter 1: Yes, Enter 0: No
	LTMAJ	6-Type I	LTMAJ	-0.155	0.110	Enter 1: Yes, Enter 0: No
	MEDWIDTH1	5-Type III	MEDWIDTH1	-0.0546	0.0249	
		6-Type III	MEDWIDTH1	-0.0106	0.0060	
	MEDTYPE1	6-Type III	MEDTYPE1	-0.3209	0.1771	Enter 0: No Median Enter 1: Painted Median Enter 2: Curbed Median Enter 3: Other

Table A.11 Variables for TWSC Intersections (Injury Accidents)

Severity Level	Model Variable	Model	Original Model Var.	Parameter		Comments
				Slope	SE	
INJURY	ADT	4	L_Q	0.539	0.145	
		5-Type I	LOG AADT1	0.8122	0.0973	AADT major road
			LOG AADT2	0.4551	0.0977	AADT minor road
		5-Type II	LOG AADT1	0.6339	0.1055	AADT major road
			LOG AADT2	0.6229	0.0870	AADT minor road
		5-Type III	LOG AADT1	1.1436	0.2763	AADT major road
			LOG AADT2	0.1357	0.1029	AADT minor road
		5-Type IV	LOG AADT1	0.9505	0.3284	AADT major road
			LOG AADT2	0.3237	0.1645	AADT minor road
		6-Type I	LOG AADT1	0.7174	0.0596	AADT major road
			LOG AADT2	0.3251	0.0569	AADT minor road
		6-Type II	LOG AADT1	0.6392	0.0638	AADT major road
			LOG AADT2	0.6584	0.0614	AADT minor road
		6-Type III	LOG AADT1	0.8498	0.2097	AADT major road
			LOG AADT2	0.2188	0.0949	AADT minor road
		6-Type IV	LOG AADT1	0.5008	0.2186	AADT major road
			LOG AADT2	0.3027	0.1341	AADT minor road
	PKTRUCK	6-Type III	PKTRUCK	-0.0253	0.0135	
		6-Type IV	PKTRUCK	-0.0520	0.0127	
	RTMAJ	5-Type I	RTMAJ*	0.3620	0.1814	Enter 1: Yes, Enter 0: No *Not included in the final model because it is not consistent with LTMAJ variable
		5-Type II	RTMAJ*	0.0451	0.1655	
	LTMAJ	6-Type I	LTMAJ	-0.258	0.132	Enter 1: Yes, Enter 0: No

Model Variables

This section defines *model variables*, the types of model variables, how they were derived and how they are used. In the second part of this section, a complete set of precise definition for all model variables is provided.

Model variables overview

Model variables are variables used either in crash modification factors (CMF) or in basic safety performance functions (BSPF), both of which are explained in the following sections of this chapter. Model variables express roadway characteristics that have an effect on the predicted number of crashes. The variable input into BSPF and CMF from this chapter should be collected according to the definitions described in this section.

There are two types of model variables used in CMF: continuous and binary.

Continuous variables measure specific roadway geometric characteristic:

- Lane width (LW)
- Shoulder width (SW)
- Inside shoulder width (ISW)
- Median width (MW)
- Access control (AC)
- Number of through lanes (NTL)
- Average degree of curve (ADCURVE)
- Average grade on section (AGRAD)
- Surrounding Land Use (URBRUR)
- Number of Lanes (LANE)
- Truck Percentage (PKTRUCK)
- Right-Turn Lane (RT MAJ)

- Left-Turn Lane (LT MAJ)
- Median Width (MEDWDTH1)
- Median Type (MEDTYPE1)

Binary variables express presence of specific roadway feature, 1- if the feature is present; 0 - if the feature is not present.

- Curb presence (CRB)
- Functional class (FC)
- Outside shoulder presence (OSP)
- Exclusive left turn lane (LT)
- Continuous left turn lane (CLTL)
- Surfaced shoulder (ST_Surfaced)
- Stabilized shoulder (ST_Stabilized)
- Combination shoulder (ST_Combined)
- Earth shoulder (ST_Earth)
- Paved Shoulder (SHOULD)

Model variables definitions

Segment model variables

Q (Annual average daily traffic, AADT) - An average daily value that represents all days of the reporting year. For two-way facilities, it is the AADT for both directions; for one-way streets, it is the directional AADT. Growth factors must be applied if the AADT is not derived from current year counts.

L (Length of section) – The length of a section measured along the centerline of a roadway. For independently aligned, divided highways, use the centerline length as the average of the lengths of the directional roadways, measured along their center lines.

Lane width (LW) - For multilane roads, it is the prevailing width of the travel lane on a section rounded to the nearest foot. The width of the travel lane is measured between the lane striping. For two-lane roads, the lane width is the distance from the centerline or delineation to where the pavement/shoulder surface changes, or to the pavement lane striping if the shoulder and pavement surface are the same. The parking lane is excluded from the lane width determination for all facility types. Where there is no delineation between the through traffic lane and the shoulder or parking lane, or where there is no centerline, estimate a reasonable split between the actual width used by traffic and the shoulder or parking lane based on state/local design guidelines.

Shoulder width (SW) – The shoulder width is the width of the shoulder, which includes rumble strips and gutter pans and does not include bike lanes and parking lanes. For two lane roads, it is the predominant width of the shoulder on a section. If there is no predominant width, enter the average width of both shoulders. For multilane roads, it is the predominant width of the outside shoulders. If there is no predominant width, enter the average width of the outside shoulders. Shoulder width should be recorded to the nearest foot.

Inside shoulder width (ISW) – The inside shoulder width is the predominant width of the inside shoulder, which includes rumble strips and gutter pans and does not include bike lanes and parking lanes. If there is no predominant width, use the average of both inside shoulders on a given section. Shoulder width should be recorded to the nearest foot.

Median width (MW) – The predominant width of the median, which includes inside shoulders, if any, measured between the inside edges of the through lanes to the nearest foot. If the median width is greater than 100 feet, enter 100. The median width does not include turning bays cut into the median.

Access Control (AC) - Measures the degree of access control on sample roadway sections. Enter 1 for full access control (preference given to through traffic movements by providing interchanges with selected public roads and by prohibiting crossing at grade and direct driveway connections); 2 for partial access control (preference given to through traffic movement. In addition to interchanges, there may be some crossings at-grade with public roads, but direct private driveway connections have been minimized through the use of frontage roads or other

local access restrictions); and 3 for no access control (include all sections that do not meet the criteria above). Control of curb cuts is not access control. If access control is not known, enter 1 for interstates, 2 for arterials, and 3 for collector and local roads.

Number of through lanes (NTL) – Determined according to the striping, if present, on multilane facilities, or according to traffic use or state/local design guidelines if no striping or only centerline striping is present. NTL is the prevailing number of through lanes in both directions carrying through traffic in the off-peak period. Exclude auxiliary lanes, such as collector-distributor lanes, weaving lanes, frontage road lanes, parking and turning lanes, acceleration/deceleration lanes, toll collection lanes and truck climbing lanes as defined in the *AASHTO Design Guide*.

Average degree of curve (ADCURVE) - Expresses the average degree of curve weighted with the length of each curve on that section. It is computed with the following formula:

$$ADCURVE = \frac{\sum_i S_i \cdot D(a)_i}{\sum_i S_i} , \quad (A.3)$$

where:

S_i = length of horizontal curve i on section,

$D(a)_i$ = degree of curve i (per 100 ft arc definition).

Average grade on section (AGRAD) – The average absolute grade on a section weighted with the length of each curve on that section. Each grade and flat segment is coded as a separate curve. The sum of all curves equals the section length. AGRAD is computed to the following formula:

$$AGRAD = \frac{\sum_i S_i \cdot G_i}{\sum_i S_i} , \quad (A.4)$$

where:

S_i = length of curve i on section,

G_i = absolute grade of curve i on section.

Curb presence (CRB) – Binary variable indicating the presence of a curb in the vicinity of the travel lane for two-lane roadways. For multilane roads this variable indicates the presence of an outside curb in the vicinity of the travel lane. Enter 1 if a curb is present in the direct vicinity of the travel lane, otherwise enter 0. If a parking lane exists and a curb is present, enter 0 for this variable.

Functional class (FC) – This binary variable indicates if a given segment is classified as a rural principal arterial. Enter 1 if the section is classified as a rural principal arterial, enter 0 otherwise. For any urban facility, enter 0.

Outside shoulder presence (OSP) – Binary variable indicating the presence of an outside shoulder. Enter 1 if the section contains an outside shoulder, otherwise enter 0. This variable is used for a UML facility.

Left turn lane (LT) – Binary variable indicating the presence of single exclusive left turn lane or multiple turning lanes that allow for simultaneous turns from all turning lanes; through movements are prohibited in these lanes. Enter 1 if the section contains a LT, otherwise enter 0.

Continuous left turn lane (CLTL) - Binary variable indicating the presence of a continuous exclusive left turning lane, from intersection to intersection on a given section; through movements are prohibited in this lane. Enter 1 if section contains CLTL, otherwise enter 0.

Surfaced shoulder (ST_Surfaced) – Binary variable indicating the presence of a paved shoulder on a section. Enter 1 if the section contains a paved shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Stabilized shoulder (ST_Stabilized) – Binary variable indicating the presence of a stabilized shoulder on a section. Enter 1 if the section contains a stabilized shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Combination shoulder (ST_Combined) – Binary variable indicating the presence of a combination shoulder on a section. Enter 1 if the section contains a combined shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Earth shoulder (ST_Earth) - Binary variable indicating the presence of an earth shoulder. Enter 1 if the section contains an earth shoulder, enter 0 otherwise. This variable is used for a UTL facility only.

Intersection model variables

Major Road Approach Volume (Q1) – Approach volumes on the major links (maximum of two) of the intersection, average daily value that represents all days of the reporting year.

Minor Road Approach Volume (Q2) – Approach volumes on the minor links of the intersection, average daily value that represents all days of the reporting year.

Surrounding Land Use (URBRUR) - Binary variable indicating whether the intersection is located in a rural or an urbanized area. Enter 1 if the intersection is located in an urban area; enter 0 if the intersection is located in a rural area.

Paved Shoulder (SHOULD) - Binary variable indicating the presence of a paved shoulder at the side of the major road of the intersection. Enter 1 if a paved shoulder exists, enter 0 otherwise.

Number of Lanes (LANE) - The number of lanes is defined as the total number of approach lanes at an intersection including through, shared, and exclusive turning lanes. The number of lanes is determined according to the striping, if present, on multilane facilities, or according to traffic use or state/local design guidelines if no striping is present.

Truck Percentage (PKTRUCK) - Variable % indicating the percentage of trucks in the intersection's peak hour traffic.

Right-Turn Lane (RT MAJ) - Binary variable indicating the presence of right-turn lanes on major roads. Enter 1 if there is a right-turn lane on the major road, otherwise enter 0.

Left-Turn Lane (LT MAJ) - Binary variable indicating the presence of left-turn lanes on at least one approach of major roads. Enter 1 if there is a left-turn lane on at least one approach of major roads, otherwise enter 0.

Median Width (MEDWDTH1) - Variable indicating the median width on major roads measured in feet. It is the predominant width of the median which includes inside shoulders, if any, measured between the inside edges of the through lanes to the nearest foot.

Median Type (MEDTYPE) - Variable indicating the type of the median on the major approach of the intersection. Enter 0 if no median exists, enter 1 for a painted median, enter 2 for a curbed median or enter 3 for other types of medians.

Appendix B   Linking Model Variables with TransCAD Variables

Segment

Linking to the HPMS Field Manual December 2000 (USDOT, FHWA) data format, which is used by TransCAD was required to use the inventory data available for all interstate, U.S., and state highways provided by FHWA. The linking process made it possible to use the HPMS data to predict the number of crashes on each inventory segment using the coded characteristics for that segment as input to BSPF and CMF.

Linking comments are provided in the “Linking comment” column. If there is no comment, a direct relationship exists between the model variable and the TransCAD variable. Model variables with direct links to the TransCAD variables are shown in the Table B.1.

Table B.1 Model variables and TransCAD variables

Model variables	TransCAD variables
LW	H54LaneW
SW	H59RShldW
ISW	H60LShldW
MW	H57MedW
AC	H55Access
NTL	H34TLanes

Direct linkage to TransCAD variables means that the value stored under the TransCAD variable in the HPMS data could be assigned to the model variable without any conversion.

Variables which required some transformation before linking them to the TransCAD variables are explained in the following text in the order mentioned in the first section of this chapter.

- *Average degree of curve* (ADCURVE) is calculated using the midpoint of all curve classes. For Class A, 2/3 of segments were assumed to be straight, thus the midpoint of class A was shifted to reflect the fact that not all segments in this class are curved.

Table B.2 shows all curves classes used in the HPMS data available in TransCAD.

Table B.2 Curves classes used in the HPMS

Item	Curve Classes			Length of Curves in Class (to 3 decimals) xx.xxx
	Curve Class	Radius Length (Metric)	Degree of Curvature (English)	
63	A	506+	0.0- 3.4	—
64	B	321- 505	3.5-5.4	—
65	C	206- 320	5.5- 8.4	—
66	D	126- 205	8.5-13.9	—
67	E	61- 125	14.0-27.9	—
68	F	<61	28+	—

The following formula is used to calculate ADCURVE using HPMS data available in TransCAD.

$$\begin{aligned}
 ADCURVE = & \frac{\frac{1}{3} \cdot 1.7 \cdot H63LCurveA + 4.45 \cdot H64LCurveB + 6.95 \cdot H65LCurveC +}{H63LCurveA + H64LCurveB + H65LCurveC +} \\
 & \frac{+ 11.2 \cdot H66LCurveD + 20.95 \cdot H67LCurveE + 28 \cdot H68LCurveF}{+ H66LCurveD + H67LCurveE + H68LCurveF},
 \end{aligned}
 \tag{B.1}$$

- *Average grade on section (AGRAD)* is calculated using the midpoint of all curve classes. No adjustment for the first class of grades was needed due to the fact that every road segment has some grade for drainage purpose.

Table B.3 shows all grade classes used in the HPMS data available in TransCAD.

Table B.3 Grade classes used in the HPMS

Item	Grade Class	Grade Classes by Gradient (Percent)	Length of Grades in Class (to 3 decimals) xx.xxx
72	A	0.0-0.4	—
73	B	0.5-2.4	—
74	C	2.5-4.4	—
75	D	4.5-6.4	—
76	E	6.5-8.4	—
77	F	8.5+	—

The following formula is used to calculate ADCURVE using the HPMS data available in TransCAD.

$$\begin{aligned}
 AGRAD = & \left(\frac{0.2 \cdot H72LGradeA + 1.45 \cdot H73LGradeB + 3.45 \cdot H74LGradeC +}{H72LGradeA + H73LGradeB + H74LGradeC +} \right. \\
 & \left. \frac{5.45 \cdot H75LGradeD + 7.45 \cdot H76LGradeE + 8.5 \cdot H77LGradeF}{+ H75LGradeD + H76LGradeE + H77LGradeF} \right),
 \end{aligned}
 \tag{B.2}$$

- The *Curb presence* (CRB) model variable is linked to the *HPMS Field Manual 2000* (USDOT, FHWA) shoulder type Item 58 data for UTL and UML as shown in Table B.4. The curb presence variable takes on a value of 1 if present and 0 if not present.

Table B.4 Curb presence

H58ShldT	Curb Variable Value
1	0
2	0
3	0
4	0
5	0
6	1

Table B.5 shows the description of TransCAD variable H58ShldT.

Table B.5 Item 58 — Shoulder type

Code	Description
1	None: No shoulders or curbs exist.
2	Surfaced shoulder exists (bituminous concrete or Portland cement concrete surface).
3	Stabilized shoulder exists (stabilized gravel or other granular material with or without admixture).
4	Combination shoulder exists (shoulder width has two or more surface types; for instance, part of the shoulder width is surfaced and a part of the width is earth, etc.).
5	Earth shoulder exists.
6	Barrier curb exists; no shoulders in front of curb.

- The *Functional Class* model (FC) variable is linked to the *HPMS Field Manual 2000* (USDOT, FHWA) functional system Item 17 (H17FC), data for RML as shown in Table B.6. The functional class variable takes on the value of 1 if section is a rural principal arterial and 0 otherwise.

Table B.6 Functional class

H17FC	Functional Class Value
2	1
6	0
7	0
8	0
9	0

Table B.7 shows the description of TransCAD variable H17FC.

Table B.7 Item 17 — Functional system code

Code	Description	Code	Description
RURAL		URBAN	
1	Principal Arterial -	11	Principal Arterial – Interstate
2	Principal Arterial - Other	12	Principal Arterial-Other Freeways &
6	Minor Arterial	14	Principal Arterial – Other
7	Major Collector	16	Minor Arterial
8	Minor Collector	17	Collector
9	Local	19	Local

For a facility to be classified as a RML, Item 17 (H17FC) has to be between 2 and 9; any other value of TransCAD variable H17FC model variable FC is equal to zero.

