Analysis of Existing Data: Prospective Views on Methodological Paradigms

DETAILS

0 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-43521-5 | DOI 10.17226/22837

AUTHORS

Jovanis, Paul P.; Shankar, Venky; Aguero-Valverde, Jonathan; Wu ,Kun-Feng; and Greenstein, Adam

The Second
STRATEGIC HIGHWAY RESEARCH PROGRAM

SHRP 2 REPORT S2-S01B-RW-1

Analysis of Existing Data

Prospective Views on Methodological Paradigms

Paul P. Jovanis, Venky Shankar, Jonathan Aguero-Valverde, Kun-Feng Wu, and Adam Greenstein
The Thomas D. Larson Pennsylvania Transportation Institute
The Pennsylvania State University

TRANSPORTATION RESEARCH BOARD
WASHINGTON, D.C.
2012
www.TRB.org

Subscriber Categories

Highways
Safety and Human Factors

The Second Strategic Highway Research Program

America's highway system is critical to meeting the mobility and economic needs of local communities, regions, and the nation. Developments in research and technology—such as advanced materials, communications technology, new data collection technologies, and human factors science—offer a new opportunity to improve the safety and reliability of this important national resource. Breakthrough resolution of significant transportation problems, however, requires concentrated resources over a short time frame. Reflecting this need, the second Strategic Highway Research Program (SHRP 2) has an intense, large-scale focus, integrates multiple fields of research and technology, and is fundamentally different from the broad, mission-oriented, discipline-based research programs that have been the mainstay of the highway research industry for half a century.

The need for SHRP 2 was identified in *TRB Special Report 260: Strategic Highway Research: Saving Lives, Reducing Congestion, Improving Quality of Life,* published in 2001 and based on a study sponsored by Congress through the Transportation Equity Act for the 21st Century (TEA-21). SHRP 2, modeled after the first Strategic Highway Research Program, is a focused, time-constrained, management-driven program designed to complement existing highway research programs. SHRP 2 focuses on applied research in four areas: Safety, to prevent or reduce the severity of highway crashes by understanding driver behavior; Renewal, to address the aging infrastructure through rapid design and construction methods that cause minimal disruptions and produce lasting facilities; Reliability, to reduce congestion through incident reduction, management, response, and mitigation; and Capacity, to integrate mobility, economic, environmental, and community needs in the planning and designing of new transportation capacity.

SHRP 2 was authorized in August 2005 as part of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU). The program is managed by the Transportation Research Board (TRB) on behalf of the National Research Council (NRC). SHRP 2 is conducted under a memorandum of understanding among the American Association of State Highway and Transportation Officials (AASHTO), the Federal Highway Administration (FHWA), and the National Academy of Sciences, parent organization of TRB and NRC. The program provides for competitive, merit-based selection of research contractors; independent research project oversight; and dissemination of research results.

SHRP 2 Report S2-S01B-RW-1

ISBN: 978-0-309-12923-7

Copyright Information

Note: SHRP 2 report numbers convey the program, focus area, project number, and publication format. Report numbers ending in "w" are published as web documents only.

Notice

The project that is the subject of this report was a part of the second Strategic Highway Research Program, conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council.

The members of the technical committee selected to monitor this project and to review this report were chosen for their special competencies and with regard for appropriate balance. The report was reviewed by the technical committee and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the Governing Board of the National Research Council.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board, the National Research Council, or the program sponsors.

The Transportation Research Board of the National Academies, the National Research Council, and the sponsors of the second Strategic Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

SHRP 2 Reports

Available by subscription and through the TRB online bookstore:

www.TRB.org/bookstore

Contact the TRB Business Office:
202-334-3213

More information about SHRP 2:

www.TRB.org/SHRP2

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

SHRP 2 STAFF

ACKNOWLEDGMENTS

This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials. It was conducted in the second Strategic Highway Research Program (SHRP 2), which is administered by the Transportation Research Board of the National Academies. The project was managed by Charles Fay, Senior Program Officer for SHRP 2 Safety.

The research team gratefully acknowledges the support provided by the Virginia Tech Transportation Institute (VTTI) and University of Michigan Transportation Institute (UMRTI) data assembly teams. In particular, the authors thank Charlie Klauer of VTTI and Scott Bogard of UMTRI for providing outstanding support and quick response to data inquiries.

FOREWORD

Charles Fay, *Senior Program Officer, Safety*

A large component of the safety research undertaken in the second Strategic Highway Research Program (SHRP 2) is aimed at reducing the injuries and fatalities that result from highway crashes. Through a naturalistic driving study (NDS) involving more than 3,000 volunteer drivers, SHRP 2 expects to learn more about how individual driver behavior interacts with vehicle and roadway characteristics. In anticipation of the large volume of data to be collected during the NDS, several projects were conducted to demonstrate that it is possible to use existing data from previous naturalistic driving studies and data from other sources to further the understanding of the risk factors associated with road crashes. More specifically, the four S01 projects, entitled Development of Analysis Methods Using Recent Data, examined the statistical relationship between surrogate measures of collisions (conflicts, critical incidents, near collisions, and roadside encroachment) and actual collisions. This report presents the results of one of these projects, undertaken by Pennsylvania State University. It documents the second phase of a two-phase project under SHRP 2 Safety Project S01B.

The primary objective of this work was to investigate structured modeling paradigms for analysis of naturalistic driving data (NDD). Five research questions were identified and various models (e.g., event-based models and categorical-outcome models) were applied to NDD to determine appropriateness for analysis and suggestions for future analyses. The following were the five research questions:

1. What is the relationship between events (e.g., crashes, near crashes, and incidents) and pre-event maneuvers? What are the contributing driver factors, environmental factors, and other factors?
2. What hierarchical structure, if any, exists in the manner in which these relationships need to be explored?
3. What kind of elucidative evidence emerges from the analysis of roadway departure crashes in terms of Questions 1 and 2? Is the illustrative hierarchy of relationships generalizable to other nonintersection crash types, such as leading vehicle crashes?
4. In terms of elucidative evidence, what types of behavioral correlates emerge? For example, are attitudinal measurements indicative of revealed behavior in terms of headway maintenance and speed reductions?
5. If elucidative evidence does in fact emerge in terms of attitudinal correlates and how their interactions vary by context, is it plausible to parse out the marginal effects of various context variables on crash risk by suitable research design?

This report will provide useful information for analysts of the SHRP 2 NDS data, as well as other naturalistic driving data sets.

CONTENTS

Executive Summary

Background

In the spring of 2007, Pennsylvania State University (Penn State) was awarded a contract to analyze existing naturalistic databases as part of the S01 Safety project within the second Strategic Highway Research Program (SHRP 2). The Penn State team proposed using data collected by the Virginia Tech Transportation Institute (VTTI) during the 100-car naturalistic driving study (Dingus, Klauer, et al. 2006) and data from the automotive collision avoidance system (Ervin et al. 2005) and the road departure crash warning (RDCW) system field operational test (LeBlanc et al. 2006) conducted by the University of Michigan Transportation Research Institute (UMTRI).

The next two subsections describe the analyses undertaken with each data set. The final section summarizes the findings of the research in terms of the five research questions identified in the original Penn State proposal and reiterated in the final Phase 1 report to SHRP 2.

Analysis of VTTI Data

Two parallel tracks were pursued in the analysis of the 100-car study data: event-based modeling and driver-based modeling. The first approach modeled the occurrence of each event in detail. The focus was on understanding the interactions of the many factors that led to event occurrence. This initiative fit nicely with the data provided by VTTI, as it allowed events to be compared at three levels (summary definitions provided by Dingus, Klauer, et al. 2006):

- Crash—any contact with an object, either moving or fixed, at any speed, in which kinetic energy is measurably transferred or dissipated;
- Near crash—a circumstance that requires a rapid, evasive maneuver by the subject vehicle, or any other vehicle, to avoid a crash; the maneuver causes the vehicle to approach the limits of its capabilities (e.g., vehicle braking greater than 0.5 *g* or steering input resulting in lateral acceleration greater than 0.4 *g*); and
- Crash-relevant incident (in this report referred to as a critical incident)—a circumstance that requires a crash avoidance response on the part of the subject.

Each of these events was identified by VTTI staff as part of the 100-car study, and the three event types were provided to Penn State in response to the team's data request. Penn State developed a structured analysis framework for these event-based data; the model specified driver attributes, the context in which the event occurred (including roadway and environmental variables), and attributes describing details about the event itself, particularly in the few seconds before and during the event. Examples of event-level variables include whether the driver was observed to be distracted just before the event and whether the vehicle crossed over the lane or road edge. One may think of these models as exploring the details of factors associated with the events.

Various model formulations were used to find variables associated with crashes and near crashes, and the attributes of vehicle motion associated with such events (e.g., vehicle over lane or road edge) that could serve as surrogate measures for crashes were investigated. If these event-related measures were shown as being positively associated with a crash or near-crash event, they were considered as potential surrogates. A set of nonincident control events was received with the original data, but it was not useful in the modeling because it contained none of the predictors used in the event analysis. The team tested the specific measures available in the data set and attempted to supplement the available vehicle kinematic data by downloading information from the NHTSA website. Unfortunately, kinematic data were only available for a small number of crashes; near crashes and critical incidents were not represented, and this approach was, therefore, abandoned.

One weakness of event analysis is that it precludes the study of drivers who experience none of the three measured events (i.e., the safest drivers). In order to include these drivers, the second analysis track conducted by Penn State with the VTTI data was a series of models of the number of events per driver. Consistent with much of the modeling in the safety field, these analyses were conducted using a set of count regression formulations (e.g., Poisson, negative binomial [NB], and zero-inflated Poisson [ZIP]) that resulted in estimates of the probability of a driver with particular attributes having 0, 1, 2, . . . , n events during the year of the 100-car study. These models allowed comparisons to be made across all drivers.

Analysis of UMTRI Data

The UMTRI data consisted of a set of drivers who experienced a series of alerts from onboard systems about potential crashes. Because no crashes were recorded in the UMTRI data, the dependent variables used in the analyses were derived from a system designed to detect excessive speed entering a curve (i.e., the curve speed warning [CSW]) and an alert triggered when the subject vehicle deviated from the lane or road edge (i.e., the lateral drift warning [LDW]).

After an initial screening of the data, the team decided to focus on the CSW alerts as they provided duration of time and thus contained more details about the driver response to the alert. Further, the curve speed event was more consistent with the road departure event covered in the VTTI analyses, and it was thought there may be some benefit from the similarity.

Two approaches were taken in the analysis of the UMTRI data. The first used a series of piecewise linear models to characterize the nature of the relationship between vehicle kinematics and CSW alert frequency and duration. The interest was in finding which kinematic variables were most correlated with the triggering of the alert. This information was used to gain insight about potential surrogates, under the assumption that the kinematic variables most associated with alert occurrence would be potentially efficacious crash surrogates to consider in subsequent research. A positive association between a kinematic variable and an alert being triggered could be an indication of a kinematic variable that might also be associated with (or potentially causing) road departure crashes. While the team acknowledges the nature of this conceptual leap, it was believed that the exploratory nature of the S01 projects would support this type of analysis. Time–series models of the kinematic data were also attempted, but as they did not yield particularly meaningful results, they are not discussed in this report.

The second approach taken with the UMTRI data was to use a cohort-based formulation to estimate the probability of a particular number of alerts being triggered for an individual driver (e.g., characterized by gender, years of driving experience, and mileage driven in particular contexts). This exposure-based analysis is based on actual miles driven under specific environmental and roadway conditions as measured by the CSW–LDW system. Because of the structure of the UMTRI data, the team was able to analyze alert frequency at a very detailed level of exposure.

One of the most important outcomes of the UMTRI modeling effort is the successful estimation of cohort models using homogeneous trip segments. This formulation takes advantage of the unique trip-by-trip data obtained in the naturalistic study, along with geographic information system (GIS)–related factors coded by UMTRI (such as road type and environmental conditions), to derive a measure of alert frequency for each trip segment. The issue of interest is the ability to truly capitalize not only on the naturalistic driver behavior data, but also on detailed GIS roadway data. Since there is a plan to collect detailed roadway data as part of SHRP 2 Safety Project S04, Acquisition of Roadway Information, the team believes this formulation merits consideration for future studies. Even though the models are estimated with alerts, there is a direct parallel to the modeling of crashes or other events of interest. In addition, researchers can very flexibly define homogeneous trip segments to match their research needs. The estimated models using the cohort formulation verify the efficacy of this approach; the findings are discussed in the response to Research Question 3.

Research Hypotheses, Findings, and Implications

The analysis of the data provided by VTTI and UMTRI was guided by the five research questions. This section states and discusses each of the five questions in sequence, specifically including the hypothesis or issue explored and a summary of what was discovered. The implications of the various findings are discussed in detail in Chapter 4.

Research Question 1

What is the relationship between events (e.g., crashes, near crashes, incidents) and pre-event maneuvers? What are the contributing driver factors, environmental factors, and other factors?

The VTTI data set was primarily used to answer this question. The general structure of the event-based models was to use predictor variables representing driver, context (i.e., roadway and environment), and event attributes. Models were estimated with context-only, driver-only, and event-only variables (and combinations of only two of these components). Resulting parameter estimates changed substantially depending on how many of the three components were represented in the model; importantly, the exclusion of any of the components led to major changes in estimated parameters (see Chapter 3). The exclusion of any of the set of variables (i.e., driver, context, or event) is likely to result in biased parameter estimates, obscuring the effect of any one variable on event occurrence. To avoid this bias, future analyses of SHRP 2 event-based data (such as in proposed research for the S08 project) should include variables representing driver, context, and event attributes. In addition, thorough tests should be conducted to explore changes in parameter values and significance. The Penn State team is concerned that parameter estimates may exhibit the same characteristics, even in data sets with large sample sizes.

The strongest variables (i.e., those showing the greatest association with crashes or near crashes) were the driver distraction variables. These variables included distractions such as those attributed to a portable electronic device, internal distractions (such as a pet), or vehicle-related distractions (such as adjusting the climate or audio controls). Although the team used distraction as a predictor variable, some distractions may be endogenous (i.e., the conditions that led to the event also led to the distraction) and may not be suitable as event predictors. A range of statistical methods to address endogeneity should be considered in these circumstances. In addition, there may be a need to explore measurement periods beyond the 5-seconds-before-event criteria used in the VTTI data base.

Special care should be exercised and perhaps specific models formulated to explore the nature of the endogeneity between distractions and other event-related measures. The team's model tests indicate that distractions as predictor variables may not be valid.

The efficacy of using categorical-outcome models (such as logit or binary hierarchical models) to compare crash and noncrash events was explored within the limits of the VTTI data by comparing crash and near-crash events (combined) with critical incidents. The Penn State team estimated a series of models that yielded generally consistent results concerning the effects of particular parameters when using a complete model specification as described above. Given a set of event-based data, it is feasible to apply well-established categorical data analysis techniques to estimate factors that differentiate between the categorical outcomes. This method implies that such a differentiation appears feasible for crashes (or other adverse events) and a sample of comparable, similarly described nonevents. Such a comparison was anticipated, but it was not achieved because the data for nonevents in the VTTI file did not contain predictor variables consistent with the events.

Gender was important in both driver- and event-based models. Many gender-related factors were revealed as main effects, but they were particularly apparent as interaction terms, especially in driver-based models. Analyses that are directly or indirectly influenced by gender should include tests of a range of main effects and interaction terms. Variables with significant promise in future modeling include level of education and years of driving experience. Several associations between number of previous crashes and violations varied with gender; these associations were not consistent, but they may warrant attention from researchers on gender issues.

A limited number of vehicle factors rose to significance; additional research is needed concerning the analysis of vehicle factors, particularly in conjunction with the gender of the driver.

Research Question 2

What hierarchical structure (statistically speaking), if any, exists in the manner in which these relationships need to be explored?

Two hierarchical models are reported with the VTTI data: one was applied to event modeling and the second to driver-based models. A third hierarchical model was estimated with the UMTRI data using a cohort approach.

Figure ES.1 shows one hierarchy successfully applied to the analysis of event data. The sketch is intended to convey that individual drivers may have any number of events; they must have at least one, but they may have more. If one were to model this with a count regression approach, each event would enter the model as if it were independent and from a different driver.

Using a hierarchical approach, driver attributes enter at the driver level, once for each driver. Event characteristics are entered as predictors for each event in which they occur. This hierarchical approach provided a conceptually justifiable approach to the modeling of complex events and was applied to both VTTI and UMTRI data. A driver-based approach presents one way to analyze drivers at a separate level from the events of interest, providing a much better depiction of the physical process being investigated.

A second hierarchical model (Figure ES.2) was used in the driver-based analysis of the VTTI data. In this structure, males and females are accounted for separately, and the model includes separate parameter estimates for each gender category.

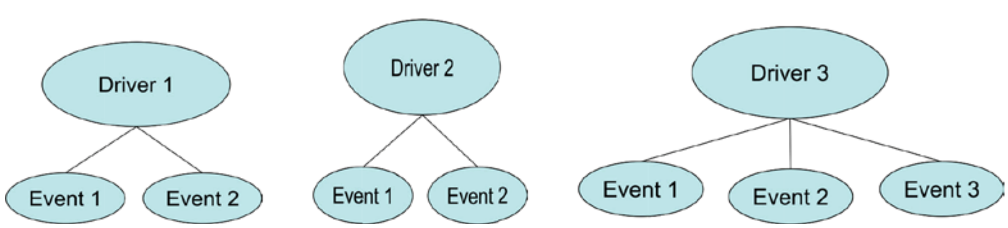

Figure ES.1. Hierarchy analysis of event data.

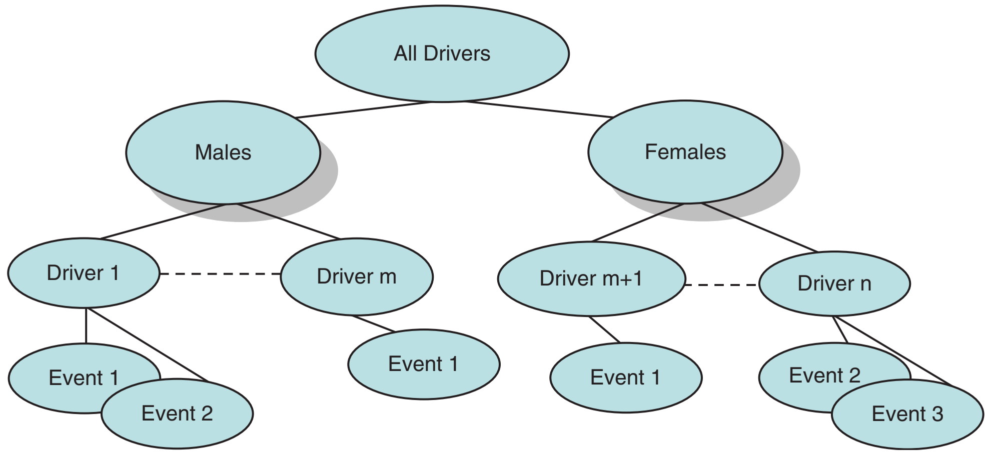

Figure ES.2. Driver-based hierarchical model.

This driver-based hierarchical model presents another example of how hierarchical approaches can be applied to naturalistic data. The benefits of obtaining gender-specific estimates of factors contributing to the risk of events are clear.

Research Question 3

What kind of elucidative evidence emerges from the analysis of roadway departure crashes in terms of Questions 1 and 2? Is the illustrative hierarchy of relationships generalizable to other nonintersection crash types such as leading vehicle crashes?

Elucidative evidence refers to evidence of the likely effect of individual predictor variables in modeling event occurrence (including crashes). The notion of elucidative evidence includes surrogate measures and their testing. Surrogates are a special type of variable that have been discussed as a general replacement for crash data; the description and interpretation of Penn State surrogate analyses are contained in the responses to this general question. Exposure requires a predictor variable reflecting time or distance of travel; exposure-based analyses of both data sets are also described in this report. Responses to this question thus provide a summary of the extent to which the modeling results provide guidance on variables to be given priority in future analysis studies. Some evidence suggests that several types of predictor variables, such as precipitating event information in VTTI models, have particularly important roles in the models.

One useful definition was articulated by Hauer in his more focused discussion of the traffic conflicts technique as a surrogate measure (Hauer and Gårder 1986): "one should be able to make inferences about the safety of an entity on the basis of a short duration 'conflict count' instead of having to wait a long time for a large number of accidents to materialize." Shankar has argued as part of this research that surrogates have a time dimension (e.g., a measure such as time to collision has a clear time dimension; time to road departure is another) (Shankar et al. 2008). In addition, Shankar argues that a surrogate should be responsive to the same interventions as a crash. An example is a curve warning system alerting the driver to unsafe conditions ahead: for a surrogate like near run-off-road crash to be valid, it must be mediated by a curve warning alert in the same way as a crash. More generally, surrogates can be considered as measures that can be substituted for crashes in a safety analysis: in the data for this project, they are typically vehicle kinematic– and event-related measures that offer some description of vehicle movement and/or position relative to the roadway.

Potential surrogates encountered in the VTTI data include the precipitating events of subject over lane or road edge and lost control. In most of the categorical event-based models these two variables were strong indicators of crash or near-crash events; in hierarchical models subject over lane or road edge was the second strongest predictor associated with a crash or near-crash event. While this measure has a strong association with crash events, this measure does not have a time dimension, so it does not directly meet Shankar's desirable criteria (Shankar 2008). Further, Hauer's rule could not be applied because the team did not have access to the comparable set of subject behavior for noncrashes. It may not be broadly applicable outside of SHRP 2's instrumented vehicles; nevertheless, it is clear that the measure has some potential as a surrogate.

The categorical models explored in this study appear to provide a useful paradigm for exploring surrogates when event-based data are available. While not directly tested with VTTI data, the Penn State team believes that kinematic measures or combinations of kinematic and roadway position measures are possible measures for future testing. For example, the subject over lane or road edge variable contained position-only information and was strongly associated with crash-related events; inclusion of longitudinal or lateral velocity and lateral position information would enhance its predictive ability.

A limitation of the categorical models deserves mention. Initial event-based models, both bivariate logistic and hierarchical, used improper speed as an event-based predictor. Successful model fit was obtained, but improvement was sought. Driver impairment 1 (drowsy, sleepy, fatigued) was substituted as a predictor and much better fit occurred overall, including reduced standard errors for several variables. While we were pleased by the improved fit, we were concerned about the apparent model instability. This may be due to the small sample size, but it may also reflect endogeneity among the predictors. As a recommendation to future SHRP 2 analysis contractors, the team suggests that care be exercised in surrogate analyses; additional empirical testing in several other sites and with other drivers should reveal more about this issue.

A method for validating events containing possible surrogates for crashes is proposed and discussed in Chapter 3. The statistical predictions from the event-based model were compared with text descriptions of the event etiology derived from video and kinematic data; the comparison showed that events originally coded by VTTI as critical incidents were statistically estimated to be crashes. It was posited that these events could be used to supplement crash data observed directly. The manipulation of the event-based models is proposed as a means of providing useful information about whether a particular critical incident or near-crash event really was similar, statistically, to a crash event in a similar context. Such a comparison is dependent on the model being correct. An additional validation technique is discussed using the cohort formulation with hierarchical models, leading to the development of safety performance functions for crashes and the surrogate measure.

The UMTRI analyses tested several kinematic measures, particularly longitudinal velocity entering curves, as a potential surrogate of event risk; in this case a CSW alert, instead of a crash event, was used. Initial tests of piecewise linear models applied to the data as a whole showed that the measure has some merit, but the models were weakened statistically by the presence of serial correlation in the observations (data were collected at 10 Hz). The team next explored tracking individual drivers through the same location multiple times to see if there were repeated behaviors or learning and to explore individual variability. The models showed different results than the aggregate. While the results were not stunning, they showed potential and are recommended over aggregate approaches.

The ability to explore context through the use of the detailed roadway data available through Google maps (i.e., by tying kinematic measures to specific road segments) should greatly enhance the findings. Tracking individual drivers repeatedly over the same route has potential for additional insight. Specifically, a range of kinematic variables can be measured at specific points of documented high crash frequency; these can be compared with a set of individual drivers' kinematic signatures through the same roads. Kinematic measures at crash locations can be compared with similar measures at low-frequency crash locations and tested for their predictive capability.

Cohort-based modeling also shows promise in quantifying context effects (this method is addressed in Research Question 5). The driver-based models using VTTI data used self-reported annual mileage as exposure. These models showed that exposure is essential to the study of the expected number of events per year for drivers. There was a strong association of exposure with the expected number of events, and the inclusion of this variable greatly improved model fit. It is clear that travel by individual drivers should be identified to the extent possible through the face camera or other technologies. The team developed a model that clearly identified drivers who were outliers with respect to the number of events they experienced. Drivers with exceptionally high, as well as low, numbers of events can be identified using this technique.

Virtually all of the event-based models showed substantial differences in the effect of distractions on event occurrence. Most generally, internal distractions (e.g., reading, moving an object in the vehicle, or dealing with a pet or insect) were most strongly associated with crash or near-crash event occurrence. Passenger-related distractions and observations of the driver talking, singing, or daydreaming also had consistent positive correlations. Interestingly, the use of a wireless device was poorly correlated to event occurrence. These findings, taken as a whole, reveal that distractions merit careful measurement in future SHRP 2 analysis efforts. Event data would be even more useful if matched with nonevent data collected from all drivers that included comparable distraction measures.

Research Question 4

In terms of elucidative evidence, what types of behavioral correlates emerge? For example, are attitudinal measurements indicative of revealed behavior in terms of headway maintenance and speed reductions?

The principal measure of behavioral correlation was the Dula Dangerous Driving Index (DDDI) (Dula and Ballard 2003) obtained by VTTI during the original 100-car data collection effort. The DDDI consists of 28 statements to which the driver is asked to respond on a 5-point Likert scale. Example test statements include “I verbally insult drivers who annoy me”; “Passengers in my car/truck tell me to calm down”; and “I will weave in and out of slower traffic.” Responses to the questions are divided into three categories: aggressive driving (AD), negative emotional (NE) driving, and risky driving (RD). Each category is intended to capture a different aspect or component of dangerous driving.

The DDDI was generally associated with an increase in crashes and near crashes in the event-based models and was also positively associated with number of events in the driver-based models. The results were not always easy to interpret or consistent with intuition. In driver-based models, for example, AD was associated with an increase in the number of events, but for females only. In the event-based models, this same component was associated with a reduction in crash or near-crash events (which could be interpreted as an increase in the likelihood of critical incidents). So, while the associations in the data were generally consistent and statistically significant (within the limits of the data), there is a concern that the findings were not as interpretable as would be desired. The testing conducted with the DDDI confirms the importance of including some measure of driver risk propensity in the remaining SHRP 2 data analyses.

In addition to the DDDI, the Life Stress Index was administered to participating primary drivers in the 100-car study. This tool attempts to measure the amount of stress present in the subject’s life by using factors such as stress at work, difficulty with personal relationships, and challenges in the family environment. The Life Stress Index was positively associated with crash and near crashes in some event-based models, but it was not a predictor in the driver-based models.

Although a metric for life stress provides some interesting data, it is not as important as driver-based risk-taking measurements. The proposed testing for the S07 projects, In-Vehicle Driving Behavior Field Study, includes a number of perceptual and cognitive tests; psychological tests include metrics for risk taking, risk perception, driver style and behavior, and thrill and adventure seeking. These data should provide more than ample measures of driver predisposition for events.

Research Question 5

If elucidative evidence does in fact emerge in terms of attitudinal correlates and how their interactions vary by context, is it plausible to parse out the marginal effects of various context variables on crash risk by suitable research design?

This question bears directly on the importance of context in the analysis of naturalistic driving data. Event modeling revealed that failure to include context-related variables will yield a model with substantially biased parameter estimates. There is no way the influence of factors such as distractions and predisposition variables can be properly assessed without the inclusion of context.

Several aspects of context were revealed to be associated with crash and near-crash outcomes. Roadway-related factors were important descriptors of context in the series of event-based models. The presence of curves was a significant factor in differentiating critical incidents from crashes and near crashes. Horizontal curves, in general, indicated a modest increase in risk. Horizontal curve presence does not show the magnitude of influence of driver behavior variables such as distractions, but it is clearly important in defining context.

Time of day, specifically dawn or dusk, was a substantial factor increasing risk and contributed importantly to the definition of context in which crash or near-crash events occurred. This variable was consistently significant and positive in all event-based models and had odds ratios (ORs) that exceeded some driver distraction and precipitating event factors. These findings are consistent with sleep- and fatigue-related studies of crash risk for both private drivers and the motor carrier industry. Future research projects conducted as part of SHRP 2 Safety Project S08, Analysis of the SHRP 2 Naturalistic Driving Study Data, need to seriously consider the identification of dawn and dusk driving as an important element of context. Comparing crashes, near crashes, and critical incidents with a sample of nonevents with comparable attributes would serve to validate these findings.