- The *Outside shoulder presence* (OSP) model variable is linked with the *HPMS Field Manual 2000* (USDOT, FHWA) shoulder type Item 58 (H58SType) data for UML facility as shown in Table B.8. The outside shoulder model variable takes on a value of 1 if a section contains an outside shoulder and 0 if it does not.

Table B.8 Outside shoulder presence

H58SType	Outside shoulder value
1	0
2	1
3	1
4	1
5	1
6	0

For a description of TransCAD variable H58SType values see Table B.5.

- The *left turn lane* (LT) and *continuous left turn lane* (CLTL) model variables are linked to the *HPMS Field Manual 2000* (USDOT, FHWA) left turning lane Item 88 (H88LTurn) data for UTL and UML as shown in Table B.9. Both variables take on the value of 1 if present and 0 if not.

Table B.9 Left turn lane and continuous left turn lane

H88LTurn	Exclusive left turn value	Continuous exclusive left turn value
0	0	0
1	1	0
2	1	1
3	1	0
4	0	0
5	0	0

TableB.10 Item 88 — Left turn lane codes

Code	Description
0	Not applicable; this is a rural section or no intersections exist on the section.
1	Turns permitted; multiple exclusive left turning lanes exist. Through movements are prohibited in these lanes. Multiple turning lanes allow for simultaneous turns from all turning lanes.
2	Turns permitted; a continuous exclusive left turning lane exists from intersection to intersection. Through movements are prohibited in this lane.
3	Turns permitted; a single exclusive left turning lane exists.
4	Turns permitted; no exclusive left turning lanes exist.
5	No left turns are permitted during the peak period.

- The ST_Surfaced, ST_Stabilized, ST_Combined, ST_Earth model variables are linked to the *HPMS Field Manual 2000* (USDOT, FHWA) shoulder type Item 58 (H58SType) data as shown in Table B.11

Table B.11 ST_Surfaced, ST_Stabilized, ST_Combined, ST_Earth

H58SType	Original model value	Slope a(i)	Transformed model variable
1	0	0	Reference case
2	3	3*model_slope	ST_Surfaced=1, 0-Otherwise
3	2	2*model_slope	ST_Stabilized=1, 0-Otherwise
4	2.5	2.5*model_slope	ST_Combination=1, 0-Otherwise
5	1	1*model_slope	ST_Earth=1, 0-otherwise
6	NA	NA	NA

Table B.12 to Table B.17 show how the model variables were linked with the TransCAD variables for all facility types. The tables are arranged by facility type in the order as specified in Section 5.1 ,Redefining Variables and Combining Slopes. The TransCAD

variable used column gives the TransCAD variable whose value is assigned to the corresponding model variable. The linking comment column provides information on the linking process, for columns where there is no comment a direct relationship exists.

Table B.12 Linking model variables to HPMS data available in TransCAD for RTL

Severity Level	Model Variable	Model	Combined Parameter	TransCAD Variable Used	Linking Comments
TOTAL	LW	1	-0.0598	H54LaneW	-
		2			
		3			
	SW	2	-0.0256	H59RShldW	-
		3			
	ADCURVE	2	0.1421	H63LCurveA - H68LCurveF	Average values based on each curve type coded in HPMS Field Manual 2000 were used. For details in section X.3.4.
IF	LW	2	0.0720	H72LGradeA - H77LGradeF	Average values based on each curve type coded in HPMS Field Manual 2000 were used. 1)
		3			
		1			
	SW	2	-0.0831	H54LaneW	-
		3			
	ADCURVE	2	0.2158	H63LCurveA - H68LCurveF	Average values based on each curve type coded in HPMS Field Manual 2000 were used. 1)
PDO	LW	2	-0.0442	H59RShldW	-
		3			
	AGRAD	3	0.0528	H72LGradeA - H77LGradeF	Average values based on each curve type coded in HPMS Field Manual 2000 were used. 1)
	AGRAD	3	0.0775	H72LGradeA - H77LGradeF	Average values based on each curve type coded in HPMS Field Manual 2000 were used. 1) 2)

Table B.13 Linking model variables to HPMS data available in TransCAD for RML

Severity Level	Final Variable	Model	Combined Parameter	TransCAD Variable Used	Linking Comments
TOTAL	LW	3	-0.1379	H54LaneW	-
	SW	1	-0.0940	H59RShldW	-
	ISW	3	-0.1329	H60LShldW	-
	MW	1	-0.0172	H57MedW	-
		2			
		3			
	AC	1	0.0044	H55Access	-
		2			
		3			
	ARHR	1	-	-	Not present in current HPMS data set
	FC	1	-0.5720	H17FC	-
IF	LW	3	-0.1472	H54LaneW	-
	ISW	2	-0.0697	H60LShldW	-
		3			
	MW	2	-0.0051	H57MedW	-
		3			
	AC	2	0.2412	H55Access	-
		3			
PDO	ISW	3	-0.0455	H60LShldW	-
	MW	2	-0.0048	H57MedW	-
		3			
	AC	2	0.1153	H55Access	-
		3			

Table B.14 Linking model variables to HPMS data available in TransCAD for RI

Severity Level	Final Variable	Model	Combined Parameter	TransCAD Variable Used	Linking Comments
TOTAL	LW	1	-0.5658	H54LaneW	-
	ISW	1	-1.3799	H60LShldW	-
	MW	1	-0.0100	H57MedW	-
IF	LW	1	-0.4179	H54LaneW	-
	ISW	1	-1.0011	H60LShldW	-
PDO	LW	1	-0.5706	H54LaneW	-
	ISW	1	-1.3338	H60LShldW	-
	MW	1	-0.0108	H57MedW	-

Table B.15 Linking model variables to HPMS data available in TransCAD for UTL

Severity Level	Final Variable	Model	Combined Parameter	TransCAD Variable Used	Linking Comments
TOTAL	LW	1	-0.1759	H54LaneW	-
		2			
	LTL	1	0.3859	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
		2			If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
	CRB	2	-0.2951	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb exists 3).
	ST Surfaced	2	-0.8889	H58ShldT	If H58ShldT=2 surfaced shoulder exists, otherwise surfaced shoulder does not exist 1).
	ST Stabilized	2	-0.5926	H58ShldT	If H58ShldT=3 stabilized shoulder exists, otherwise stabilized shoulder does not exist 1).
	ST Combination	2	-0.7408	H58ShldT	If H58ShldT=4 combination shoulder exists, otherwise combination shoulder does not exist 1).
	ST Earth	2	-0.2963	H58ShldT	If H58ShldT=5 earth shoulder exists, otherwise earth shoulder does not exist 1).
IF	LW	1	-0.1527	H54LaneW	-
		2			
	LTL	1	0.4522	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
		2			
	CRB	2	-0.2320	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb exists 3).
	ST Surfaced	2	-0.7707	H58ShldT	If H58ShldT=2 surfaced shoulder exists, otherwise surfaced shoulder does not exist 1).
	ST Stabilized	2	-0.5138	H58ShldT	If H58ShldT=3 stabilized shoulder exists, otherwise stabilized shoulder does not exist 1).
	ST Combination	2	-0.6423	H58ShldT	If H58ShldT=4 combination shoulder exists, otherwise combination shoulder does not exist 1).
	ST Earth	2	-0.2569	H58ShldT	If H58ShldT=5 earth shoulder exists, otherwise earth shoulder does not exist 1).
PDO	LW	1	-0.1658	H54LaneW	-
		2			
	LTL	1	0.3783	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
		2			
	CRB	2	-0.2686	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb exists 3).
	ST Surfaced	2	-0.8547	H58ShldT	If H58ShldT=2 surfaced shoulder exists, otherwise surfaced shoulder does not exist 1).
	ST Stabilized	2	-0.5698	H58ShldT	If H58ShldT=3 stabilized shoulder exists, otherwise stabilized shoulder does not exist 1).
	ST Combination	2	-0.7123	H58ShldT	If H58ShldT=4 combination shoulder exists, otherwise combination shoulder does not exist 1).
	ST Earth	2	-0.2849	H58ShldT	If H58ShldT=5 earth shoulder exists, otherwise earth shoulder does not exist 1).
	SW	1	0.0563	H59RShldW	-
		2			
	LW	1	-0.1527	H54LaneW	-
		2			
	LTL	1	0.4522	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
		2			
	CRB	2	-0.2320	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb exists 3).
	ST Surfaced	2	-0.7707	H58ShldT	If H58ShldT=2 surfaced shoulder exists, otherwise surfaced shoulder does not exist 1).
	ST Stabilized	2	-0.5138	H58ShldT	If H58ShldT=3 stabilized shoulder exists, otherwise stabilized shoulder does not exist 1).
	ST Combination	2	-0.6423	H58ShldT	If H58ShldT=4 combination shoulder exists, otherwise combination shoulder does not exist 1).
	ST Earth	2	-0.2569	H58ShldT	If H58ShldT=5 earth shoulder exists, otherwise earth shoulder does not exist 1).
	SW	2	0.0754	H59RShldW	-
	LW	2	-0.1658	H54LaneW	-
	LTL	2	0.3783	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
	CRB	2	-0.2686	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb exists 3).
	ST Surfaced	2	-0.8547	H58ShldT	If H58ShldT=2 surfaced shoulder exists, otherwise surfaced shoulder does not exist 1).
	ST Stabilized	2	-0.5698	H58ShldT	If H58ShldT=3 stabilized shoulder exists, otherwise stabilized shoulder does not exist 1).
	ST Combination	2	-0.7123	H58ShldT	If H58ShldT=4 combination shoulder exists, otherwise combination shoulder does not exist 1).
	ST Earth	2	-0.2849	H58ShldT	If H58ShldT=5 earth shoulder exists, otherwise earth shoulder does not exist 1).
	SW	2	0.0524	H59RShldW	-

Table B.16 Linking model variables to HPMS data available in TransCAD for UML

Severity Level	Final Variable	Model	Combined Parameter	TransCAD Variable Used	Linking Comments
TOTAL	LW	4	-0.1541	H54LaneW	-
	NTL	1	-1.0260	H34TLanes	-
		2			-
	AC	2	0.5632	H55Access	-
		4			-
	-	1	-	-	Data not provided in current HPMS data. Slope adjusted from per km to per mile
	-	1	-	-	Data not provided in current HPMS data. Slope adjusted from per km to per mile
	CRB	4	-0.2990	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb does not exist 3).
	OSP	3	-0.6310	H58ShldT	If H58ShldT=1,6 outside shoulder does not exist, otherwise shoulder does exist 1).
	CLTL	2	-1.9360	H88LTurn	If H88LTurn=2 continuous exclusive left turn exists otherwise no continuous exclusive left turn exist 2).
IF	LTL	3	-0.3058	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
		4			
	LW	4	-0.1521	H54LaneW	-
	NTL	2	-1.0950	H34TLanes	-
	AC	2	0.6485	H55Access	-
		4			-
	CRB	4	-0.3766	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb does not exist 3).
	OSP	3	-0.5250	H58ShldT	If H58ShldT=1,6 outside shoulder does not exist, otherwise shoulder does exist 1).
	CLTL	2	-2.3580	H88LTurn	If H88LTurn=2 continuous exclusive left turn exists otherwise no continuous exclusive left turn exist 2).
	LTL	3	-0.3558	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
		4			
PDO	SW	2	-0.2050	H60LShldW	-
	LW	4	-0.1525	H54LaneW	-
	NTL	2	-0.9490	H34TLanes	-
	AC	2	0.5231	H55Access	-
		4			-
	CRB	4	-0.2786	H58ShldT	If H58ShldT=6 shoulder curb exists, otherwise shoulder curb exists 3).
	OSP	3	-0.6690	H58ShldT	If H58ShldT=1,6 outside shoulder does not exist, otherwise shoulder does exist 1).
	CLTL	2	-1.7670	H88LTurn	If H88LTurn=2 continuous exclusive left turn exists otherwise no continuous exclusive left turn exist 2).
	LTL	3	-0.3052	H88LTurn	If H88LTurn=1,2,3 exclusive left turn exists otherwise no exclusive left turn exist 2).
		4			

Table B.17 Linking model variables to HPMS data available in TransCAD for UI

Severity Level	Final Variable Description	Model	Combined Parameter	TransCAD Variable Used	Linking Comments
TOTAL	LW	1	-0.6507	H54LaneW	-
	ISW	1	-0.2438	H60LShldW	-
	MW	2	-0.0176	H57MedW	-
IF	LW	1	-0.9166	H54LaneW	-
	ISW	1	-0.1783	H60LShldW	-
PDO	LW	1	-0.7974	H54LaneW	-
	ISW	1	-0.2544	H60LShldW	-
	MW	2	-0.0156	H57MedW	-

Intersections

Linking to the *HPMS Field Manual 2000* (USDOT, FHWA) data format which is used by TransCAD was required to use all the inventory data available. The linking process made it possible to use the HPMS data to predict the number of crashes on each inventory intersection using the coded characteristics for that intersection as input to the crash prediction models. The model variables with a direct link to the TransCAD variables are shown in Table B.18.

Table B.18 Model variables with direct link to TransCAD variables

Model Variable	Variable Symbol	Transcad Variable
Annual Average Daily	AADT	Item 33
Truck Percentage	PKTRUCK	Item 83
Median Width	MEDWDTH1	Item 57

Direct linkage to TransCAD variables means that the value stored under the TransCAD variable in the HPMS data could be assigned to the model variable without any conversion. Variables which required some transformation before linking them to the TransCAD variables are presented in Table B.19.

Table B.19 Model variables with no direct link to TransCAD variables

Model Variable	Variable Symbol	TransCAD Variable
Surrounding Land Use	URBRUR	Item 13
Number of Lanes	LANE	Item 34
Right-Turn Lane	RT MAJ	Item 89
Left-Turn Lane	LT MAJ	Item 88
Median Type	MEDTYPE	Item 56
Paved Shoulder	SHOULD	Item 58

- The Surrounding Land Use (URBRUR) variable is linked to the *HPMS Field Manual 2000* (USDOT, FHWA) Rural/Urban Designation Item 13, as shown in Table B.20 and Table B.21. The surrounding land use variable takes on the value of 1 if the intersection is located in an urban area and 0 otherwise.

Table B.20 Surrounding land use linked to HPMS Field Manual 2000, Item 13

Item 13	Surrounding Land Use
1	0
2	1
3	1
4	1

Table B.21 Item 13 - Rural/Urban designation

Code	Description
1	Rural Area
2	Small Urban Area (Population 5,000 to 49,999)
3	Small Urbanized Area (Population 50,000 to 199,999)
4	Large Urbanized Area (Population 200,000 or More)

- The Number of Lanes (LANE) variable (used only for signalized intersections) can be linked to the *HPMS Field Manual 2000* (USDOT, FHWA) Number of Through Lanes, Item 34. However, in order to use this variable, we needed to add the number of through, shared, and exclusive turning lanes to obtain the total number of lanes.
- The Right Turn Lane (RT MAJ) variable can be linked to the *HPMS Field Manual 2000* (USDOT, FHWA) Right Turning Lanes, Item 89, and
- 6.23. The right-lane turn variable takes on the value of 1 if a right-turn lane exists on major road or 0 otherwise.

Table B.22 Right turn lane linked to HPMS Field Manual 2000, Item 89

Item 89	Right Turn Lane
0	0
1	1
2	1
3	1
4	0
5	0

Table B.23 Item 89 Right turn lane codes

Code	Description
0	Not applicable; this is a rural section or no intersections exist on the section.
1	Turns permitted; multiple exclusive right turning lanes exist. Through movements are prohibited in these lanes. Multiple turning lanes allow for simultaneous turns from all turning lanes.
2	Turns permitted; a continuous exclusive right turning lane exists from intersection to intersection. Through movements are prohibited in this lane.
3	Turns permitted; a single exclusive right turning lane exists.
4	Turns permitted; no exclusive right turning lanes exist.
5	No right turns are permitted during the peak period.

The Left Turn Lane (LT MAJ) variable can be linked to the *HPMS Field Manual 2000* (USDOT, FHWA) Left Turning Lanes, Item 88, as shown in on major roads or 0 otherwise. The left-lane turn variable takes on the value of 1 if a left-turn lane exists on major roads or 0 otherwise.