Run-off-road crashes were consistently and negatively associated with increased traffic levels; this seems like a plausible association, as drivers are more likely to have crashes, near crashes, and critical incidents under more congested traffic conditions. This association was not as strong as the associations with the other variables.

In cohort-based models formulated with the UMTRI data, context was generally more strongly associated with event outcome (i.e., CSW alerts) than driver-based variables. This general finding supports the emphasis on context that has stimulated much discussion during recent research symposia. Interestingly, the hierarchical model described in Chapter 3 identifies variability between drivers as a major factor in explaining CSW alert frequency. Taken together, these findings support the concept that context and driver attributes are complementary and closely linked.

The cohort-based approach enables the researcher to use naturalistic driving data to examine both driver and context factors in a consistent exposure framework. Such research is only possible with the detailed data available from a naturalistic driving database, such as the UMTRI RDCW data set, which provides data on individual drivers monitored through a series of contexts.

Cohort analysis represents a breakthrough in analysis paradigms for naturalistic data. The driver is tracked through a roadway network defined as homogeneous based on the needs of the analysis team. Once segments are defined, events (using appropriate screening criteria) can be allocated to the segments. The analyst can make the segment designation as fine or coarse as roadway and roadside data allow. This framework provides the measurement of the driver's behavior throughout the driver's travel, as well as in the seconds immediately preceding or following a crash.

A range of statistical methods was used to provide examples of how the cohort-based data structure can be used. These are intended to assist future SHRP 2 safety studies by providing guidance about data manipulation and variable formulation.

CHAPTER 1

Introduction

Background of Naturalistic Driving Studies

Naturalistic driving experiments have been conducted for several years and include studies of drivers in their own vehicles and a series of technology tests to assess the safety consequences of advanced in-vehicle technologies. Generally, the 100-car study conducted by VTTI provided naturalistic data to make causal crash assessments, with a focus on the few seconds before and after crashes and events of interest (near crashes and critical incidents). UMTRI conducted an evaluation of a roadway departure and curve warning system as part of the U.S. Intelligent Transportation Systems program.

While both studies were extensive, they did not focus on analysis methods per se. The S01 project sought to fill the analysis gap by focusing on analysis of existing data with an eye toward developing analysis tools for the 2,500-car study.

This report describes examples of analyses conducted in exploring paradigms for naturalistic driving data analysis. The authors build on and begin to test some of the paradigms identified in the Phase 1 report and by Shankar and associates (2008). This report has four parts. The literature review that concludes this first chapter examines the literature concerning naturalistic driving studies and hierarchical statistical methods applied to traffic safety. Chapter 2 describes the research methodology and available data, and Chapter 3 discusses the models estimated from the data. Chapter 4 summarizes the findings and their implications for SHRP 2 Safety projects and recommends future research.

Literature Review

Naturalistic Driving Studies

Stutts et al. (2005) unobtrusively collected video data from 70 volunteer participants driving their own vehicles over a period of 1 week. This study provided some of the first naturalistic data on drivers' exposure to potentially distracting events. The data were analyzed by the bootstrap percentile method of Mooney and Duval (1993) and provided some evidence that distractions can negatively affect driving performance, as measured by higher percentages of drivers having no hands on the steering wheel, their eyes directed inside rather than outside the vehicle, and their vehicles wandering in the travel lane or crossing into another travel lane.

Teams of researchers at Virginia Tech have collected one of the most extensive sets of naturalistic driving data, including naturalistic driving of light-duty vehicles (Dingus et al. 2006; Klauer et al. 2006; Lee et al. 2004; Neale et al. 2005). In addition, VTTI has collected data and conducted numerous fatigue and drowsiness studies as part of a series of studies of driver drowsiness systems (Dingus, Neale, et al. 2006; Hanowski et al. 2005; Hanowski, Hickman, Fumero, et al. 2007; Hanowski, Hickman, Wierwille, et al. 2007). The 100-car naturalistic driving study database contains many extreme cases of driving behavior and performance, including severe fatigue, impairment, judgment error, risk taking, willingness to engage in secondary tasks, aggressive driving, and traffic violations (Neale et al. 2005). The data set includes approximately 2,000,000 vehicle miles, almost 43,000 hours of data, 241 primary and secondary drivers, 12 to 13 months of data collection for each vehicle, and data from a highly capable instrumentation system that included five channels of video and vehicle kinematic sensors.

Driver inattention was analyzed using the driving data set, and risk (ORs) was calculated using both crash and near-crash data, as well as normal baseline driving data, for various sources of inattention (Klauer et al. 2006). The risk percentages were also calculated to estimate the percentage of crashes and near crashes occurring in the population that resulted from inattention.

Among the research involving naturalistic truck driving data (Dingus, Neale, et al. 2006) were cluster analysis studies of distraction-related incidents (Hanowski et al. 2005). Two primary findings were that single drivers drive significantly more aggressively than do team drivers, and that the frequency of critical incidents and fatigue-related critical incidents varied significantly by the hour of the day. The relationship between sleep quantity and involvement in critical incidents (crashes, near crashes, or crash-relevant conflicts) was studied using detailed sleep and driving data (Hanowski, Hickman, Fumero et al. 2007). Interactions between light-duty and heavy-duty vehicles used the same data for another targeted study (Hanowski, Hickman, Wierwille, et al. 2007).

In addition to the previous light-duty vehicle and truck studies, research was conducted at VTTI concerning collision warning systems (McLaughlin et al. 2008). Seventy-three events were collected during actual driving. Data from the host vehicle, such as speed, yaw, acceleration, different control states (e.g., brake pedal, turn signals), and measures of driver attention, were also collected.

Other researchers have used more limited naturalistic driving data sets to assess a range of safety and operations questions. One study developed a model of lane-change duration for improving microscopic traffic flow simulation (Tijerina et al. 1999). Another study collected real-world driving data from a small sample of drivers to identify periods of drowsiness and inattention and validate drowsy driver detection algorithms. Data on driver exposure to environmental factors and encounters with driving conflicts, near crashes, and actual crashes were used to characterize driver and vehicle performance, as well as the driving environment (Toledo and Zohar 2007).

Researchers at UMTRI conducted a series of naturalistic driving studies as part of a series of field operational tests for the U.S. Intelligent Transportation Systems program (NHTSA 2005; Bogard et al. 1998; daSilva and Najm 2006; Ervin et al. 2005; Sayer et al. 2005; Sayer 2006). Systems tested included integrated forward collision warning and adaptive cruise control. Targeted studies of distraction and behavior (Ervin et al. 2005) found that variability of steering angle, mean and variability of lane position, mean and variability of throttle position, and variability of speed were affected by contextual factors such as road type, road curvature, and road condition. Conversing with passengers was the most common secondary behavior (15.3%), followed by grooming (6.5%) and using cellular phones (5.3%). This study found that the use of a cellular phone, eating or drinking, and grooming resulted in increased steering variance but did not affect lane position or speed variance. Another study (Sayer et al. 2005) quantified subjective reliability and performance of an in-vehicle warning system as a function of age, gender, weather conditions, light levels, and roadway classifications.

UMTRI's *Road Departure Crash Warning System Field Operational Test: Methodology and Results* (LeBlanc et al. 2006) evaluated the effects of CSW alerts on vehicle lateral acceleration in curves by exploring how drivers travel through curves and how they respond when presented with curve speed alert warnings. The study found that lateral acceleration was higher during the day than at night and also higher for right turns than left turns (both when evaluating lateral acceleration on a time-based average of the value when it exceeded a specific driver's 90th percentile value and when the average of the maximum value for individual curves exceeded a driver's 90th percentile value) (LeBlanc et al. 2006, pp. 8–28). However, no driver had a 90th percentile value of lateral acceleration greater than the 8.2 m/s^2 nominal threshold for CSW alerts, and only two drivers had 90th percentile values greater than 7.2 m/s^2 (LeBlanc et al. 2006, pp. 8–24). A combination of these two evaluations found that the availability of CSW alerts did not have a dramatic effect on a driver's chosen lateral acceleration (mathematically and physically related to longitudinal speed). From a fundamental perspective, drivers may choose speeds at which to traverse horizontal roadway curves based on their comfort level of lateral acceleration. They may not want to feel excessive force while traversing curves and will decrease their longitudinal speed to maintain comfortable conditions.

Spacek (2005) examined the different types of paths drivers take while traversing horizontal curves and compared the frequency of the different types to the best possible path, which involves following the centerline of the lane perfectly through the entire curve. Using this ideal behavior as a baseline, there were five categories for comparison: normal (slightly cutting into the inside of the curve for a portion of traversal), correcting (reaching outside of the curve and overcompensating by turning harder toward the inside of the curve), cutting (strong cutting into the inside of the curve to counteract centripetal acceleration—a conscious process), swinging (starting toward the outside of the lane and finishing closer to the inside of the lane), and drifting (behavior opposite of swinging). The results of the study showed that, excluding undefined paths taken, cutting and normal behavior ranked first and second, respectively, in terms of percentage of track types taken for most curve radii (with a slight reversal of ranking for right-hand curves with radii greater than 65 m).

Wilson et al. (2007) provided an in-depth analysis and evaluation of the UMTRI study (LeBlanc et al. 2006). Most of the major findings involving CSW alerts involved acceleration characteristics and speed approaching and during curve traversal. Vehicle speed approaching a curve was a major change-in-speed factor for triggering and reacting to alerts. There was a general positive correlation between approach speed and acceleration upon alerts being triggered. CSW alerts on ramps were mostly analyzed for exit ramps, as an alert is much more likely on exit ramps than entrance ramps due to the higher travel speed expected on limited access roadways. The beginning curvature of exit ramps played a role in actually eliciting CSW alerts; if the ramp began with curves with larger radii and progressed to smaller radii, the system would be able to detect over time that it was a ramp and not another roadway classification. However, some false alerts could be triggered if the initial curve on the ramp had a relatively small radius.

The Design Quality Assurance Bureau of the New York State Department of Transportation (2003) discussed issues surrounding superelevation and how it relates to speed choices while traversing curves (Bonneson 2000). It is important to note that typical passenger cars will skid before rolling over during a turning movement, especially if the roadway surface is wet. Since a passenger car (the Nissan Altima) was used in the UMTRI study, it was assumed that any situation that triggered an alert may have resulted in a skid if necessary response maneuvers were not undertaken. Friction allows deceleration and steering forces to be transmitted from the tires to the roadway surface. The friction factor is used in place of the more common coefficient of friction as a ratio of the lateral forces that the pavement can resist from the vehicle. Changes in speed can reduce this friction factor, thus reducing the friction available for cornering, making curve traversal more difficult. This friction factor depends on vehicle speed and weight, tire conditions, and pavement conditions. However, speed is the most important variable in determining the friction factor, as it is the only variable that truly determines if a vehicle can safely traverse a curve under prevailing conditions. This makes speed likely the most important kinematic variable that should be evaluated in this analysis.

Interpreting adaptation to alerts through speed changes does not necessarily account for the effects of traffic conditions, as certain roadway types have widely varying traffic volumes throughout the day. If higher volumes exist and headways decrease substantially, speed output in the data set will be influenced and may affect model results. Fitzpatrick et al. (2000) discussed speed prediction by recording speeds on two-lane rural highways; they only included vehicles that were at free-flow speeds (headway >5 s). Regression models were run to relate speed to several geometric variables, including horizontal curve radius (some models did not include horizontal curve radius as a predictor). Models that included radius as a predictor (always in the form of $1/R$, where R = radius) had adjusted R^2 values above 0.5. Regardless of vertical geometry, $1/R$ had a strong correlation with speed (85th percentile speed) and was always significant, sometimes being the only significant predictor. Sufficiently large horizontal radii were considered by the researchers to be a condition that drivers would deem insufficiently severe to require speed reductions. If vertical alignment was considered an important factor, it was considered the controlling factor in speed decisions for radii greater than 800 m. Sharp drops in speeds occurred for radii less than 250 m.

Fitzpatrick et al. (2000) also modeled acceleration in and near horizontal curves based on assumptions used in previous FHWA research:

- All acceleration and deceleration occur outside the limits of the horizontal curve.
- Acceleration and deceleration rates are constant and equal to 0.85 m/s^2.

Tangents had to be kept above a certain minimum (244 m), and grades were kept close to 0% (maximum 5% downgrade or upgrade) to maintain conditions that allowed vehicles to reach the maximum possible speed on the tangent approaching the curve without other factors coming into play. Free-flow conditions were required (minimum headway of 5 s) as well. They found that speeds did not drop substantially until vehicles reached a point 200 m from the point of curvature of the given curve. They also found that the average acceleration rate in the 200-m zone before the curve was –0.1143 m/s^2, which was significantly different from the previously assumed value of –0.85 m/s^2. This rate ranged from 0.01 to 0.54 m/s^2, which was primarily affected by curve radius. Acceleration within curve limits was found to be –0.0724 m/s^2, which was significantly lower than the assumed value of –0.85 m/s^2. The maximum observed positive and negative acceleration rates were significantly different from –0.85 m/s^2. Thus, the assumptions from previous research were not appropriate for this particular study. Using these results, models were developed that showed that the drop in speed approaching a curve was generally inversely related to curve radius. Negative acceleration may begin to occur at various points upstream of curves, which may be due to site-specific differences. Thus, analysis of speed choices and acceleration rates at horizontal roadway curves should be carefully performed and possibly take into account site-specific differences, especially those surrounding curve radii and approach tangent lengths.

Different forms of adaptation can occur on curves based on the presence of CSW alerts, personal comfort levels with respect to speed and how it affects the centripetal force the driver feels, comfort, curve perception, and the driver's ability to maintain a reasonable path while traversing curves. However, most of the analyses of adaptation involve the effects of longitudinal speed choices. Longitudinal speed appears to be the primary factor in assessing the level of danger approaching and traversing horizontal roadway curves, but it can be indirectly related to other factors mentioned here.

The literature (see Table 1.1) largely confirms the rationale for the SHRP 2 naturalistic driving project (S08): the literature analyzing naturalistic data is limited. Very little attention is paid to identifying paradigms (i.e., frameworks or organized structures) for analysis and the development of methods that evolve from those paradigms. The exploration of analysis paradigms is the focus of the Penn State research.

Table 1.1. Related Reports from Literature Review

Reference	Research Objective	Analysis Method
Stutts et al. 2005	Study nature of driver distractions	Frequency of distractions; bootstrap analyses of lane wanderings, lane encroachments, and sudden braking associated with distractions
Hanowski et al. 2005	Study driver distraction in commercial vehicle operations	Cluster analysis; cross-classification analysis
Dingus, Neale, et al. 2006	Collect and analyze naturalistic data applied to driving fatigue	Hazard analysis combined with analysis of variance
Hanowski, Hickman, Fumero, et al. 2007	Identify and analyze light vehicle–heavy vehicle interactions	Descriptive comparisons of percentages of critical incidents by category
Hanowski, Hickman, Wierwille, et al. 2007	Quantify and analyze sleep of commercial vehicle drivers and associations with crashes	Matched paired *t*-test
McLaughlin et al. 2008	Analyze collision avoidance systems using naturalistic data	Method constructed to test collision avoidance systems based on driver reaction and vehicle kinematics
Bogard et al. 1998	Characterize safety and comfort issues of driver interactions with adaptive cruise control	Histogram; descriptive statistics analysis
Tijerina et al. 1999	Identify periods of driver drowsiness and inattention and validate drowsy driver detection algorithms	Drowsy detection algorithm developed by Wierwille
Dingus et al. 2006	Collect large-scale naturalistic driving data; define a near crash using quantitative measure Characterize driver behavior (e.g., driver inattention) and roadway environment as they relate to incidents, near crashes, and crashes Characterize changes in driver behavior over time with consideration of rear-end conflict and lane change as contributing factors	Range/range rate approach applied to quantify a near crash (Kiefer et al. 2003); risk ratio applied to driver behavior change over time (Greensberg et al. 1993); estimation of Poisson rate per million vehicle miles traveled in relation to Heinrich triangle by scenarios (Heinrich et al. 1980)
Lee et al. 2004	Characterize and analyze nature and severity of lane changes	Analysis of variance (ANOVA); chi-square analyses
Klauer et al. 2006	Characterize driver inattention using driving data collected in the 100-car naturalistic driving study	Odds ratios
Ervin et al. 2005	Analyze impact of integrated forward collision warning (FCW) and adaptive cruise control (ACC) systems on driver safety and acceptance	Paired *t*-test; ANOVA
Sayer et al. 2005	Determine frequency and conditions under which drivers engage in secondary behaviors; explore relationship between behaviors and driving performance using UMTRI RDCW field operational test data	Mixed-model analysis of variance; autoregressive integrated moving average model (ARIMA)
LeBlanc et al. 2006	Analyze suitability of RDCW system, which combines LDW and CSW functions	Descriptive statistics analysis
Sullivan et al. 2007	Examine how driver behavior is influenced by the reliability of an in-vehicle warning system using data derived from UMTRI RDCW field operational test	Mixed-model analysis of variance

Hierarchical Modeling Methods Applied to Road Safety

There are a number of advantages to applying hierarchical methods to naturalistic data. First, safety data frequently have natural hierarchies. For example, it is well known that males and females have important differences in crash etiology and outcome. One natural hierarchy is thus a gender differentiation with respect to the vehicle operator. This allows one to estimate, separately for males and females, the effect of predictor variables on the dependent variable (crash, near crash, or critical incident). The hierarchy of the model allows one to structure the analysis to reflect more closely what actually happens on the road. If one considers the crash event itself, a natural hierarchy is to consider event and context variables at one level. These represent details of the situation at the immediate time surrounding the crash event. A second level of consideration could be driver attributes, such as years with a driver's license (i.e., driving experience), which operates over a longer time period. These examples illustrate the value of hierarchy: to better represent the reality one is seeking to model.

Yet another advantage of hierarchical models is that they allow flexibility in model structure; this feature is particularly appealing when initially exploring new data sets for which model structure is unknown or not well defined. The Phase 1 report noted that in this context, hierarchical methods allow better exploration of variable effects with alternative structures. This is a common theme touched on in Chapter 3.

Computational advances have facilitated the use of specific hierarchical methods, but from the perspective of this report the principal advantages are flexibility and facilitated data exploration.

There have been several applications of hierarchical models in the analysis of crash frequency and severity level (Goldstein 1995; Rasbash et al. 2002; Sullivan et al. 2007; Wolfinger and O'Connell 1993), and there is an additional application to law enforcement (Yannis et al. 2008). Iterative generalized least squares (Jones and Jorgensen 2003) was employed to fit a binary logistic regression model in which the response variable indicated whether each casualty survived with serious injuries (response = 0) or died (response = 1). When normality is assumed, however, the full Bayesian estimation and empirical Bayes, treating the prior distribution as known, are the same as iterative generalized least squares. Data were analyzed using a hierarchy of casualties (Level 1), within accidents (Level 2), and within municipalities (Level 3).

Another study (Goldstein 1995) compared the efficiency of multilevel logistic models (MLMs) by using maximum likelihood, generalized estimating equation models, and logistic models using simulated French road crash data between 1996 and 2000. The hierarchical structure was modeled as the probability that an occupant (Level 1) in a car (Level 2) during a crash (Level 3) died; the response variable, severity, is treated as binary. The MLM was the most efficient model, while both generalized estimating equation models and logistic models underestimated parameters and confidence intervals (CIs). MLM estimates were obtained through the iterative process of restrictive iterative generalized least squares and penalized quasi-likelihood implemented using MLwiN software (Rasbash et al. 2000).

Hierarchical binomial logistic models (Rasbash et al. 2000) were used to solve the suspected heterogeneity in underlying causal mechanisms associated with different crash types. This study built models for angle, rear-end, and sideswipe crashes with the response variable of crash probabilities and data consisting of two levels: Level 1 consisted of crash-level characteristics, and Level 2 consisted of intersection-level characteristics from 91 two-lane rural intersections in the State of Georgia. The estimation of multilevel binomial logistic models was performed using a GLIMMIX macro in SAS software. The GLIMMIX macro employs a pseudolikelihood (Kim et al. 2007; Wolfinger and O'Connell 1993). This study suggested that incorporating naturalistic driving data (personal characteristics such as driver attentiveness, reaction times, vision, and aggressiveness and vehicle data such as braking characteristics, mass, steering characteristics, and tire condition) into the models may improve prediction accuracy. A hierarchical binomial logistic model has been developed (Wolfinger and O'Connell 1993) with a two-level specification: the response variable is dichotomous for high (fatal or severe) and low (slight or no) injury severity at the individual level (Level 1), and the crash level (Level 2) includes various crash features such as street lighting and road surface conditions.

Multilevel negative binomial (NB) models have been used to capture the spatial variation of the effect of alcohol enforcement intensification (Yannis et al. 2008). The response variable is the number of road accidents with casualties, and the explanatory variables include the alcohol controls in Level 1 and socioeconomic parameters such as population in Level 2 (different regions). The parameter estimates were also obtained through the iterative process of restrictive iterative generalized least squares and quasi-likelihood implemented using MLwiN software (Rasbash et al. 2000).

Summary

In summary, a series of naturalistic driving studies has been conducted for both light-duty and heavy-duty vehicles. Some modeling has been conducted with multiple predictor variables, but no hierarchical models have been applied to the data. Existing hierarchical applications include several studies of injury severity and two studies, one focusing on enforcement and the other on crash type, but no applications to naturalistic data were found. Shankar and associates (2008) argue that naturalistic data analysis would benefit greatly from the application of hierarchical methods because, among other reasons

- The functional form for models is not well documented,
- Sample size limitations may hinder frequentist approaches, and
- Driver, event, and context variables are known in the data, but their interrelationship in crash modeling is largely untested.

Chapters 2 and 3 focus on empirical data analysis: Chapter 2 describes the data and analysis approaches used with each data set, and Chapter 3 summarizes the results of the modeling. Chapter 4 provides an overall summary of the study linked specifically to the five Penn State research questions and their implications for the SHRP 2 Safety program. The last portion of Chapter 4 summarizes lessons learned along with suggestions for future research.

CHAPTER 2

Research Approach

Overview

The analysis of the data provided by VTTI and UMTRI was guided by the following five research questions:

1. *What is the nature of the relationship between events (e.g., crashes, near crashes, incidents), and pre-event maneuvers? What are the contributing driver, environmental factors, and other factors?* There are many findings and implications to share from analyses concerning this question.
2. *What hierarchical structure (statistically speaking), if any, exists in the manner in which these relationships need to be explored?* Two specific hierarchical models are reported, both using VTTI data: one was applied to event modeling and the second to driver-based models. A series of comparisons between hierarchical models, estimated using Bayesian methods and frequentist models, which are estimated using typical maximum likelihood principles, is presented.
3. *What kind of elucidative evidence emerges from the analysis of roadway departure crashes in terms of Questions 1 and 2? Is the illustrative hierarchy of relationships generalizable to other nonintersection crash types such as leading vehicle crashes?* Elucidative evidence refers to evidence of the likely effect of individual predictor variables in modeling event occurrence (including crashes). Surrogates are a special type of variable that have been discussed as a general replacement for crash data; the description and interpretation of Penn State surrogate analyses are contained in the responses to this general question. Exposure requires a predictor variable reflecting time or distance of travel; exposure-based analyses of both data sets are described in Chapter 3. Responses to this question thus provide a summary of the extent to which the modeling results provide guidance on variables to be given priority in future analysis studies.
4. *In terms of elucidative evidence, what types of behavioral correlates emerge? For example, are attitudinal measurements indicative of revealed behavior in terms of headway maintenance and speed reductions?* Several behavioral correlates (also referred to in several SHRP 2 safety symposia as crash predisposition measures) have emerged as factors of interest. The responses to this question describe the work in this area.
5. *If elucidative evidence does in fact emerge in terms of attitudinal correlates and how their interactions vary by context, is it plausible to parse out the marginal effects of various context variables on crash risk by suitable research design?* This question bears directly on the importance of context in the analysis of naturalistic data. This section summarizes the findings and discusses their implications for SHRP 2 scheduled projects (specifically S04 and S08).

Analysis of VTTI Data

Two parallel tracks were pursued in the analysis of the 100-car study data. The first approach modeled the occurrence of each event in detail and focused on understanding the interaction of the many factors that led to event occurrence. This initiative fit nicely with the data provided by VTTI, as it allowed the team to compare events at three levels (summary definitions provided by Dingus, Klauer, et al. 2006):

- Crash event—any contact with an object, either moving or fixed, at any speed, in which kinetic energy is measurably transferred or dissipated;
- Near crash—a circumstance that requires a rapid, evasive maneuver by the subject vehicle, or any other vehicle, to avoid a crash; the maneuver causes the vehicle to approach the limits of its capabilities (e.g., vehicle braking greater than 0.5 g or steering input resulting in lateral acceleration greater than 0.4 g); and
- Crash-relevant incident (in this report referred to as a critical incident)—a circumstance that requires a crash avoidance response on the part of the subject.

Each of these events was identified by VTTI staff as part of the 100-car study, and the three event types were provided to Penn State in response to the team's data request. Penn State developed a structured analysis framework for these event-based data; the model specified driver attributes, the context in which the event occurred (including roadway and environmental variables), and attributes describing details about the event itself, particularly in the few seconds before and during the event. Examples of event-level variables include whether the driver was observed to be distracted just before the event and whether the vehicle crossed over the lane or road edge. One may think of these models as exploring the details of factors associated with the events.

Various model formulations were used to find variables associated with crashes and near crashes, and the attributes of vehicle motion associated with such events (e.g., vehicle over lane or road edge) that could serve as surrogate measures for crashes were investigated. If these event-related measures were shown as being positively associated with crash or near-crash events, they were considered as potential surrogates. The team tested the specific measures available in the data set and attempted to supplement the available vehicle kinematic data by downloading information from the NHTSA website. Unfortunately, kinematic data were only available for a small number of crashes; near crashes and critical incidents were not represented, and this approach was, therefore, abandoned.

One weakness of event analysis is that it precludes the study of drivers who experience none of the three measured events (i.e., the safest drivers). In order to include these drivers, the second analysis track conducted by Penn State with the VTTI data was a series of models of the number of events per driver. Consistent with much of the modeling in the safety field, these analyses were conducted using a set of count regression formulations (e.g., Poisson, negative binomial [NB], and zero-inflated Poisson [ZIP]) that resulted in estimates of the probability of a driver with particular attributes having 0, 1, 2, . . . , *n* events during the year of the 100-car study. These models allowed comparisons to be made across all drivers.

Analysis of UMTRI Data

The UMTRI data consisted of a set of drivers who experienced a series of alerts from onboard systems about potential crashes. Because there were no crashes during the study, the dependent variables used in the analyses were derived from a system designed to detect excessive speed entering a curve (i.e., CSW) and an alert triggered when the subject vehicle deviated from the lane or road edge (i.e., LDW).

After an initial screening of the data, the team decided to focus on the CSW alerts as they provided alert duration data and thus contained more details about the driver response to the alert. Further, the curve speed event was more consistent with the road departure event covered in the VTTI analyses, and it was thought there may be some benefit from the similarity.

Two approaches were taken in the analysis of the UMTRI data. The first was to use a series of piecewise linear models to characterize the nature of the relationship between vehicle kinematics and CSW alert frequency and duration. The interest was in finding which kinematic variables were most correlated with the triggering of the alert. This information was used to gain insight about potential surrogates, under the assumption that the kinematic variables most associated with alert occurrence would be potentially good crash surrogates to consider in subsequent research. A positive association between a kinematic variable and an alert could be an indication of a kinematic variable that might also be associated with (or potentially causing) road departure crash occurrence. While the team acknowledges the nature of this conceptual leap, it was believed that the exploratory nature of the SHRP 2 S01 projects would support this type of analysis. Time–series models of the kinematic data were also attempted, but they did not yield particularly meaningful results and are not discussed in this report.

The second approach taken with the UMTRI data was to use a cohort-based formulation to estimate the probability of a particular number of alerts being triggered for an individual driver (e.g., characterized by gender, years of driving experience, and mileage driven in particular contexts). This formulation is based on actual miles driven under specific environmental and roadway conditions as measured by the CSW–LDW system. Because of the structure of the UMTRI data, the team was able to analyze alert frequency at a very detailed level of exposure.