Table B.24 Left turn lane linked to HPMS Field Manual 2000, Item 89

Item 88	Left Turn Lane
0	0
1	1
2	1
3	1
4	0
5	0

Table B.25 Item 88 Left turn lane codes

Code	Description
0	Not applicable; this is a rural section or no intersections exist on the section.
1	Turns permitted; multiple exclusive left turning lanes exist. Through movements are prohibited in these lanes. Multiple turning lanes allow for simultaneous turns from all turning lanes.
2	Turns permitted; a continuous exclusive left turning lane exists from intersection to intersection. Through movements are prohibited in this lane.
3	Turns permitted; a single exclusive left turning lane exists.
4	Turns permitted; no exclusive left turning lanes exist.
5	No left turns are permitted during the peak period.

- The Median Type (MEDTYPE) variable can be linked to the *HPMS Field Manual 2000* (USDOT, FHWA) Median Type, Item 56. In the original model, the median type variable takes on the value of 0 if no median exists, 1 for a painted median, 2 for a curbed median, or 3 for other types of medians. However, in Item 56, the median type variable takes on the value of 1 for a curbed median, 2 for a positive barrier, 3 for an unprotected median, and 4 if no median exists. Therefore, we needed to assign each case of the original model variable to the corresponding TransCAD variable, by estimating a new slope for each case according to the original model's slope, as it is shown in Table B.26.

Table B.26 Median type linked to HPMS Field Manual 2000, Item 56

Item 56	Median Type Value	Slope Estimation	Final Slope
1	2	$2 * (\text{Model Slope}) = 2 * (-0.3209)$	-0.6418
2,3	1,3	$[(1+3)/2] * (\text{Model Slope}) = 2 * (-0.3209)$	-0.6418

In this case, the Median Type variable takes on the value of 0 if no median exists or 1 in every other case with slope (-0.6418), as it is shown in Table B.27.

Table B.27 Median type linked to HPMS Field Manual 2000, Item 56

Item 56	Median Type	Slope
1	1	-0.6418
2	1	
3	1	
4	0	

- The Paved Shoulder (SHOULD) variable is linked to the *HPMS Field Manual 2000* (USDOT, FHWA) Shoulder Type Item 58, as shown in Table B.28.
- 6.28 and 6.29. The paved shoulder variable takes on the value of 1 if a paved shoulder exists at the side of the major road of the intersection, or 0 otherwise.

Table B.28 Paved shoulder linked to HPMS Field Manual 2000, Item 58

Item 58	Paved Shoulder
1	0
2	1
3	0
4	0
5	0
6	0

Table B.29 Item 58 - Shoulder type

Code	Description
1	None: No shoulders or curbs exist.
2	Surfaced shoulder exists (bituminous concrete or Portland cement concrete surface).
3	Stabilized shoulder exists (stabilized gravel or other granular material with or without admixture).
4	Combination shoulder exists (shoulder width has two or more surface types; for instance, part of the shoulder width is surfaced and a part of the width is earth, etc.).
5	Earth shoulder exists.
6	Barrier curb exists; no shoulders in front of curb.

Table B.30 to Table B.39 summarize for every model (signalized intersections (total and injury accidents) and TWSC intersections (total and injury accidents)) the final variables with their slopes and the linking to the variables available in TransCAD.

Table B.30 Linking model variables to HPMS data available in TransCAD for signalized intersections (Total Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
TOTAL	LOG AADT1	5	0.4499	Item 33	AADT major road
	LOG AADT2		0.2699		AADT minor road
	URBRUR	1	0.0949	Item 13	Enter 1: Urban Enter 0: Rural
	SHOULD	1	-0.1102	Item 58	Enter 1: Paved Shoulder Enter 0: Otherwise
	LANE	3	0.0673	Item 34	Total Number of Lanes
	PKTRUCK	4	0.0315	Item 83	Truck % (Peak Hour)

Table B.31 Linking model variables to HPMS data available in Transcad for signalized intersections (Injury Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
INJURY	LOG AADT1	5	0.6697	Item 33	AADT major road
	LOG AADT2		0.2509		AADT minor road
	URBRUR	3	-0.2677	Item 13	Enter 1: Urban Enter 0: Rural
	LANE	3	0.0916	Item 34	Total Number of Lanes
	PKTRUCK	4	0.0323	Item 83	Truck % (Peak Hour)

Table B.32 Linking model variables to HPMS data available in TransCAD for Type I - TWSC intersections (Total Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
TOTAL	LOG AADT1	6-Type I	0.7001	Item 33	AADT major road
	LOG AADT2		0.3785		AADT minor road
	URBRUR	2	0.1193	Item 13	Enter 1: Urban Enter 0: Rural
	SHOULD	2	-0.1679	Item 58	Enter 1: Paved Shoulder Enter 0: Otherwise
	RT MAJ	6-Type I	-0.1887	Item 89	Enter 1: Yes, Enter 0: No
	LT MAJ	6-Type I	-0.155	Item 88	Enter 1: Yes, Enter 0: No

Table B.33 Linking model variables to HPMS data available in TransCAD for Type II - TWSC intersections (Total Accidents)

Severity Level	Model Variable	Model	Parameter	Transcad Variable used	Linking Comments
TOTAL	LOG AADT1	6-Type II	0.7079	Item 33	AADT major road
	LOG AADT2		0.5153		AADT minor road
	URBRUR	2	0.1193	Item 13	Enter 1: Urban Enter 0: Rural
	SHOULD	2	-0.1679	Item 58	Enter 1: Paved Shoulder Enter 0: Otherwise

Table B.34 Linking model variables to HPMS data available in Transcad for Type III - TWSC intersections (Total Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
TOTAL	LOG AADT1	6-Type III	0.8877	Item 33	AADT major road
	LOG AADT2		0.3228		AADT minor road
	URBRUR	2	0.1193	Item 13	Enter 1: Urban Enter 0: Rural
	SHOULD	2	-0.1679	Item 58	Enter 1: Paved Shoulder Enter 0: Otherwise
	MEDWDTH1	6-Type III	-0.0106	Item 89	
	MEDTYPE	6-Type III	0.6418*	Item 88	*Slope Adjusted (See Error! Reference source not found. and Table B.27.)

Table B.35 Linking model variables to HPMS data available in Transcad for Type IV - TWSC intersections (Total Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
TOTAL	LOG AADT1	6-Type IV	0.7350	Item 33	AADT major road
	LOG AADT2		0.2390		AADT minor road
	URBRUR	2	0.1193	Item 13	Enter 1: Urban Enter 0: Rural
	SHOULD	2	-0.1679	Item 58	Enter 1: Paved Shoulder Enter 0: Otherwise
	PKTRUCK	6-Type IV	-0.0479	Item 83	Truck % (Peak Hour)

Table B.36 Linking model variables to HPMS data available in TransCAD for Type I - TWSC intersections (Injury Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
INJURY	LOG AADT1	6-Type I	0.7174	Item 33	AADT major road
	LOG AADT2		0.3251		AADT minor road
	LT MAJ	6-Type I	-0.258	Item 88	Enter 1: Yes, Enter 0: No

Table B.37 Linking model variables to HPMS data available in TransCAD for Type II - TWSC intersections (Injury Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
INJURY	LOG AADT1	6-Type II	0.6392	Item 33	AADT major road
	LOG AADT2		0.6584		AADT minor road

Table B.38 Linking model variables to HPMS data available in TransCAD for Type III - TWSC intersections (Injury Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
INJURY	LOG AADT1	6-Type III	0.8498	Item 33	AADT major road
	LOG AADT2		0.2188		AADT minor road
	PKTRUCK	6-Type III	-0.0253	Item 83	Truck % (Peak Hour)

Table B.39 Linking model variables to HPMS data available in TransCAD for Type IV - TWSC intersections (Injury Accidents)

Severity Level	Model Variable	Model	Parameter	TransCAD Variable used	Linking Comments
INJURY	LOG AADT1	6-Type IV	0.5008	Item 33	AADT major road
	LOG AADT2		0.3027		AADT minor road
	PKTRUCK	6-Type IV	-0.0520	Item 83	Truck % (Peak Hour)

Preliminary Set of Models per Facility

In this section, the preliminary set of models for all facility types, including roadway segments and intersections, will be presented. These models have been derived from the selected models being discussed in the previous section of this chapter, after redefining the original model variables and recalculating the corresponding slopes. The preliminary models per facility are shown below.

Segments

Model for RTL

Table B.40 Parameter Estimates for RTL

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.0831	-0.0524	-0.0598
SW	-0.0442	-0.0105	-0.0256
ADCURV	0.2158	-	0.1421
AGRAD	0.0528	0.0775	0.0720

Model for RML

Table B.41 Parameter Estimates for RML

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.1472	-	-0.1379
SW	-	-	-0.0940
ISW	-0.0697	-0.0455	-0.1329
MW	-0.0051	-0.0048	-0.0172
AC	0.2412	0.1153	0.0044
FC	-	-	-0.5720

Model for RI**Table B.42 Parameter Estimates for RI**

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.4179	-0.5706	-0.5658
ISW	-1.0011	-1.3338	-1.3799
MW	-	-0.0108	-0.0100

Model for UTL**Table B.43 Parameter Estimates for UTL**

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.1527	-0.1658	-0.1759
LTL	0.4522	0.3783	0.3859
CRB	-0.2320	-0.2686	-0.2951
ST_Surfaced	-0.7707	-0.8547	-0.8889
ST_Stabilized	-0.5138	-0.5698	-0.5926
ST_combination	-0.6423	-0.7123	-0.7408
ST_Earth	-0.2569	-0.2849	-0.2963
CELT	-	-	-1.5440
SW	0.0754	0.0524	0.0563

Model for UML**Table B.44 Parameter Estimates for UML**

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.1521	-0.1525	-0.1541
NTL	-1.0950	-0.9490	-1.0260
AC	0.6485	0.5231	0.5632
CRB	-0.3766	-0.2786	-0.2990
OSP	-0.5250	-0.6690	-0.6310
CLTL	-2.3580	-1.7670	-1.9360
LTL	-0.3558	-0.3052	-0.3058
SW	-0.2050	-	-

Model for UI**Table B.45 Parameter Estimates for UI**

Variables	Injury Crashes	PDO Crashes	Total
	Coeff.	Coeff.	Coeff.
LW	-0.9166	-0.7974	-0.6507
ISW	-0.1783	-0.2544	-0.2438
MW	-	-0.0156	-0.0176

Intersections

Model for Signalized Intersections

Table B.46 Parameter Estimates for Signalized Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.6697	0.4499
LOG AADT2	0.2509	0.2699
URBRUR	-0.2677	0.0949
SHOULD	-	-0.1102
LANE	0.0916	0.0673
PKTRUCK	0.0323	0.0315

Model for TWSC Intersections

➤ **Type 1- Three-Legged stop controlled intersections of two-lane roads**

Table B.47 Parameter Estimates for Type 1 - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.7174	0.7001
LOG AADT2	0.3251	0.3785
URBRUR	-	0.1193
SHOULD	-	-0.1679
RTMAJ	-	-0.1887
LTMAJ	-0.258	-0.155

➤ **Type II- Four-Legged stop controlled intersections of two-lane roads**

Table B.48 Parameter Estimates for Type1I - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.6392	0.7079
LOG AADT2	0.6584	0.5153
URBRUR	-	0.1193
SHOULD	-	-0.1679

➤ **Type III- Three-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads**

Table B.49 Parameter Estimates for Type 1II - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.8498	0.8877
LOG AADT2	0.2188	0.3228
URBRUR	-	0.1193
SHOULD	-	-0.1679
MEDWIDTH1	-	-0.0106
MEDTYPE	-	0.6418
PKTRUCK	-0.0253	-

- **Type IV- Four-Legged stop controlled intersections with two lanes on the minor and four lanes on major roads**

Table B.50 Parameter Estimates for Type1V - TWSC Intersection

Variables	Injury Crashes	Total
	Coeff.	Coeff.
LOG AADT1	0.5008	0.7350
LOG AADT2	0.3027	0.2390
URBRUR	-	0.1193
SHOULD	-	-0.1679
PKTRUCK	-0.0520	-0.0479

Appendix C New Indiana Safety Performance Functions

Segments

Summary of Data

Description of Segment Types

Segment types

1. rural two lane
2. rural multi lane
3. rural interstate
4. urban two lane
5. urban multi lane
6. urban interstate

Segment type =1 Rural Two Lanes

Injury and Fatal Crashes

Full Model

Negative binomial regression

Number of obs	=	11112
LR chi2(17)	=	2654.38
Prob > chi2	=	0.0000
Pseudo R2	=	0.0884

Log likelihood = -13690.378

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.6565624	.0162525	40.40	0.000	.624708	.6884167
logaad	.8298617	.0260598	31.84	0.000	.7787854	.8809379
h54lanew	-.0660017	.0229594	-2.87	0.004	-.1110013	-.0210021
h55access	-.0227761	.2153222	-0.11	0.916	-.4447999	.3992477
h57medw	.0046631	.0117203	0.40	0.691	-.0183083	.0276345
h59rshldw	-.0425169	.0071097	-5.98	0.000	-.0564516	-.0285821
h60lshldw	-.1504826	.1796898	-0.84	0.402	-.5026682	.201703
adcurve	.0308616	.0092631	3.33	0.001	.0127063	.0490169
agrade	.019704	.0106763	1.85	0.065	-.0012213	.0406292
crb	.3033295	.2396877	1.27	0.206	-.1664498	.7731087
fc	-.193731	.0473926	-4.09	0.000	-.2866187	-.1008433
int_l	-13.61319	471.0177	-0.03	0.977	-936.791	909.5646
int_r	-.1115607	.7953878	-0.14	0.888	-1.670492	1.447371
int_lr	-.6656085	.8665617	-0.77	0.442	-2.364038	1.032821
celt	.6279817	1.15748	0.54	0.587	-1.640637	2.896601
st_improved	.2474078	.056345	4.39	0.000	.1369735	.357842
sig	.5095343	.0717098	7.11	0.000	.3689858	.6500828
_cons	-5.721108	.6945791	-8.24	0.000	-7.082458	-4.359758
/lnalpha	.316012	.0319736			.253345	.378679
alpha	1.371647	.0438564			1.288328	1.460354

Likelihood-ratio test of alpha=0: chibar2(01) = 4805.15 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression      Number of obs   =      11112
                                LR chi2(2)           =      2495.43
                                Prob > chi2           =      0.0000
Log likelihood = -13769.851      Pseudo R2        =      0.0831

```

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.6397727	.0155615	41.11	0.000	.6092727	.6702726
logaad	.735771	.0203711	36.12	0.000	.6958443	.7756977
_cons	-5.656101	.1664465	-33.98	0.000	-5.98233	-5.329872
/lnalpha	.3573614	.0315141			.2955948	.4191279
alpha	1.429552	.0450511			1.343926	1.520635

Likelihood-ratio test of alpha=0: chibar2(01) = 5039.68 Prob>=chibar2 = 0.000

Best Model