The team believes the successful estimation of the models predicting the number of alerts using homogeneous trip segments is one of the most important outcomes of the UMTRI modeling effort. This formulation takes advantage of the unique trip-by-trip information in the naturalistic study, along with GIS-related factors coded by UMTRI (such as road type and environmental conditions), to derive a measure of alert frequency in each trip segment. The issue of interest is the ability to truly capitalize not only on the naturalistic driver behavior data, but also on detailed GIS roadway data. Since there is a plan to collect detailed roadway data as part of the scheduled SHRP 2 Safety Project S04, the Penn State team believes this formulation merits consideration for future studies. Even though the models are estimated with alerts, there is a direct parallel to the modeling of crashes or other events of interest. In addition, researchers can flexibly define homogeneous trip segments to match their research needs. The Penn State team discussed this approach during several SHRP 2–sponsored research symposia. The estimated models using the cohort formulation verify the efficacy of this approach; the findings contribute to answering Research Question 3.

Analysis Plan for VTTI Data

Figure 2.1 is an overview of the analysis conducted with the VTTI data, including separate analysis streams for driver-based and event-based models. This differentiation in modeling approach was identified in the proposal for the current study and reflects the authors' view of the most sensible way to approach exploration of the data set. The driver-based models estimate the number of events expected of all drivers in the data set, including drivers with zero observed events. The VTTI data set lacks details concerning all the contexts in which the exposure to crash risk occurred, including the many miles driven with no events. While these data are available in concept within the original 100-car naturalistic data set, they were not provided to the Penn State team.

Conversely, the event-based models estimate the probability of having an event in a given context for drivers with events; these models do not include the best drivers in the data set (that is, those with no events of interest). As discussed in the background section of this report, this omission should not necessarily be the case in all naturalistic data sets. Penn State's proposal anticipated that data would be available for event-based analyses that included nonevent observations (so-called control epochs in the VTTI data), but because these epochs contained no context information, they were useless for modeling. Consistent with the desire to explore multiple analysis paradigms, both driver- and event-based analyses are included in this report.

The driver-based models, both frequentist (those applying classical maximum likelihood principles with asymptotic

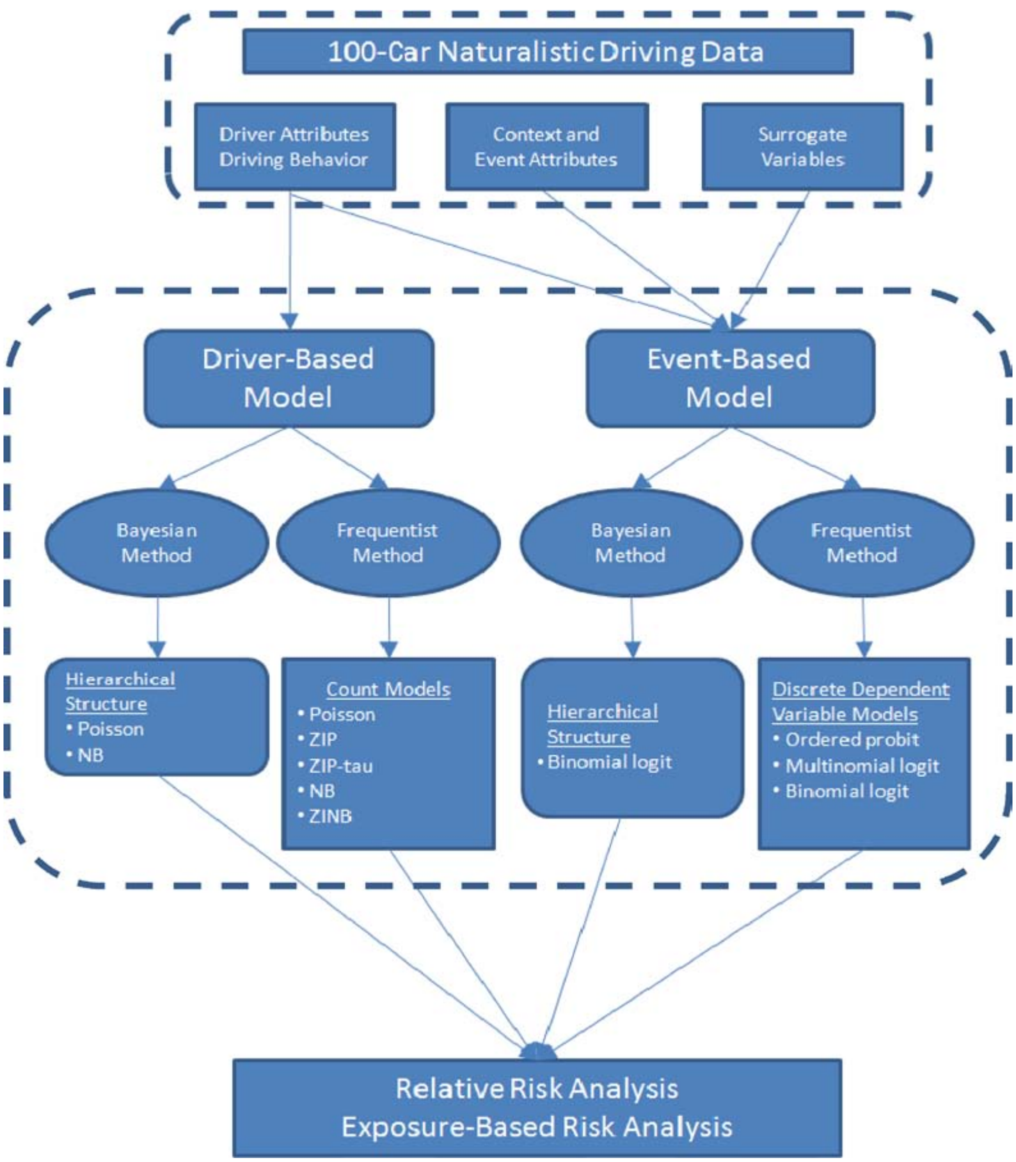

Figure 2.1. Overview of modeling design for VTTI data.

normality assumptions) and Bayesian, are fundamentally count regressions that estimate the probability that a driver with given attributes has 0, 1, 2, 3, . . . , *n* events during the 1-year duration of the Virginia Tech study. Hierarchical models are estimated using Bayesian methods. Count models under consideration during the study include Poisson, NB, ZIP, zero-inflated NB (ZINB), and other models.

Event-based models include (by definition) crashes, near crashes, and critical incidents. As with driver-based models, a range of model forms was considered, including probit, logit (binary, ordered, and multivariate), and hierarchical versions of these using Bayesian formulations.

Figure 2.1 illustrates that the driver-based models use driver attributes in the data set along with driving behavior (classified as crash, near crash, and critical incident). The event-based models use context and event variables as predictors along with driver attributes in a search for valid surrogates for crashes. The last box in the figure calls for the undertaking of relative risk and exposure-based designs. These studies could not be undertaken with measured exposure using the VTTI data because data were not available for the processed data set that Penn State received. Instead, the team used as exposure the subject-estimated annual mileage obtained during driver interviews. Thus, the exposure-based risk analysis shown in Figure 2.1 represents the driver-based VTTI modeling, which included self-reported annual miles driven for each primary driver. Exposure-based models and relative risk analyses using measured travel in different contexts were developed using UMTRI data and are described in that section of the report.

Table 2.1 shows the summary statistics for the driver-based model. Only the statistically significant covariates included in the final models are presented in the table summary. Driver attributes are presented for all the drivers and by gender.

Table 2.1. Summary Statistics for Variables Used in VTTI Driver-Based Models

Driver Group	Variable	Mean	SD	Min	Max
All drivers	Number of events	2.37	5.06	0	28
	Gender (male)	0.60	0.49	0	1
	Drivers with BS degree or above	0.63	0.49	0	1
	Scaled Dula Dangerous Driving Index (DDDI) aggressive driving (AD) score	6.23	1.16	4.0	9.1
	Scaled DDDI risky driving (RD) score	10.38	1.29	7.2	14.9
	Driving experience	18.73	14.41	1.5	52
	Past violations	1.35	1.31	0	5
	Total mileage	**11,369**	**5,726**	**12**	**23,980**
Males	Number of events	1.72	5.03	0	28
	BS degree or above	0.47	0.50	0	1
	Scaled DDDI AD score	3.87	3.30	0	9.1
	Scaled DDDI RD score	6.22	5.16	0	13.1
	Driving experience	12.36	14.73	0	52
	Past violations	0.67	1.09	0	5
	Total mileage	**7,445**	**7,461**	**0**	**23,980**
Females	Number of events	0.65	1.63	0	10
	BS degree or above	0.16	0.37	0	1
	Scaled DDDI AD score	2.36	2.99	0	8.1
	Scaled DDDI RD score	4.16	5.24	0	14.9
	Driving experience	6.37	12.25	0	51
	Past violations	0.67	1.20	0	5
	Total mileage	**3,924**	**6,021**	**0**	**21,564**

Note: SD = standard deviation; Min = minimum; Max = maximum.

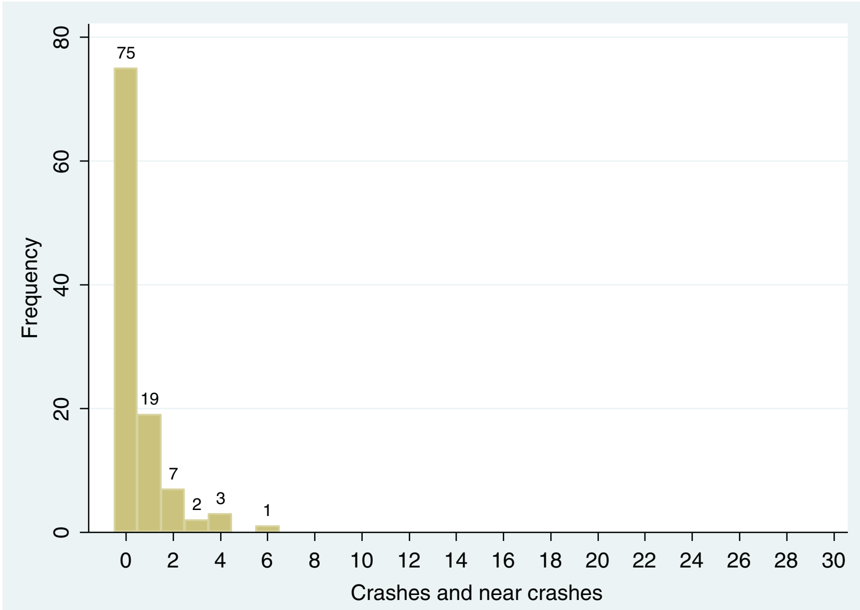

Figure 2.2. Frequency distribution of crashes and near crashes.

Characteristics of Dependent Variables

To discuss the effects of driver attributes on the number of events during a time period, simple relationships must be formulated between them, such as number of events per driver as some function (f) of his or her attributes:

$$\text{number of crashes per person during a period of time} = f(\text{driver attributes})$$

Although the dependent variable on the left-hand side of this simple equation is not difficult to obtain, this subsample of run-off-road–related events includes only 17 crashes, which presents a problem for model significance and also fails to utilize the information from the other 180 events (30 near crashes and 150 critical incidents). Therefore, two distinct dependent variables are considered. The first combines crashes and near crashes; the second combines crashes, near crashes, and critical incidents.

Figure 2.2 presents the frequency distribution for crashes and near crashes for the 1-year study, and Figure 2.3 presents the frequency distribution for all events (that is, crashes, near crashes, and critical incidents). In preliminary modeling,

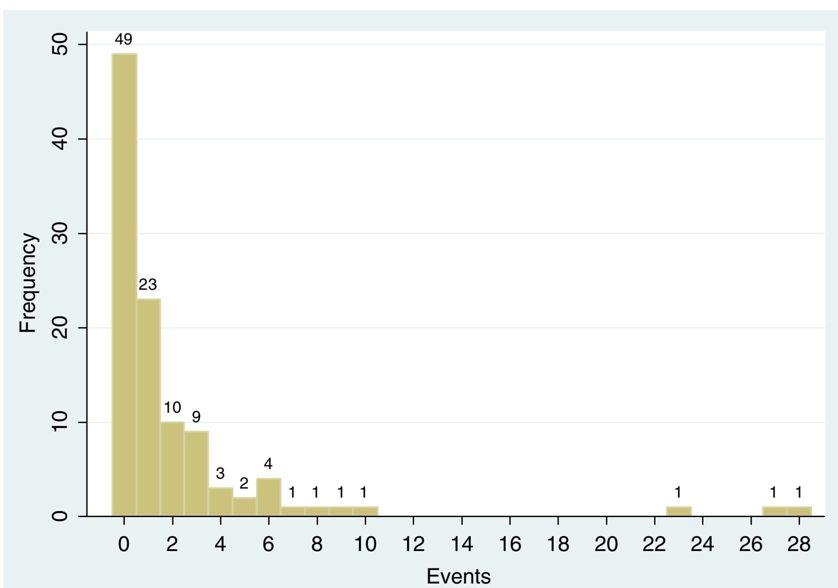

Figure 2.3. Frequency distribution of all events (crashes, near crashes, and critical incidents).

Table 2.2. Correlation Between All Primary VTTI Variables

	EVENTS	PACCIDENT	AGE	GENDER	EDU	DRIYEAR	CRASH	CYEAR	PVIOLATION
EVENTS	1.000	0.072	–0.070	0.050	0.056	–0.039	0.712	0.096	–0.019
PACCIDENT	0.072	1.000	0.400	0.696	0.400	0.228	0.087	0.289	–0.024
AGE	–0.070	0.400	1.000	0.676	0.326	0.758	–0.094	0.085	–0.073
GENDER	0.050	0.696	0.676	1.000	0.572	0.421	0.047	0.201	–0.030
EDU	0.056	0.400	0.326	0.572	1.000	0.182	0.073	0.080	–0.042
DRIYEAR	–0.039	0.228	0.758	0.421	0.182	1.000	–0.067	0.086	–0.069
CRASH	0.712	0.087	–0.094	0.047	0.073	–0.067	1.000	0.089	–0.108
CYEAR	0.096	0.289	0.085	0.201	0.080	0.086	0.089	1.000	0.089
PVIOLATION	–0.019	–0.024	–0.073	–0.030	–0.042	–0.069	–0.108	0.089	1.000

EVENTS = crash, near crash, or critical incident; PACCIDENT = past accident; EDU = educational level; DRIYEAR = number of years driving; CRASH = number of crashes experienced by the subject during the study; CYEAR = vehicle age (years); PVIOLATION = past violation.

models were fit by both dependent variables. The data appear overdispersed, with a large number of zero-event drivers and a few drivers with high counts of events.

Characteristics of Predictor Variables

The first step in checking the predictor variable data was to scrutinize the correlations between all variables in hand to avoid multicollinearity and to get a rough sketch of the overall data, as shown in Table 2.2. The high correlation between driver age and driving experience (0.76) was expected and is summarized in Figure 2.4. Dropping age instead of driving experience may improve model results, since driving experience usually reflects driving skill more directly than age (Shinar 2007).

Education level should not necessarily be considered a continuous variable, as nonlinear relationships may exist. Hence, education levels were initially tested as categorical, as shown in Figure 2.5. The education levels 4, 6, 7, and 8 constituted less than 20% of the total. Education levels were combined to reduce the number of categories to three: some college attended, bachelor's degree, and professional (master's, PhD, or other) degree. Descriptive statistics for the grouped education variable are shown in Table 2.3.

Crash Predisposition Measures

The Dula Dangerous Driving Index (DDDI) was used to measure drivers' self-reported likelihoods of dangerous driving. Each DDDI scale—DDDI total, aggressive driving (AD), negative emotional (NE) driving, and risky driving (RD)—had tests of internal reliability and evidence of construct validity of the scales as part of initial scale development and testing

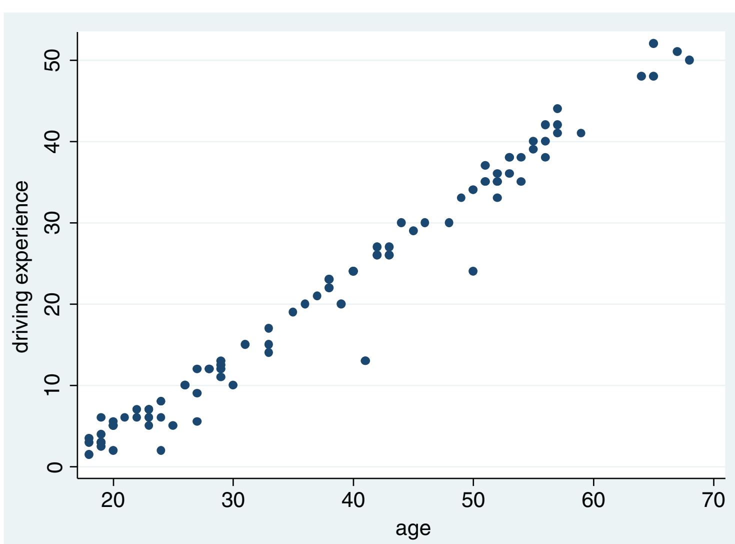

Figure 2.4. Plot of driving experience against driver age.

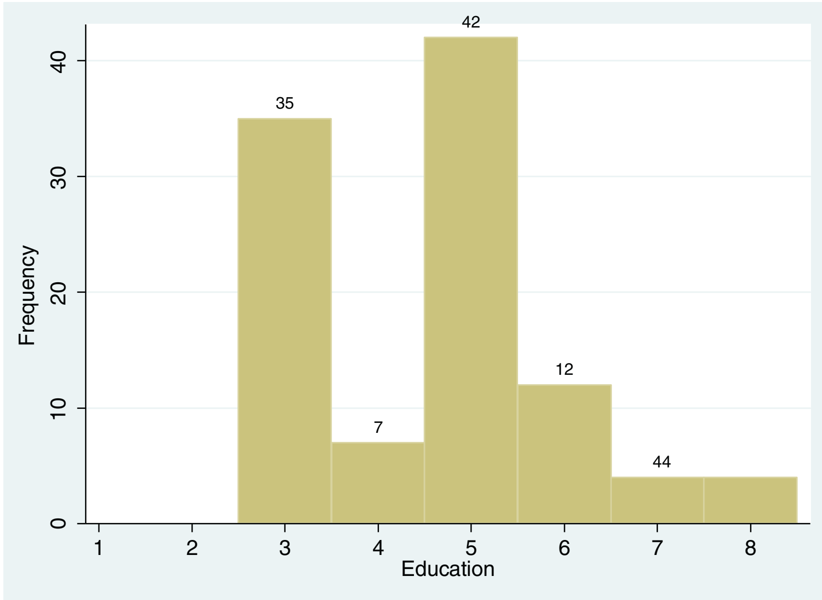

Figure 2.5. Histogram of education levels.

(Dula and Ballard 2003). Participants responded to the items using the following 5-point Likert scale: A = never, B = rarely, C = sometimes, D = often, and E = always. In order to quantify the DDDI, numerical values were assigned to each response (1 through 5 for A through E, respectively). The higher the score per driver, the more dangerous that person's driving behavior was considered to be. To further account for the inconsistency of driver responses, DDDI, AD, RD, and NE were rescaled by deflating the scores for each driver by their specific mean.

Two additional indices measuring driver risk predisposition were included in the analysis (see Dingus et al. 2006 for additional information). The Driver Stress Inventory uses a 10-point Likert scale to obtain information about drivers' general attitudes toward driving on a variety of roadways and in traffic congestion. The Life Stress Inventory contains information about the types of stress that the subject may have experienced in the past year (e.g., ill relative, marital or relationship problems, work performance); evidence suggests that these types of stresses can predispose an individual to have an elevated crash risk. These indices were used in both event- and driver-based models to assess correlations with event occurrence.

Table 2.3. Descriptive Statistics of Grouped Education Levels for 100 Cases

Variable	Mean	SD	Min	Max
Some college	0.4	0.492	0	1
BS degree	0.4	0.492	0	1
Postgraduate college	0.2	0.402	0	1

Driver-Based Analysis

The naturalistic driving environment is a complex web of interactions between various measurable factors that represent both physical infrastructure and human behavior and attributes. Modeling of subjects offers the potential to capture interactions and characterization of heterogeneity in the driving environment. To answer Research Question 1 (What is the nature of the relationship between events [e.g., crashes, near crashes, incidents], and pre-event maneuvers? What are the contributing driver factors, environmental factors, and other factors?), it is crucial to determine what kinds of drivers tend to have higher counts of events that potentially increase crash probability; this focus on driver characteristics leads to the idea of driver-based models. These models are used to examine the relationships between driver attributes and risky driving events by using driver characteristics such as gender, driving experience, and other socioeconomic variables.

A significant amount of research has been conducted on the application of Poisson and NB distributions (Jovanis and Chang 1986; Miaou 1994; Shankar et al. 1995; Poch and Mannering 1996; Milton and Mannering 1996; Lord et al. 2005, 2007) to predict crash frequencies. The Poisson model is only appropriate if the mean and variance of crash frequencies are approximately equal, but the NB model can be applied if the data are overdispersed (i.e., the variance of the data is significantly greater than the mean).

Let λ represent the expected number of events per driver during a period of time as a function of β, a set of estimated parameters, and x_i, a set of crash contributing factors (Jovanis and Chang 1986). The Poisson regression finds maximum likelihood estimates of the β parameters:

$$\ln(\lambda_i) = \beta x_i + \varepsilon_i \tag{1}$$

In traffic safety, crash counts are often overdispersed. Thus, the use of the NB distribution to represent the distribution of crash counts is considered. In Equation 1, exp (ε_i) is a gamma-distributed variable with mean 1 and variance α. If the number of crashes is conditioned on exp (ε_i), the resulting probability distribution for a given count y_i is

$$p(y_i) = \frac{\Gamma(y_i + \theta)}{\Gamma(\theta) y_i!} \left(\frac{\theta}{\theta + \lambda_i}\right)^{\theta} \left(\frac{\lambda_i}{\theta + \lambda_i}\right)^{y_i} \tag{2}$$

where $\theta = \alpha^{-1}$, Γ represents the gamma function, and α is an overdispersion parameter. When $\alpha > 0$, there is overdispersion of the distribution about the mean. The NB distribution can capture overdispersion that occurs as a result of unobserved heterogeneity in crash data.

In the context of crashes, the likelihood of having a large proportion of zero frequencies is high, implying the importance of zero-inflated models. Since their formal introduction by Lambert (1992), the use of these models has grown and can be found in numerous fields. Crash frequencies can be modeled as belonging in two states (Shankar et al. 1997). One state occurs when the entity of interest is inherently safe (theoretical zero-crash state). In the second state, crash frequencies follow some known distribution. ZIP and ZINB models can handle this dual-state phenomenon (Miaou 1994; Shankar et al. 1997).

An overabundance of zeros in a crash count distribution may reflect true lifetime proportions or may arise as a result of partial observability, which poses methodological challenges. Shankar (2004) pointed out that if partial observability and overdispersion are suspected, NB variants of the ZIP model are plausible. With probability p_i, the Poisson process, with probability $1 - p_i$, contributes in combination to the apparent excess zero problems. Thus,

$$p(y_i = 0) = p_i + (1 - p_i)e^{-\lambda_i} \tag{3}$$

and

$$p(y_i = k) = (1 - p_i)\frac{e^{-\lambda_i}(\lambda_i)^k}{k!} \tag{4}$$

where k is the number of crashes with mean λ_i. Combining Equations 3 and 4 provides the ZIP model of crash frequency. In the present case, the team used the logit model to estimate the proportion of observations with a zero frequency and used count regression for the other frequencies.

The Vuong statistic (Vuong 1989) is often used as a measure of whether the ZIP or ZINB model fits the modeled data better. Shankar (1997) proposed a decision guideline for model selection among Poisson, NB, ZIP, and ZINB models using the Vuong statistic and α, as shown in Table 2.4.

Table 2.4. Decision Guideline for Model Selection

	***t*-Statistic for Overdispersion Parameter α**	
Vuong Statistic	**< |1.96|**	**> |1.96|**
< |1.96|	ZIP or NB	NB
> 1.96	ZIP	ZINB

Gender differences in crash experience and etiology are well established in the safety literature. Hierarchical structures such as the one illustrated in Figure 2.6 can be used to explore these differences in the VTTI data. An advantage of hierarchical models is that they can capture driver differences over time and space, depending on how the data are clustered. This allows the

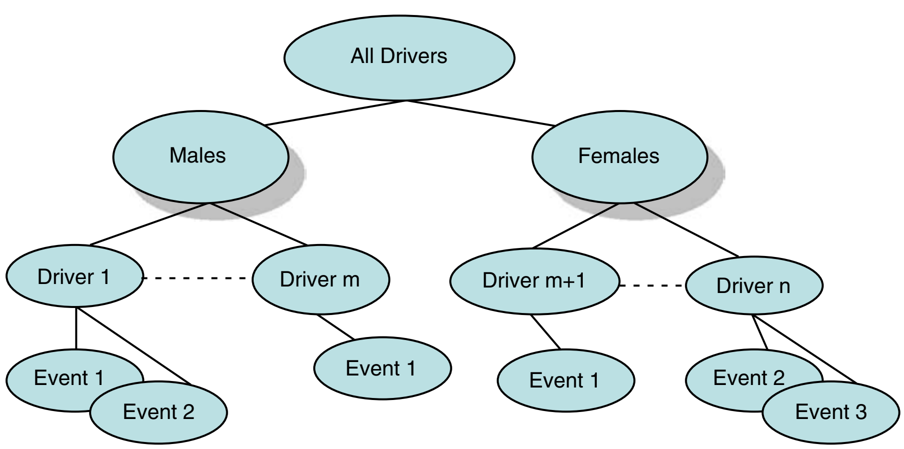

Figure 2.6. Hierarchy for driver-based model.

investigation of individuals (driver-level parameters) themselves as random effects. In the driver-based model, the drivers are the units of analysis, but they are aggregated by gender to explore gender-specific differences between drivers.

Frequentist models, such as the NB, use predictors (such as a male dummy variable or gender-based interaction terms) that estimate the difference in probabilities of having *Y* events between males and females; the hierarchical model in Figure 2.6 shows the effect of a predictor on male crash probability and female crash probability individually. The effect of the attribute on male event probabilities is thus estimated separately from the effect on female probabilities. This produces a model that is much more interpretable.

For the driver-based model, the levels are defined by gender, creating two groups. The response variable is the number of events for each driver participating in the study. The number of events is modeled as a Poisson distribution (see Aguero-Valverde and Jovanis 2008 for a similar formulation):

$$y_{ij} \sim \text{Poisson}(\theta_{ij}) \tag{5}$$

where y_{ij} is the observed number of events for driver *i* of gender *j*, and θ_{ij} is the expected Poisson rate. The Poisson rate is modeled as a function of the covariates following a lognormal distribution, as shown in Equation 6:

$$\log(\theta_{ij}) = \beta_{0j} + \sum_{k=1}^{K} \beta_{jk} X_{ijk} + v_{ij} \tag{6}$$

where

β_{0j} = intercept for gender *j*,
β_{jk} = coefficient for *k* covariate and gender *j*,
X_{ijk} = value of *k* covariate for event *i* of gender *j*, and
v_{ij} = random effects at Level 1.

The random effects represented by v_{ij} capture the extra-Poisson heterogeneity among drivers.

At the second stage, the coefficients (β_{jk}), including the intercepts, are modeled using very noninformative normal priors:

$$\beta_{jk} \sim \text{N}(0, 1000) \tag{7}$$

Now, the prior distribution for the Level 1 random effects is given by

$$v_{ij} \sim \text{N}(0, \tau_v^{-1}) \tag{8}$$

where τ_v is the inverse of the variance, also known as precision. The precision has a gamma prior:

$$\tau_v \sim \text{Gamma}(0.001, 0.001) \tag{9}$$

with a mean of 1 and a variance of 1,000.

The driver-based model seeks to explore the relationship between driver attributes and the expected number of crashes or risky events per driver during a period of time. Count models such as Poisson, NB, ZIP, ZINB, Bayesian multilevel Poisson, and Bayesian multilevel NB models are well suited to handle VTTI data.

Event-Based Analysis

A series of frequentist models was estimated with a wide range of predictor variables. In order to search for consistency in modeled predictor effects, three sets of models were compared. The first set of models is binary logit, in which the base alternative was a critical incident and the other alternative was a crash or near crash; positive parameters in this type of model reflect an increase in the likelihood of moving from a critical incident to a crash or near crash. The second set of models is multinomial logit, in which there are three categories: the baseline is again a critical incident, one category is a near-crash event, and the third is a crash event; positive parameters reflect an increase in the likelihood of the category of outcome compared with a critical incident. A third set of models using an ordered logit formulation was estimated.

Estimation and comparison of all three models was chosen as a basic approach because of the limited experience modeling naturalistic data. The team believes that greater confidence can be accumulated about the utility of the naturalistic driving analysis paradigms if consistent results are obtained across the methods. The three logit models used in this study are very commonly used in transportation analysis (see also Washington et al. 2003).

One can think of these models as reflecting a conditional analysis: a study of factors contributing to crashes and near crashes compared with those contributing to critical incidents, given that an event has occurred. This has a rough parallel in most models of injury severity in crashes: the model is an estimate of injury severity given that a crash has occurred. In both cases, the models do not provide an estimate of the probability of an event (VTTI data) or the probability of an injury (crash severity analysis) because of a lack of appropriate exposure data.

Bayesian models offer considerable additional flexibility in event-based analyses and assist in answering Research Question 2 (What hierarchical structure [statistically speaking], if any, exists in the manner in which these relationships need to be explored?). Figure 2.7 presents the hierarchy for the Bayesian event-based models. Events and their attributes are included at the first level, and driver attributes are represented at the second level.

In the Bayesian formulation, outcomes are modeled as Bernoulli trials, in which a crash or near crash is considered a success:

$$y_{ij} \sim Bernoulli(p_{ij}) \tag{10}$$

where p_{ij} is the probability of success for event *i* of driver *j* defined by Equation 11.

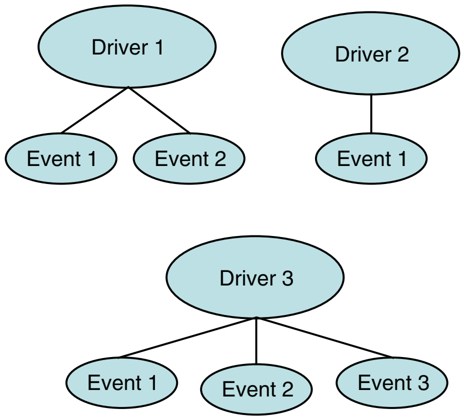

Figure 2.7. Hierarchy used for Bayesian event-based models with VTTI data.

$$\operatorname{logit}\left(p_{ij}\right)=\log\left(\frac{p_{ij}}{1-p_{ij}}\right)=\alpha+\sum_{k=1}^{K}\beta_k X_{ijk}+\sum_{l=1}^{L}\gamma_l Z_{jl} \tag{11}$$

where

α = intercept,
β_k = coefficient for k Level 1 (event level) covariate,
X_{ijk} = k covariate for i event of j driver,
γ_l = coefficient for l Level 2 (driver level) covariate, and
Z_{jl} = l covariate for driver j.

Very noninformative normal priors are used for all of the coefficients including the intercept [$\sim N$ (0, 1000)].

Analysis Plan for UMTRI Data

The UMTRI data provided the opportunity to work with a rich set of vehicle kinematic variables, but the limitation was that there were no crashes in the data set. As a result, the Penn State team used the occurrence of an alert (either an LDW or CSW) as the dependent variable. This decision immediately created the challenge of learning about this measure, as there is a very limited literature of its analysis. As a result, a building block approach was taken with the alert data. Using the UMTRI technical report as an initial guide, the team first explored general attributes of the data before settling on an analysis plan. As the team worked with the data, it became clear that the CSW alerts, because of their relation to road curvature, had a closer association with lateral and longitudinal vehicle kinematics. The LDW alerts were more closely associated with vehicle position within a lane. The team thus chose to focus on CSW alerts analyses, given the more likely application to road departure crashes, at least those occurring on curves.

Rather than select specific model forms, which could have confounded the identification of promising variables, the team chose a piecewise linear modeling approach and explored several formulations. The goal in these analyses was to identify factors that were associated with the triggering of alerts in the hope of identifying these variables as good prospects for future study in SHRP 2. It was also expected that a cohort-based formulation might have particular advantages for future data analysis. This twofold objective, the identification of promising variables and testing of alternative analysis structures, particularly the cohort formulation, motivated the research plan.

Kinematic Models

The Penn State team conducted a careful structured analysis of the kinematic data received from UMTRI by using the steps outlined in Figure 2.8. Initially, the data received from UMTRI needed to be organized in such a way that event counts could be defined. Alert durations and alert counts, both aggregated

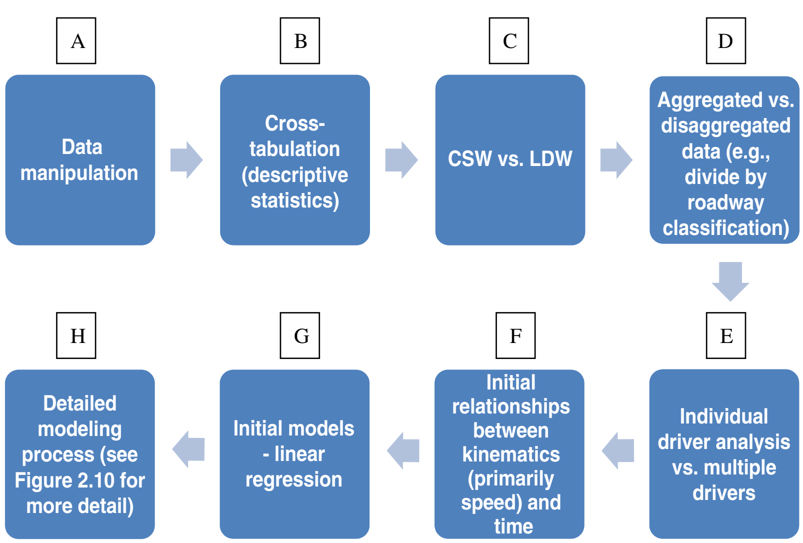

Figure 2.8. Flow chart showing analysis process for UMTRI data.

and by driver, needed to be obtained (Step A). Several cross-tabulations were performed to determine the number of alerts occurring under certain circumstances (Step B). Kinematic analysis depended on which type of alert contained the data that were usable for modeling.

All LDW alerts in the data set essentially had no duration (they were listed as instantaneous). Kinematics do not play a crucial role in the occurrence of LDW alerts; they are solely based on the threshold of lateral displacement from the center-line of the travel lane. In the case of LDW alerts, there would likely be a much weaker relationship between various vehicle kinematics and longitudinal speed than was seen during preliminary analysis of CSW alert data. Kinematics were deemed to be a desirable attribute of potential crash surrogates, a focus of S01 activities, so it was decided that CSW alerts would be used for initial analysis (Step C). LDW alerts were included as part of cohort-based model development.

The overall approach to the CSW models involved a consideration of the relationship between longitudinal speed and combinations of exogenous and kinematic factors. However, the primary focus was on the effect of changes in vehicle kinematics on vehicle longitudinal speed. The research team focused on CSW alert and vehicle kinematics to learn more about one type of road departure event: those occurring on curves. Since kinematic data were not available through the VTTI data set, the team hoped to learn more about this issue through the UMTRI data analyses. The proposal stated that the goal was exploratory and explicitly said the team would not "compare models" across data sets. The team understood the data to be different and has sought to take advantage of those differences.

Early in the analysis, through careful review of the UMTRI report and communication with UMTRI researchers, it became clear that driver adaptation to CSW needed to be considered. One can think about driver behavior on curves by constructing a diagram such as Figure 2.9. A theoretical baseline is considered constant deceleration while approaching and moving through the curve (solid line with triangles). Another possible driver behavior is to approach a curve while slowing and then to decelerate as needed to safely navigate the curve (dash–dot line with crosses). One possible driver adaptation to the CSW is to approach the curve without decelerating, waiting for the system to provide an alert, and then decelerating more rapidly (dashed line with diamonds). With the CSW engaged, drivers may approach curves at a constant speed until the alert is triggered, but the prealert speed may be lower than that observed during the RDCW-disabled period (dotted line with squares). Only empirical testing will determine which of the suggested models is observed.

To determine an initial relationship between kinematics and time, individual drivers were randomly sampled, and longitudinal speed was taken as the first kinematic variable (Steps D and E of Figure 2.8). The relationship between speed and time was deemed useful for modeling. Additional relationships were developed between speed and other kinematic variables in the initial models (Step F). Other models were subsequently developed to look at the relationship in a more detailed manner, mostly on the aggregate level (Steps G and H).

To continue to learn more about the data set, the team evaluated candidate dependent measurements, including alert frequency, alert duration, and alerts per trip for all drivers. Alert counts were then divided by alert ID, from which average durations were found. Alert counts, classified by alert ID, were then obtained for different exogenous factors (e.g., headlamp status, turn-signal status, and windshield wiper status) provided in the RDCW alert data set. Examples of the organized data can be seen in Tables 2.5 through 2.8.

The total number of alerts, 2,605, was used as the sample size for analysis. However, this count only provided a glimpse of the relationship between the exogenous factors and alert frequency and duration and did not take into account the kinematic variables included in the data set. Different time periods of observation in the data set were identified by using a combination of headlight status and solar zenith angle (i.e., the angle of the sun relative to the horizon, which was used to identify light or darkness conditions for driving). This variable was transformed into binary form and included in all models.

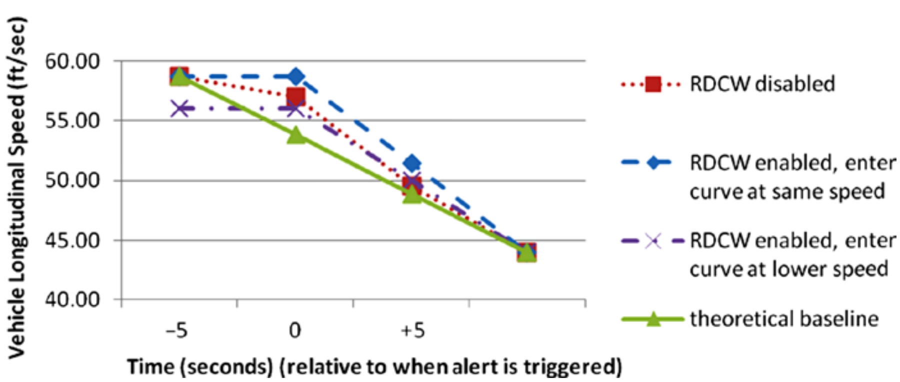

Figure 2.9. Conceptual model of driver speed adaptation to CSW alerts.

Table 2.5. CSW Alert Counts and Duration Summary (by ID)

Number and Duration of CSW Alerts	Alert Type		All CSW Alerts
	Cautionary	Imminent	
No. of alerts (%)	1,867 (71.67)	738 (28.33)	2,605 (100)
Average duration (s)	2.242	4.632	2.919
SD of duration (s)	7.923	15.494	—

Table 2.6. CSW Alert Counts by Wiper Status

	No. of CSW Alerts by Type		
Wiper Status	Alert 5 Cautionary	Alert 6 Imminent	Total
0 (Off)	1,730	683	2,413
1 (Low)	43	15	58
2 (High)	6	4	10
4 (Intermittent)	88	36	124
Total	1,867	738	2,605

Table 2.7. CSW Alert Counts by Headlamp Status

Headlamp Status	No. of CSW Alerts by Type		Total
	Alert 5 Cautionary	Alert 6 Imminent	
0 (Off)	1,234	491	1,725
2 (Low)	609	236	845
3 (High)	24	11	35
Total	1,867	738	2,605

Weather status was estimated by using wiper status. Table 2.9 shows the counts of alerts (regardless of ID) classified between daylight status, wiper status, and headlamp status.

The Penn State team sought to explore changes in driver behavior through curves with and without the CSW alert system activated. To determine if differences existed in driver longitudinal and lateral speed behavior between the first week (no alerts provided to driver) and Weeks 2 to 4 (alerts provided), 41 drivers were sampled by plotting longitudinal and lateral speed versus time for one randomly selected alert. Sampling individual drivers with randomly selected alerts provided some evidence of differences between Week 1 and Weeks 2 to 4, so speed changes for all drivers were modeled to determine differences between the two time periods. One may think of these analyses as part of Steps E, F, and G in Figure 2.8. It was noted that driver adaptation was possibly occurring, so this consideration was included in all the analyses conducted.

Table 2.8. CSW Alert Counts by System State (Disabled or Enabled)

(a) With Duration Summary

Number and Duration of CSW Alerts	System State		All CSW Alerts
	Disabled	Enabled	
No. of alerts (%)	694 (26.64)	1,911 (73.36)	2,605 (100)
Average duration (s)	3.057	2.869	2.919
SD of duration (s)	12.495	10.759	—

(b) By Headlamp Status

Headlamp Status	System State		Total
	Disabled	Enabled	
0 (Off)	462	1,263	1,725
2 (Low)	220	625	845
3 (High)	12	23	35
Total	694	1,911	2,605

(c) By Wiper Status

	System State		
Wiper Status	Disabled	Enabled	Total
0 (Off)	638	1,775	2,413
1 (Low)	11	47	58
2 (High)	3	7	10
4 (Intermittent)	42	82	124
Total	694	1,911	2,605

Table 2.9. Count of Alerts by Daylight, Wiper, and Headlamp Status

Wiper Status	Daylight				Total
	Light		Dark		
	Headlamps				
	On	Off	On	Off	
0 (Off)	280	1,630	496	7	2,413
1 (Low)	21	22	13	2	58
2 (High)	5	3	2	0	10
4 (Intermittent)	48	61	15	0	124
Total	354	1,716	526	9	2,605

The overall approach to the modeling considered the relationship between longitudinal speed and combinations of exogenous and kinematic factors. The motivation was to seek kinematic variables that may be particularly good surrogates (e.g., by identifying variables that may influence longitudinal speed entering curves). In order to explore a variable's utility as a surrogate, it was necessary to understand how it was related to other vehicle kinematics. Early data analysis quickly led to a focus on longitudinal speed entering curves as a promising variable. Much of the piecewise linear modeling focused on this variable.

The modeling process is shown in Figure 2.10. Regimes refer to the number of separate pieces of longitudinal speed that were modeled. Single-regime models assumed that one equation could be used to estimate the longitudinal speed for each tenth of a second from 5 s before the alert occurred through the completion of the alert. Two-regime models broke this time into two pieces or regimes, and three-regime models into three separate pieces. The first step was to determine basic relationships between longitudinal speed, time, certain important exogenous factors, and several kinematic variables. It was deemed necessary to use interaction terms for the kinematic variables to see how they affected each other.

The first models developed were single-regime linear regression models using both main effect kinematic variables and first-order interaction terms among the variables. After studying the results of these models, the team determined that kinematic characteristics changed too much over time, so time periods for each model were divided into two, and eventually three, regimes. The goal here was to better understand the nature of the relationship between the triggering of the alert for excessive curve speed entry and the detailed vehicle dynamics. Recall that the alert is being used as a substitute for a crash event. The team wanted to better understand how the kinematic measures interacted before the alert and during the alert as the driver responded. Recall also that kinematic data were not obtained from VTTI; the time-based regimes were an attempt to learn something about kinematics during an event of interest and model it.

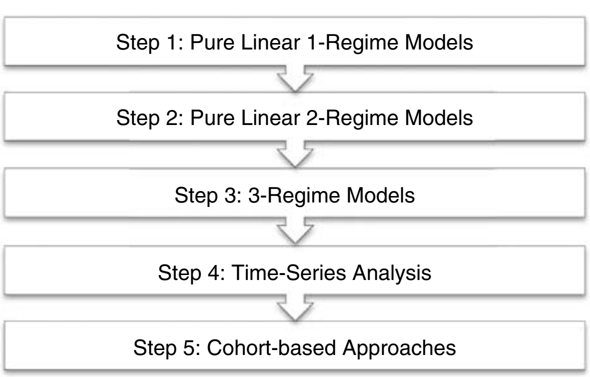

Figure 2.10. Flow chart depicting modeling of longitudinal speed entering curve in UMTRI data.

In all these models, variables were input at their collected rate of 10 Hz. This means that there is serial correlation present within the data because multiple observations were made on the same event, closely spaced in time. The Penn State team recognized this as an analysis issue, but sought to learn more about the nature of the interrelationship between the kinematic variables. A thorough review of the literature revealed few useful references that could improve the analysis plan. The team was initially more interested in the associations between variables during alert events and less concerned about the variables' statistical significance. For this reason, modeling activities were continued, but the team was mindful of the need to return to the correlation issue in the future.

Similar arguments can be made about the endogeneity present in the data. Many of the kinematic variables are the result of driver perception and feedback while negotiating and approaching curves. As such, they are part of the same physical and psychological process undertaken by the driver during the driving task; they are not independent predictor variables. Issues of endogeneity and serial correlation were explored through the unsuccessful testing of time–series models. Cohort-based approaches were used to fully integrate exposure with event occurrence.

Cohort-Based Approaches

Research Questions 1 to 3 focus on the identification of surrogates and the evaluation of behavioral and contextual correlates of surrogates. Research Question 5 (If elucidative evidence does emerge in terms of attitudinal correlates and how their interactions vary by context, is it plausible to parse out the marginal effects of various context variables on crash risk by suitable research design?) focuses on the definition of relative risk and exposure-based risk, especially vis-à-vis context. Such assessments can be made using a cohort-based approach. The cohort design can be used to formulate an exposure-based model relating potential risk factors to several possible outcomes. The cohort design is well suited to account for measures of exposure such as time at risk or distance traveled under specific driving conditions. These measures can be readily obtained from naturalistic studies if the data are suitably structured after collection.

Data Structure

The proposed cohort analyses begin, in general, with a driver as the unit of analysis; the driver is followed over multiple

Table 2.10. Initial Cohort-Based Data Structure for UMTRI CSW Alerts

Outcome (0/1)	Length	Time	Context (all context variables needed)			Driver Attributes (as many as needed)

trips throughout the course of the study. Each driver is associated with specific attributes that are constant, such as age, gender, driver attitudinal measures, and vehicle type and characteristics. Other variables can change throughout the course of the study and within each trip (e.g., roadway type, roadway characteristics, environmental factors, driver distraction, driver impairment, and driving speed). A subset of these variables can be used to define a cohort—that is, a trip segment that is homogeneous with respect to the variables of interest. Travel time and/or distance may thus be accumulated during the study for individual drivers in each defined context (i.e., a homogeneous trip segment).

Travel undertaken in each homogeneous trip segment would then be aggregated to determine total exposure and total number of events within a cohort. A cohort thus represents a set of drivers, by type, who experience travel over defined homogeneous trip segments characterized by time or distance of travel. The number of events of interest (e.g., crashes or other events) occurring for a cohort is thus accumulated across identical drivers, retaining the number of events and/or the time between events for each driver.

This concept is illustrated in Table 2.10, a sample table that contains the initial cohort data in which a particular outcome (i.e., an event or nonevent) occurs after some period of time or length of travel. The context and driver attributes are selected by the researcher depending on the issues to be explored.

Table 2.11, a second sample table, shows how the individual outcomes can be grouped, if needed, for each cohort. Each unique combination of driver and context variables is now listed with the cumulative time or distance—a measure of exposure to risk. Notice that each cohort includes the sum of individual trip segments and their outcomes. Each driver's outcomes are aggregated and matched to context. The sum of the "1" values in the Outcome column in Table 2.10 are the number of events of interest for that cohort. The length and time variables from Table 2.10 are also summed to derive the total time and total distance for each driver in each context. Note that the trips without an event of interest (i.e., outcome zero) are summed and included in the corresponding total distance and time for each cohort. A dummy variable designation is employed for the context variables and driver attributes.

This structure allows for the computation of exposure-based risk (addressing Research Question 3). At the choice of the analyst, the cohort data can remain in the individual trip form of Table 2.10 with essentially a 0, 1 outcome (and the implied use of categorical dependent variables to be modeled), or the data can be aggregated as suggested in Table 2.11, and a count regression approach can be used to estimate the number of events in each cohort.

There is also flexibility in the definition of the events of interest. In the present study, alerts were used as the dependent variable because they were available in the UMTRI data set. In the

Table 2.11. Summed Event Outcomes by Context and Driver Attributes with Exposure Measures

No. of Outcomes (count)	Total Length (vehicle mile)	Total Time (vehicle hour)	Context (all context variables needed)			Driver Attributes (as many as needed)

larger data set available in SHRP 2 Safety Project S07, crashes could certainly be used, or even crashes of a specific type such as roadway departure or intersection-related crashes.

For the UMTRI data, the Penn State team demonstrates both the categorical-outcome models using logistic regression and survival models along with count regression models using data formulated as shown in Table 2.11. This data set allows the estimation of a count regression model of the probability of having Y events during the study period. An estimation is formulated of the mean of the underlying probability distribution (such as Poisson and NB).

Once this basic structure is obtained, several additional analyses may be undertaken beyond the basic count regression:

1. The week of the study can be included from the UMTRI data to test driver adaptation with the RDCW system installed. This information is not required with general naturalistic driving data, but it provides an opportunity to test for learning. As with context and driver attributes, there would be dummy variables to describe each week of the study, with Week 1 as the baseline.
2. A case–control formulation is possible from the basic data (Table 2.10); each row in the data set is either a case ($Y = 1$) or a control ($Y = 0$). While there may be large variability in the data, such a model can be formulated and estimated using different random samples of controls.

Table 2.12 shows the structure used to define cohorts in the UMTRI data on the basis of rural or urban settings, roadway functional classification, ramp presence, and lighting conditions. The model would provide an estimate (through parameter values) of the effect of each of these factors on the outcome measure (e.g., CSW or LDW alerts).

Risk and Relative Risk

Risk was calculated for each cohort as the number of total alerts, CSW alerts, and LDW alerts (each analyzed separately) divided by the total exposure (time) to the specific environment (see Table 2.13 for example calculations of risk and relative risk, using RDCW system data). In particular, each cohort was compared with the baseline cohort (Cohort 3) to determine the relative risk. Relative risk (RR) is a ratio of the probability of the event occurring in the exposed group versus a nonexposed group:

$$RR = \frac{P_{\text{exposed}}}{P_{\text{non-exposed}}}$$

These basic calculations aimed to satisfy the proposal commitment to estimate exposure-based risk and relative risk.

Cohort-Based Count Regression and Event Analysis

The initial analysis involved the use of traditional NB count regressions to show how both context- and driver-related variables affect the likelihood of alert occurrence. The first set of sample models included all drivers, but the data were disaggregated by the roadway functional classification used for the homogeneous trip segment data set. Two functional classes were used to illustrate these models: Functional Class 1—limited access (limited-access freeway); and Functional Class 3—nonlimited access (minor surface). The response variable was either the number of LDW or CSW alerts (not the total number of alerts). The eight models estimated are summarized in Table 2.14.

The initial predictors considered included the following:

- Context variables: ramp (for nonlimited access), urban/rural, day/night, wet/dry (based on windshield wiper use), and RDCW disabled/enabled status; and
- Driver variables: gender, education, years of driving experience, last year's mileage driven, use of glasses or contacts, and smoker/nonsmoker.

It may not be appropriate or useful to include kinematic variables in a specification or model of this type because the averages of kinematic values over homogeneous trip segments may not represent what is actually happening during the course of the traversal of the entire segment. In addition, the aggregation of average values for each kinematic variable may be problematic because they may be affected by factors that could be used to redefine homogenous trip segments, but they are not included in the data set. For example, suppose additional variables were included in the data set, including curve and tangent presence, presence of an intersection, traffic volume, and grade. These could be used to redefine homogeneous trip segments. Once the segments are redefined, average speeds would more accurately reflect travel speeds on each segment.

The UMTRI event-based models used binary logit structure, including single- and multilevel specification, similar to those used in the VTTI event-based models.

Summary

This chapter describes the model structures applied to the VTTI and UMTRI data sets in order to identify prospective views of methodological paradigms. For each data set, the Penn State team described why it developed the specific model paradigms and how the paradigms related to the proposed research questions. The next chapter presents a summary of the results of the empirical investigation.

Table 2.12. Cohort Structure

Setting	Functional Classification	Ramp	Day/ Night	Cohort No.	Total Alerts	CSW	LDW	Segment Time (vehicle hour)	Segment Distance (mi)
Urban	FC1, limited access	yes	day	1	NA	NA	NA	NA	NA
			night	2	NA	NA	NA	NA	NA
		no	day	3	2,399	146	2,253	4,778.16	31,379.245
			night	4	945	49	896	1,466.784	9,693.391
	FC2, limited access	yes	day	5	NA	NA	NA	NA	NA
			night	6	NA	NA	NA	NA	NA
		no	day	7	1,074	56	1,018	1,494.564	9,444.842
			night	8	526	23	503	572.268	3,827.514
	FC3, limited access	yes	day	9	NA	NA	NA	NA	NA
			night	10	NA	NA	NA	NA	NA
		no	day	11	28	3	25	42.94	228.528
			night	12	12	2	10	18.264	116.138
	FC1, nonlimited access	yes	day	13	248	134	114	153.08	760.026
			night	14	100	36	64	45.076	264.462
		no	day	15	0	0	0	0.821	1.446
			night	16	0	0	0	1.84	1.537
	FC2, nonlimited access	yes	day	17	292	178	114	119.327	548.112
			night	18	121	58	63	38.313	198.587
		no	day	19	324	124	200	975.078	3,346.723
			night	20	93	24	69	215.723	818.44
	FC3, nonlimited access	yes	day	21	448	353	95	203.914	801.668
			night	22	117	77	40	60.693	264.53
		no	day	23	902	184	718	3,750.3	10,563.31
			night	24	388	52	336	994.524	3,264.723
	FC4, nonlimited access	yes	day	25	234	204	30	103.732	320.584
			night	26	74	61	13	34.36	123.676
		no	day	27	1,653	228	1,425	5,972.7	17,149.845
			night	28	640	65	575	1,569.57	4,797.187
	FC5, nonlimited access	yes	day	29	5	4	1	3.19	7.571
			night	30	0	0	0	0.516	2.028
		no	day	31	234	198	36	3,583.62	5,779.455
			night	32	77	59	18	1,086.198	1,684.258
	No functional class	no	day	65	NA	NA	NA	NA	NA
			night	66	NA	NA	NA	NA	NA
			day	67	400	0	400	5,694.9	21,623.717
			night	68	135	0	135	1,230.654	3,559.623

(continued on next page)

Table 2.12. Cohort Structure (continued)

Setting	Functional Classification	Ramp	Day/ Night	Cohort No.	Total Alerts	CSW	LDW	Segment Time (vehicle hour)	Segment Distance (mi)
Rural	FC1, limited access	yes	day	33	NA	NA	NA	NA	NA
			night	34	NA	NA	NA	NA	NA
		no	day	35	163	9	154	680.61	5,053.614
			night	36	87	2	85	136.287	1,025.625
	FC2, limited access	yes	day	37	NA	NA	NA	NA	NA
			night	38	NA	NA	NA	NA	NA
		no	day	39	169	6	163	433.368	3,079.085
			night	40	63	0	63	101.293	755.244
	FC3, limited access	yes	day	41	NA	NA	NA	NA	NA
			night	42	NA	NA	NA	NA	NA
		no	day	43	0	0	0	0.141	0.773
			night	44	NA	NA	NA	NA	NA
	FC1, nonlimited access	yes	day	45	11	8	3	3.223	18.804
			night	46	0	0	0	0.003	0.026
		no	day	47	NA	NA	NA	NA	NA
			night	48	NA	NA	NA	NA	NA
	FC2, nonlimited access	yes	day	49	7	6	1	2.653	12.312
			night	50	0	0	0	0.674	3.746
		no	day	51	50	0	50	244.535	1,471.637
			night	52	6	0	6	13.984	89.434
	FC3, nonlimited access	yes	day	53	23	21	2	6.844	20.91
			night	54	2	1	1	0.554	2.293
		no	day	55	215	44	171	644.496	3,331.434
			night	56	41	6	35	85.865	433.793
	FC4, nonlimited access	yes	day	57	25	22	3	7.433	26.621
			night	58	9	7	2	2.717	9.983
		no	day	59	407	90	317	872.292	4,041.153
			night	60	165	13	152	250.577	1,106.478
	FC5, nonlimited access	yes	day	61	0	0	0	0.246	0.804
			night	62	NA	NA	NA	NA	NA
		no	day	63	52	46	6	300.024	929.395
			night	64	5	5	0	67.803	213.677
	No functional class	no	day	69	NA	NA	NA	NA	NA
			night	70	NA	NA	NA	NA	NA
			day	71	63	0	63	355.584	1,672.599
			night	72	2	0	2	25.996	36.334