```

Negative binomial regression      Number of obs   =      11112
                                LR chi2(8)           =      2630.76
                                Prob > chi2           =      0.0000
Log likelihood = -13702.188      Pseudo R2        =      0.0876

```

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.6693466	.0159631	41.93	0.000	.6380595	.7006338
logaad	.8388607	.0259481	32.33	0.000	.7880034	.8897181
h54lanew	-.0778716	.0227444	-3.42	0.001	-.1224498	-.0332934
h59rshldw	-.0303854	.0063803	-4.76	0.000	-.0428906	-.0178802
adcurve	.0293257	.0092519	3.17	0.002	.0111923	.0474591
agrade	.0195783	.0106601	1.84	0.066	-.001315	.0404717
fc	-.1812517	.0472557	-3.84	0.000	-.2738712	-.0886322
sig	.4802645	.0707198	6.79	0.000	.3416564	.6188727
_cons	-5.548371	.2419086	-22.94	0.000	-6.022503	-5.074239
/lnalpha	.3188769	.0319677			.2562214	.3815325
alpha	1.375582	.0439742			1.292039	1.464527

Likelihood-ratio test of alpha=0: chibar2(01) = 4820.89 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

Negative binomial regression

Number of obs = 11112

LR chi2(17) = 3434.86

Prob > chi2 = 0.0000

Pseudo R2 = 0.0677

Log likelihood = -23660.047

	pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength		.597931	.0122354	48.87	0.000	.57395	.621912
logaadt		.7740052	.020512	37.73	0.000	.7338023	.8142081
h54lanew		-.0425993	.018434	-2.31	0.021	-.0787292	-.0064694
h55access		.196824	.1850875	1.06	0.288	-.1659409	.5595889
h57medw		.0133142	.0118214	1.13	0.260	-.0098552	.0364837
h59rshldw		-.0261133	.0058267	-4.48	0.000	-.0375334	-.0146932
h60lshldw		-.4377233	.1939198	-2.26	0.024	-.8177992	-.0576474
adcurve		.0199533	.0068761	2.90	0.004	.0064764	.0334302
agrade		.0206902	.0087222	2.37	0.018	.0035949	.0377855
crb		.3280148	.1867073	1.76	0.079	-.0379248	.6939545

fc	-.2218966	.0409713	-5.42	0.000	-.3021989	-.1415943
int_l	-17.23132	1496.303	-0.01	0.991	-2949.932	2915.469
int_r	.0680057	.7226631	0.09	0.925	-1.348388	1.484399
int_lr	.2572229	.7305645	0.35	0.725	-1.174657	1.689103
celt	-.2237147	1.02018	-0.22	0.826	-2.22323	1.775801
st_improved	.0590814	.0443779	1.33	0.183	-.0278977	.1460605
sig	.6053931	.0619623	9.77	0.000	.4839493	.7268369
_cons	-4.862529	.5926923	-8.20	0.000	-6.024184	-3.700873

/lnalpha	.3491664	.0201657			.3096424	.3886904
-----						
alpha	1.417885	.0285926			1.362938	1.475048

Likelihood-ratio test of alpha=0: chibar2(01) = 2.5e+04 Prob>=chibar2 = 0.000						

Minimum Model

Negative binomial regression	Number of obs	=	11112
	LR chi2(2)	=	3209.83
	Prob > chi2	=	0.0000
Log likelihood = -23772.558	Pseudo R2	=	0.0632

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
-----					
loglength	.5652281	.0115868	48.78	0.000	.5425184 .5879377
logaadt	.7099907	.0161893	43.86	0.000	.6782604 .7417211
_cons	-4.240588	.1306907	-32.45	0.000	-4.496737 -3.984439

/lnalpha	.3823765	.0199683			.3432395 .4215136
-----					
alpha	1.465764	.0292687			1.409506 1.524267

Likelihood-ratio test of alpha=0: chibar2(01) = 2.6e+04 Prob>=chibar2 = 0.000					

Best Model

Negative binomial regression	Number of obs	=	11112
	LR chi2(10)	=	3426.22
	Prob > chi2	=	0.0000
Log likelihood = -23664.364	Pseudo R2	=	0.0675

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
-----					
loglength	.6007767	.012012	50.01	0.000	.5772335 .6243198
logaadt	.77586	.0204613	37.92	0.000	.7357566 .8159633
h54lanew	-.0470426	.0181727	-2.59	0.010	-.0826604 -.0114248
h59rshldw	-.023304	.0051487	-4.53	0.000	-.0333952 -.0132127
h60lshldw	-.2886184	.1000415	-2.88	0.004	-.4846962 -.0925405
adcurve	.0196329	.0068757	2.86	0.004	.0061567 .0331091
agrade	.0204652	.0087198	2.35	0.019	.0033746 .0375558
crb	.3176103	.1778864	1.79	0.074	-.0310406 .6662611
fc	-.2150087	.0408014	-5.27	0.000	-.2949779 -.1350394
sig	.5986014	.0612662	9.77	0.000	.4785218 .718681
_cons	-4.195272	.1925646	-21.79	0.000	-4.572692 -3.817853

/lnalpha	.349859	.0201651	.3103362	.3893818
alpha	1.418867	.0286116	1.363884	1.476068

Likelihood-ratio test of alpha=0: $\text{chibar2}(01) = 2.5\text{e}+04$ Prob>=chibar2 = 0.000

All Crashes

Full Model

```

Negative binomial regression      Number of obs   =      11112
                                LR chi2(17)          =      3601.22
                                Prob > chi2           =      0.0000
Log likelihood = -25926.14       Pseudo R2        =      0.0649

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.6001072	.0118157	50.79	0.000	.5769489	.6232655
logaadt	.7760357	.0199065	38.98	0.000	.7370196	.8150518
h54lanew	-.041538	.0181169	-2.29	0.022	-.0770465	-.0060295
h55access	.1820966	.1823303	1.00	0.318	-.1752641	.5394573
h57medw	.0094969	.0101486	0.94	0.349	-.010394	.0293879
h59rshldw	-.0298205	.0057436	-5.19	0.000	-.0410777	-.0185633
h60lshldw	-.3382431	.1571211	-2.15	0.031	-.6461949	-.0302914
adcurve	.0217094	.0067246	3.23	0.001	.0085295	.0348893
agrade	.020401	.0085586	2.38	0.017	.0036265	.0371754
crb	.3567757	.1837357	1.94	0.052	-.0033396	.716891
fc	-.2216185	.0405826	-5.46	0.000	-.3011589	-.1420781
int_l	-17.97706	1904.221	-0.01	0.992	-3750.183	3714.228
int_r	-.0105297	.7251842	-0.01	0.988	-1.431865	1.410805
int_lr	.09409	.7364664	0.13	0.898	-1.349358	1.537538
celt	-.1142759	1.02624	-0.11	0.911	-2.12567	1.897118
st_improved	.1174066	.043323	2.71	0.007	.0324951	.2023182
sig	.6157934	.0616094	10.00	0.000	.4950412	.7365456
_cons	-4.631913	.5832988	-7.94	0.000	-5.775158	-3.488668
/lnalpha	.3738191	.0188971			.3367814	.4108568
alpha	1.453274	.0274627			1.400433	1.508109

Likelihood-ratio test of alpha=0: chibar2(01) = 3.5e+04 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression      Number of obs   =      11112
                                LR chi2(2)          =      3365.99
                                Prob > chi2           =      0.0000
Log likelihood = -26043.751       Pseudo R2        =      0.0607

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.5716717	.0112545	50.80	0.000	.5496134	.5937301
logaadt	.709278	.0156865	45.22	0.000	.6785329	.740023
_cons	-3.978413	.1264001	-31.47	0.000	-4.226153	-3.730673
/lnalpha	.4058032	.0187256			.3691017	.4425048
alpha	1.500507	.0280979			1.446435	1.556601

Likelihood-ratio test of alpha=0: chibar2(01) = 3.6e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 11112
 LR chi2(10) = 3587.54
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0647

Log likelihood = -25932.978

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.6054753	.0116276	52.07	0.000	.5826857	.6282649
logaadt	.7794708	.0198642	39.24	0.000	.7405376	.818404
h54lanew	-.0498018	.0178678	-2.79	0.005	-.0848221	-.0147814
h59rshldw	-.0234672	.0050995	-4.60	0.000	-.033462	-.0134725
h60lshldw	-.2483078	.0909027	-2.73	0.006	-.4264739	-.0701417
adcurve	.0210632	.0067258	3.13	0.002	.0078809	.0342455
agrade	.0199348	.0085555	2.33	0.020	.0031663	.0367034
crb	.3014846	.1759873	1.71	0.087	-.0434442	.6464133
fc	-.2107371	.0404264	-5.21	0.000	-.2899714	-.1315028
sig	.601789	.0609431	9.87	0.000	.4823428	.7212353
_cons	-3.939073	.188736	-20.87	0.000	-4.308989	-3.569158
/lnalpha	.3748871	.0188986			.3378464	.4119277
alpha	1.454827	.0274942			1.401925	1.509725

Likelihood-ratio test of alpha=0: chibar2(01) = 3.5e+04 Prob>=chibar2 = 0.000

Segment type =2 Rural Multilanes

Injury and Fatal Crashes

Full Model

```

Negative binomial regression      Number of obs   =      1313
                                LR chi2(16)          =      350.56
                                Prob > chi2           =      0.0000
Log likelihood = -2030.6163      Pseudo R2        =      0.0795

```

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.6365264	.0465658	13.67	0.000	.5452592	.7277937
logaad	1.169453	.0919102	12.72	0.000	.9893118	1.349593
h34tlanes	.2284121	.74001	0.31	0.758	-1.221981	1.678805
h54lanew	-.1407937	.1662427	-0.85	0.397	-.4666235	.1850361
h55access	-.0976884	.1170383	-0.83	0.404	-.3270793	.1317024
h57medw	-.0075943	.0031712	-2.39	0.017	-.0138098	-.0013789
h59rshldw	-.0302679	.0320925	-0.94	0.346	-.093168	.0326323
h60lshldw	.0191539	.0338875	0.57	0.572	-.0472643	.085572
adcurve	-.0121071	.0508679	-0.24	0.812	-.1118064	.0875922
agrade	-.0645177	.0473673	-1.36	0.173	-.1573559	.0283206
crb	-.5150349	1.231382	-0.42	0.676	-2.928499	1.898429
fc	-.4646218	.128161	-3.63	0.000	-.7158128	-.2134309
int_l	.412623	1.72563	0.24	0.811	-2.96955	3.794796
celt	-15.60778	2224.708	-0.01	0.994	-4375.955	4344.74
st_improved	.1540194	.320376	0.48	0.631	-.473906	.7819448
sig	.5526772	.1244454	4.44	0.000	.3087687	.7965856
_cons	-8.447109	3.591604	-2.35	0.019	-15.48652	-1.407695
/lnalpha	.6491953	.0705964			.510829	.7875617
alpha	1.914	.1351214			1.666672	2.19803

Likelihood-ratio test of alpha=0: chibar2(01) = 1869.29 Prob>=chibar2 = 0.000

Minimum Model

```
Negative binomial regression      Number of obs   =      1313
                                LR chi2(2)           =      274.76
                                Prob > chi2          =      0.0000
Log likelihood = -2068.5136       Pseudo R2        =      0.0623
```

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.5496504	.0438139	12.55	0.000	.4637766	.6355241
logaadt	1.105639	.084821	13.03	0.000	.9393926	1.271885
_cons	-9.474218	.7932963	-11.94	0.000	-11.02905	-7.919386
/lnalpha	.7873657	.0674829			.6551015	.9196298
alpha	2.1976	.1483005			1.925338	2.508362

Likelihood-ratio test of alpha=0: chibar2(01) = 2184.82 Prob>=chibar2 = 0.000

Best Model

```
Negative binomial regression      Number of obs   =      1313
                                LR chi2(5)           =      344.74
                                Prob > chi2          =      0.0000
Log likelihood = -2033.5224       Pseudo R2        =      0.0781
```

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.6507616	.0455751	14.28	0.000	.561436	.7400872
logaadt	1.157185	.0869378	13.31	0.000	.9867905	1.32758
h57medw	-.0071121	.0024337	-2.92	0.003	-.0118821	-.0023422
fc	-.4911811	.1112981	-4.41	0.000	-.7093213	-.2730409
sig	.5565941	.1226578	4.54	0.000	.3161892	.7969989
_cons	-9.457434	.800493	-11.81	0.000	-11.02637	-7.888496
/lnalpha	.6600304	.0701666			.5225064	.7975543
alpha	1.934851	.1357619			1.686249	2.220105

Likelihood-ratio test of alpha=0: chibar2(01) = 1929.69 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

```

Negative binomial regression      Number of obs   =      1313
                                LR chi2(16)          =      348.45
                                Prob > chi2           =      0.0000
Log likelihood = -3307.2152      Pseudo R2         =      0.0500

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.5515146	.0391157	14.10	0.000	.4748493	.62818
logaad	.9649339	.0787803	12.25	0.000	.8105274	1.11934
h34tlanes	.601721	.7238137	0.83	0.406	-.8169277	2.02037
h54lanew	-.128966	.153841	-0.84	0.402	-.4304888	.1725568
h55access	-.1033983	.1041989	-0.99	0.321	-.3076245	.1008278
h57medw	-.0022432	.0028655	-0.78	0.434	-.0078594	.0033731
h59rshldw	-.009347	.0269654	-0.35	0.729	-.0621983	.0435043
h60lshldw	-.0201166	.0305165	-0.66	0.510	-.0799278	.0396947
adcurve	-.0126294	.0482269	-0.26	0.793	-.1071524	.0818936
agrade	-.0980971	.0400266	-2.45	0.014	-.1765477	-.0196465
crb	-.7779455	1.090723	-0.71	0.476	-2.915723	1.359832
fc	-.3117953	.1133393	-2.75	0.006	-.5339363	-.0896543
int_l	.977791	1.474036	0.66	0.507	-1.911266	3.866848
celt	-1.037367	2.128998	-0.49	0.626	-5.210126	3.135392
st_improved	-.3447217	.2718071	-1.27	0.205	-.8774537	.1880104
sig	.6249094	.1190584	5.25	0.000	.3915592	.8582596
_cons	-6.821091	3.457606	-1.97	0.049	-13.59787	-.0443089
/lnalpha	.7010095	.0499531			.6031033	.7989158
alpha	2.015787	.1006948			1.827782	2.223129

```

Likelihood-ratio test of alpha=0:  chibar2(01) = 8373.14 Prob>=chibar2 = 0.000

```


Minimum Model

Negative binomial regression

Number of obs = 1313
 LR chi2(2) = 276.48
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0397

Log likelihood = -3343.1998

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.4736104	.036458	12.99	0.000	.4021541	.5450668
logaadt	.9369828	.0676866	13.84	0.000	.8043194	1.069646
_cons	-6.696189	.629436	-10.64	0.000	-7.92986	-5.462517
/lnalpha	.7770914	.048995			.681063	.8731197
alpha	2.175136	.1065707			1.975977	2.394369

Likelihood-ratio test of alpha=0: chibar2(01) = 9190.16 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 1313
 LR chi2(6) = 338.90
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0487

Log likelihood = -3311.992

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.5525139	.0380088	14.54	0.000	.4780182	.6270097
logaadt	.9644909	.0756163	12.76	0.000	.8162856	1.112696
h54lanew	-.1943794	.1393652	-1.39	0.163	-.4675301	.0787713
fc	-.2797943	.0964178	-2.90	0.004	-.4687696	-.0908189
st_improved	-.5342219	.1789099	-2.99	0.003	-.8848789	-.183565
sig	.6297749	.1175802	5.36	0.000	.3993219	.8602279
_cons	-4.076367	1.696473	-2.40	0.016	-7.401394	-.7513398
/lnalpha	.7099	.0498688			.612159	.8076411
alpha	2.033788	.1014226			1.844409	2.242612

Likelihood-ratio test of alpha=0: chibar2(01) = 8519.36 Prob>=chibar2 = 0.000

All Crashes

Full Model

Negative binomial regression

Number of obs = 1313

LR chi2(16) = 371.51

Prob > chi2 = 0.0000

Pseudo R2 = 0.0493

Log likelihood = -3578.7935

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.5634066	.0380686	14.80	0.000	.4887935	.6380196
logaad	1.006033	.0778789	12.92	0.000	.8533932	1.158673
h34tlanes	.5577887	.7257927	0.77	0.442	-.8647388	1.980316
h54lanew	-.1200613	.1522443	-0.79	0.430	-.4184546	.1783319
h55access	-.1044452	.10401	-1.00	0.315	-.308301	.0994106
h57medw	-.0034451	.002848	-1.21	0.226	-.0090271	.0021368
h59rshldw	-.0063211	.0268327	-0.24	0.814	-.0589123	.0462701
h60lshldw	-.0158505	.0303981	-0.52	0.602	-.0754296	.0437286
adcurve	-.0145728	.0472835	-0.31	0.758	-.1072468	.0781012
agrade	-.0974676	.0395631	-2.46	0.014	-.1750099	-.0199253
crb	-.7368672	1.079468	-0.68	0.495	-2.852585	1.37885
fc	-.3628122	.1130265	-3.21	0.001	-.5843401	-.1412844
int_l	.910698	1.474953	0.62	0.537	-1.980157	3.801553
celt	-1.206511	2.133197	-0.57	0.572	-5.3875	2.974479
st_improved	-.2955143	.2709456	-1.09	0.275	-.8265579	.2355294
sig	.6014474	.1180112	5.10	0.000	.3701497	.8327452
_cons	-6.875773	3.448997	-1.99	0.046	-13.63568	-.1158625
/lnalpha	.7084037	.0477582			.6147993	.8020081
alpha	2.030747	.0969849			1.849285	2.230015

Likelihood-ratio test of alpha=0: chibar2(01) = 1.1e+04 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs = 1313

LR chi2(2) = 296.67

Prob > chi2 = 0.0000

Pseudo R2 = 0.0394

Log likelihood = -3616.212

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.4863591	.0359313	13.54	0.000	.4159351	.5567831
logaad	.9686064	.0675413	14.34	0.000	.8362278	1.100985
_cons	-6.730863	.6282129	-10.71	0.000	-7.962137	-5.499588
/lnalpha	.7835712	.0468438			.6917592	.8753833
alpha	2.189277	.1025539			1.997226	2.399795

Likelihood-ratio test of alpha=0: chibar2(01) = 1.2e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 1313

LR chi2(6) = 361.80

Prob > chi2 = 0.0000

Pseudo R2 = 0.0481

Log likelihood = -3583.6475

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.5737391	.0374559	15.32	0.000	.5003268	.6471514
logaadt	.9881004	.074402	13.28	0.000	.8422752	1.133926
h57medw	-.0033895	.0022909	-1.48	0.139	-.0078795	.0011005
fc	-.3266571	.0983842	-3.32	0.001	-.5194867	-.1338276
st_improved	-.3401561	.1903637	-1.79	0.074	-.7132622	.03295
sig	.6034941	.1166053	5.18	0.000	.3749519	.8320363
_cons	-6.345661	.655873	-9.68	0.000	-7.631149	-5.060174
/lnalpha	.7167536	.0476935			.6232761	.8102311
alpha	2.047775	.0976655			1.865028	2.248428

Likelihood-ratio test of alpha=0: chibar2(01) = 1.1e+04 Prob>=chibar2 = 0.000

Segment type =3 Rural Interstate

Injury and Fatal Crashes

Full Model

Negative binomial regression

Number of obs	=	307
LR chi2(8)	=	349.97
Prob > chi2	=	0.0000
Pseudo R2	=	0.2101

Log likelihood = -657.85096

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.064465	.051416	20.70	0.000	.9636917	1.165239
logaad	.8899865	.1353313	6.58	0.000	.6247421	1.155231
h34tlanes	.0157128	.1292233	0.12	0.903	-.2375603	.2689859
h57medw	.0005056	.0034725	0.15	0.884	-.0063003	.0073115
h59rshldw	-.13426	.0693478	-1.94	0.053	-.2701791	.0016591
h60lshldw	-.017328	.026916	-0.64	0.520	-.0700823	.0354263
adcurve	-.1306029	.0863501	-1.51	0.130	-.299846	.0386403
agrade	-.1214118	.0875324	-1.39	0.165	-.2929722	.0501487
_cons	-6.879631	1.817008	-3.79	0.000	-10.4409	-3.31836
/lnalpha	-1.496056	.1908278			-1.870072	-1.12204
alpha	.2240119	.0427477			.1541126	.3256148

Likelihood-ratio test of alpha=0: chibar2(01) = 113.44 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs	=	307
LR chi2(2)	=	341.44
Prob > chi2	=	0.0000
Pseudo R2	=	0.2050

Log likelihood = -662.11779

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.077114	.0506324	21.27	0.000	.9778767	1.176352
logaad	1.059796	.1135468	9.33	0.000	.8372486	1.282344
_cons	-10.15081	1.166152	-8.70	0.000	-12.43643	-7.865197
/lnalpha	-1.40557	.1828614			-1.763972	-1.047168
alpha	.2452273	.0448426			.1713629	.3509301

Likelihood-ratio test of alpha=0: chibar2(01) = 133.12 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 307

LR chi2(3) = 345.23

Prob > chi2 = 0.0000

Pseudo R2 = 0.2073

Log likelihood = -660.22328

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.062939	.0503372	21.12	0.000	.9642804	1.161598
logaadt	.9454565	.1260006	7.50	0.000	.6985	1.192413
h59rshldw	-.1347467	.0686336	-1.96	0.050	-.2692661	-.0002273
_cons	-7.574624	1.738935	-4.36	0.000	-10.98287	-4.166373
/lnalpha	-1.45598	.1869055			-1.822308	-1.089652
alpha	.2331717	.0435811			.1616522	.3363335

Likelihood-ratio test of alpha=0: chibar2(01) = 123.27 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

```

Negative binomial regression                                Number of obs   =          307
                                                           LR chi2(8)      =          367.73
                                                           Prob > chi2     =          0.0000
Log likelihood = -1097.6559                               Pseudo R2       =          0.1435

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.066191	.0437474	24.37	0.000	.980448	1.151935
logaad	.7978151	.1344661	5.93	0.000	.5342664	1.061364
h34tlanes	.0264226	.1303037	0.20	0.839	-.228968	.2818132
h57medw	.00074	.0033619	0.22	0.826	-.0058493	.0073292
h59rshldw	-.076627	.0661201	-1.16	0.246	-.2062201	.052966
h60lshldw	.0052949	.0258213	0.21	0.838	-.0453138	.0559037
adcurve	-.0363901	.064233	-0.57	0.571	-.1622844	.0895043
agrade	.0276451	.078657	0.35	0.725	-.1265198	.18181
_cons	-5.278729	1.777617	-2.97	0.003	-8.762795	-1.794664
/lnalpha	-.8325556	.1103106			-1.04876	-.6163508
alpha	.4349363	.0479781			.3503718	.5399111

Likelihood-ratio test of alpha=0: chibar2(01) = 1477.26 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression                                Number of obs   =          307
                                                           LR chi2(2)      =          365.91
                                                           Prob > chi2     =          0.0000
Log likelihood = -1098.5649                               Pseudo R2       =          0.1428

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.069217	.0427846	24.99	0.000	.9853604	1.153073
logaad	.8616069	.1075687	8.01	0.000	.650776	1.072438
_cons	-6.554372	1.093055	-6.00	0.000	-8.69672	-4.412025
/lnalpha	-.8245025	.1095029			-1.039124	-.6098807
alpha	.4384531	.0480119			.3537643	.5434157

Likelihood-ratio test of alpha=0: chibar2(01) = 1599.34 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 307
 LR chi2(2) = 365.91
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.1428

Log likelihood = -1098.5649

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.069217	.0427846	24.99	0.000	.9853604	1.153073
logaadt	.8616069	.1075687	8.01	0.000	.650776	1.072438
_cons	-6.554372	1.093055	-6.00	0.000	-8.69672	-4.412025
/lnalpha	-.8245025	.1095029			-1.039124	-.6098807
alpha	.4384531	.0480119			.3537643	.5434157

Likelihood-ratio test of alpha=0: chibar2(01) = 1599.34 Prob>=chibar2 = 0.000

All Crashes

Full Model

```

Negative binomial regression                                Number of obs   =          307
                                                           LR chi2(8)      =        370.28
                                                           Prob > chi2     =         0.0000
Log likelihood = -1151.5597                               Pseudo R2       =         0.1385

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.071244	.0433609	24.71	0.000	.9862578	1.156229
logaadt	.8160068	.1335166	6.11	0.000	.554319	1.077695
h34tlanes	.0053521	.1309384	0.04	0.967	-.2512825	.2619866
h57medw	-.000133	.0033692	-0.04	0.969	-.0067365	.0064705
h59rshldw	-.0758203	.0662052	-1.15	0.252	-.2055801	.0539395
h60lshldw	-.006161	.02585	-0.24	0.812	-.0568261	.0445042
adcurve	-.0534074	.0643411	-0.83	0.407	-.1795137	.0726989
agrade	.0230101	.0784942	0.29	0.769	-.1308358	.176856
_cons	-5.075105	1.772597	-2.86	0.004	-8.54933	-1.600879
/lnalpha	-.8028942	.1078537			-1.014284	-.5915048
alpha	.4480304	.0483217			.3626622	.5534938

Likelihood-ratio test of alpha=0: chibar2(01) = 1765.77 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression                                Number of obs   =          307
                                                           LR chi2(2)      =        368.29
                                                           Prob > chi2     =         0.0000
Log likelihood = -1152.5579                               Pseudo R2       =         0.1378

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.07193	.042358	25.31	0.000	.9889095	1.15495
logaadt	.8819571	.1059471	8.32	0.000	.6743046	1.08961
_cons	-6.575203	1.076089	-6.11	0.000	-8.684299	-4.466107
/lnalpha	-.7968649	.1072608			-1.007092	-.5866376
alpha	.4507399	.0483467			.3652796	.5561943

Likelihood-ratio test of alpha=0: chibar2(01) = 1882.63 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 307

LR chi2(2) = 368.29

Prob > chi2 = 0.0000

Pseudo R2 = 0.1378

Log likelihood = -1152.5579

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	1.07193	.042358	25.31	0.000	.9889095	1.15495
logaadt	.8819571	.1059471	8.32	0.000	.6743046	1.08961
_cons	-6.575203	1.076089	-6.11	0.000	-8.684299	-4.466107
/lnalpha	-.7968649	.1072608			-1.007092	-.5866376
alpha	.4507399	.0483467			.3652796	.5561943

Likelihood-ratio test of alpha=0: chibar2(01) = 1882.63 Prob>=chibar2 = 0.000

Segment type =4 Rural Two Lanes

Injury and Fatal Crashes

Full Model

Negative binomial regression

Number of obs = 3221

LR chi2(16) = 665.99

Prob > chi2 = 0.0000

Pseudo R2 = 0.0549

Log likelihood = -5734.5988

	if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength		.4395649	.028079	15.65	0.000	.384531	.4945987
logaad		.7267777	.0480615	15.12	0.000	.632579	.8209765
h54lanew		-.0033962	.053925	-0.06	0.950	-.1090872	.1022948
h55access		1.05971	.3263036	3.25	0.001	.4201666	1.699253
h57medw		.0060146	.010392	0.58	0.563	-.0143532	.0263825
h59rshldw		.0071564	.0131815	0.54	0.587	-.0186789	.0329917
h60lshldw		-.1264105	.1985028	-0.64	0.524	-.5154688	.2626478
adcurve		.0056928	.0099657	0.57	0.568	-.0138395	.0252251
agrade		-.0109119	.0168408	-0.65	0.517	-.0439193	.0220956
crb		.3020756	.1148991	2.63	0.009	.0768775	.5272737
int_l		.5998803	.1984187	3.02	0.003	.2109869	.9887738
int_r		.4052575	.4428821	0.92	0.360	-.4627755	1.27329
int_lr		.2337473	.5186306	0.45	0.652	-.7827501	1.250245
celt		-.4722432	.2056598	-2.30	0.022	-.8753291	-.0691574
st_improved		-.1060341	.0913319	-1.16	0.246	-.2850413	.0729732
sig		.5226896	.0608785	8.59	0.000	.40337	.6420092
_cons		-8.50352	1.191058	-7.14	0.000	-10.83795	-6.169089
/lnalpha		.3774375	.0427005			.293746	.4611289
alpha		1.458542	.0622805			1.341443	1.585863

Likelihood-ratio test of alpha=0: chibar2(01) = 3892.06 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs	=	3221
LR chi2(2)	=	521.49
Prob > chi2	=	0.0000
Pseudo R2	=	0.0430

Log likelihood = -5806.8491

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.4186577	.0269683	15.52	0.000	.3658009	.4715146
logaad	.87573	.0451544	19.39	0.000	.787229	.964231
_cons	-6.585447	.4100693	-16.06	0.000	-7.389168	-5.781726
/lnalpha	.4705824	.041213			.3898064	.5513584
alpha	1.600926	.065979			1.476695	1.735609

Likelihood-ratio test of alpha=0: chibar2(01) = 4419.54 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs	=	3221
LR chi2(7)	=	661.73
Prob > chi2	=	0.0000
Pseudo R2	=	0.0545

Log likelihood = -5736.7279

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.4284651	.0268698	15.95	0.000	.3758013	.4811288
logaad	.7296002	.0455081	16.03	0.000	.640406	.8187944
h55access	1.078472	.3194025	3.38	0.001	.4524548	1.704489
crb	.3375575	.1103233	3.06	0.002	.1213278	.5537872
int_l	.5867443	.1942629	3.02	0.003	.2059961	.9674925
celt	-.4492604	.201815	-2.23	0.026	-.8448106	-.0537102
sig	.5337275	.060294	8.85	0.000	.4155535	.6519015
_cons	-8.695573	1.028538	-8.45	0.000	-10.71147	-6.679676
/lnalpha	.3788378	.0426859			.2951749	.4625006
alpha	1.460586	.0623464			1.343361	1.58804

Likelihood-ratio test of alpha=0: chibar2(01) = 3896.10 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

```

Negative binomial regression                                Number of obs   =       3221
                                                           LR chi2(16)     =       805.05
                                                           Prob > chi2      =       0.0000
Log likelihood = -9216.5714                               Pseudo R2       =       0.0418

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3952815	.0241574	16.36	0.000	.3479339	.4426291
logaadt	.599821	.0372112	16.12	0.000	.5268884	.6727536
h54lanew	.1593416	.0463091	3.44	0.001	.0685775	.2501057
h55access	.7426066	.2448547	3.03	0.002	.2627003	1.222513
h57medw	.0070048	.0086159	0.81	0.416	-.009882	.0238916
h59rshldw	.0038715	.0118959	0.33	0.745	-.0194439	.027187
h60lshldw	-.2776649	.1585962	-1.75	0.080	-.5885077	.033178
adcurve	.0065509	.0085606	0.77	0.444	-.0102277	.0233294
agrade	-.0075569	.0148555	-0.51	0.611	-.0366732	.0215594
crb	.2772353	.1055218	2.63	0.009	.0704164	.4840542
int_l	.2949304	.197411	1.49	0.135	-.0919879	.6818488
int_r	.0352962	.4258649	0.08	0.934	-.7993837	.8699761
int_lr	-.3732273	.487	-0.77	0.443	-1.32773	.5812752
celt	-.147061	.2047812	-0.72	0.473	-.5484248	.2543028
st_improved	-.3135027	.0813823	-3.85	0.000	-.4730091	-.1539964
sig	.6334919	.0554201	11.43	0.000	.5248706	.7421133
_cons	-7.021588	.9473852	-7.41	0.000	-8.878428	-5.164747
/lnalpha	.404035	.0301613			.34492	.46315
alpha	1.497856	.0451772			1.411877	1.589072

Likelihood-ratio test of alpha=0: chibar2(01) = 1.9e+04 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression                                Number of obs   =       3221
                                                           LR chi2(2)      =       507.44
                                                           Prob > chi2      =       0.0000
Log likelihood = -9365.3734                               Pseudo R2       =       0.0264

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3415076	.0243267	14.04	0.000	.2938282	.389187
logaadt	.7449011	.0368708	20.20	0.000	.6726357	.8171666
_cons	-4.216033	.3365777	-12.53	0.000	-4.875713	-3.556352
/lnalpha	.5152734	.029295			.4578563	.5726906
alpha	1.674096	.0490427			1.580682	1.773031

Likelihood-ratio test of alpha=0: chibar2(01) = 2.2e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 3221

LR chi2(8) = 790.08

Prob > chi2 = 0.0000

Log likelihood = -9224.0542

Pseudo R2 = 0.0411

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3884012	.0239705	16.20	0.000	.34142	.4353825
logaadt	.6111326	.0363614	16.81	0.000	.5398656	.6823996
h55access	.744692	.2438747	3.05	0.002	.2667064	1.222678
h60lshldw	-.1502799	.081864	-1.84	0.066	-.3107304	.0101705
crb	.2797449	.1043436	2.68	0.007	.0752351	.4842546
int_1	.1568695	.0703061	2.23	0.026	.0190721	.294667
st_improved	-.3378241	.0499984	-6.76	0.000	-.4358191	-.2398291
sig	.6307819	.0550527	11.46	0.000	.5228806	.7386833
_cons	-5.229869	.78232	-6.69	0.000	-6.763188	-3.69655
/lnalpha	.406996	.0301814			.3478416	.4661504
alpha	1.502298	.0453414			1.416008	1.593847