Table 2.13. Risk and Relative Risk for Each Cohort

Cohort	Risk (time)		RR (time)		Risk (distance)		RR (distance)	
	CSW	LDW	CSW	LDW	CSW	LDW	CSW	LDW
3	0.0306	0.4715	1.000	1.000	0.0047	0.0718	1.000	1.000
4	0.0334	0.6109	1.093	1.296	0.0051	0.0924	1.086	1.287
7	0.0375	0.6811	1.226	1.445	0.0059	0.1078	1.274	1.501
8	0.0402	0.8790	1.315	1.864	0.0060	0.1314	1.292	1.830
11	0.0699	0.5822	2.287	1.235	0.0131	0.1094	2.821	1.524
12	0.1095	0.5475	3.584	1.161	0.0172	0.0861	3.701	1.199
13	0.8754	0.7447	28.648	1.579	0.1763	0.1500	37.894	2.089
14	0.7986	1.4198	26.137	3.011	0.1361	0.2420	29.257	3.371
15	0.0000	0.0000	0.000	0.000	0.0000	0.0000	0.000	0.000
16	0.0000	0.0000	0.000	0.000	0.0000	0.0000	0.000	0.000
17	1.4917	0.9554	48.819	2.026	0.3248	0.2080	69.798	2.897
18	1.5138	1.6443	49.544	3.487	0.2921	0.3172	62.772	4.418
19	0.1272	0.2051	4.162	0.435	0.0371	0.0598	7.963	0.832
20	0.1113	0.3199	3.641	0.678	0.0293	0.0843	6.303	1.174
21	1.7311	0.4659	56.655	0.988	0.4403	0.1185	94.639	1.650
22	1.2687	0.6591	41.520	1.398	0.2911	0.1512	62.561	2.106
23	0.0491	0.1915	1.606	0.406	0.0174	0.0680	3.744	0.947
24	0.0523	0.3379	1.711	0.717	0.0159	0.1029	3.423	1.433
25	1.9666	0.2892	64.361	0.613	0.6363	0.0936	136.766	1.303
26	1.7753	0.3784	58.102	0.802	0.4932	0.1051	106.007	1.464
27	0.0382	0.2386	1.249	0.506	0.0133	0.0831	2.857	1.157
28	0.0414	0.3663	1.355	0.777	0.0135	0.1199	2.912	1.669
29	1.2540	0.3135	41.041	0.665	0.5283	0.1321	113.546	1.840
30	0.0000	0.0000	0.000	0.000	0.0000	0.0000	0.000	0.000
31	0.0553	0.0100	1.808	0.021	0.0343	0.0062	7.363	0.087
32	0.0543	0.0166	1.778	0.035	0.0350	0.0107	7.529	0.149
35	0.0132	0.2263	0.433	0.480	0.0018	0.0305	0.383	0.424
36	0.0147	0.6237	0.480	1.323	0.0020	0.0829	0.419	1.154
39	0.0138	0.3761	0.453	0.798	0.0019	0.0529	0.419	0.737
40	0.0000	0.6220	0.000	1.319	0.0000	0.0834	0.000	1.162
43	0.0000	0.0000	0.000	0.000	0.0000	0.0000	0.000	0.000
45	2.4823	0.9309	81.240	1.974	0.4254	0.1595	91.438	2.222
46	0.0000	0.0000	0.000	0.000	0.0000	0.0000	0.000	0.000
49	2.2616	0.3769	74.016	0.799	0.4873	0.0812	104.737	1.131
50	0.0000	0.0000	0.000	0.000	0.0000	0.0000	0.000	0.000
51	0.0000	0.2045	0.000	0.434	0.0000	0.0340	0.000	0.473
52	0.0000	0.4290	0.000	0.910	0.0000	0.0671	0.000	0.934
53	3.0686	0.2922	100.425	0.620	1.0043	0.0956	215.847	1.332

(continued on next page)

Table 2.13. Risk and Relative Risk for Each Cohort (continued)

Cohort	Risk (time)		RR (time)		Risk (distance)		RR (distance)	
	CSW	LDW	CSW	LDW	CSW	LDW	CSW	LDW
54	1.8044	1.8044	59.052	3.827	0.4361	0.4361	93.724	6.074
55	0.0683	0.2653	2.234	0.563	0.0132	0.0513	2.839	0.715
56	0.0699	0.4076	2.287	0.864	0.0138	0.0807	2.973	1.124
57	2.9598	0.4036	96.867	0.856	0.8264	0.1127	177.618	1.570
58	2.5767	0.7362	84.329	1.561	0.7012	0.2003	150.705	2.790
59	0.1032	0.3634	3.377	0.771	0.0223	0.0784	4.787	1.093
60	0.0519	0.6066	1.698	1.286	0.0117	0.1374	2.525	1.913
61	0.0000	0.0000	0.000	0.000	0.0000	0.0000	0.000	0.000
63	0.1533	0.0200	5.018	0.042	0.0495	0.0065	10.638	0.090
64	0.0737	0.0000	2.413	0.000	0.0234	0.0000	5.029	0.000
67	0.0000	0.0702	0.000	0.149	0.0000	0.0185	0.000	0.258
68	0.0000	0.1097	0.000	0.233	0.0000	0.0379	0.000	0.528
71	0.0000	0.1772	0.000	0.376	0.0000	0.0377	0.000	0.525
72	0.0000	0.0769	0.000	0.163	0.0000	0.0550	0.000	0.767

Table 2.14. Division of Single-Level Model Types by Exposure Measure, Functional Class, and Alert Type

Distance as exposure	Functional Class 1—limited access	CSW	1
		LDW	2
	Functional Class 3—nonlimited access	CSW	3
		LDW	4
Time as exposure	Functional Class 1—limited access	CSW	5
		LDW	6
	Functional Class 3—nonlimited access	CSW	7
		LDW	8

CHAPTER 3

Data Description and Modeling Results

This chapter describes each of the data sets analyzed by the Penn State team along with the results of the analyses performed on those data. The chapter begins with a discussion of the VTTI data and models followed by a discussion of the data from UMTRI and the models estimated from those data.

VTTI Driver-Based Data and Models

Table 3.1 provides an overview of the data provided by VTTI that were used by the Penn State team. As noted in Chapters 1 and 2, the sample size was much less than the team expected, and in some cases, variables requested were not provided.

Events are identified by using a screening technique developed by VTTI (Dingus et al. 2006); the Penn State team received event data and other attributes directly from VTTI. Review of video in and around the vehicle was used by VTTI to develop variables typically recorded on police accident reports (e.g., crash type and assessments of precipitating event, driver distraction, and impairment), but in this case observed on the video. In addition, the driver-related attributes typically collected at the time of subject recruitment were subsequently used in models to explore their association with event occurrence. Variables in the data set included measures of demographics, physiological attributes, and often measures of possible crash predisposition, such as indices of driver aggression or life stress.

Specific variables derived from the data are defined in Table 3.1. All data were originally collected by VTTI and assembled into a database for Penn State. The basic dependent variable was the identification of event by type (i.e., crash, near crash, or critical incident). This dependent variable was used individually in event-based models as a categorical outcome or as a count by type in driver-based models. Precipitating event attributes included a designation of vehicle loss of control by the driver when driver actions resulted in the vehicle being over the lane or edge line. These variable descriptors were taken from descriptions in police accident reports or NHTSA databases to facilitate subsequent analyses. Driver impairment is based on observation of the driver on video, not on any in-vehicle technology.

Driver distraction was carefully categorized and included distraction from wireless device use; vehicle-related activities (e.g., adjusting heat and radio); passenger-related activities (e.g., talking and interacting with a passenger); talking to self, singing, and other activities; internal distractions such as daydreaming or being lost in thought; dining (includes eating or drinking); and a collection of other distractions occurring in small numbers individually. Each of these distraction categories was only coded once for each driver and was dichotomous.

Context variables were also observed from the video, including the four variables shown in Table 3.1. Traffic density was estimated in five categories based on the density of traffic observed around the vehicle at the time of the event.

Driver attributes included objective variables such as age and gender along with crash predisposition measures. The Dula Dangerous Driving Index (DDDI) was intended to provide a detailed description of the type of dangerous driving participants may engage in, as well as a total danger index (Dula and Ballard 2003). The Life Stress Index was used to describe the level of stress the driver was experiencing caused by issues such as job problems and family difficulties (Dingus et al. 2006). Such issues do not pertain directly to driving, but they are thought to be associated with crash risk. All driver attribute data were collected at the time of subject recruitment.

Comparisons with Poisson and Negative Binomial Distributions

A preliminary check of model fit was conducted by plotting a dependent variable with an assumed Poisson distribution by using the sample mean, since Poisson distribution can be described by only one parameter, λ. An NB distribution is also plotted by using the sample mean and variance, since an

Table 3.1. Summary of VTTI Variables Used in Modeling

Variable Type	Definition	Source	Comment
Event of interest	Crash; near crash; critical incident	Observed from video (see text)	Crashes include events recorded on police accident reports; others are new information available only from naturalistic studies.
Event Attributes			
Precipitating event • lost control • subject over lane/road edge	Event immediately preceding crash	Video	Observed from video in naturalistic studies; categorical dichotomous variable.
Driver impairment • suspected drug or alcohol impairment • fatigued/sleepy	Suspected alcohol/drug involvement Suspected fatigue/ sleepiness/drowsiness	Video	Alcohol/drug involvement observed in naturalistic study; required much judgment.
Driver distraction • wireless device • vehicle-related • passenger-related • talking, singing, etc. • internal distraction • dining • other	Distraction by category	Video	Distraction is observed in video.
Context			
Road, environment, and traffic conditions at time of event • presence of curve • day/night/dusk • road surface condition • traffic density	Presence of a road element or environmental condition at time of event	Video	The context within which the event occurred is observed through the use of video; categorical variable, typically dichotomous.
Driver Attributes			
Demographic • gender • age • years driving	Self-reported demographic data	Self-reported survey	Obtained through self-reports as recorded on questionnaires before the initiation of driving in the instrumented vehicle.
Psychological (measures of crash predisposition) • Dula Dangerous Driving Index (DDDI) • Life Stress Index • Driving Stress Inventory	Measures of personality, life stress and/or risk acceptance at time of initiation into study	Self-reported through use of specific tools before driving	Specific predisposition used in total; components of DDDI used to separate negative emotion, aggressive driving, and risky driving. DDDI also used with individual scale adjustment.
Estimated exposure	Estimated number of miles driven in previous year	Self-reported during subject screening	Obtained during subject intake surveys.

NB distribution needs to be described by two parameters, α and β, which can be substituted by the sample mean and variance. The values of the overdispersion parameter α, shown in Figure 3.1 and Figure 3.2, are 2.55 and 2.15, respectively. The values of α suggest that no matter which dependent variable is used, the NB distribution fits better than the Poisson distribution as a result of the violation of the assumption of the mean being equal to variance. Poisson models tend to underpredict zeros, and NB models overpredict zeros a bit. Accordingly, one may also consider the use of ZIP and ZINB models, which can improve model goodness of fit.

Systematic Model Testing

A series of count regression models were tested in groups. In general, a Poisson regression, NB regression, and ZIP model were tested with the same set of variables. Early testing using this approach indicated that parameter estimates were not consistent

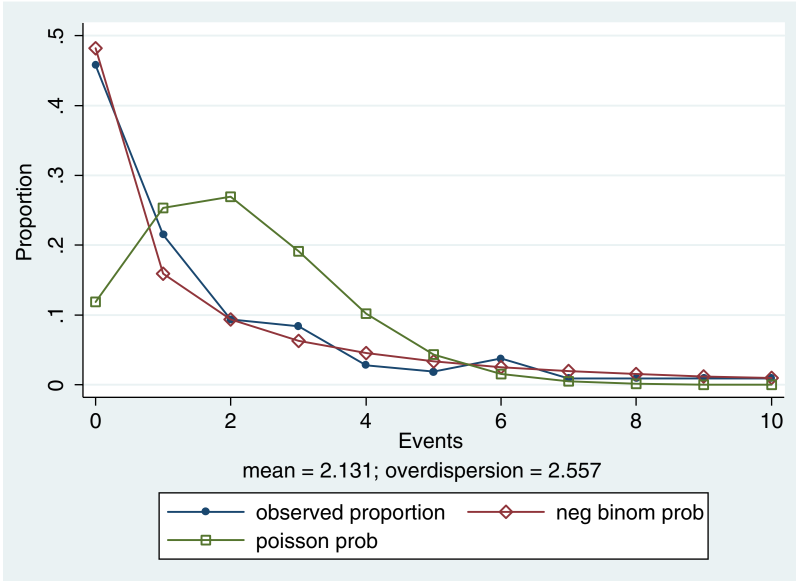

Figure 3.1. Observed total number of events compared with Poisson and NB distributions.

across the three basic regression types. As a result, several sets of models were constructed, and model fit was assessed for each set; the findings of the models, including model fit criteria, are summarized in Table 3.2. The model sets are described as follows:

- Set A: Predictors are main effects for objective data (e.g., age, gender, years driving). This is the starting point for most modeling, but it proved inadequate in the present endeavor as the three model types did not yield consistent parameter estimates, levels of significance, or goodness of fit.
- Set B: Predictors included a series of gender interaction terms for each of the variables used in Set A. The objective was to explore gender differences, which were expected from the literature. Overall consistency of the models improved, but there were still differences in model significance.
- Set C: Predictors included Set A plus a series of predisposition variables (DDDI, Life Stress Index, and Driving Stress Inventory) as main effects. Some predisposition variables were significant; main effects proved a poor specification for these data.

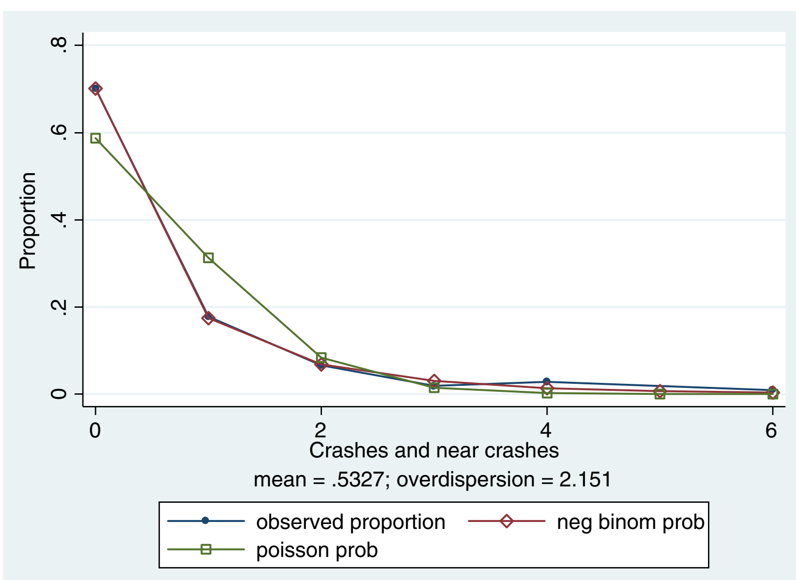

Figure 3.2. Observed crashes and near crashes compared with Poisson and NB distributions.

Table 3.2. Overview of Driver-Based Count Regression Models

Model Set	Predictors Included in Testing	*N*	NB Models: Predicted Zeros (%)[a]	NB Models: α[b](_)	NB Models: Pearson Statistic[c]	NB Models: Log Likelihood	Note
A	Gender, number years driving, education beyond bachelor's degree, bachelor's degree, number of previous violations, number of previous crashes, age of vehicle	83	62.62	1.68 (.39)	1.49	–151.54	All primary predictors are main effects
B	DRYRF, DRYRM, PGRADM, PGRADF, BSM, BSF, PVIOF, PVIOM, PACCM, PACCF, CYRF, CYRM	83	59.77	1.49 (.35)	1.38	–147.56	Gender difference consideration: Break down primary predictors into interaction terms
C	Gender, years of driving experience, post bachelor's, bachelor's degree, previous violations, previous accidents, miles driven in last year, scaled DDDI AD score, scaled DDDI NE score, car age, Life Stress Index score	83	54.72	1.21 (.30)	1.29	–142.41	Consider driver's attitude toward driving, such as DDDI scores, Driving Stress Inventory, and Life Stress Inventory
D	DRYRF, DRYRM, PGRADM, PGRADF, BSM, BSF, PVIOF, PVIOM, PACCM, PACCF, CYRF, CYRM, miles driven in previous year	83	52.72	1.12 (.28)	1.12	–139.26	Consider miles driven in a year
E	LOWDRM1, CYRF, BSM, PVIOF, miles driven in previous year, ADADF	83	51.08	1.04 (.26)	0.96	–138.31	Final model, constant α
F	LOWDRM1, CYRF, BSM, PVIOF, miles driven in previous year, ADADF	83	NA	NA	NA	–132.30	Final model with parameterized α

[a] Observed percentage of zeros is 47%.
[b] α is overdispersion parameter; significantly greater than zero indicates overdispersion.
[c] Pearson chi-square dispersion statistic; sum of model Pearson residuals divided by degrees of freedom.

- Set D: This set combined Set B with miles driven per year, a combination that provided a dramatic improvement in overall fit.
- Set E: Main effects and interactions were included in this parsimonious model.
- Set F: This parsimonious model had a parameterized estimate for α using a linear model.

Four evaluation criteria are listed in Table 3.2: the percentage of zeros predicted, the value of α, the Pearson dispersion statistic, and the log likelihood. This summary shows how the team systematically evaluated count regression model quality.

Several trends are apparent in the data. The log likelihood generally improves with the smallest value (best fit) occurring with the NB model enhanced by a parameterized estimation for α. This improvement in overall fit is obtained with six predictors for the NB portion and another three for estimating α (the model with parameterized α is summarized in Table 3.3). The Pearson statistic shows continued improvement and α continues to drop steadily, indicating that overdispersion is becoming less of a problem. Note also that as the model explains more variability in the data, the value of the α parameter declines. Finally, the percentage of zeros predicted is gradually getting closer to the actual level of 47%, indicating a generally better fit in this important attribute.

One can quickly see that the first four predictors represent variables interacting with gender. Inexperienced males have an elevated number of events expected, while having a college degree reduces the expected number. For females, the number of years driving and the number of previous violations are negatively correlated with the expected number of events, and the scaled AD score is positively associated. All predictors are significant at conventional levels. The team used a comfortable $p = .20$ as the cutoff to allow the inclusion of variables of potential interest that may fail conventional tests because of the small sample size. Several variables contribute to the estimation of overdispersion, including miles driven, years of driving experience, and scaled AD score.

Additional Discussion of Frequentist Models: Elasticity

To provide some insight into the implication of parameter estimation results, elasticities were computed to determine the marginal effects of the independent variables (Shankar et al. 1995). Elasticity provides an estimate of the impact of a

Table 3.3. NB Driver-Based Model with Parameterized α, Best Overall Driver-Based Model

Variable	**Coefficient**	**SE**	***Z***	***p*-Value**	**95% CI**
Males with <10 years driving experience	1.757	0.331	5.310	.000	(1.108, 2.405)
Years driving for females	–0.324	0.109	–2.980	.003	(–0.537, –0.111)
Having a college degree for males	–0.610	0.315	–1.930	.053	(–1.228, 0.008)
Number of previous violations for females	–0.537	0.170	–3.150	.002	(–0.871, –0.203)
Miles driven in previous year	0.000	0.000	5.650	.000	(0.000, 0.000)
Scaled AD score for females	0.393	0.165	2.390	.017	(0.070, 0.715)
Constant	–1.907	0.494	–3.860	.000	(–2.875, –0.939)
Variables that Parameterize the Dispersion Parameter	**Coefficient**	**SE**	***Z***	***p*-Value**	**95% CI**
Scaled AD index (ADAD)	1.241	0.485	2.560	.011	(0.290, 2.191)
Years driving experience	0.063	0.034	1.870	.061	(–0.003, 0.130)
Miles driven in previous year	0.000	0.000	1.480	.138	(0.000, 0.000)
Constant	–11.615	4.506	–2.580	.010	(–20.447, –2.784)

Note: The model is based on 83 observations. SE = standard error; likelihood ratio (LR).
LR chi-squared (6) = 39.42
probability > chi-squared = 0
pseudo R^2 = 0.1297
log likelihood = –132.30

variable on the expected frequency and is interpreted as the effect of a 1% change in the variable on the expected frequency λ_i. The elasticity of frequency λ_i is defined as

$$E_{x_{ik}}^{\lambda_i} = \frac{\partial \lambda_i}{\lambda_i} \times \frac{x_{ik}}{\partial x_{ik}} = \beta_k x_{ik} \tag{12}$$

where

E = elasticity,
x_{ik} = value of kth independent variable for observation i,
β_k = estimated parameter for kth independent variable, and
λ_i = expected frequency for observation i.

Note that elasticities are computed for each observation i. It is common to report a single elasticity as the average elasticity over all values of i.

The elasticity shown in Equation 12 is only appropriate for continuous variables such as highway lane width, distance from the outside shoulder edge to roadside features, and vertical curve length (Shankar et al. 1995). It is not a valid evaluator for binary categorical indicator variables. A pseudoelasticity can be computed to estimate an approximate elasticity for indicator variables. The pseudoelasticity gives the incremental change in frequency caused by changes in the indicator variables. The pseudoelasticity for indicator variables is computed as

$$E_{x_{ik}}^{\lambda_i} = \frac{\dfrac{\partial \lambda_i}{\lambda_i}}{\dfrac{\partial x_{ik}}{x_{ik}}} = \frac{\left(e^0 - e^{\beta_k}\right)/e^{\beta_k}}{(0-1)/1} = \frac{e^{\beta_k} - 1}{e^{\beta_k}} \tag{13}$$

For example, the average event frequency λ_i for driver i increases 0.71% if the driver is male and has less than 10 years of driving experience, compared with males with more than 10 years of driving experience, assuming the error terms are independent of x_{ik} and remain unchanged (and the model is correct). The elasticities from the final NB model are shown in Table 3.4.

The elasticity gives an indication of the effect of a predictor on the outcome (expected number of events). A quick scan of Table 3.4 shows that males with at least a college degree are substantially safer than their counterparts; males with less than 10 years' experience have an increase in expected events.

A 1% increase in driving mileage results in a 0.14% increase in expected number of events for both males and females per 1,000 miles driven. Mileage driven per year represents driver exposure to events. More exposure results in a higher probability of crashes and thus a higher expected number of events.

Table 3.4. Elasticities from Final NB Model

Variable	**Elasticity**
Males' driving experience <10 years	0.71
Interaction between females and age of car	–0.36
Males with bachelor's degree or above	–2.11
Females' past violations	–0.74
Mileage divided by 1,000	0.14
Scaled DDDI aggressive driving (AD) score	0.40

The scaled DDDI for aggressive driving (AD) showed a 0.4% increase in the expected number of events for a 1% change in the index. Recall that this index represented strong positive responses to questions such as: "I verbally insult drivers who annoy me" and "I deliberately use my car/truck to block drivers who tailgate me." Although ongoing research debates the effects of aggressiveness on the probability of crashes and the design of questionnaires that intend to quantify driver aggressiveness, results show the existence of an association between driver aggressiveness and the number of events.

A 1% increase in females' past violations results in a 0.74% decrease in the expected number of critical events compared with males. This result could be interpreted as a type of learning effect.

A 1% increase in the interaction between females and vehicle age results in a 0.36% decrease in the expected number of events compared with the interaction between males and vehicle age. This result is difficult to interpret and may be a variable that represents another phenomenon. Despite many attempts to remove this variable, it persisted.

Multilevel Driver-Based Modeling

The multilevel driver-based model in Table 3.5 has a two-level specification. The response variable is the number of events that each individual had in a year, and the explanatory variables include individual socioeconomic characteristics at Level 1 and gender at Level 2. Models were estimated using the open-source software OpenBUGS. The first 1,000 model iterations were discarded as burn-in. The next 100,000 iterations were used to obtain summary statistics of the posterior distribution of parameters. Convergence was assessed by visual inspection of the Monte Carlo–Markov chains. The number of iterations was selected such that the Monte Carlo error for each parameter in the model would be less than 10% of the parameter's SD.

In Table 3.5, the variable names appear in the first column followed by the estimated parameter value, or the mean, and its SD in the next two columns, respectively. The hierarchical modeling structure (full Bayes) produces 5% and 95% credible set estimates, instead of the CIs normally produced in frequentist estimation. Parameters with 5% to 95% credible set values that do not include zero are generally accepted as significant. A single asterisk (*) indicates a significant variable, and a double asterisk (**) indicates a variable that is marginally significant (i.e., has a credible set of 10% to 90% that does not include zero). Elasticities were also calculated based on the coefficients from the 100,000 iterations and Equations 9 and 10, as shown in Table 3.6. Therefore, the credible sets for all elasticities of predictors are also available.

Table 3.5. Estimates for Multilevel Driver-Based Model (NB)

Variable	Mean	SD	Percentile					
			2.5%	5%	10%	90%	95%	97.5%
Intercept (F)*	–13.71	8.26	–29.32	–27.42	–24.44	–3.14	–0.39	2.63
Intercept (M)*	–18.97	6.93	–33.17	–30.74	–27.94	–10.29	–7.95	–5.98
Scaled DDDI aggressive driving (AD) (F)**	0.59	0.38	–0.18	–0.06	0.09	1.07	1.20	1.31
Scaled DDDI AD (M)	0.01	0.26	–0.51	–0.42	–0.32	0.34	0.43	0.52
Scaled DDDI risky driving (RD) (F)	0.34	0.32	–0.32	–0.21	–0.08	0.73	0.87	0.98
Scaled DDDI RD (M)**	0.41	0.28	–0.12	–0.04	0.06	0.77	0.88	0.98
Bachelor's degree or above (F)	0.08	0.87	–1.62	–1.33	–1.02	1.19	1.52	1.83
Bachelor's degree or above (M)*	–1.24	0.65	–2.53	–2.31	–2.06	–0.42	–0.17	0.04
Years of driving experience (F)**	–0.05	0.03	–0.11	–0.10	–0.09	–0.01	0.00	0.01
Years of driving experience (M)*	–0.06	0.02	–0.11	–0.10	–0.09	–0.03	–0.02	–0.02
Mileage driven in past year (F)*	0.87	0.55	–0.13	0.03	0.18	1.63	1.81	1.99
Mileage driven in past year (M)*	1.77	0.56	0.74	0.89	1.07	2.49	2.73	2.93
Past violations (F)*	–0.60	0.29	–1.21	–1.10	–0.98	–0.23	–0.13	–0.05
Past violations (M)	0.06	0.22	–0.38	–0.31	–0.22	0.33	0.41	0.49
sigma^2.v*	1.71	0.59	0.85	0.94	1.06	2.49	2.81	3.13

*Significant at 95%, **significant at 90%.
Note: M = male; F = female.
Dbar (posterior mean of the deviance) = 204.7
Dhat (point estimate of deviance) = 159
DIC (deviance information criterion) = 250.3
pD = 45.62 (pD [effective number of parameters] = Dbar – Dhat)

Table 3.6. Elasticity Estimates from Multilevel NB Model

Variable	Mean	SD	Percentile					
			2.5%	5%	10%	90%	95%	97.5%
E.adad (F)**	1.16	0.80	–0.37	–0.13	0.14	2.17	2.51	2.79
E.adad (M)	–0.73	0.86	–2.38	–2.13	–1.82	0.37	0.70	0.96
E.adrd (F)**	1.07	0.83	–0.62	–0.32	0.02	2.12	2.40	2.66
E.adrd (M)**	1.79	1.34	–0.89	–0.42	0.09	3.49	3.98	4.39
E.BSabove (F)	–0.83	1.95	–5.65	–3.92	–2.57	0.48	0.60	0.68
E.BSabove (M)	–0.97	1.10	–3.73	–3.04	–2.36	0.16	0.33	0.46
E.exp (F)*	–0.31	0.18	–0.68	–0.61	–0.54	–0.08	–0.02	0.03
E.exp (M)*	–0.79	0.27	–1.34	–1.24	–1.14	–0.46	–0.38	–0.30
E.mil (F)*	2.71	1.47	0.09	0.46	0.89	4.58	5.10	5.76
E.mil (M)*	10.42	3.65	3.98	4.84	5.91	15.21	16.81	18.34
E.pvio (F)*	–0.35	0.17	–0.69	–0.63	–0.56	–0.14	–0.09	–0.04
E.pvio (M)	0.10	0.10	–0.11	–0.07	–0.03	0.23	0.26	0.29

Note: E. = elasticity.

Both females' and males' total mileage driven in 1 year are inherently significant, supporting the argument that higher exposure increases the likelihood of events. Moreover, both males and females with more driving experience had a reduced expected number of events. Female scaled aggressive driving (AD) scores and male risky driving (RD) scores are marginally significant, implying that a higher female scaled AD score and male scaled RD score increase the expected number of events.

This model provides two additional findings. First, males with at least a college degree have fewer events than males without bachelor's degrees; this effect is not significant for females. Second, females who had more traffic violations in the past had fewer events in a year.

The interested reader can compare the parameter values and the elasticities in this hierarchical model (Tables 3.5 and 3.6, respectively) with the NB model in Tables 3.3 and 3.4. There are changes in sign and magnitude, but the greatest difference is that one can now assess the effect of the variable on men and women separately.