Likelihood-ratio test of alpha=0: chibar2(01) = 1.9e+04 Prob>=chibar2 = 0.000

All Crashes

Full Model

```

Negative binomial regression                                Number of obs   =       3221
                                                           LR chi2(16)     =       819.52
                                                           Prob > chi2     =       0.0000
Log likelihood = -9934.6753                               Pseudo R2      =       0.0396

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.4014117	.0236816	16.95	0.000	.3549966	.4478269
logaadt	.6075043	.0363815	16.70	0.000	.5361978	.6788108
h54lanew	.1446936	.0447938	3.23	0.001	.0568994	.2324878
h55access	.7693594	.2390911	3.22	0.001	.3007494	1.237969
h57medw	.007862	.0084854	0.93	0.354	-.008769	.024493
h59rshldw	.0042938	.0117709	0.36	0.715	-.0187768	.0273643
h60lshldw	-.2743433	.1579269	-1.74	0.082	-.5838743	.0351877
adcurve	.0059427	.0084435	0.70	0.482	-.0106062	.0224917
agrade	-.0068061	.0147041	-0.46	0.643	-.0356255	.0220134
crb	.2781293	.1045736	2.66	0.008	.0731688	.4830899
int_l	.3653542	.1943307	1.88	0.060	-.015527	.7462354
int_r	.1268243	.4214895	0.30	0.763	-.69928	.9529286
int_lr	-.2781087	.4819683	-0.58	0.564	-1.222749	.6665317
celt	-.2208236	.2014292	-1.10	0.273	-.6156176	.1739704
st_improved	-.2736209	.0805636	-3.40	0.001	-.4315227	-.1157192
sig	.6135134	.0549799	11.16	0.000	.5057549	.7212719
_cons	-6.759998	.9247617	-7.31	0.000	-8.572498	-4.947499
/lnalpha	.4028484	.0288514			.3463008	.4593961
alpha	1.49608	.043164			1.413828	1.583118

Likelihood-ratio test of alpha=0: chibar2(01) = 2.5e+04 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression                                Number of obs   =       3221
                                                           LR chi2(2)     =       541.85
                                                           Prob > chi2     =       0.0000
Log likelihood = -10073.514                               Pseudo R2      =       0.0262

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3541243	.0237304	14.92	0.000	.3076135	.400635
logaadt	.744633	.0356222	20.90	0.000	.6748148	.8144512
_cons	-3.951469	.3249589	-12.16	0.000	-4.588376	-3.314561
/lnalpha	.5003496	.0281679			.4451417	.5555576
alpha	1.649298	.0464572			1.560711	1.742913

Likelihood-ratio test of alpha=0: chibar2(01) = 2.9e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 3221

LR chi2(8) = 804.92

Prob > chi2 = 0.0000

Pseudo R2 = 0.0389

Log likelihood = -9941.9781

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3944245	.0234944	16.79	0.000	.3483762	.4404727
logaad	.6175736	.0354154	17.44	0.000	.5481607	.6869865
h55access	.7736177	.2381396	3.25	0.001	.3068727	1.240363
h60lshldw	-.1297019	.0793413	-1.63	0.102	-.2852079	.0258041
crb	.2836979	.1036034	2.74	0.006	.080639	.4867568
int_l	.1621542	.0697679	2.32	0.020	.0254116	.2988967
st_improved	-.2920547	.0494336	-5.91	0.000	-.3889428	-.1951666
sig	.6129425	.0545969	11.23	0.000	.5059345	.7199504
_cons	-5.1398	.7634528	-6.73	0.000	-6.636141	-3.64346
/lnalpha	.4056937	.0288697			.3491102	.4622772
alpha	1.500343	.0433144			1.417805	1.587685

Likelihood-ratio test of alpha=0: chibar2(01) = 2.5e+04 Prob>=chibar2 = 0.000

Segment type =5 Urban Multilanes

Injury and Fatal Crashes

Full Model

Negative binomial regression

Number of obs = 2836

LR chi2(17) = 526.20

Prob > chi2 = 0.0000

Pseudo R2 = 0.0371

Log likelihood = -6822.6344

	if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength		.377408	.0311799	12.10	0.000	.3162964	.4385195
logaad		.8462935	.0523731	16.16	0.000	.7436442	.9489428
h34tlanes		-.0291165	.0519806	-0.56	0.575	-.1309966	.0727636
h54lanew		-.089788	.0740944	-1.21	0.226	-.2350104	.0554343
h55access		.3036946	.0720691	4.21	0.000	.1624418	.4449474
h57medw		-.0037667	.0018157	-2.07	0.038	-.0073254	-.0002081
h59rshldw		.0205861	.0174894	1.18	0.239	-.0136925	.0548647
h60lshldw		-.0058821	.0176833	-0.33	0.739	-.0405407	.0287766
adcurve		-.0164299	.0151842	-1.08	0.279	-.0461904	.0133306
agrade		-.0522402	.0225095	-2.32	0.020	-.0963581	-.0081224
crb		-.0768988	.0983511	-0.78	0.434	-.2696635	.1158658
int_l		.054498	.1178179	0.46	0.644	-.1764208	.2854168
int_r		.145754	.2300359	0.63	0.526	-.3051081	.596616
int_lr		.1311505	.2037478	0.64	0.520	-.268188	.5304889
celt		.122724	.1332861	0.92	0.357	-.138512	.3839599
st_improved		-.3523057	.1544857	-2.28	0.023	-.6550921	-.0495194
sig		.4696721	.0570695	8.23	0.000	.3578179	.5815263
_cons		-6.010627	.999031	-6.02	0.000	-7.968691	-4.052562
/lnalpha		.5307281	.0354878			.4611732	.6002829
alpha		1.70017	.0603353			1.585934	1.822634

Likelihood-ratio test of alpha=0: chibar2(01) = 1.0e+04 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs = 2836
 LR chi2(2) = 330.65
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0233

Log likelihood = -6920.4092

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3271412	.0315209	10.38	0.000	.2653613	.3889211
logaadt	.8099365	.0506072	16.00	0.000	.7107481	.9091249
_cons	-6.037825	.5052075	-11.95	0.000	-7.028013	-5.047636
/lnalpha	.6304507	.0345005			.5628309	.6980705
alpha	1.878457	.0648078			1.755635	2.009871

Likelihood-ratio test of alpha=0: chibar2(01) = 1.1e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 2836
 LR chi2(7) = 514.51
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0363

Log likelihood = -6828.4762

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3826168	.0309435	12.37	0.000	.3219688	.4432649
logaadt	.8315057	.0489887	16.97	0.000	.7354897	.9275218
h55access	.3014135	.0704868	4.28	0.000	.1632618	.4395652
h57medw	-.0036742	.0015561	-2.36	0.018	-.0067241	-.0006244
celt	.1703366	.0828312	2.06	0.040	.0079904	.3326828
st_improved	-.1584882	.0656838	-2.41	0.016	-.2872262	-.0297502
sig	.4808626	.0562473	8.55	0.000	.3706199	.5911052
_cons	-7.123673	.5406811	-13.18	0.000	-8.183389	-6.063958
/lnalpha	.5362479	.035442			.4667829	.605713
alpha	1.70958	.0605909			1.594855	1.832558

Likelihood-ratio test of alpha=0: chibar2(01) = 1.0e+04 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

```

Negative binomial regression      Number of obs   =      2836
                                LR chi2(17)          =      572.45
                                Prob > chi2           =      0.0000
Log likelihood = -9955.6432      Pseudo R2        =      0.0279

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3759811	.0301293	12.48	0.000	.3169286	.4350335
logaadt	.798726	.0483853	16.51	0.000	.7038925	.8935595
h34tlanes	-.0197359	.049552	-0.40	0.690	-.116856	.0773843
h54lanew	.007625	.07027	0.11	0.914	-.1301016	.1453517
h55access	.3859812	.0692417	5.57	0.000	.25027	.5216924
h57medw	-.0031167	.0018082	-1.72	0.085	-.0066606	.0004272
h59rshldw	.0295945	.0168318	1.76	0.079	-.0033952	.0625843
h60lshldw	-.0056941	.0173485	-0.33	0.743	-.0396965	.0283083
adcurve	.0025863	.0143303	0.18	0.857	-.0255005	.030673
agrade	-.0663624	.0215435	-3.08	0.002	-.1085868	-.024138
crb	.0221121	.0944239	0.23	0.815	-.1629554	.2071796
int_l	-.0412066	.1154116	-0.36	0.721	-.2674091	.184996
int_r	.0025667	.225965	0.01	0.991	-.4403166	.4454501
int_lr	.1136841	.1978968	0.57	0.566	-.2741866	.5015547
celt	.1618352	.1301508	1.24	0.214	-.0932557	.4169261
st_improved	-.5661498	.1467366	-3.86	0.000	-.8537483	-.2785512
sig	.4752833	.0550457	8.63	0.000	.3673958	.5831709
_cons	-5.663051	.9499193	-5.96	0.000	-7.524858	-3.801243
/lnalpha	.5722595	.0284255			.5165466	.6279725
alpha	1.772267	.0503776			1.676229	1.873808

Likelihood-ratio test of alpha=0: chibar2(01) = 4.6e+04 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression      Number of obs   =      2836
                                LR chi2(2)          =      307.83
                                Prob > chi2           =      0.0000
Log likelihood = -10087.951      Pseudo R2        =      0.0150

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.305321	.0315777	9.67	0.000	.2434299	.3672122
logaadt	.7850194	.0488086	16.08	0.000	.6893563	.8806825
_cons	-4.567765	.4859023	-9.40	0.000	-5.520116	-3.615414
/lnalpha	.6712328	.0277854			.6167743	.7256912
alpha	1.956648	.0543663			1.852941	2.066159

Likelihood-ratio test of alpha=0: chibar2(01) = 5.2e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 2836

LR chi2(7) = 559.09

Prob > chi2 = 0.0000

Pseudo R2 = 0.0273

Log likelihood = -9962.321

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3768881	.0299425	12.59	0.000	.3182018	.4355743
logaad	.7893174	.0457253	17.26	0.000	.6996974	.8789373
h55access	.3693784	.0681878	5.42	0.000	.2357327	.5030241
h57medw	-.0022605	.001541	-1.47	0.142	-.0052807	.0007598
celt	.1201212	.0807769	1.49	0.137	-.0381987	.2784411
st_improved	-.3397092	.0642512	-5.29	0.000	-.4656392	-.2137792
sig	.4833995	.0541895	8.92	0.000	.3771901	.589609
_cons	-5.585711	.4984116	-11.21	0.000	-6.56258	-4.608842
/lnalpha	.5772635	.0283931			.5216139	.632913
alpha	1.781158	.0505727			1.684744	1.883088

Likelihood-ratio test of alpha=0: chibar2(01) = 4.6e+04 Prob>=chibar2 = 0.000

All Crashes

Full Model

```

Negative binomial regression
Log likelihood = -10631.769
Number of obs   =      2836
LR chi2(17)    =      577.87
Prob > chi2     =      0.0000
Pseudo R2      =      0.0265

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.374312	.0295953	12.65	0.000	.3163063	.4323177
logaadt	.8078003	.0477557	16.92	0.000	.7142008	.9013997
h34tlanes	-.0221463	.049614	-0.45	0.655	-.1193879	.0750953
h54lanew	-.0179894	.070292	-0.26	0.798	-.1557592	.1197803
h55access	.3622922	.0688883	5.26	0.000	.2272736	.4973108
h57medw	-.0034581	.0017838	-1.94	0.053	-.0069542	.000038
h59rshldw	.0291022	.0167019	1.74	0.081	-.0036328	.0618373
h60lshldw	-.0041961	.017053	-0.25	0.806	-.0376194	.0292273
adcurve	-.0016266	.0142465	-0.11	0.909	-.0295492	.0262961
agrade	-.0629758	.0214297	-2.94	0.003	-.1049772	-.0209744
crb	.0004276	.094123	0.00	0.996	-.18405	.1849052
int_l	-.0164647	.1149035	-0.14	0.886	-.2416715	.2087421
int_r	.0350828	.2247442	0.16	0.876	-.4054078	.4755734
int_lr	.1148717	.1972582	0.58	0.560	-.2717473	.5014907
celt	.1516611	.1297254	1.17	0.242	-.102596	.4059181
st_improved	-.5347741	.1459418	-3.66	0.000	-.8208148	-.2487334
sig	.4773932	.0546962	8.73	0.000	.3701907	.5845957
_cons	-5.134075	.9451715	-5.43	0.000	-6.986577	-3.281573
/lnalpha	.5719184	.0276632			.5176995	.6261373
alpha	1.771663	.0490099			1.678163	1.870372

Likelihood-ratio test of alpha=0: chibar2(01) = 5.8e+04 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression
Log likelihood = -10758.902
Number of obs   =      2836
LR chi2(2)     =      323.61
Prob > chi2     =      0.0000
Pseudo R2      =      0.0148

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3109903	.0309785	10.04	0.000	.2502735	.3717071
logaadt	.7891551	.0479935	16.44	0.000	.6950897	.8832206
_cons	-4.349303	.4780646	-9.10	0.000	-5.286293	-3.412314
/lnalpha	.6631818	.0271126			.610042	.7163216
alpha	1.940958	.0526245			1.840509	2.04689

Likelihood-ratio test of alpha=0: chibar2(01) = 6.7e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 2836

LR chi2(7) = 565.45

Prob > chi2 = 0.0000

Pseudo R2 = 0.0259

Log likelihood = -10637.98

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.3764297	.0294097	12.80	0.000	.3187879	.4340716
logaad	.7968965	.0450785	17.68	0.000	.7085442	.8852487
h55access	.3476778	.0678147	5.13	0.000	.2147634	.4805922
h57medw	-.0026542	.0015247	-1.74	0.082	-.0056426	.0003342
celt	.1328172	.0804761	1.65	0.099	-.024913	.2905474
st_improved	-.3009946	.0636949	-4.73	0.000	-.4258343	-.176155
sig	.4858323	.0538566	9.02	0.000	.3802754	.5913893
_cons	-5.363138	.4931617	-10.88	0.000	-6.329717	-4.396559
/lnalpha	.5763449	.0276376			.5221761	.6305137
alpha	1.779522	.0491818			1.685692	1.878575

Likelihood-ratio test of alpha=0: chibar2(01) = 5.9e+04 Prob>=chibar2 = 0.000

Segment type =6 Urban Interstate

Injury and Fatal Crashes

Full Model

Negative binomial regression

Number of obs = 473

LR chi2(8) = 298.71

Prob > chi2 = 0.0000

Pseudo R2 = 0.1057

Log likelihood = -1264.1367

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.7718566	.051563	14.97	0.000	.670795	.8729182
logaad	1.302133	.1518686	8.57	0.000	1.004476	1.599789
h34tlanes	.035862	.060966	0.59	0.556	-.0836292	.1553532
h57medw	-.0031613	.003048	-1.04	0.300	-.0091352	.0028127
h59rshldw	.0193902	.1597767	0.12	0.903	-.2937663	.3325468
h60lshldw	-.0059509	.0193364	-0.31	0.758	-.0438496	.0319478
adcurve	.0050799	.0568016	0.09	0.929	-.1062493	.116409
agrade	-.1691864	.058946	-2.87	0.004	-.2847184	-.0536543
_cons	-12.13298	2.483945	-4.88	0.000	-17.00142	-7.264537
/lnalpha	-.1048193	.0900277			-.2812703	.0716317
alpha	.9004872	.0810688			.7548243	1.07426

Likelihood-ratio test of alpha=0: chibar2(01) = 1494.22 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs = 473

LR chi2(2) = 289.32

Prob > chi2 = 0.0000

Pseudo R2 = 0.1023

Log likelihood = -1268.8319

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.7674191	.0508452	15.09	0.000	.6677642	.8670739
logaad	1.396418	.0916918	15.23	0.000	1.216705	1.576131
_cons	-13.12111	.9979538	-13.15	0.000	-15.07706	-11.16515
/lnalpha	-.081577	.0897542			-.257492	.094338
alpha	.9216618	.082723			.7729878	1.098931

Likelihood-ratio test of alpha=0: chibar2(01) = 1524.85 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 473
 LR chi2(2) = 289.32
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.1023

Log likelihood = -1268.8319

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.7674191	.0508452	15.09	0.000	.6677642	.8670739
logaadt	1.396418	.0916918	15.23	0.000	1.216705	1.576131
_cons	-13.12111	.9979538	-13.15	0.000	-15.07706	-11.16515
/lnalpha	-.081577	.0897542			-.257492	.094338
alpha	.9216618	.082723			.7729878	1.098931

Likelihood-ratio test of alpha=0: chibar2(01) = 1524.85 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

Negative binomial regression

Number of obs = 473

LR chi2(8) = 346.06

Prob > chi2 = 0.0000

Log likelihood = -1878.1302

Pseudo R2 = 0.0844

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.8494923	.0488595	17.39	0.000	.7537294	.9452551
logaad	1.296366	.1451133	8.93	0.000	1.011949	1.580783
h34tlanes	-.0044908	.0590541	-0.08	0.939	-.1202347	.1112531
h57medw	-.0021478	.0029523	-0.73	0.467	-.0079343	.0036387
h59rshldw	-.102571	.1468767	-0.70	0.485	-.390444	.185302
h60lshldw	.0335912	.0192534	1.74	0.081	-.0041447	.0713271
adcurve	-.0131762	.056469	-0.23	0.816	-.1238535	.0975011
agrade	-.1607849	.0558033	-2.88	0.004	-.2701573	-.0514124
_cons	-9.437773	2.318171	-4.07	0.000	-13.9813	-4.894241
/lnalpha	.0252868	.072341			-.1164989	.1670725
alpha	1.025609	.0741936			.8900311	1.18184

Likelihood-ratio test of alpha=0: chibar2(01) = 8065.85 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs = 473

LR chi2(2) = 334.09

Prob > chi2 = 0.0000

Log likelihood = -1884.1133

Pseudo R2 = 0.0814

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.8714416	.0482459	18.06	0.000	.7768813	.9660019
logaad	1.392564	.085276	16.33	0.000	1.225426	1.559702
_cons	-11.60195	.9240401	-12.56	0.000	-13.41303	-9.790861
/lnalpha	.0560956	.0717227			-.0844782	.1966694
alpha	1.057699	.075861			.9189917	1.217342

Likelihood-ratio test of alpha=0: chibar2(01) = 8562.52 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 473
 LR chi2(2) = 334.09
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0814

Log likelihood = -1884.1133

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.8714416	.0482459	18.06	0.000	.7768813	.9660019
logaadt	1.392564	.085276	16.33	0.000	1.225426	1.559702
_cons	-11.60195	.9240401	-12.56	0.000	-13.41303	-9.790861
/lnalpha	.0560956	.0717227			-.0844782	.1966694
alpha	1.057699	.075861			.9189917	1.217342

Likelihood-ratio test of alpha=0: chibar2(01) = 8562.52 Prob>=chibar2 = 0.000

All Crashes

Full Model

```

Negative binomial regression                                Number of obs   =          473
                                                           LR chi2(8)      =          341.17
                                                           Prob > chi2     =          0.0000
Log likelihood = -1976.7419                                Pseudo R2       =          0.0794

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.8418522	.0487101	17.28	0.000	.7463822	.9373221
logaadt	1.291005	.1446881	8.92	0.000	1.007422	1.574589
h34tlanes	.0092827	.0598703	0.16	0.877	-.1080609	.1266262
h57medw	-.0023113	.0029489	-0.78	0.433	-.008091	.0034685
h59rshldw	-.0688157	.1491611	-0.46	0.645	-.3611661	.2235347
h60lshldw	.0251077	.0194158	1.29	0.196	-.0129466	.0631621
adcurve	-.0089868	.0563094	-0.16	0.873	-.1193511	.1013776
agrade	-.1639105	.0557062	-2.94	0.003	-.2730926	-.0547283
_cons	-9.522892	2.345003	-4.06	0.000	-14.11901	-4.92677
/lnalpha	.0479917	.070394			-.0899779	.1859614
alpha	1.049162	.0738547			.9139514	1.204376

Likelihood-ratio test of alpha=0: chibar2(01) = 1.0e+04 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression                                Number of obs   =          473
                                                           LR chi2(2)      =          330.37
                                                           Prob > chi2     =          0.0000
Log likelihood = -1982.1401                                Pseudo R2       =          0.0769

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.8582897	.0480111	17.88	0.000	.7641896	.9523897
logaadt	1.397176	.0852215	16.39	0.000	1.230145	1.564207
_cons	-11.44343	.9228982	-12.40	0.000	-13.25228	-9.634585
/lnalpha	.074267	.0699147			-.0627632	.2112972
alpha	1.077094	.0753047			.9391658	1.235279

Likelihood-ratio test of alpha=0: chibar2(01) = 1.1e+04 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 473
 LR chi2(2) = 330.37
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0769

Log likelihood = -1982.1401

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
loglength	.8582897	.0480111	17.88	0.000	.7641896	.9523897
logaadt	1.397176	.0852215	16.39	0.000	1.230145	1.564207
_cons	-11.44343	.9228982	-12.40	0.000	-13.25228	-9.634585
/lnalpha	.074267	.0699147			-.0627632	.2112972
alpha	1.077094	.0753047			.9391658	1.235279

Likelihood-ratio test of alpha=0: chibar2(01) = 1.1e+04 Prob>=chibar2 = 0.000

Intersection models

Summary of Data

Intersection type description

- 1.-4 TWSC
Leg3: 3-leg intersections
Maj2: 2-lane on major roads
- 5. AWSC
- 6. Signalized
- 7.-8. Signal: signalized

Intersection type =1-4 TWSC

Injury and Fatal Crashes

Full Model

Negative binomial regression

Number of obs	=	444
LR chi2(10)	=	82.19
Prob > chi2	=	0.0000
Pseudo R2	=	0.0506

Log likelihood = -770.35598

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4423814	.1513019	2.92	0.003	.1458351	.7389277
logh2aadt	.2532037	.1313878	1.93	0.054	-.0043117	.5107191
h3urbrur	-.1034621	.2255357	-0.46	0.646	-.545504	.3385799
h4should	.0768806	.1938787	0.40	0.692	-.3031146	.4568759
h5rtmaj	1.526586	1.264705	1.21	0.227	-.9521897	4.005361
h6ltmaj	-1.443204	.7289724	-1.98	0.048	-2.871963	-.0144442
h7lane	-.0693262	.1217254	-0.57	0.569	-.3079036	.1692513
h9medwdth1	-.0045784	.0060083	-0.76	0.446	-.0163545	.0071977
leg3	-.5944251	.2312945	-2.57	0.010	-1.047754	-.1410963
maj2	-.323762	.4202247	-0.77	0.441	-1.147387	.4998633
_cons	-4.127835	1.083811	-3.81	0.000	-6.252066	-2.003603
/lnalpha	.562882	.1146216			.3382277	.7875363
alpha	1.755725	.2012441			1.40246	2.197975

Likelihood-ratio test of alpha=0: chibar2(01) = 582.25 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs	=	444
LR chi2(2)	=	69.98
Prob > chi2	=	0.0000
Pseudo R2	=	0.0431

Log likelihood = -776.46169

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.2938811	.1247482	2.36	0.018	.0493792	.538383
logh2aadt	.3340705	.114457	2.92	0.004	.1097389	.5584021
_cons	-4.448522	.6973536	-6.38	0.000	-5.81531	-3.081734
/lnalpha	.6252098	.1118783			.4059324	.8444872
alpha	1.868638	.20906			1.500701	2.326784

Likelihood-ratio test of alpha=0: chibar2(01) = 648.54 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 444
 LR chi2(3) = 77.95
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0480

Log likelihood = -772.47614

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4110659	.1303294	3.15	0.002	.155625	.6665068
logh2aadt	.2136236	.1209122	1.77	0.077	-.02336	.4506072
leg3	-.5476222	.1973582	-2.77	0.006	-.9344371	-.1608073
_cons	-4.179837	.6864259	-6.09	0.000	-5.525207	-2.834467
/lnalpha	.5827797	.1138644			.3596095	.8059498
alpha	1.79101	.2039323			1.43277	2.238822

Likelihood-ratio test of alpha=0: chibar2(01) = 602.75 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

Negative binomial regression

Number of obs	=	444
LR chi2(10)	=	161.03
Prob > chi2	=	0.0000
Pseudo R2	=	0.0621

Log likelihood = -1215.7451

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5457868	.1163602	4.69	0.000	.317725	.7738485
logh2aadt	.1577581	.1001464	1.58	0.115	-.0385252	.3540414
h3urbrur	.2241401	.1755593	1.28	0.202	-.1199499	.56823
h4should	.204132	.1613838	1.26	0.206	-.1121744	.5204384
h5rtmaj	.4149689	.9827757	0.42	0.673	-1.511236	2.341174
h6ltmaj	-.1975487	.508395	-0.39	0.698	-1.193985	.7988873
h7lane	-.009348	.0940402	-0.10	0.921	-.1936633	.1749674
h9medwdth1	-.0081834	.0048231	-1.70	0.090	-.0176366	.0012697
leg3	-.3525171	.1838806	-1.92	0.055	-.7129164	.0078823
maj2	-.0090712	.3247527	-0.03	0.978	-.6455749	.6274325
_cons	-3.91842	.8388204	-4.67	0.000	-5.562478	-2.274363
/lnalpha	.2469001	.0830675			.0840907	.4097095
alpha	1.280051	.1063307			1.087728	1.50638

Likelihood-ratio test of alpha=0: chibar2(01) = 2282.24 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs	=	444
LR chi2(2)	=	147.64
Prob > chi2	=	0.0000
Pseudo R2	=	0.0569

Log likelihood = -1222.4391

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4420964	.0962409	4.59	0.000	.2534676	.6307252
logh2aadt	.2654395	.0864916	3.07	0.002	.095919	.4349599
_cons	-4.084484	.5244966	-7.79	0.000	-5.112478	-3.056489
/lnalpha	.2897659	.0817437			.1295512	.4499806
alpha	1.336115	.1092189			1.138317	1.568282

Likelihood-ratio test of alpha=0: chibar2(01) = 2567.64 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 444
 LR chi2(4) = 157.61
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0608

Log likelihood = -1217.4569

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.6022592	.1083896	5.56	0.000	.3898196	.8146988
logh2aadt	.1922771	.0926462	2.08	0.038	.0106938	.3738603
h9medwdth1	-.0074153	.0026335	-2.82	0.005	-.0125769	-.0022537
leg3	-.3084182	.1624267	-1.90	0.058	-.6267686	.0099323
_cons	-4.617199	.5743193	-8.04	0.000	-5.742844	-3.491554
/lnalpha	.2566546	.0828446			.0942822	.4190271
alpha	1.292599	.1070848			1.09887	1.520482

Likelihood-ratio test of alpha=0: chibar2(01) = 2319.05 Prob>=chibar2 = 0.000

All Crashes

Full Model

```

Negative binomial regression
Log likelihood = -1324.3654
Number of obs   =      444
LR chi2(10)     =     149.65
Prob > chi2     =     0.0000
Pseudo R2      =     0.0535

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5131803	.1159161	4.43	0.000	.285989	.7403717
logh2aadt	.1857275	.1013881	1.83	0.067	-.0129895	.3844445
h3urbrur	.1637665	.1769823	0.93	0.355	-.1831124	.5106455
h4should	.1701395	.1600415	1.06	0.288	-.1435361	.483815
h5rtmaj	.5348025	.9994217	0.54	0.593	-1.424028	2.493633
h6ltmaj	-.3343231	.5191005	-0.64	0.520	-1.351741	.6830953
h7lane	-.0270808	.0958046	-0.28	0.777	-.2148543	.1606927
h9medwdth1	-.0069235	.0049041	-1.41	0.158	-.0165355	.0026884
leg3	-.4138757	.1848687	-2.24	0.025	-.7762116	-.0515398
maj2	-.0865807	.3314367	-0.26	0.794	-.7361846	.5630232
_cons	-3.372182	.844654	-3.99	0.000	-5.027673	-1.71669
/lnalpha	.2949939	.0788156			.1405182	.4494696
alpha	1.343118	.1058586			1.15087	1.567481

Likelihood-ratio test of alpha=0: chibar2(01) = 3124.90 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression
Log likelihood = -1329.6847
Number of obs   =      444
LR chi2(2)      =     139.01
Prob > chi2     =     0.0000
Pseudo R2      =     0.0497

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4048608	.0961051	4.21	0.000	.2164982	.5932234
logh2aadt	.2889522	.0873239	3.31	0.001	.1178005	.4601039
_cons	-3.657072	.5296076	-6.91	0.000	-4.695084	-2.619061
/lnalpha	.3264258	.0779628			.1736215	.47923
alpha	1.386005	.1080569			1.189605	1.614831

Likelihood-ratio test of alpha=0: chibar2(01) = 3438.26 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 444
 LR chi2(4) = 147.37
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0527

Log likelihood = -1325.5058

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.551668	.1082847	5.09	0.000	.3394338	.7639021
logh2aadt	.2059767	.0937491	2.20	0.028	.0222318	.3897215
h9medwdth1	-.0055775	.0026329	-2.12	0.034	-.0107378	-.0004172
leg3	-.3651507	.1623335	-2.25	0.024	-.6833186	-.0469828
_cons	-3.979133	.577967	-6.88	0.000	-5.111928	-2.846339
/lnalpha	.3011393	.0786859			.1469177	.4553608
alpha	1.351398	.1063359			1.158259	1.576742

Likelihood-ratio test of alpha=0: chibar2(01) = 3162.07 Prob>=chibar2 = 0.000

Intersection type=5   AWSC

Injury and Fatal Crashes

Full Model

Negative binomial regression

Log likelihood = -279.24748

Number of obs = 172
 LR chi2(7) = 29.08
 Prob > chi2 = 0.0001
 Pseudo R2 = 0.0495

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4726311	.301648	1.57	0.117	-.118588	1.06385
logh2aadt	.3449075	.2155623	1.60	0.110	-.0775867	.7674018
h3urbrur	.8948626	.516739	1.73	0.083	-.1179273	1.907653
h4should	-.5387379	.3097844	-1.74	0.082	-1.145904	.0684283
h6ltmaj	-1.148277	1.03593	-1.11	0.268	-3.178661	.8821079
h7lane	-.2096285	.2495087	-0.84	0.401	-.6986565	.2793995
h9medwdth1	-.0502765	.0556554	-0.90	0.366	-.159359	.058806
_cons	-5.272142	1.961996	-2.69	0.007	-9.117584	-1.426701
/lnalpha	.0983684	.2202796			-.3333717	.5301084
alpha	1.103369	.2430497			.7165038	1.699116

Likelihood-ratio test of alpha=0:   chibar2(01) =   87.19   Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Log likelihood = -283.84954

Number of obs = 172
 LR chi2(2) = 19.88
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0338

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5252702	.2768828	1.90	0.058	-.0174101	1.067951
logh2aadt	.2850114	.1739492	1.64	0.101	-.0559227	.6259455
_cons	-6.056456	1.706246	-3.55	0.000	-9.400635	-2.712276
/lnalpha	.2090014	.2111818			-.2049074	.6229102
alpha	1.232447	.2602703			.8147228	1.864346

Likelihood-ratio test of alpha=0:   chibar2(01) =   104.89   Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 172
 LR chi2(5) = 27.39
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0466

Log likelihood = -280.09148

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5736829	.277044	2.07	0.038	.0306866	1.116679
logh2aadt	.2358315	.1718371	1.37	0.170	-.100963	.5726261
h3urbrur	.7442036	.4573103	1.63	0.104	-.1521081	1.640515
h4should	-.5322093	.3010366	-1.77	0.077	-1.12223	.0578115
h9medwdth1	-.0521012	.0512253	-1.02	0.309	-.1525009	.0482985
_cons	-6.048931	1.734574	-3.49	0.000	-9.448633	-2.649228
/lnalpha	.1308593	.2161371			-.2927617	.5544802
alpha	1.139807	.2463547			.7461999	1.741036

Likelihood-ratio test of alpha=0: chibar2(01) = 94.74 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

```

Negative binomial regression      Number of obs   =      172
                                LR chi2(7)           =      44.45
                                Prob > chi2           =      0.0000
Log likelihood = -460.15925      Pseudo R2        =      0.0461

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4631441	.2187249	2.12	0.034	.0344511	.8918371
logh2aadt	.4186691	.1709983	2.45	0.014	.0835185	.7538197
h3urbrur	.0755359	.439158	0.17	0.863	-.7851981	.9362698
h4should	.0363209	.224283	0.16	0.871	-.4032657	.4759075
h6ltmaj	.3543048	.8247642	0.43	0.667	-1.262203	1.970813
h7lane	-.1703344	.1984509	-0.86	0.391	-.559291	.2186222
h9medwdth1	-.0539142	.0323085	-1.67	0.095	-.1172377	.0094093
_cons	-4.663483	1.374412	-3.39	0.001	-7.35728	-1.969685
/lnalpha	-.1019952	.1472066			-.3905148	.1865244
alpha	.9030339	.1329325			.6767084	1.205054

Likelihood-ratio test of alpha=0: chibar2(01) = 385.32 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression      Number of obs   =      172
                                LR chi2(2)           =      37.17
                                Prob > chi2           =      0.0000
Log likelihood = -463.79914      Pseudo R2        =      0.0385

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5122477	.2034617	2.52	0.012	.1134701	.9110254
logh2aadt	.3501771	.1389564	2.52	0.012	.0778276	.6225266
_cons	-5.168209	1.217788	-4.24	0.000	-7.555031	-2.781388
/lnalpha	-.0566645	.1456678			-.342168	.2288391
alpha	.9449111	.1376431			.7102288	1.25714

Likelihood-ratio test of alpha=0: chibar2(01) = 403.05 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 172
 LR chi2(3) = 42.97
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0445

Log likelihood = -460.89895

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5572414	.2017429	2.76	0.006	.1618325	.9526503
logh2aadt	.3299104	.1377816	2.39	0.017	.0598634	.5999574
h9medwdth1	-.0555329	.0295954	-1.88	0.061	-.1135388	.0024729
_cons	-5.374577	1.20863	-4.45	0.000	-7.743449	-3.005705
/lnalpha	-.089241	.1464992			-.3763742	.1978922
alpha	.9146252	.1339919			.6863455	1.218831

Likelihood-ratio test of alpha=0: chibar2(01) = 392.64 Prob>=chibar2 = 0.000

All Crashes

Full Model

Negative binomial regression

Number of obs = 172
 LR chi2(7) = 44.41
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0426

Log likelihood = -499.26244

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4557483	.2182363	2.09	0.037	.0280129	.8834837
logh2aadt	.3981405	.1677066	2.37	0.018	.0694417	.7268393
h3urbrur	.3050749	.4385538	0.70	0.487	-.5544747	1.164624
h4should	-.0863839	.2266651	-0.38	0.703	-.5306394	.3578716
h6ltmaj	.008037	.8358691	0.01	0.992	-1.630236	1.64631
h7lane	-.1768864	.1967716	-0.90	0.369	-.5625516	.2087789
h9medwdth1	-.0595612	.0344684	-1.73	0.084	-.1271181	.0079957
_cons	-4.163229	1.373244	-3.03	0.002	-6.854739	-1.47172
/lnalpha	-.0559583	.1375819			-.3256139	.2136973
alpha	.9455786	.1300945			.7220839	1.238248