Discussion of Outliers

As a by-product of running the hierarchical driver-based model using a Poisson lognormal model, the team was able to observe individual drivers' random effects. Any random effect that has a mean significantly different from zero for a driver identifies that person as an outlier or a substantial deviation from the sampled driver population. This deviation can be interpreted as the driver coming from a different population of drivers than the majority of drivers in the data set. The chance that each of the five drivers (2, 4, 15, 16, and 55) listed in Table 3.7 is identified as an outlier is at least 95%. For example, Driver 55 had 28 events at his age of 59 years. Statistical outliers thus may reflect some type of selection bias or model misspecification.

Table 3.7. Identification of Outliers: Mean and SD of Expected Number of Events for Driver-Based Hierarchical Model

Driver ID	Mean	SD	Percentile					
			2.5%	5%	10%	90%	95%	97.5%
Driver 2	1.78	0.92	0.03	0.31	0.63	2.96	3.33	3.66
Driver 4	–1.75	0.82	–3.45	–3.15	–2.81	–0.73	–0.45	–0.21
Driver 15	1.80	0.60	0.64	0.84	1.05	2.57	2.80	3.01
Driver 16	2.01	0.62	0.83	1.02	1.23	2.80	3.05	3.26
Driver 55	2.59	0.61	1.42	1.61	1.83	3.36	3.60	3.81

The conclusion is that there are substantial advantages to using the Bayesian approach, one of which is to identify and quantify individual driver heterogeneity. In this case, the drivers so identified may be considered as sampled from a different population of drivers. Four of the drivers (2, 15, 16, and 55) had higher underlying event risk. Driver 2 was an inexperienced male driver with high annual mileage who had a lower than expected event risk. No generalities can be made from a single observation; however, this result shows that drivers with low expected event risk, as well as high-risk drivers, may be identified by this method.

VTTI Data: Event-Based Models

Comparing Different Single-Level Models: Effects of Omitting Predictors

In many respects the easiest way to begin to understand the modeling conducted on the 100-car study events is to present a series of straightforward examples. The Penn State team conducted initial screening of binary, multinomial, and ordered logit models to assess their ease of interpretation and overall quality of prediction. The binary logit model resulted in the best goodness of fit, as indicated by an Akaike Information Criterion (AIC) of 205.32. The AIC values for the multinomial and ordered logit models were 296.62 and 293.11, respectively. The binary logit model is, in many respects, the most straightforward to interpret. Thus, the focus of the analysis was based on the use of binary logit models.

Binary logit models were estimated to develop an understanding of the effects of omitting variables. One model included only driving environment variables, and the other included a combination of driving environment variables and driver attributes. The third model, with the identical specification, was a model with the previous two sets of variables and an added set reflecting event attributes. The data used in the analysis are summarized in Table 3.8, and the results of the modeling are summarized in Table 3.9. Additional tests were conducted with other pairs of the three sets of variables, but discussion of these three models is sufficient to identify trends.

Table 3.9 shows the parameter mean and SD for each of the three models. Next to these values is the percentage difference with respect to the last, fully specified model. The next column for each model shows the odds ratio (OR) and, in parentheses, the difference in the OR between the full model and the other two.

The first model considers context only; this is a familiar model to many road safety analysts because context variables form the primary variables typically included in a safety performance function, which is fundamental to contemporary road safety management. Notice that by adding driver attributes, the parameters in the first model change substantially: some change sign and some, such as curve, change in level of significance as well. Using a Chi-square test to compare the two models results in a significant difference being found with the calculated Chi-square equal to 8.82 (resulting level of significance is $p = 0.0030$).

The situation becomes even more dramatic when event attributes are added: ORs double or triple for some predictors. The model's goodness of fit is dramatically improved, as shown by both the log likelihood and pseudo R^2 values. Tests between the full model and the model with context and driver attributes again show significant improvement with a Chi-square value of 10.73 and level of significance, $p = 0.0011$.

Why did the team undertake this exercise? The primary reason was to point out the difficulty in modeling a data set with almost no prior history. Naturalistic data really are unique, and this experience points out one of the challenges in their analysis. Many other transportation fields provide researchers with a sufficient history to know where to begin and what pitfalls to avoid. In naturalistic analysis researchers are virtually starting from scratch and have to evolve rules from their own experience. This experience is in some ways similar to what was learned in the driver-based models: develop models carefully. In this case, the message is to be sure all three components of the model are present, or ORs and other estimation output are likely to be biased.

For the purposes of discussion, one can assume that the fully specified binary logit model shown in Table 3.9 is a reasonable comparison to a hierarchical model, as specified in Chapter 2.

Multilevel Event-Based Model

Table 3.10 presents a summary of an event-based hierarchical model. The table shows all variables included in the hierarchy. Parameter values and SEs are included along with ORs (for significant variables only). At Level 1, the event-based data set presents two types of variables: event attributes (occurrences inside the car) and driving environment (occurrences outside the car). Level 2 models driver attributes, representing the varying effects of predisposition (DDDI and Life Stress Index values) and years of driving experience.

Notice that the two precipitating events are significant (loss of control is marginally significant). Loss of control includes excessive speed and a loss of control with poor road conditions. Unfortunately, the team was unable to make further inferences concerning lane versus road edge departures (or left- versus right-side departures) because of limitations in sample size for crashes and near crashes. Both parameters are positive, indicating that these behaviors increase crash and near-crash likelihood compared with that for a critical incident.

With the modeling results for a binary logit model (frequentist estimated) and a hierarchical model now available, the pattern of parameter significance and magnitude can be discussed. Several interpretations of the distraction variables

Table 3.8. Variable Definitions for Tests of Event-Based Omitted Variable Bias

Group	Variable Name	Definition	Variable Type	Mean	SD	Min	Max
Dependent variable	Event outcome	Crash/near crash (1); critical incident (0)	Binary	0.25	0.44	0	1
Event attributes	Precipitating Factor 1: Loss of control	Lose control of vehicle as a result of vehicle failures, poor road conditions, excessive speed (from GES critical event)	Binary	0.37	0.23	0	1
	Precipitating Factor 2: Subject over lane line/ road edge	System detected vehicle over the lane line or roadway edge (from GES critical event)	Binary	0.50	0.25	0	1
	Driver Impairment 1: Drowsy, sleepy, asleep, and fatigue	Driver appears to show these characteristics (all are based on GES "driver distracted by" variable)	Binary	0.21	0.16	0	1
	Distraction 1: Wireless device	Distraction related to locating or operating a wireless device	Binary	0.10	0.08	0	1
	Distraction 2: Vehicle related	Adjusting climate control, radio, audio devices, etc.	Binary	0.04	0.04	0	1
	Distraction 3: Passenger related	Distraction attributable to passenger in vehicle	Binary	0.06	0.05	0	1
	Distraction 4: Talking/ singing/daydreaming	Self-evident definition	Binary	0.04	0.03	0	1
	Distraction 5: Internal distraction	Reading, moving object, handling insect or pet	Binary	0.06	0.05	0	1
	Distraction 6: Dining	Includes eating and drinking	Binary	0.02	0.02	0	1
	Distraction 7: Other	Smoking, external distraction, personal hygiene, driving-related inattention to forward roadway	Binary	0.11	0.10	0	1
Driving context	Alignment	Curve (1); tangent (0)	Binary	0.31	0.22	0	1
	Lighting	Dawn/dusk (1); day (0)	Binary	0.06	0.06	0	1
	Surface condition	Dry (1); wet/icy/snowy (0)	Binary	0.19	0.15	0	1
	Traffic density	Not free flow (1); free flow (0)	Binary	0.25	0.19	0	1
Driver attributes	DDDI AD Index	Scale reflecting intent to harm	Continuous	11.44	1.41	7	23
	DDDI NE Index	Scale reflecting negative emotions during driving	Continuous	21.19	1.41	11	34
	DDDI RD Index	Scale reflecting risky driving	Continuous	19.28	0.71	12	31
	Driver experience	Number of years with license	Continuous	16.167	7.071	1.5	52
	Life Stress Index	Scale reflecting stress in one's life	Continuous	180.2	45.3	0	560

are available. The baseline for this set of variables is no distraction. Variables were extracted from video observations of drivers during the events of interest, including

- Distraction 1: wireless device—related to locating or operating a wireless device.
- Distraction 2: vehicle related—adjusting climate control, radio, audio devices, and other vehicle devices.
- Distraction 3: passenger related—attributable to a passenger in the subject's vehicle.
- Distraction 4: talking, singing, or daydreaming—self-evident definition.
- Distraction 5: internal distraction—reading, a moving object in the vehicle, dealing with an insect or pet.
- Distraction 6: dining—includes eating or drinking.
- Distraction 7: other—smoking, external distraction, personal hygiene, and driving-related inattention to forward roadway were aggregated into this category as a result of sample size constraints.

Among the distractions, internal and vehicle- and passenger-related distractions are significant, and talking/singing/daydreaming is marginally significant. All have positive coefficients, indicating that they increase the likelihood of a crash

Table 3.9. Summary of Initial Estimated Binary Logit Event-Based Models

Type	Variable	Context Only				Context and Driver Attributes				Fully Specified		
		Parameter		Coeff Diff (%)	OR (% Point Diff)	Parameter		Coeff Diff (%)	OR (% Point Diff)	Parameter		OR
		Coeff	SD			Coeff	SD			Coeff	SD	
	Intercept	–1.155	0.231	NA	NA	0.903	1.172	NA	NA	–1.672	1.808	NA
Event attributes	Precipitating Event 1: Loss of control									1.135	1.023	3.111
	Precipitating Event 2: Subject over lane line/road edge									2.269	0.998	9.670
	Driver Impairment 1: Drowsy/sleepy/asleep/fatigued									1.571	0.618	4.811
	Distraction 1: Wireless device									0.780	0.775	2.181
	Distraction 2: Vehicle related	NA				NA				2.224	0.940	9.244
	Distraction 3: Passenger related									1.794	0.848	6.013
	Distraction 4: Talking/singing/daydreaming									1.996	1.089	7.360
	Distraction 5: Internal distraction									3.086	0.985	21.889
	Distraction 6: Dining									1.879	1.306	6.547
	Distraction 7: Other									1.248	0.740	3.483
Context	Alignment 1: Curve	0.644	0.331	–30.89	1.904 (–63.6)	0.475	0.371	–0.06	1.608 (–93.2)	0.932	0.475	2.540
	Lighting 1: Dawn/dusk	1.105	0.609	–51.82	3.019 (–688.6)	0.743	0.657	–50.13	2.102 (–780.3)	2.293	0.743	9.905
	Surface condition 1: Wet/icy/snowy	0.078	0.411	–91.53	1.081 (–142.6)	–0.171	0.463	–118.61	0.843 (–166.4)	0.919	0.621	2.507
	Traffic density 1: Not free flow	–1.633	0.554	–25.34	0.196 (8.4)	–1.995	0.594	–8.77	0.136 (2.4)	–2.187	0.686	0.112
Driver attributes	DDDI AD score					–0.099	0.043	30.39	0.906 (1.9)	–0.120	0.050	0.887
	DDDI NE score					–0.095	0.056	19.19	0.909 (–0.6)	–0.089	0.070	0.915
	DDDI RD score	NA				0.049	0.075	–33.79	1.050 (0.2)	0.047	0.092	1.048
	Driver experience					–0.033	0.015	49.33	0.968 (2.6)	–0.060	0.020	0.942
	Life Stress Index					0.004	0.001	–17.03	1.004 (0)	0.004	0.002	1.004
Fit Statistics	Log likelihood	–117.802				–103.424				–82.818		
	Pseudo R^2	0.079				0.192				0.349		

Note: Coeff = Coefficient; Diff = Differential.

*Significant difference between context with driver attributes and context only: chi-square = 8.82, p = .0030.

**Significant difference between full specification, single level and context with driver attributes: chi-square = 10.73, p = .0011.

Table 3.10. Multilevel Event-Based Model

Type	Variable	Parameter		Percentile		OR
		Mean	SD	2.5%	97.5%	
Level 1 Covariates: Event attributes	Intercept	–1.77	1.88	–5.58	1.86	NA
	Precipitating Event 1: Loss of control	1.51	1.13	–0.50	3.91	9.08
	Precipitating Event 2: Subject over lane line/road edge*	2.83	1.11	0.90	5.21	33.42
	Driver Impairment 1: Drowsy/sleepy/asleep/fatigued*	1.29	0.58	0.17	2.46	4.29
	Distraction 1: Wireless device	0.25	0.78	–1.31	1.74	NA
	Distraction 2: Vehicle related*	1.95	0.97	–0.01	3.81	11.1
	Distraction 3: Passenger related*	1.51	0.86	–0.16	3.21	6.58
	Distraction 4: Talking/singing/daydreaming**	1.67	1.14	–0.64	3.85	9.82
	Distraction 5: Internal distraction*	3.13	1.00	1.24	5.18	38.82
	Distraction 6: Dining	1.05	1.44	–1.83	3.82	NA
	Distraction 7: Other	1.12	1.05	–0.97	3.18	NA
Level 1 Covariates: Driving environment	Alignment 1: Curve*	1.10	0.51	0.11	2.11	3.38
	Lighting 1: Dawn/dusk*	2.42	0.79	0.89	3.99	15.12
	Surface condition 1: Wet/icy/snowy**	0.82	0.65	–0.43	2.10	2.79
	Traffic density 1: Not free flow*	–2.38	0.69	–3.83	–1.11	0.12
Level 2 Covariates: Driver attributes	DDDI AD Index*	–13.67	5.31	–24.35	–3.41	0.87
	DDDI NE Index	–10.83	7.07	–24.88	2.70	NA
	DDDI RD Index	6.13	9.21	–11.83	24.39	NA
	Years of driving experience*	–0.06	0.02	–0.11	–0.03	0.94
	Life Stress Index*	0.50	0.17	0.18	0.83	1.67

*Significant at 10% level.
**Significant at 20% level.

or near crash compared with the likelihood of an incident. Using a wireless device (at least at the time of data collection) is not significant. Driver impairment, including drowsiness, sleepiness, and fatigue, is also significant and positive.

Three context variables are significant determinants of event likelihood. Non-free-flow traffic density is negative in sign, indicating that this reduces the likelihood of a run-off-road crash or near crash and increases the likelihood of a critical incident. The presence of a curve and conditions at dawn or dusk both increase run-off-road crash and near-crash likelihood. The variable wet/icy/snowy is marginally significant and positive.

Life stress and years of driving experience are significant in differentiating events. A high score on the life stress test increases the likelihood of a crash or near crash compared with the likelihood of an incident, while drivers with more years of (self-reported) driving experience have a reduced likelihood of a crash or near crash compared with the likelihood for a critical incident.

Several driver-level variables are significant. The Life Stress Inventory asks drivers to mark each event that occurred during the past year. There are 42 events, including such items as personal injury or illness, change in financial state, and change in social activities. Each of these items has a weight; the variable entered in the model is the sum of the weights for each driver for each of the items checked. Based on the structure of the inventory, it measures general stress in someone's life.

The DDDI consists of 28 statements to which the driver is asked to respond on a 5-point Likert scale (never, rarely, sometimes, often, and always). Each of the categories of response is assigned an integer from 1 to 5. Example index statements include but are not limited to the following: "I verbally insult drivers who annoy me"; "Passengers in my car/truck tell me to calm down"; and "I will weave in and out of slower traffic." The responses to the questions are divided into the three categories of aggressive driving (AD), negative emotional (NE) driving, and risky driving (RD) (Dula and Ballard 2003). Each captures a different aspect or component of dangerous driving. The AD component is intended to reflect behavior intended to harm other living beings, either physically or emotionally. A positive response to the first example statement mentioned above represents AD, to the second statement NE, and to the last statement, RD. The value of the predictor variable is the sum of the rating responses to each question in each of the three DDDI components. The model indicates that those who are aggressive or who have negative emotions while driving are more likely to have critical incidents (i.e., less likely to have crashes or near crashes). This finding needs to be verified with a larger data set using nonevents as the baseline. The point here is that these predisposition measures need to be included in models of this type because they appear to be associated with event outcomes.

A few words are in order concerning the use of event-based models to test possible surrogates. Potential surrogates include the precipitating events of subject over lane or road edge and loss of control. These two variables were derived by the VTTI data coders as part of the original 100-car data set. In most event models they were strong indicators of crash or near-crash events in the categorical models; in hierarchical models subject over lane or road edge was the second-strongest predictor associated with the prediction of a crash or near-crash event. Although this measure is strongly associated with crash events, it lacks a time dimension, which is one of the desirable surrogate criteria proposed by Shankar and colleagues (2008). In their discussion of the traffic conflicts technique as a surrogate measure, Hauer and Gårder (1986) commented that "one should be able to make inferences about the safety of an entity on the basis of a short duration 'conflict count' instead of having to wait a long time for a large number of accidents to materialize." This suggestion could not be applied because the team did not have access to the comparable set of subject behaviors for noncrashes. Were such data available, the hierarchical model could be formulated to test the association between this measure and crashes. The application of this measure outside of SHRP 2's instrumented vehicles is as yet uncertain, but it is clear that it has some potential as a surrogate.

These models have important implications for SHRP 2 program concerns to identify useful surrogate measures. The categorical models explored in this study appear to be a useful paradigm to explore surrogates when they include event-based data. While kinematic measures or combinations of kinematic and roadway position measures were not directly tested with VTTI data, the Penn State team believes they are possible measures for future testing. The subject over lane or road edge variable contained position-only information and was very strongly associated with crash-related events; the team believes that the inclusion of longitudinal or lateral velocity and lateral position information would enhance this variable's predictive ability.

A limitation of the categorical models deserves mention. Initial event-based models, both bivariate logistic and hierarchical, used improper speed as an event-based predictor. Successful model fit was obtained, but improvement was sought. Driver Impairment 1 (drowsy, sleepy, fatigued) was substituted as a predictor and a much better fit occurred overall, including reduced SEs for several variables. While the team was pleased by the improved fit, there was concern about the apparent model instability. Such instability may be the result of the small sample size, but it may also reflect endogeneity among the predictors. As a recommendation to future SHRP 2 analysis contractors, the team suggests that care be exercised in surrogate analyses; additional empirical testing in several sites and with other drivers should reveal more about this issue.

A Method of Identifying Event Validity in Surrogate Testing

Background

One of the principal goals and challenges of the SHRP 2 Safety program is to develop procedures to identify crash surrogates. One useful definition was articulated by Hauer and Gårder (1986) in their focused discussion of the traffic conflicts technique as a surrogate measure (quoted above). Additional attributes of surrogates as having a time dimension and being responsive to countermeasures in the same way as in an actual crash have been proposed as part of the present research (Shankar et al. 2008). More generally, surrogates can be considered as measures that can be substituted for crashes in a safety analysis: in the data for this project, they are typically vehicle kinematic– and event-related measures that offer some description of vehicle movement and/or position relative to the roadway.

In concept, one would like to test and explore these issues in a naturalistic data set of many crashes, near crashes, and critical incidents. With a large naturalistic data set of 100 or more vehicles measured over 2 years or more (as in the S07 project), potentially thousands of observations of each candidate surrogate (e.g., thousands of measures of individual lateral accelerations at curves) would be available. It may be useful to explore whether the events containing the surrogate measures are similar to crash events. If there were a way to test for similarity, then researchers might be able to obtain a large enough and more valid set of surrogate measurements.

So, the goal here is to develop a way to validate surrogates. The Penn State team wants to see if the observations they have do the best job of identifying safety problems. The specific test of validity proposed is to use the event-based model to predict the probability of a crash event. Observations of the surrogate measure (e.g., a vehicle kinematic measurement) would then be screened to include only those involved in events predicted to be crashes by the model. Of course, such an analysis is contingent on the model's correctness. This method is offered as a promising way to improve future surrogate analysis.

Proposed Method

The proposed method takes advantage of the event-based models developed from the VTTI data analysis for road departure crashes. One of the important factors derived from the event-based models was the presence of context variables as important predictors of crashes or near crashes. These variables included the presence of a horizontal curve, dawn or dusk, road surface conditions, traffic conditions (free-flow or non-free-flow), and presence of driver distraction at the time of the event. Other combinations of context variables could be used, but these are of particular interest because they were among the most significant predictors in the event-based models.

The basis of the method is to work with the predicted outcomes from the model that differentiates the two event groups. Output from the hierarchical models was chosen because the team believed that such output is more valid from a strictly statistical standpoint. Figure 3.3 summarizes all the observations. Events are denoted with a number after a letter code as follows: I for incident, NC for near crash, and C for crash. The left half of the figure shows results for events taking place on tangent sections (curve = 0), and the right half for events occurring on horizontal curves. The *y* axis represents the predicted probability for each event. Any event with a predicted probability above 0.5 is considered to have been predicted to have that outcome. For example, crashes C130 and C133 occur on tangent sections and are predicted as crashes with a probability close to 76%. Incidents I19 and I195 also occur on tangent sections but are predicted as crashes. This is the first example of a validated event that can be used as a source of a surrogate observation (e.g., the value coded for the variable exceed road or lane edge); if kinematic measures were available as candidate surrogates, the lateral vehicle position or the longitudinal or lateral speed when the event took place could be used.

Events I19 and I195 are incidents, but they are predicted to be crashes. Thus they may be considered statistically close to crash events even though crashes were avoided. A surrogate may be selected from these events with additional validity as well. What is of interest is that researchers now have a statistical method to identify and quantify these promising events.

While the figure is a bit cluttered, several promising incidents occurring on curves are readily identifiable (e.g., I32, I31, and I205). Events NC624, NC198, and C13 are among several events correctly predicted as crashes or near crashes.

Figure 3.3 provides an aggregate perspective, as it contains only the context variable of presence of curve. All six variables listed above can be used to create more specific, well-defined contexts to determine how many valid events are identified in each context.

One particular advantage of naturalistic data such as the VTTI data is the ability it offers the researcher to use the narrative to compare the etiology of incidents and crashes in each context. The narrative, derived from analysis of video after the event is identified by kinematic screening criteria, can be used to verify if the etiology of the incident in a context was actually similar to that of crashes in the same context. Such verification can be thought of as additional validation for the event in question.

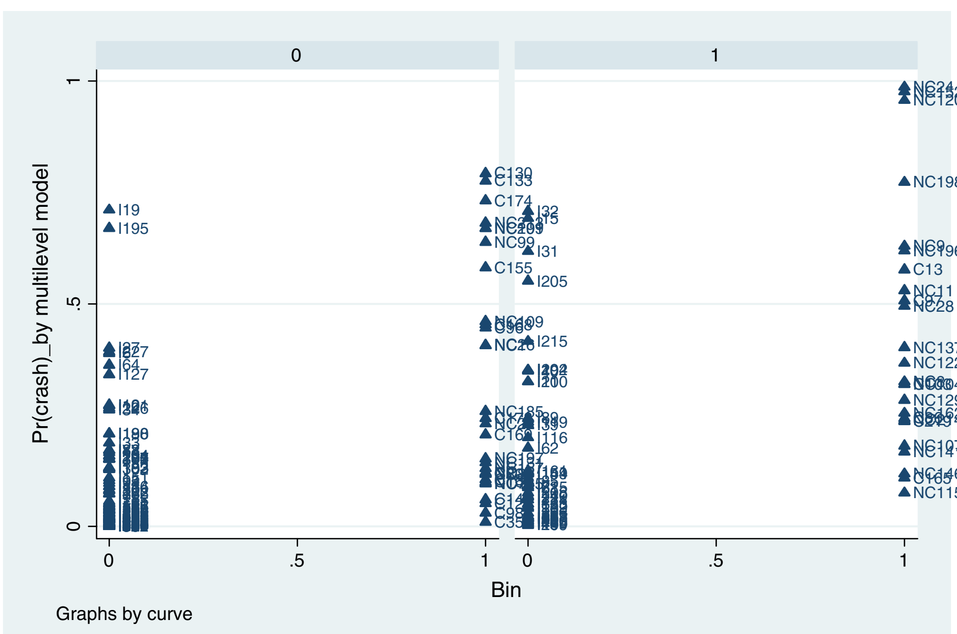

Figure 3.3. Predicted events for tangent (left) and curve (right) sections.

Figure 3.4. Example 1: Using predicted probabilities to identify crash surrogate events.

Applications to More Specific Contexts

EXAMPLE 1

Figure 3.4 shows the predicted probabilities for three event cases that occurred during a crash scenario with the following conditions: tangent road section, dawn or dusk lighting conditions, free-flow traffic, dry road, and vehicle over road edge. For this context, critical incident I195 was predicted with high probability as a crash. Although the model was structured to differentiate crashes and near crashes from critical incidents, a by-product of the model estimation is the identification of events that are predicted to be similar to crashes or near crashes. These three events were validated as being similar by comparing the narratives written by data coders at VTTI during data assembly. All three events involved a driver falling asleep and nearly running off the road. Other scenarios and contexts yielded different numbers of crashes, near crashes, and critical incidents. There is now a structured statistical method that offers promise in using naturalistic data to identify events that are similar to crashes; once identified as such, measures strongly correlated with the event outcome can be tested as surrogates.

EXAMPLE 2

Figure 3.5 illustrates another context in which the nature of the relationships is less clear. This context includes the following

Figure 3.5. Example 2: Using predicted probabilities to identify crash surrogate events.

attributes: horizontal roadway curve, daylight, free-flow traffic conditions, dry surface, no loss of control, and vehicle off road edge. Incident I15 was identified as a promising event, and crash C97 and near crashes NC120, NC198, and NC9 were correctly predicted. There are a host of incidents that were correctly predicted but were likely to be poor events for the purposes because they were not similar to crashes. Several crashes and near crashes were also predicted in the range of 0.0 to 0.4 (not well described by the model). This case illustrates that different contexts have differing numbers of events, and the ability to predict varies substantially.

Other Promising Events Identified and Their Corresponding Narratives

The following six cases evaluate I19, I195, I15, I31, I205, and I32 as potential useful events.

Case 1 (Figure 3.6) involves the following context: straight alignment, daylight, free-flow traffic, dry surface, no driver impairment, no loss of control, and vehicle over edge of road. The following three narratives compare I19 to C174 and NC99:

- I19: Subject driver is reading, and, as a result, she loses control of the car. She has to steer to the left in order to avoid any kind of conflict (internal distraction).
- C174: Subject driver is holding a cup in her right hand and turning right at an intersection. She cuts the corner and hits the curb on the right side.
- NC99: Subject driver is reaching for what appears to be a cell phone charger. She takes both hands off the wheel and looks away from the road to obtain the object. The vehicle drifts to the right and nearly hits a boat loaded on a trailer that is parked on the right side of the road.

Case 2 (Figure 3.7) involves the following context: straight alignment, dawn or dusk, free-flow traffic, dry surface, fatigued driver, no loss of control, and vehicle over edge of road. The following narratives compare I195 to NC199 and NC201 (these three events belong to the same male, age 59 years):

- I195: Subject falls asleep behind the wheel and drifts toward the right edge of the road. He suddenly wakes up and jerks the wheel to the left to get back in his lane.
- NC199: Subject driver falls asleep while driving, and the vehicle runs off the road to the right.
- NC201: Subject driver falls asleep while driving, and the vehicle runs off the road to the right.

Case 3 (Figure 3.8) involves the following context: curve, daylight, free-flow traffic, dry surface, no fatigue, no loss of control, and vehicle over edge of road. The narratives below compare I15 to C97, NC120, NC198, and NC9:

- I15: Subject is distracted and drifts over the left side of her lane. She has to steer right to avoid hitting the median.
- C97: Subject driver pulls over to park along the right side of the road and hits the curb as he is parking.
- NC120: Subject driver is looking at a piece of paper as he drives under an overpass. The road curves to the left and the vehicle veers left and nearly hits the left median.

Figure 3.6. Predicted crash probabilities for Case 1.

Figure 3.7. Predicted crash probabilities for Case 2.

- NC198: Subject driver is looking at or for an unknown object near the passenger seat. The road curves to the right and the vehicle goes off the right edge of the road.
- NC9: Subject driver is talking to a passenger in the adjacent seat while driving on a single-lane road. The road forks at an interchange, and the subject driver moves from the lane on the right to the lane on the left, nearly hitting the median in between the two.

These examples are sufficient to illustrate the method. The team hopes that if a valid model is developed, the screening of valid events will help in the identification of the surrogate measured within those contexts, eliminating the need to go back to narratives for additional assurance. Automating the process through the use of an event-based model promises potential time savings and accurate surrogate identification.

Figure 3.8. Predicted crash probabilities for Case 3.

UMTRI Data: Kinematic Models

Table 3.11 is a glossary of interaction term variable acronyms used in UMTRI kinematic models. They include explicit recognition of positive (e.g., PlanoffPpi as positive lane offset interacting with positive pitch) and negative (e.g., Nlaspro, which is negative lateral speed interacting with positive roll) measures so the team could better understand the vehicle's movement when alerts were triggered. This table is included to help the reader track the detailed discussion of the models that follows. The text generally follows the summary of model structure described for UMTRI data in Chapter 2.

Single-Regime Models

The first single-regime models developed relationships between speed (both longitudinal and lateral) and main effect kinematic variables. Different combinations of predictor variables were tested, particularly exploring the inclusion and exclusion of steer angle and roadway classification. It was determined during the course of this initial step in the modeling process that lateral speed proved to be a poor potential surrogate: models showed poor fit, and parameter estimates had little to no effect on the dependent variable (see Table 3.12 as an example).

The model in Table 3.13, however, has a respectable goodness of fit and reasonable parameter estimates. All parameters are significant. It is hard to distinguish the validity of coefficient signs (positive versus negative) because they may be partly based on curve direction. Using directional lateral speed and lane offset (Table 3.13) instead of the general forms improves their significance and interpretation. Overall model fit significantly improves as well, even with steering angle removed. Roadway classification has a substantial effect on longitudinal speed (see Table 3.14), and all road classes are significant except for Road Class 6 (ramps).

Kinematic variables were formed into interacting variables to determine further how they affect longitudinal speed, since it seemed plausible that changes in different vehicle kinematics are correlated (e.g., yaw and roll are correlated, since yaw occurs when the wheels are turned, thus causing the car to undergo slight roll). An example of strong correlation between yaw and steering angle can be seen in Table 3.15 (entries of high correlation are in bold).

The strong correlation between steer and yaw is expected. When a vehicle enters a horizontal roadway curve, the driver is required to turn the wheels to maintain position on the roadway, thus changing the steering angle from zero. When a vehicle undergoes a turn, it experiences rotation about its vertical axis (perpendicular to the roadway surface), also known as yaw. Thus, a change in steering angle will result in a change in yaw in the same direction. Because this relationship exists, models can be developed without having to include all

Table 3.11. Glossary for Kinematic Interaction Terms

Kinematic Interaction Terms	
PlanoffPlas	Positive lane offset, positive lateral speed
NlanoffNlas	Negative lane offset, negative lateral speed
PlanoffNlas	Positive lane offset, negative lateral speed
NlanoffPlas	Negative lane offset, positive lateral speed
PlanoffPy	Positive lane offset, positive yaw rate
PlanoffNy	Positive lane offset, negative yaw rate
NlanoffPy	Negative lane offset, positive yaw rate
NlanoffNy	Negative lane offset, negative yaw rate
PlanoffPpi	Positive lane offset, positive pitch rate
PlanoffNpi	Positive lane offset, negative pitch rate
NlanoffPpi	Negative lane offset, positive pitch rate
NlanoffNpi	Negative lane offset, negative pitch rate
PlanoffPro	Positive lane offset, positive roll angle
PlanoffNro	Positive lane offset, negative roll angle
NlanoffPro	Negative lane offset, positive roll angle
NlanoffNro	Negative lane offset, negative roll angle
PlasPy	Positive lateral speed, positive yaw rate
PlasNy	Positive lateral speed, negative yaw rate
NlasNy	Negative lateral speed, negative yaw rate
NlasPy	Negative lateral speed, positive yaw rate
PlasPpi	Positive lateral speed, positive pitch rate
NlasPpi	Negative lateral speed, positive pitch rate
NlasNpi	Negative lateral speed, negative pitch rate
PlasNpi	Positive lateral speed, negative pitch rate
PlasPro	Positive lateral speed, positive roll angle
NlasPro	Negative lateral speed, positive roll angle
NlasNro	Negative lateral speed, negative roll angle
PlasNro	Positive lateral speed, negative roll angle
PyPro	Positive yaw rate, positive roll angle
PyNro	Positive yaw rate, negative roll angle
NyNro	Negative yaw rate, negative roll angle
NyPro	Negative yaw rate, positive roll angle
PyPpi	Positive yaw rate, positive pitch rate
PyNpi	Positive yaw rate, negative pitch rate
NyNpi	Negative yaw rate, negative pitch rate
NyPpi	Negative yaw rate, positive pitch rate
ProPpi	Positive roll angle, positive pitch rate
ProNpi	Positive roll angle, negative pitch rate
NroNpi	Negative roll angle, negative pitch rate
NroPpi	Negative roll angle, positive pitch rate

Table 3.12. Lateral Speed Model, Linear Regression

Variable Name	Coefficient	Std. Err.	*t*-statistic	*p*-value
Speed	0.0025	0.0000	52.2700	0.0000
Yaw rate	0.0059	0.0004	13.7600	0.0000
Pitch rate	0.0007	0.0003	2.3800	0.0170
Roll angle	0.0020	0.0002	13.2900	0.0000
Steer angle	–0.0014	0.0001	–17.2400	0.0000
Lane offset	0.0690	0.0004	155.0500	0.0000
Constant	–0.0632	0.0013	–49.0100	0.0000

Number of obs = 1,391,799; Prob > F = 0.0000; R-squared = 0.019; Adj R-squared = 0.019.

kinematic variables. When analyzing model results, it can be assumed that longitudinal speed would be affected similarly by yaw rate and steering angle.

Tables 3.16, Table 3.17a, and Table 3.17b summarize the best models established for the pure linear, single-regime models. Both models include the variables measurement duration (a measure of time; measurement duration = 0 at the beginning of observation, and measurement duration = 5 at the time when the alert is triggered), dark or light, RDCW disabled or enabled, and roadway classification. Both models also have several kinematic variables as interaction terms. Model (a) does not include interaction terms for lane offset, instead using only main effects. Table 3.16 shows the goodness-of-fit results from the two models. The parameter estimates are shown in Table 3.17a and b.

Table 3.13. Model 1: Longitudinal Speed, Single Regime, Directional Kinematics (Except Lateral Speed and Lane Offset), Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Lateral speed	0.4487	9.56	0.0000
Lane offset	0.1497	5.27	0.0000
Positive pitch	–0.7406	–29.11	0.0000
Negative pitch	–0.5802	–24.91	0.0000
Positive roll	1.2721	115.08	0.0000
Negative roll	1.6482	145.07	0.0000
Positive yaw	–0.9798	–34.00	0.0000
Negative yaw	–0.3201	–14.12	0.0000
PSteer	–0.3897	–74.92	0.0000
NSteer	–0.4188	–116.02	0.0000
Constant	47.4985	1327.18	0.0000

Number of obs = 336,548; Prob > F = 0.0000; R-squared = 0.3182; Adj R-squared = 0.3182.

Table 3.14. Model 2: Longitudinal Speed, Single Regime, and Directional Kinematics (All) with Roadway Classification, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Positive lateral speed	0.7148	9.68	0.0000
Negative lateral speed	0.8419	14.56	0.0000
Positive lane offset	–1.9080	–38.64	0.0000
Negative lane offset	–1.6185	–39.95	0.0000
Positive pitch	–0.5164	–23.87	0.0000
Negative pitch	–0.3892	–19.66	0.0000
Positive roll	1.1603	125.88	0.0000
Negative roll	1.5100	167.13	0.0000
Positive yaw	–2.5338	–260.63	0.0000
Negative yaw	–2.0422	–204.29	0.0000
Road class: Limited access	12.0936	11.42	0.0000
Road class: Major surface	–7.4495	–7.04	0.0000
Road class: Minor surface	–4.1978	–3.97	0.0000
Road class: Local road	–7.3286	–6.92	0.0000
Road class: Ramp	1.2477	1.18	0.2380
Constant	48.4004	45.76	0.0000

Number of obs = 336,547; Prob > F = 0.0000; R-squared = 0.5079; Adj R-squared = 0.5079.

Yaw and lane offset have the biggest effect on longitudinal speed in the single-regime model without interaction terms. Because the test vehicles had rigid bodies, roll would have less effect than yaw on longitudinal speed changes for kinematic models. Yaw will have the same (mathematical) relationship to speed and steering angle regardless of vehicle body rigidity. Roadway classification plays an important role in affecting longitudinal speed because of the difference in design speeds between roadways of different classifications. The interaction terms with the greatest effect on longitudinal speed are those

Table 3.15. Correlation Between Steer and Yaw

	Positive Yaw	Negative Yaw	Positive Steer Angle	Negative Steer Angle
Positive yaw	1			
Negative yaw	–0.27	1		
Positive steer angle	**0.9777**	–0.25	1	
Negative steer angle	–0.2656	**0.9381**	–0.2469	1.0000

Table 3.16. Goodness of Fit, Kinematic Models

(a) Lane Offset Main Effects

Source	SS	df	MS
Model	28965404.1	34	851923.7
Residual	32225550.1	336513	95.76317
Total	61190954.2	336547	181.8199

Number of observations = 336,548
$F(34,336513) = 8655.86$
probability > F = 0
$R^2 = 0.4734$
adjusted $R^2 = 0.4733$
root MSE = 9.7859

(b) Lane Offset Interactions

Source	SS	df	MS
Model	29386711.1	48	612223.2
Residual	31804243	336499	94.51512
Total	61190954.2	336547	181.8199

Number of observations = 336,548
$F(48,336499) = 59,56.3$
probability > F = 0
$R^2 = 0.4802$
adjusted $R^2 = 0.4802$
root MSE = 9.7219

combining lane offset and yaw, lane offset and pitch, lateral speed and yaw, and lateral speed and roll; the latter two sets of terms (i.e., lateral speed and yaw and lateral speed and roll) can be directly related in fundamental kinematics. Both models in Tables 3.17a and b had good fit and generally significant parameters.

The use of roadway classification improved the goodness of fit of the single-regime models, showing that roadway classification may play an integral role in determining other relationships between vehicle kinematics and longitudinal speed. The goodness of fit for the non-interaction-term models was substantially better with roadway classification included (both interaction term models included roadway classification).

Two-Regime Models

The eight models summarized in Tables 3.18 through 3.26 show the initial approach to the second step in the flow chart in Figure 2.10—that is, pure linear, two regimes. These models include additional predictors, dark and roadway classification. The kinematic variables are separated into positive and negative based on directionality of measurement. Tables 3.22 through 3.25 include a variable called measurement duration. The first regime is defined as occurring when measurement duration is between 0 and 5 s: this is the time before the alert is triggered

Table 3.17a. Model 3: Longitudinal Speed, Single Regime, Interaction Kinematic Variables, Except Lane Offset Model

Variable Name	Coefficient	*t*	*p* > *t*
Constant	62.1109	1048.44	0
Measurement duration	–0.2674	–62.97	0
Dark	0.798	18.28	0
RDCW system disabled	0.9835	25.56	0
Road class: Unknown	–11.7686	–18.52	0
Road class: Major surface	–20.1951	–321.36	0
Road class: Minor surface	–16.9241	–262.02	0
Road class: Local	–20.1031	–283	0
Road class: Ramp	–11.0104	–189.6	0
PlasPy	–1.9049	–25.47	0
PlasNy	–2.01	–28.7	0
NlasNy	–1.2461	–24.79	0
NlasPy	–1.566	–36.64	0
PlasPpi	–1.6472	–17.32	0
NlasPpi	–0.6317	–8.73	0
NlasNpi	–0.7383	–9.88	0
PlasNpi	–1.4035	–13.01	0
PlasPro	2.0415	35.49	0
NlasPro	1.1364	29.88	0
NlasNro	1.5697	43.32	0
PlasNro	1.819	29.73	0
PyPro	–0.5621	–6.07	0
PyNro	–0.0686	–144.55	0
NyNro	–0.6218	–3.13	0.002
NyPro	–0.055	–72.34	0
PyPpi	–0.3765	–20.98	0
PyNpi	–0.3984	–18.47	0
NyNpi	–0.3637	–13.64	0
NyPpi	–0.3645	–19.75	0
NroNpi	0.3786	20.99	0
NroPpi	0.3412	22.99	0
ProPpi	0.2945	21	0
ProNpi	0.3131	17.17	0
PlaneOff	–2.2706	–45.14	0
NlaneOff	–1.8736	–47.31	0

Table 3.17b. Model 3: Longitudinal Speed, Single Regime, Interaction Kinematic Variables, All; Linear Regression

Variable Name	Coefficient	*t*	*p* > *t*	Variable Name	Coefficient	*t*	*p* > *t*
Constant	61.2216	1081.46	0	PyPpi	–0.187	–12.27	0
Measurement duration	–0.279	–65.24	0	PyNpi	–0.2215	–12.09	0
Dark	0.758	17.43	0	NyNpi	–0.265	–11.03	0
RDCW system disabled	0.886	23.22	0	NyPpi	–0.2412	–13.41	0
Road class: Unknown	–11.1077	–18.09	0	NroNpi	0.2509	16.03	0
Road class: Major surface	–19.878	–320.61	0	NroPpi	0.2096	15.93	0
Road class: Minor surface	–16.6159	–259.31	0	ProPpi	0.231	16.34	0
Road class: Local	–19.7223	–278.23	0	ProNpi	0.2524	14.81	0
Road class: Ramp	–10.9782	–191.14	0	PlanoffPlas	–0.1476	–0.79	0.427
PlasPy	–1.4205	–20.6	0	NlanoffNlas	–0.5811	–14	0
PlasNy	–1.6799	–25.34	0	PlanoffNlas	–1.7864	–9.79	0
NlasNy	–0.7116	–13.26	0	NlanoffPlas	–0.7072	–4.66	0
NlasPy	–1.0807	–23.28	0	PlanoffPy	–1.175	–31.41	0
PlasPpi	–0.7441	–7.97	0	PlanoffNy	–0.9226	–21.19	0
NlasPpi	0.3035	3.87	0	NlanoffPy	–0.6296	–20.97	0
NlasNpi	0.1211	1.55	0.121	NlanoffNy	–0.758	–23.68	0
PlasNpi	–0.595	–5.83	0	PlanoffPpi	–1.4029	–25.31	0
PlasPro	1.6342	28.72	0	PlanoffNpi	–1.1881	–21.54	0
NlasPro	0.6117	14.65	0	NlanoffPpi	–0.9522	–20.17	0
NlasNro	1.0995	27.07	0	NlanoffNpi	–0.8805	–19.84	0
PlasNro	1.2776	21.95	0	PlanoffPro	0.5273	14.49	0
PyPro	–0.5421	–5.94	0	PlanoffNro	1.0945	33.86	0
PyNro	–0.0678	–119.14	0	NlanoffPro	0.6313	25.37	0
NyNro	–0.6098	–3.35	0.001	NlanoffNro	0.5448	19.64	0
NyPro	–0.0478	–54.34	0				

(recall that each observation of an alert or pseudoalert begins 5 s before the alert and continues until 5 s after the alert is extinguished). The second regime is the magnitude of the defined measurement duration beyond 5 s. This is the duration of time after the alert is triggered until 5 s after the alert turns off—that is, it is the end of observation for each specific alert.

In Table 3.18, the fit of Model 4 is good, as indicated by an adjusted R^2 value of 0.47. Roadway classification has a negative effect on speed, since freeway is the baseline class (RC0 is insignificant due to a small sample size). Directional lane offset variables should be combined, as drivers would be assumed to decrease longitudinal speed in order to increase ease of repositioning their vehicles laterally regardless of offset direction. Roll parameters have positive signs (although lower coefficient absolute values than most other kinematics) because they imply larger values of longitudinal speed (higher speed relates to greater lateral force, which is translated to the vehicle body through roll). Yaw would tend to decrease longitudinal speed as a result of additional friction between the tires and the roadway when yaw does not equal zero. Lateral speed is associated with increases in longitudinal speed, since vehicles with higher speeds tend to have more difficulty maintaining lane position.

R^2 in Model 5 (Table 3.19) is about the same as in Model 4, indicating good model fit. The effects of the kinematic variables on longitudinal speed are similar to but slightly stronger than those in Model 4 (the signs are the same, but the absolute values are generally higher). The alert system being on may have caused drivers to react differently approaching curves, even before alerts were triggered. As in previous models, roadway classification tends to decrease longitudinal speed, given the baseline roadway classification of freeway. RC0 was dropped due to the absence of observations in this regime.

Table 3.18. Model 4: Longitudinal Speed, Two-Regime, Week 1, 5 s Before Alert, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	1.2521	10.04	0.0000
Positive yaw	–3.1605	–85.18	0.0000
Negative yaw	–1.6041	–36.75	0.0000
Positive roll	0.6950	19.74	0.0000
Negative roll	1.7320	54.31	0.0000
Positive pitch	–0.6966	–10.04	0.0000
Negative pitch	–0.5804	–8.70	0.0000
Positive lane offset	–1.7872	–11.22	0.0000
Negative lane offset	–1.8011	–13.66	0.0000
Positive lateral speed	1.8389	8.24	0.0000
Negative lateral speed	0.7882	4.40	0.0000
Road class: Unknown	2.8306	0.83	0.4060
Road class: Major surface	–18.7711	–117.72	0.0000
Road class: Minor surface	–16.1927	–99.69	0.0000
Road class: Local	–17.5117	–79.02	0.0000
Road class: Ramp	–9.1532	–63.32	0.0000
Constant	62.0448	443.54	0.0000

Number of obs = 34,700; Prob > F = 0.0000; R-squared = 0.4026; Adj R-squared = 0.4024.

Table 3.19. Model 5: Longitudinal Speed, Two-Regime, Weeks 2 to 4, 5 s Before Alert, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	0.8362	10.98	0.0000
Positive yaw	–2.256558	–112.88	0.0000
Negative yaw	–2.245391	–85.35	0.0000
Positive roll	1.1166	52.02	0.0000
Negative roll	1.0615	56.77	0.0000
Positive pitch	–0.6801366	–16.42	0.0000
Negative pitch	–0.5676225	–14.23	0.0000
Positive lane offset	–2.385719	–24.56	0.0000
Negative lane offset	–2.647468	–32.4	0.0000
Positive lateral speed	2.1194	13.77	0.0000
Negative lateral speed	1.7488	14.48	0.0000
Road class: Unknown	(dropped)	NA	NA
Road class: Major surface	–17.74598	–176.17	0.0000
Road class: Minor surface	–15.44258	–148.85	0.0000
Road class: Local	–19.32087	–163.8	0.0000
Road class: Ramp	–9.120569	–98.1	0.0000
Constant	61.4877	664.57	0.0000

Number of obs = 95,550; Prob > F = 0.0000; R-squared = 0.4719; Adj R-squared = 0.4718.

Table 3.20. Model 6: Longitudinal Speed, Two-Regime, Week 1, After Alert Triggered, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	0.4697	4.85	0.0000
Positive yaw	–2.742213	–129.57	0.0000
Negative yaw	–2.213393	–85.57	0.0000
Positive roll	1.4140	61.70	0.0000
Negative roll	1.8732	94.54	0.0000
Positive pitch	–0.43522	–8.4	0.0000
Negative pitch	–0.2240527	–4.97	0.0000
Positive lane offset	–1.900215	–16.88	0.0000
Negative lane offset	–0.7797679	–8.55	0.0000
Positive lateral speed	0.2729	1.71	0.0880
Negative lateral speed	0.0749	0.60	0.5490
Road class: Unknown	(dropped)	NA	NA
Road class: Major surface	–20.70069	–140.53	0.0000
Road class: Minor surface	–17.37473	–123.56	0.0000
Road class: Local	–17.57807	–96.2	0.0000
Road class: Ramp	–11.50564	–97.97	0.0000
Constant	59.3494	462.94	0.0000

Number of obs = 55,916; Prob > F = 0.0000; R-squared = 0.545; Adj R-squared = 0.5448.

Table 3.21. Model 7: Longitudinal Speed, Two-Regime, Weeks 2 to 4, After Alert Triggered, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	1.0133	17.67	0.0000
Positive yaw	–2.463306	–180.99	0.0000
Negative yaw	–1.949263	–160.11	0.0000
Positive roll	1.2893	108.46	0.0000
Negative roll	1.6317	128.17	0.0000
Positive pitch	–0.3776544	–12.21	0.0000
Negative pitch	–0.2904763	–10.49	0.0000
Positive lane offset	–1.93061	–27.09	0.0000
Negative lane offset	–1.431001	–24.74	0.0000
Positive lateral speed	–0.1336889	–1.25	0.2100
Negative lateral speed	0.8682	10.43	0.0000
Road class: Unknown	–12.0463	–11.28	0.0000
Road class: Major surface	–20.07718	–220.41	0.0000
Road class: Minor surface	–16.29487	–179.14	0.0000
Road class: Local	–19.30426	–203.35	0.0000
Road class: Ramp	–11.32301	–146.28	0.0000
Constant	58.1756	706.91	0.0000

Number of obs = 150,382; Prob > F = 0.0000; R-squared = 0.5184; Adj R-squared = 0.5184.

Table 3.22. Model 8: Longitudinal Speed, Two-Regime, Week 1, 5 s Before Alert, Add Measurement Duration, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	1.2525	10.04	0.0000
Positive yaw	–3.164436	–85.22	0.0000
Negative yaw	–1.609261	–36.83	0.0000
Positive roll	0.6985	19.82	0.0000
Negative roll	1.7355	54.38	0.0000
Positive pitch	–0.6952267	–10.02	0.0000
Negative pitch	–0.5789266	–8.67	0.0000
Positive lane offset	–1.772985	–11.12	0.0000
Negative lane offset	–1.785805	–13.53	0.0000
Positive lateral speed	1.8408	8.25	0.0000
Negative lateral speed	0.7782	4.35	0.0000
Road class: Unknown	2.6442	0.78	0.4380
Road class: Major surface	–18.76191	–117.64	0.0000
Road class: Minor surface	–16.18095	–99.58	0.0000
Road class: Local	–17.49641	–78.93	0.0000
Road class: Ramp	–9.127808	–63	0.0000
Measurement duration	–0.0925925	–2.57	0.0100
Constant	62.2482	387.28	0.0000

Number of obs = 34,700; Prob > F = 0.0000; R-squared = 0.4727; Adj R-squared = 0.4725.

Table 3.23. Model 9: Longitudinal Speed, Two-Regime, Weeks 2 to 4, 5 s Before Alert, Add Measurement Duration, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	0.8362	10.98	0.0000
Positive yaw	–2.256553	–112.72	0.0000
Negative yaw	–2.245389	–85.35	0.0000
Positive roll	1.1166	52.00	0.0000
Negative roll	1.0614	56.68	0.0000
Positive pitch	–0.68014	–16.41	0.0000
Negative pitch	–0.5676268	–14.23	0.0000
Positive lane offset	–2.385723	–24.56	0.0000
Negative lane offset	–2.64747	–32.4	0.0000
Positive lateral speed	2.1194	13.76	0.0000
Negative lateral speed	1.7488	14.48	0.0000
Road class: Unknown	(dropped)	NA	NA
Road class: Major surface	–17.74599	–176.11	0.0000
Road class: Minor surface	–15.4426	–148.78	0.0000
Road class: Local	–19.32089	–163.68	0.0000
Road class: Ramp	–9.120596	–97.91	0.0000
Measurement duration	0.0001	0.00	0.9960
Constant	61.4875	591.98	0.0000

Number of obs = 95,550; Prob > F = 0.0000; R-squared = 0.4719; Adj R-squared = 0.4718.

In Table 3.20, Model 6 has an improved R^2 (0.54) compared with the previous models. The coefficient for the constant is slightly lower than in Models 4 and 5, implying that longitudinal speed through the curve would be lower after an alert was triggered. Lateral speed is insignificant, regardless of direction. The signs for virtually all kinematic variables are the same as in Models 4 and 5, but the general effect of the variables decreases. This may be the result of anticipated drops in speed while traversing curves. Roadway classification has a generally lesser effect, but only marginally less.

In Table 3.21, R^2 is 0.518 for Model 7, very close to that in the previous model. Positive lateral speed becomes insignificant and changes sign. The effects of yaw and roll on longitudinal speed become stronger, while the effect of most other kinematic variables decreases. Roadway classification increases in general significance while having more overall effect (greater absolute value of coefficients). The Model 7 constant is slightly lower than in Model 6, as drivers would be expected to decrease their speed more quickly than in the system-disabled period.

In Table 3.22, the R^2 value for Model 8 is 0.47, similar to the same model without the measurement duration variable. Measurement duration is marginally significant and contributes little to the model's goodness of fit. All signs for kinematic variables are the same as in the models without measurement duration, and the coefficients are similar. The constant is higher than for all other two-regime models, as drivers would be expected to enter curves at higher speeds, relying on the alert system to warn them when deceleration is necessary.

In Table 3.23, R^2, variable signs, coefficients, and significances for Model 9 are similar to the model without measurement duration. Measurement duration has virtually no effect on longitudinal speed. Its *t*-statistic is not actually zero, since the *p*-value is not exactly one. STATA, the software package used to run these models, will display a zero for a *t*-statistic if it is close enough to zero based on the number of decimal places with zeros. The actual *t*-statistic is

$$\frac{\textit{coefficient}}{\textit{standard error}} = \frac{0.0001049}{0.0218084} = 0.0048$$

In Table 3.24, the R^2 in Model 10 is slightly higher than in the comparable model without measurement duration (Model 6). As in Model 9, coefficient values are similar and the constant is slightly higher; measurement duration is significant and has a negative effect on longitudinal speed.

In Table 3.25, the R^2 for Model 11 is 0.526, slightly higher than in the model without measurement duration. Positive

Table 3.24. Model 10: Longitudinal Speed, Two-Regime, Week 1, After Alert, Add Measurement Duration, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	0.3240	3.35	0.0010
Positive yaw	–2.7082	–128.32	0.0000
Negative yaw	–2.1695	–84.09	0.0000
Positive roll	1.4102	61.84	0.0000
Negative roll	1.9022	96.31	0.0000
Positive pitch	–0.4412	–8.56	0.0000
Negative pitch	–0.2392	–5.33	0.0000
Positive lane offset	–1.9257	–17.19	0.0000
Negative lane offset	–0.5608	–6.15	0.0000
Positive lateral speed	0.1085	0.68	0.4950
Negative lateral speed	–0.0466	–0.37	0.7080
Road class: Unknown	(dropped)	NA	NA
Road class: Major surface	–20.8376	–142.07	0.0000
Road class: Minor surface	–17.5749	–125.40	0.0000
Road class: Local	–17.7594	–97.61	0.0000
Road class: Ramp	–11.5181	–98.58	0.0000
Measurement duration	–0.2314	–23.88	0.0000
Constant	61.3664	401.15	0.0000

Number of obs = 55,916; Prob > F = 0.0000; R-squared = 0.5496; Adj R-squared = 0.5494.

Table 3.25. Model 11: Longitudinal Speed, Two-Regime, Weeks 2 to 4, After Alert, Add Measurement Duration, Linear Regression

Variable Name	Coefficient	*t*-statistic	*p*-value
Dark	0.9294	16.32	0.0000
Positive yaw	–2.3985	–176.70	0.0000
Negative yaw	–1.9214	–158.84	0.0000
Positive roll	1.3072	110.74	0.0000
Negative roll	1.6316	129.14	0.0000
Positive pitch	–0.3959	–12.90	0.0000
Negative pitch	–0.3001	–10.93	0.0000
Positive lane offset	–1.9495	–27.57	0.0000
Negative lane offset	–1.4182	–24.71	0.0000
Positive lateral speed	–0.1298	–1.23	0.2200
Negative lateral speed	0.7935	9.61	0.0000
Road class: Unknown	–10.8223	–10.21	0.0000
Road class: Major surface	–20.1622	–222.99	0.0000
Road class: Minor surface	–16.4879	–182.46	0.0000
Road class: Local	–19.2835	–204.68	0.0000
Road class: Ramp	–11.2679	–146.66	0.0000
Measurement duration	–0.2897	–48.01	0.0000
Constant	60.6840	625.89	0.0000

Number of obs = 150,382; Prob > F = 0.0000; R-squared = 0.5257; Adj R-squared = 0.5256.

lateral speed is the only insignificant variable. All variables in this model have the same signs and similar effects as in the model without measurement duration.

Table 3.26 shows the summary of deceleration results from the two-regime models. There is clear evidence of driver adaptation to the CSW technology. During the first week (pseudoalerts), drivers approached curves during the first 5 s of measurement at a deceleration rate of 0.093 mph/0.1 s. After activation of the system (Weeks 2 to 4), this same 5-s time period had virtually no deceleration (–0.0001 mph/0.1 s). Changes were also observed in deceleration after an alert was triggered (compared with deceleration during the pseudo-alert). In this case, drivers decelerated at a rate of 0.231 mph/0.1 s compared with 0.29 mph/0.1 s. Taken in combination, this driver adaptation indicates that when the CSW is engaged, drivers approach the curve at a constant speed and then decelerate relatively rapidly compared with a decelerating entry and less rapid deceleration without the technology. One interpretation is that the drivers are relying on the system to warn them of an unsafe curve entry rather than approaching curves more cautiously. It is recognized that the origins of these parameters are models that include all drivers, a form of aggregate analysis. The next steps were to construct similar models for individual drivers or smaller groups of drivers.