Likelihood-ratio test of alpha=0: chibar2(01) = 543.30 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs = 172
 LR chi2(2) = 36.02
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0345

Log likelihood = -503.45709

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5084206	.2027074	2.51	0.012	.1111214	.9057197
logh2aadt	.327662	.1360887	2.41	0.016	.060933	.5943911
_cons	-4.719761	1.201158	-3.93	0.000	-7.073987	-2.365534
/lnalpha	-.0076512	.1362505			-.2746972	.2593949
alpha	.992378	.135212			.7598022	1.296145

Likelihood-ratio test of alpha=0: chibar2(01) = 570.24 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 172
 LR chi2(3) = 42.50
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0407

Log likelihood = -500.21944

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5550991	.2005056	2.77	0.006	.1621153	.9480828
logh2aadt	.3072953	.1348243	2.28	0.023	.0430446	.571546
h9medwdth1	-.0597076	.029994	-1.99	0.047	-.1184948	-.0009204
_cons	-4.938678	1.189145	-4.15	0.000	-7.26936	-2.607996
/lnalpha	-.0418936	.1369786			-.3103667	.2265795
alpha	.9589718	.1313586			.7331781	1.254302

Likelihood-ratio test of alpha=0: chibar2(01) = 555.46 Prob>=chibar2 = 0.000

Intersection type=6   Signalized

Injury and Fatal Crashes

Full Model

Negative binomial regression

Log likelihood = -1046.6165

Number of obs = 375
 LR chi2(8) = 99.83
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0455

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.3962037	.1337775	2.96	0.003	.1340046	.6584028
logh2aadt	.2979757	.081054	3.68	0.000	.1391128	.4568385
h3urbrur	.1291639	.1416962	0.91	0.362	-.1485555	.4068833
h4should	-.1216985	.1192352	-1.02	0.307	-.3553951	.1119982
h5rtmaj	.582573	.3596773	1.62	0.105	-.1223815	1.287528
h6ltmaj	.1453091	.1507302	0.96	0.335	-.1501167	.4407349
h7lane	.0518715	.049848	1.04	0.298	-.0458289	.1495718
h9medwdth1	.0000256	.0029134	0.01	0.993	-.0056846	.0057358
_cons	-4.929078	1.078272	-4.57	0.000	-7.042452	-2.815704
/lnalpha	-.1056394	.0941903			-.290249	.0789702
alpha	.899749	.0847476			.7480773	1.082172

Likelihood-ratio test of alpha=0:   chibar2(01) = 1105.96   Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Log likelihood = -1050.6074

Number of obs = 375
 LR chi2(2) = 91.84
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0419

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4806187	.1123663	4.28	0.000	.2603848	.7008527
logh2aadt	.3334419	.0731493	4.56	0.000	.190072	.4768118
_cons	-5.719572	.8716568	-6.56	0.000	-7.427988	-4.011157
/lnalpha	-.0773671	.0933252			-.2602811	.1055468
alpha	.92555	.0863771			.7708349	1.111318

Likelihood-ratio test of alpha=0:   chibar2(01) = 1167.25   Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 375
 LR chi2(3) = 95.55
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0436

Log likelihood = -1048.7565

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.451646	.1125018	4.01	0.000	.2311464	.6721456
logh2aadt	.3449878	.0727826	4.74	0.000	.2023365	.4876392
h5rtmaj	.6400565	.3596997	1.78	0.075	-.064942	1.345055
_cons	-5.557619	.8680493	-6.40	0.000	-7.258965	-3.856274
/lnalpha	-.0923581	.093917			-.2764322	.0917159
alpha	.9117785	.0856315			.7584851	1.096053

Likelihood-ratio test of alpha=0: chibar2(01) = 1123.26 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

```

Negative binomial regression                                Number of obs   =          375
                                                           LR chi2(8)      =          112.07
                                                           Prob > chi2     =           0.0000
Log likelihood = -1538.8347                               Pseudo R2       =           0.0351

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5469754	.1318615	4.15	0.000	.2885317	.8054192
logh2aadt	.3460279	.0758316	4.56	0.000	.1974006	.4946551
h3urbrur	.1963396	.1384681	1.42	0.156	-.075053	.4677322
h4should	-.1424871	.1180777	-1.21	0.228	-.3739151	.0889409
h5rtmaj	.3953299	.3736774	1.06	0.290	-.3370643	1.127724
h6ltmaj	.2797678	.1515269	1.85	0.065	-.0172194	.576755
h7lane	-.0385624	.0491271	-0.78	0.432	-.1348497	.0577248
h9medwdth1	-.0005156	.0029091	-0.18	0.859	-.0062173	.005186
_cons	-4.992405	1.064938	-4.69	0.000	-7.079646	-2.905165
/lnalpha	.0354043	.0770225			-.115557	.1863655
alpha	1.036038	.0797982			.8908698	1.204863

Likelihood-ratio test of alpha=0: chibar2(01) = 5844.51 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression                                Number of obs   =          375
                                                           LR chi2(2)      =           98.34
                                                           Prob > chi2     =           0.0000
Log likelihood = -1545.6963                               Pseudo R2       =           0.0308

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5101309	.1122821	4.54	0.000	.2900621	.7301997
logh2aadt	.3297667	.0687014	4.80	0.000	.1951144	.464419
_cons	-4.580393	.873943	-5.24	0.000	-6.29329	-2.867496
/lnalpha	.0734726	.0763524			-.0761753	.2231205
alpha	1.076239	.0821734			.9266538	1.249971

Likelihood-ratio test of alpha=0: chibar2(01) = 6253.19 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 375
 LR chi2(4) = 108.33
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0340

Log likelihood = -1540.7027

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4739056	.1155551	4.10	0.000	.2474218	.7003894
logh2aadt	.3119422	.0687118	4.54	0.000	.1772696	.4466147
h3urbrur	.2141565	.1311715	1.63	0.103	-.0429349	.4712478
h6ltmaj	.2889628	.1515881	1.91	0.057	-.0081444	.5860701
_cons	-4.288103	.8914171	-4.81	0.000	-6.035249	-2.540958
/lnalpha	.0456098	.0768541			-.1050214	.1962411
alpha	1.046666	.0804406			.9003053	1.21682

Likelihood-ratio test of alpha=0: chibar2(01) = 5983.04 Prob>=chibar2 = 0.000

All Crashes

Full Model

```

Negative binomial regression                                Number of obs   =          375
                                                           LR chi2(8)      =        110.63
                                                           Prob > chi2     =         0.0000
Log likelihood = -1621.4968                               Pseudo R2       =         0.0330

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5170271	.1303467	3.97	0.000	.2615522	.772502
logh2aadt	.3300044	.0747056	4.42	0.000	.183584	.4764248
h3urbrur	.1816818	.1377711	1.32	0.187	-.0883446	.4517082
h4should	-.1398106	.1169177	-1.20	0.232	-.368965	.0893438
h5rtmaj	.4216781	.3710819	1.14	0.256	-.3056291	1.148985
h6ltmaj	.258414	.1505183	1.72	0.086	-.0365965	.5534245
h7lane	-.0180231	.0488173	-0.37	0.712	-.1137033	.0776571
h9medwdth1	-.0003268	.0028903	-0.11	0.910	-.0059917	.0053381
_cons	-4.437866	1.051386	-4.22	0.000	-6.498544	-2.377188
/lnalpha	.0288958	.0752391			-.1185702	.1763617
alpha	1.029317	.0774449			.8881895	1.192869

Likelihood-ratio test of alpha=0: chibar2(01) = 7315.89 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression                                Number of obs   =          375
                                                           LR chi2(2)      =         99.12
                                                           Prob > chi2     =         0.0000
Log likelihood = -1627.2506                               Pseudo R2       =         0.0296

```

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.5096687	.1104111	4.62	0.000	.293267	.7260705
logh2aadt	.3264915	.0674429	4.84	0.000	.1943058	.4586771
_cons	-4.326168	.8636731	-5.01	0.000	-6.018937	-2.6334
/lnalpha	.0598092	.0747287			-.0866563	.2062748
alpha	1.061634	.0793345			.9169922	1.229091

Likelihood-ratio test of alpha=0: chibar2(01) = 7767.97 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 375
 LR chi2(4) = 107.45
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0320

Log likelihood = -1623.0858

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4747764	.1140717	4.16	0.000	.2512001	.6983528
logh2aadt	.3106225	.0676072	4.59	0.000	.1781149	.4431301
h3urbrur	.1904052	.130356	1.46	0.144	-.0650879	.4458982
h6ltmaj	.26579	.1504013	1.77	0.077	-.0289911	.560571
_cons	-4.040787	.886076	-4.56	0.000	-5.777464	-2.30411
/lnalpha	.0373452	.0751065			-.1098608	.1845513
alpha	1.038051	.0779644			.8959588	1.202679

Likelihood-ratio test of alpha=0: chibar2(01) = 7473.43 Prob>=chibar2 = 0.000

Intersection type =7-8 Interstate

Injury and Fatal Crashes

Full Model

Negative binomial regression

Number of obs	=	93
LR chi2(7)	=	66.79
Prob > chi2	=	0.0000
Pseudo R2	=	0.1181

Log likelihood = -249.37704

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.6653626	.2195438	3.03	0.002	.2350647	1.09566
logh2aadt	.5248106	.1664924	3.15	0.002	.1984915	.8511296
h3urbrur	.3635077	.2218997	1.64	0.101	-.0714077	.7984231
h4should	-1.269954	.580472	-2.19	0.029	-2.407659	-.1322502
h7lane	-.1289906	.0829408	-1.56	0.120	-.2915516	.0335704
h9medwdth1	.0070512	.0062221	1.13	0.257	-.0051439	.0192463
signal	.824961	.2820482	2.92	0.003	.2721566	1.377765
_cons	-8.418964	2.076043	-4.06	0.000	-12.48793	-4.349994
/lnalpha	-.8857918	.2216792			-1.320275	-.4513085
alpha	.4123875	.0914177			.2670618	.6367943

Likelihood-ratio test of alpha=0: chibar2(01) = 121.49 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs	=	93
LR chi2(2)	=	45.45
Prob > chi2	=	0.0000
Pseudo R2	=	0.0804

Log likelihood = -260.04274

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.3170194	.1967697	1.61	0.107	-.068642	.7026808
logh2aadt	.6542444	.13528	4.84	0.000	.3891005	.9193884
_cons	-7.476724	1.67047	-4.48	0.000	-10.75078	-4.202663
/lnalpha	-.5221781	.1908228			-.8961839	-.1481723
alpha	.593227	.1132012			.4081241	.8622825

Likelihood-ratio test of alpha=0: chibar2(01) = 266.36 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 93
 LR chi2(6) = 65.53
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.1159

Log likelihood = -250.0044

if3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.6346833	.2176803	2.92	0.004	.2080378	1.061329
logh2aadt	.4993272	.1650966	3.02	0.002	.1757439	.8229105
h3urbrur	.3295428	.2191702	1.50	0.133	-.100023	.7591085
h4should	-.8841181	.470939	-1.88	0.060	-1.807142	.0389054
h7lane	-.143995	.0823553	-1.75	0.080	-.3054085	.0174185
signal	.913997	.2703194	3.38	0.001	.3841808	1.443813
_cons	-7.702876	1.972601	-3.90	0.000	-11.5691	-3.83665
/lnalpha	-.8819487	.2227545			-1.31854	-.4453579
alpha	.4139754	.0922149			.2675257	.640595

Likelihood-ratio test of alpha=0: chibar2(01) = 120.54 Prob>=chibar2 = 0.000

PDO Crashes

Full Model

```

Negative binomial regression
Log likelihood = -365.78373
Number of obs   =      93
LR chi2(7)      =     65.34
Prob > chi2     =     0.0000
Pseudo R2      =     0.0820

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.6908934	.1930123	3.58	0.000	.3125964	1.06919
logh2aadt	.5545224	.1612838	3.44	0.001	.238412	.8706328
h3urbrur	.163051	.2241588	0.73	0.467	-.2762922	.6023942
h4should	-1.423715	.5861392	-2.43	0.015	-2.572527	-.274903
h7lane	-.048226	.0771789	-0.62	0.532	-.1994937	.1030418
h9medwdth1	.0138042	.0059734	2.31	0.021	.0020966	.0255119
signal	.37645	.3059463	1.23	0.219	-.2231937	.9760938
_cons	-8.369649	1.801602	-4.65	0.000	-11.90072	-4.838573
/lnalpha	-.5953381	.1603048			-.9095297	-.2811465
alpha	.5513761	.0883882			.4027136	.7549177

Likelihood-ratio test of alpha=0: chibar2(01) = 864.13 Prob>=chibar2 = 0.000

Minimum Model

```

Negative binomial regression
Log likelihood = -371.33442
Number of obs   =      93
LR chi2(2)      =     54.24
Prob > chi2     =     0.0000
Pseudo R2      =     0.0681

```

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.471963	.1680427	2.81	0.005	.1426053	.8013206
logh2aadt	.5536618	.1253719	4.42	0.000	.3079374	.7993861
_cons	-6.889425	1.366234	-5.04	0.000	-9.567195	-4.211655
/lnalpha	-.4708113	.1559292			-.776427	-.1651956
alpha	.6244954	.0973771			.4600468	.8477279

Likelihood-ratio test of alpha=0: chibar2(01) = 1139.99 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 93
 LR chi2(4) = 62.74
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0787

Log likelihood = -367.08221

pdo3	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.6589827	.1771431	3.72	0.000	.3117887	1.006177
logh2aadt	.6150622	.1360725	4.52	0.000	.3483651	.8817594
h4should	-1.520353	.5864639	-2.59	0.010	-2.669801	-.3709047
h9medwdth1	.0158712	.0055686	2.85	0.004	.004957	.0267854
_cons	-8.865674	1.539906	-5.76	0.000	-11.88384	-5.847513
/lnalpha	-.5607669	.15835			-.8711272	-.2504067
alpha	.5707712	.0903816			.4184796	.7784841

Likelihood-ratio test of alpha=0: chibar2(01) = 1032.17 Prob>=chibar2 = 0.000

All Crashes

Full Model

Negative binomial regression

Number of obs = 93
 LR chi2(7) = 68.68
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0816

Log likelihood = -386.38567

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.6843033	.1904803	3.59	0.000	.3109688	1.057638
logh2aadt	.5477278	.1552516	3.53	0.000	.2434402	.8520154
h3urbrur	.2106238	.2187948	0.96	0.336	-.2182062	.6394538
h4should	-1.432064	.5701308	-2.51	0.012	-2.5495	-.3146279
h7lane	-.0607158	.0757345	-0.80	0.423	-.2091527	.087721
h9medwdth1	.0130122	.0057608	2.26	0.024	.0017213	.0243032
signal	.4690075	.2989242	1.57	0.117	-.1168731	1.054888
_cons	-7.88031	1.769142	-4.45	0.000	-11.34776	-4.412856
/lnalpha	-.634327	.1562067			-.9404865	-.3281675
alpha	.5302923	.0828352			.3904378	.7202424

Likelihood-ratio test of alpha=0: chibar2(01) = 1076.45 Prob>=chibar2 = 0.000

Minimum Model

Negative binomial regression

Number of obs = 93
 LR chi2(2) = 55.20
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0656

Log likelihood = -393.12994

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.4461731	.1681521	2.65	0.008	.1166011	.7757452
logh2aadt	.5728572	.121395	4.72	0.000	.3349273	.810787
_cons	-6.550892	1.371489	-4.78	0.000	-9.238961	-3.862824
/lnalpha	-.4876427	.1515676			-.7847098	-.1905756
alpha	.6140722	.0930735			.4562521	.8264833

Likelihood-ratio test of alpha=0: chibar2(01) = 1498.68 Prob>=chibar2 = 0.000

Best Model

Negative binomial regression

Number of obs = 93
 LR chi2(5) = 67.32
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0800

Log likelihood = -387.06979

total	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logh1aadt	.6662422	.1723448	3.87	0.000	.3284526	1.004032
logh2aadt	.5311176	.1390937	3.82	0.000	.258499	.8037362
h4should	-1.492732	.567271	-2.63	0.009	-2.604562	-.3809008
h9medwdth1	.0126076	.0056043	2.25	0.024	.0016234	.0235918
signal	.5045686	.3003222	1.68	0.093	-.0840521	1.093189
_cons	-7.820434	1.546089	-5.06	0.000	-10.85071	-4.790156
/lnalpha	-.6176568	.155445			-.9223233	-.3129903
alpha	.5392064	.0838169			.3975942	.731257

Likelihood-ratio test of alpha=0: chibar2(01) = 1142.92 Prob>=chibar2 = 0.000