Adding measurement duration affected the models slightly, but there were no drastic changes in coefficients and goodness of fit. Lateral speed had little effect on longitudinal speed for any two-regime model. When measurement duration was added to the models, the R^2 value increased slightly for most models. The effect of specific kinematic factors tended to be relatively constant across the models. Because the baseline classification was freeway, all roadway classification variables had negative coefficients. Measurement duration played a more important role in the after-alert-was-triggered models; the effect of time on change in longitudinal speed should be more noticeable after an alert is triggered (drivers would be expected to decrease speed more significantly after an alert is triggered to decrease the degree of danger).

Table 3.26. Summary of Deceleration Results for Two-Regime Models

	mph/0.1 s	
	Week 1	**Week 2–4**
5 s before alert (0 ≤ measurement duration ≤ 5)	0.0926	–0.0001
After alert triggered	0.2314	0.2897

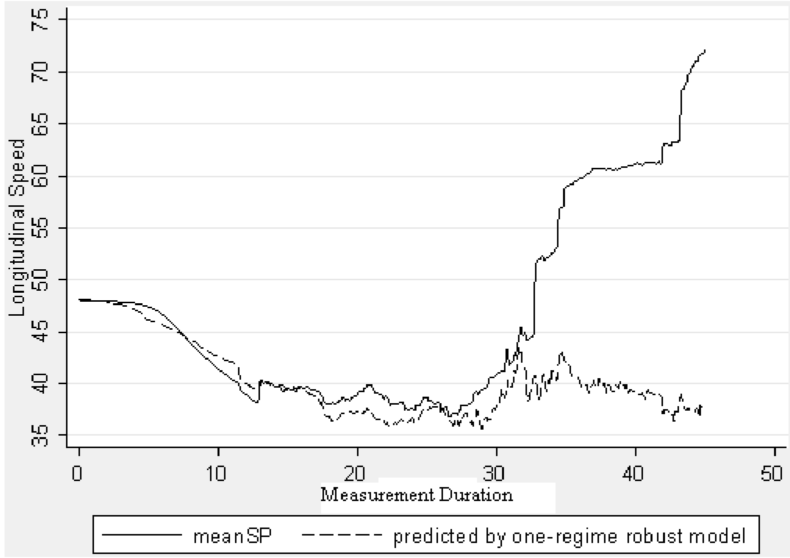

Figure 3.9. Longitudinal speed versus measurement duration (solid line is mean of data points).

Three-Regime Models

The two-regime model assumes that one model can be used to characterize driver longitudinal speed for alert durations as long as 15 to 20 s. There are relatively few of these long-duration events, but they may need a different model, and they may also influence the estimation of the models of short-to-moderate duration. In addition, few of the observations 20 s from the alert trigger were accurately estimated in the two-regime model. Figure 3.9 summarizes model fit to the data for the best two-regime model.

Thus, the next step was to create a model with three regimes. All models included the measurement duration variable. As in the two-regime models, there was a division at measurement duration of 5 s and, in addition, a second division at measurement duration of 20 s. The estimation results for the three-regime model are shown in Tables 3.27 through 3.29.

One noticeable feature of the models in Tables 3.28 and 3.29 is the substantial change in the estimated value for measurement duration: drivers with the system accelerate more aggressively after 20 s than drivers without the system. This difference may reflect a greater confidence in system users that they have, in fact, exited the curve, but it nevertheless demonstrates driver adaptation.

A Chow test was performed to determine if there were significant changes in parameters as a whole for the time periods 5 s before the alert and the time after the alert was triggered. Table 3.30 shows that each pair of models differed on the basis that the parameter coefficients were sufficiently different. This test confirms that driver behavior changes with the system activated compared with the system not activated.

Cohort-Based Approach

The cohort design can be used to formulate an exposure-based model relating potential risk factors to several possible outcomes. The cohort design is well-suited to account for measures of exposure such as time at risk or distance traveled under specific driving conditions. Survival analysis, count regression, and logistic regression are suitable statistical methods to analyze data from a cohort design. The Penn State team estimated several models with both CSW and LDW alerts as predictors. Only CSW alert findings are reported because they are principally the same as the LDW alert findings and CSW alerts are, as discussed above, more correlated with roadway departure events.

The count regression models for CSW alerts highlight the effect of roadway-related context. The team estimated and compared two models: one for limited access roads (UMTRI Functional Class 1) and the other for nonlimited access roads (UMTRI Functional Class 3). Logistic regression modeling highlights the importance of driver variables and the effect that driver variables have on model fit, parameter significance, magnitude, and sign. Interestingly, the results are rather different from those for the event-based models developed with the VTTI data.

Count Regression

The initial analysis involved the use of negative binomial (NB) count regressions to show how both context and driver-related variables affect the likelihood of alert occurrence. In the first set of sample models, the data were segmented by roadway functional classification. In the second set of sample models, multi-

Table 3.27. Model 12: Longitudinal Speed, Three-Regime, 5 s Before Alert, Linear Regression

Model 12a Week 1				Model 12b Weeks 2–4			
Variable Name	**Coefficient**	***t*-statistic**	***p*-value**	**Variable Name**	**Coefficient**	***t*-statistic**	***p*-value**
Dark	1.2517	9.11	0.0000	Dark	0.8362	10.60	0.0000
Positive yaw	–3.1645	–53.23	0.0000	Positive yaw	–2.2566	–80.07	0.0000
Negative yaw	–1.6091	–21.79	0.0000	Negative yaw	–2.2454	–41.04	0.0000
Positive roll	0.6983	13.46	0.0000	Positive roll	1.1166	30.34	0.0000
Negative roll	1.7354	37.36	0.0000	Negative roll	1.0614	44.56	0.0000
Positive pitch	–0.6950	–10.68	0.0000	Positive pitch	–0.6801	–17.40	0.0000
Negative pitch	–0.5788	–9.32	0.0000	Negative pitch	–0.5676	–15.12	0.0000
Positive lane offset	–1.7731	–12.07	0.0000	Positive lane offset	–2.3857	–23.22	0.0000
Negative lane offset	–1.7836	–14.11	0.0000	Negative lane offset	–2.6475	–32.64	0.0000
Positive lateral speed	1.8396	7.84	0.0000	Positive lateral speed	2.1194	13.21	0.0000
Negative lateral speed	0.7765	4.90	0.0000	Negative lateral speed	1.7488	14.84	0.0000
Road class: Major surface	–18.7643	–119.89	0.0000	Road class: Major surface	–17.7460	–178.11	0.0000
Road class: Minor surface	–16.1835	–94.49	0.0000	Road class: Minor surface	–15.4426	–136.77	0.0000
Road class: Local	–17.4991	–85.12	0.0000	Road class: Local	–19.3209	–159.24	0.0000
Road class: Ramp	–9.1299	–58.20	0.0000	Road class: Ramp	–9.1206	–91.98	0.0000
Measurement duration	–0.0932	–2.58	0.0100	Measurement duration	0.0001	0.00	0.9960
Constant	62.2522	376.14	0.0000	Constant	61.4875	561.09	0.0000

Number of obs = 34,700; Prob > F = 0.0000; R-squared = 0.4727; Adj R-squared = 0.4725.

Number of obs = 95,500; Prob > F = 0.0000; R-squared = 0.4719; Adj R-squared = 0.4718.

Table 3.28. Model 13: Longitudinal Speed, Three-Regime, Alert Triggered to 20 s, Linear Regression

Model 13a Week 1				Model 13b Weeks 2–4			
Variable Name	**Coefficient**	***t*-statistic**	***p*-value**	**Variable Name**	**Coefficient**	***t*-statistic**	***p*-value**
Dark	0.3660	3.37	0.0010	Dark	1.0610	17.61	0.0000
Positive yaw	–2.7340	–96.83	0.0000	Positive yaw	–2.3950	–137.45	0.0000
Negative yaw	–2.1470	–62.77	0.0000	Negative yaw	–1.8980	–98.45	0.0000
Positive roll	1.4360	49.50	0.0000	Positive roll	1.3430	82.35	0.0000
Negative roll	1.9510	75.38	0.0000	Negative roll	1.6760	108.50	0.0000
Positive pitch	–0.4370	–8.75	0.0000	Positive pitch	–0.4380	–15.07	0.0000
Negative pitch	–0.2220	–5.00	0.0000	Negative pitch	–0.3330	–12.58	0.0000
Positive lane offset	–1.9350	–17.25	0.0000	Positive lane offset	–1.9080	–27.28	0.0000
Negative lane offset	–0.7630	–8.84	0.0000	Negative lane offset	–1.4710	–26.92	0.0000
Positive lateral speed	0.3260	2.08	0.0370	Positive lateral speed	–0.1220	–1.21	0.2260
Negative lateral speed	0.1840	1.69	0.0910	Negative lateral speed	0.6800	8.70	0.0000
Road class: Major surface	–20.8720	–130.05	0.0000	Road class: Major surface	–20.0790	–210.70	0.0000
Road class: Minor surface	–17.5940	–116.31	0.0000	Road class: Minor surface	–16.5850	–172.11	0.0000
Road class: Local	–17.7790	–94.20	0.0000	Road class: Local	–19.3090	–189.80	0.0000
Road class: Ramp	–11.6630	–81.92	0.0000	Road class: Ramp	–11.4410	–132.71	0.0000
Measurement duration	–0.4190	–28.81	0.0000	Measurement duration	–0.5480	–63.66	0.0000
Constant	62.8520	340.11	0.0000	Constant	62.7250	562.56	0.0000

Number of obs = 537,880; Prob > F = 0.0000; R-squared = 0.5539; Adj R-squared = 0.5539.

Number of obs = 145,707; Prob > F = 0.0000; R-squared = 0.5347; Adj R-squared = 0.5347.

Table 3.29. Model 14: Longitudinal Speed, Three-Regime, 20+ s, Linear Regression

Model 14a Week 1				Model 14b Weeks 2–4			
Variable Name	**Coefficient**	***t*-statistic**	***p*-value**	**Variable Name**	**Coefficient**	***t*-statistic**	***p*-value**
Dark	–3.6790	–11.22	0.0000	Dark	–1.1040	–2.75	0.0060
Positive yaw	–1.7680	–17.69	0.0000	Positive yaw	–1.4590	–19.13	0.0000
Negative yaw	0.4560	1.32	0.1860	Negative yaw	–1.1830	–9.32	0.0000
Positive roll	0.3010	1.34	0.1800	Positive roll	1.1090	11.53	0.0000
Negative roll	1.5030	13.46	0.0000	Negative roll	0.8630	10.68	0.0000
Positive pitch	0.0380	0.15	0.8780	Positive pitch	0.3610	2.50	0.0120
Negative pitch	–0.2570	–1.42	0.1570	Negative pitch	0.5820	5.07	0.0000
Positive lane offset	–2.7010	–5.51	0.0000	Positive lane offset	–2.6160	–9.16	0.0000
Negative lane offset	0.3910	1.14	0.2550	Negative lane offset	–2.3950	–9.48	0.0000
Positive lateral speed	–7.2000	–6.73	0.0000	Positive lateral speed	–3.9790	–6.81	0.0000
Negative lateral speed	–3.6780	–4.45	0.0000	Negative lateral speed	1.6290	2.57	0.0100
Road class: Major surface	–12.6560	–14.53	0.0000	Road class: Major surface	–28.0880	–16.84	0.0000
Road class: Minor surface	–12.1850	–17.31	0.0000	Road class: Minor surface	–4.7560	–5.97	0.0000
Road class: Local	–8.0200	–8.73	0.0000	Road class: Local	–14.6600	–18.70	0.0000
Road class: Ramp	–14.8210	–21.85	0.0000	Road class: Ramp	–9.6080	–13.95	0.0000
Measurement duration	0.3190	6.32	0.0000	Measurement duration	0.7820	22.95	0.0000
Constant	48.7330	31.49	0.0000	Constant	34.3720	27.52	0.0000

Number of obs = 2,158; Prob > F = 0.0000; R-squared = 0.5554; Adj R-squared = 0.5554.

Number of obs = 4,766; Prob > F = 0.0000; R-squared = 0.5171; Adj R-squared = 0.5171.

Table 3.30. Chow Test for Comparison of Parameter Estimates

	Models 12a and 12b	**Models 13a and 13b**
SSR pooled	12252145.2	16207567.5
SSR Model 1	3218375.28	4537057.93
SSR Model 2	8950406.52	11583691.7
k	17	17
n1	34700	53788
n2	95550	145707
Numerator	4903.729412	5106.933529
Denominator	94.07076341	81.24429421
F	52.4856042	63.04938272
Degree N	17	17
Degree D	130250	199495
Result	$F > 1.96$	$F > 1.96$
	$p = 0$	$p = 0$

level specification was applied to cluster driver attributes at a second, separate level. Note that road class was used, but additional dimensions beyond road class could have been specified, such as day/night or wet/dry conditions. The cohort may be defined quite flexibly, using any variable that is an attribute of the road, environment, and/or driver (and is, of course, continuously measured as part of the naturalistic data).

Single-Level Models

The single-level models segment the data by functional class and alert type. Initial context-related predictors include ramp presence (for nonlimited access roads only), urban/rural settings, day/night, dry/wet conditions (based on the use of windshield wipers), and RDCW system disabled/enabled state. Driver predictors included gender, education, years of driving experience, last year's mileage driven, use of glasses or contacts, and whether or not the driver is a smoker. Two-way interaction terms were tested for both context and driver attributes. Note that the structure of the model bears a strong similarity to the event-based models estimated using VTTI data. The VTTI models were able to capture only those attributes immediately surrounding the event. The cohort formulation includes many of the same variables, but the cohort models include exposure measured on the same scale as context, which is important in obtaining a broader view of the effect of context throughout the driver's travel.

Table 3.31 summarizes the model results for limited access segments. Factors increasing the number of alerts on limited access segments include exposure in the form of distance, dry daytime conditions, urban settings, being male, high mileage in the previous year if the driver is a female, and being a male with a bachelor's degree or above; all other predictors decrease alert frequency. Six of the predictors in this model are insignificant. Higher driving experience generally drives down CSW occurrence, and wet conditions would likely decrease alert frequency as drivers tend to decrease travel speeds in wet weather. The use of the NB model over the Poisson model was warranted based on alpha and its likelihood ratio test. A chi-square test of the model compared to only a constant term yielded a test statistic value of 47.89 with a corresponding significance probability less than 0.001.

The following factors increase CSW alert frequency on Functional Class 3 roads: distance as a form of exposure, daytime conditions, urban roadways (both ramps and nonramps), rural ramps (insignificant), males with more mileage driven in the last year (although gender–mileage interactions are insignificant), and males with a bachelor's degree or above (although the gender–education interaction term is generally insignificant, at least marginally). Dry conditions also increase alert frequency based on the fact that the wet variable has a negative coefficient. Overdispersion results show that the NB regression is a better choice than the Poisson regression. These results are summarized in Table 3.32. Once again a chi-square test of the model compared to a constant term should that the model was significant with a significance probability less than 0.001.

Initially, the team expected to see a positive correlation between speed and CSW alert counts. However, the result showed that such a correlation does not always occur. Checking scatter plots of speed against CSW count, it was observed that some speeds were below 18 mph, which is the minimum speed required to trigger CSW alerts. This observation illustrates the need to carefully define *homogeneous* when using naturalistic data with a cohort data structure. There is tremendous power in the method, but only if recognized in the collection of the original data set.

The count regression approach with cohort structure can be used to explore crash surrogate measures and their utility in safety analyses. One extension of the models in Tables 3.30 and 3.31 is the inclusion of potential crash surrogates as predictor variables. A count of the frequency of a surrogate occurrence can be used as a predictor, and its association and significance can be tested against the dependent measure.

Table 3.31. CSW NB Regression, Functional Class 1: Limited Access, Distance as Exposure

CSW	Coefficient	SE	*z*	*p* > *z*	95% CI
Miles driven	0.027	0.006	4.700	<0.001	(0.016, 0.039)
RDCW disabled	–0.362	0.236	–1.540	0.125	(–0.825, 0.100)
Nighttime	baseline	NA	NA	NA	NA
Daytime wet	–0.402	0.376	–1.070	0.284	(–1.139, 0.334)
Daytime dry	1.131	0.366	3.090	0.002	(0.413, 1.849)
Urban	1.296	0.389	3.330	0.001	(0.534, 2.058)
Male	0.777	0.717	1.080	0.279	(–0.629, 2.183)
Female last year's mileage (per 1,000 mi)	0.080	0.033	2.400	0.017	(0.0145, 0.145)
Male last year's mileage (per 1,000 mi)	–0.014	0.023	–0.600	0.550	(–0.059, 0.035)
Male driving experience (years)	–0.031	0.011	–2.830	0.005	(–0.052, –0.009)
Female driving experience (years)	–0.043	0.012	–3.650	<0.001	(–0.066, –0.020)
Female with bachelor's degree or above	–0.532	0.399	–1.330	0.183	(–1.314, 0.251)
Male with bachelor's degree or above	0.325	0.315	1.030	0.301	(–0.291, 0.942)
Constant	–2.807	0.613	–4.580	<0.001	(–4.009, –1.606)
Alpha	1.220	0.348	NA	NA	(0.697, 2.133)

Number of observations = 405
log likelihood = –275.43729
LR chi-squared (12) = 98.54
pseudo R^2 = 0.1517
LR test of alpha = 0
chi-bar squared (01) = 47.89

Table 3.32. CSW NB Regression, Functional Class 3: Nonlimited Access, Distance As Exposure

CSW	Coefficient	SE	z	p > z	95% CI
Miles driven	0.087	0.016	5.440	<0.001	(0.055, 0.118)
RDCW disabled	–0.762	0.133	–5.730	<0.001	(–1.023, –0.502)
Day	1.054	0.150	7.030	<0.001	(0.760, 1.348)
Wet	–1.535	0.175	–8.780	<0.001	(–1.878, –1.193)
Rural nonramp	baseline	NA	NA	NA	NA
Urban ramp	2.184	0.270	8.090	<0.001	(1.655, 2.714)
Rural ramp	0.493	0.377	1.310	0.191	(–0.246, 1.231)
Urban nonramp	1.068	0.254	4.200	<0.001	(0.569, 1.566)
Male	–0.542	0.417	–1.300	0.194	(–1.359, 0.275)
Female last year's mileage	–0.001	0.017	–0.060	0.950	(–0.034, 0.032)
Male last year's mileage	0.015	0.011	1.330	0.183	(–0.007, 0.037)
Male driving experience (years)	–0.016	0.006	–2.790	0.005	(–0.027, –0.005)
Female driving experience (years)	–0.023	0.006	–4.090	<0.001	(–0.035, –0.012)
Female with bachelor's degree or above	–0.178	0.207	–0.860	0.390	(–0.583, 0.228)
Male with bachelor's degree or above	0.285	0.163	1.750	0.080	(–0.034, 0.605)
Constant	–1.702	0.411	–4.140	<0.001	(–2.508, –0.895)
Alpha	0.987	0.143	NA	NA	(0.742, 1.312)

Number of observations = 900
log likelihood = –832.47787
LR chi-squared (14) = 320.99
pseudo R^2 = 0.1616
LR test of alpha = 0
chi-bar squared (01) = 199.53

This would show an association between a surrogate and an event of interest such as crashes.

Another way to explore surrogate measures is to use them as dependent variables. The variable is entered as a count on a segment similar to the way crashes would be entered for an identification of sites with promise (see Aguero-Valverde and Jovanis 2008 for a recent example of the standard sites with promise formulation). Bivariate Poisson–log normal or similar formulations within a Bayes hierarchical structure (Aguero-Valverde and Jovanis 2010) can be used with crash and surrogate frequency as the dependent variables. Using a common specification, the researcher could explore differences in the significance of predictors. Of even greater utility would be the development of safety performance functions for both crashes and surrogate measures. One could then compare the sites with promise developed for the two safety performance functions. A test of the validity of a surrogate would be its ability to identify the same sites with promise. The ability to validate a surrogate in this way is of particular importance in that one application of surrogates is to identify risky locations without waiting for years of crash data. Specific crash types such as roadway departure could be paired with and tested against relevant surrogates (such as lateral accelerations) to obtain more targeted evaluations. The cohort formulation would allow the validity of surrogates to be tested using safety performance functions. This concept could be explored as part of the SHRP 2 S08 projects.

Multilevel Models

Since the output from models including either distance or time as exposure were consistent using the single-level structure, the team only considered models including distance as a form of exposure as analysis examples for further multilevel formulation. The goal of this model development is to demonstrate the application of hierarchical models to cohort-structured data.

Figure 3.10 summarizes the application of the multilevel approach to CSW alerts on limited access roads (Functional Class 1). The first equation in Figure 3.10 says that the number of CSW alerts obeys the NB distribution. The predictors used here are those used in the best single-level models. The second equation says that the expected number of CSW alerts (log π) is a function of miles driven (miles), RDCW disabled

$$
\begin{aligned}
&\mathrm{CSW}_{ij} \sim \text{-ve Binomial}(\pi_{ij})\\
&\log(\pi_{ij}) = \beta_{0j}\mathrm{cons} + 0.018(0.005)\mathrm{miles}_{ij} + -0.467(0.201)\mathrm{rdcwdisabled}_{ij} + -0.814(0.368)\mathrm{wetday}_{ij} +\\
&\qquad 0.631(0.244)\mathrm{dryday}_{ij} + -1.640(0.669)\mathrm{wetnight}_{ij} + \beta_{6j}\mathrm{gender}_j + \beta_{7j}\mathrm{Fbsabove}_j + \beta_{8j}\mathrm{Mbsabove}_j +\\
&\qquad \beta_{9j}\mathrm{Fmile}_j + \beta_{10j}\mathrm{Mmile}_j + \beta_{11j}\mathrm{Fexp}_j + \beta_{12j}\mathrm{Mexp}_j + 1.379(0.362)\mathrm{urban}_{ij}\\
&\beta_{0j} = -2.891(0.856) + u_{0j}\\
&\beta_{6j} = 1.283(1.054) + u_{6j}\\
&\beta_{7j} = -0.352(0.554) + u_{7j}\\
&\beta_{8j} = 0.238(0.454) + u_{8j}\\
&\beta_{9j} = 0.093(0.045) + u_{9j}\\
&\beta_{10j} = -0.019(0.029) + u_{10j}\\
&\beta_{11j} = -0.040(0.017) + u_{11j}\\
&\beta_{12j} = -0.032(0.015) + u_{12j}
\end{aligned}
$$

$$
\begin{bmatrix} u_{0j}\\ u_{6j}\\ u_{7j}\\ u_{8j}\\ u_{9j}\\ u_{10j}\\ u_{11j}\\ u_{12j} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix}
0.977(0.327) & & & & & & & \\
0 & 0.000(0.000) & & & & & & \\
0 & 0 & 0.000(0.000) & & & & & \\
0 & 0 & 0 & 0.000(0.000) & & & & \\
0 & 0 & 0 & 0 & 0.000(0.000) & & & \\
0 & 0 & 0 & 0 & 0 & 0.000(0.000) & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0.000(0.001) & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.000(0.000)
\end{bmatrix}
$$

$$
\mathrm{var}(\mathrm{CSW}_{ij} \mid \pi_{ij}) = \pi_{ij} + \pi_{ij}^{2}/v
$$

Figure 3.10. Multilevel NB model: Functional Class 1, limited access, CSW.

status (rdcwdisabled), the interaction between daylight and the use of windshield wipers (wetday), the interaction between daylight and no use of windshield wipers (dryday), the interaction between night and the use of windshield wipers (wetnight), gender (1 if male), females with bachelor's degrees or above (Fbsabove), males with bachelor's degrees or above (Mbsabove), the interaction between females and last year's miles driven in thousands (Fmiles), the interaction between males and last year's miles driven in thousands (Mmiles), the interaction between females and years of driving experience (Fexp), the interaction between males and years of driving experience (Mexp), and urban/rural settings. The link function used here is logarithm.

The unit of the first level is the context combination (cohort). The coefficients of the predictors in the first level are shown in the second equation (with SEs in parentheses). The variable miles (miles traveled in the homogeneous trip segment) indicates the exposure measured directly. Greater exposure results in a higher expected number of alerts triggered. The negative sign of rdcwdisabled implies that the number of CSW alerts triggered during Weeks 2 to 4 is greater than in Week 1, resulting from higher exposure (3 weeks with system enabled versus 1 week with system disabled). The baseline for the group of interacting variables (wiper use and day/night) is the interaction between night and the absence of wiper usage; wet reduces the expected number of alerts triggered in both daytime and nighttime conditions.

The unit of the second level is the individual driver. The coefficients for the second-level predictors, such as gender, bachelor's degree or above, last year's miles driven, and years of driving experience, are shown after the second equation. Specifically, it is a random intercept and random slope model formulation (i.e., both the intercept and the slope vary randomly across the subjects). Thus, the third equation says that the mean constant term for all drivers is −2.891 (SE is 0.856), and their variance is 0.977 (with SE of 0.327; these values are shown in the covariance matrix for all random effects after the second-level predictors). The SE for the coefficient (0.856) is used to construct the CI for the estimated parameter. The variance for the random intercepts (0.977) indicates how the intercepts vary across individual subjects. In other words, the individual subject intercept varies about this mean (−2.891) with a variance estimated as 0.977 (SE is 0.327). Similarly, the fourth equation says that the mean gender effect is 1.283 with SE of 1.054, suggesting that males tend to have higher numbers of CSW alerts. The individual subject slopes do not vary about this mean on Functional Class 1 because the value for the variance of gender (u_{61}) is 0.000.

Concerning multilevel model random effect covariance, it can be assumed either that the second-level predictors are independent of each other or that they are correlated to each other. While a more generalized setup can be used to specify a correlated covariance matrix, the team had to assume independence (hence values under the diagonal were restricted to being zeros) because of computational difficulties. Along the diagonal, the variance for the random constant term is 0.977, which is much greater than that for other random effects, implying that individual drivers provide the main source of variation. This finding illustrates that the multilevel approach applied to cohort-based event data can potentially identify driver-related factors that would be difficult or impossible to detect using other typical approaches (such as those applied to the VTTI data).

All predictors have the same sign and similar magnitudes as the single-level model for expected numbers of CSW alerts triggered and can be interpreted similarly; however, the SEs in the single-level models were underestimated. The following factors increase CSW alert frequency on Functional Class 1 roads: exposure in the form of distance, dry daytime conditions, urban settings, being male, higher previous mileage as a female, and being a male with a bachelor's degree or above; all other predictors decrease alert frequency. The variables of being male, being of either gender with a bachelor's degree or above, and higher previous mileage as a male in this model are insignificant.

Logistic Regression Models

The UMTRI data were also applied to logit models, using a single-level structure, to compare alert events with nonalerts, which can only be done using a homogenous trip segment structure. Thus, these models were structured as event-based models using homogeneous trip segment data. There are 331,641 homogeneous trip segments for all drivers in the data set. Each of the 2,605 CSW and 10,452 LDW alerts were matched to the segments on which they occurred. These logit models can be used to form the basis of a case–control study, which matches cases and controls (noncases) within the limits of certain confounding factors (e.g., posted speed limit and average annual daily traffic).

Compared with the baseline functional class, all other functional classes except Functional Class 3 (limited access, which was insignificant) increase a driver's odds of having a CSW alert. Ramps substantially increase alert odds, while daytime conditions slightly increase the odds. Urban settings, wet conditions (based on windshield wiper use), driving with the RDCW system disabled, higher minimum segment speeds, and longer segment distances decrease CSW odds, although urban settings and windshield wiper use are insignificant. Higher maximum segment speed and higher numbers of brake applications on a segment slightly increase CSW alert odds, although brake applications are insignificant. The pseudo R^2 indicates a moderately reasonable fit for this model (see Table 3.33).

Table 3.33. Single-Level Cohort-Based Logit Model, CSW

CSW	Coefficient	OR	SE for OR	z	p > z	95% CI for OR
Functional Class 1: Limited access	baseline	1.00	NA	NA	NA	NA
Functional Class 2: Limited access	0.42	1.528	0.203	3.190	0.001	(1.178, 1.983)
Functional Class 3: Limited access	0.57	1.760	0.804	1.240	0.216	(0.719, 4.309)
Functional Class 1: Nonlimited access	1.11	3.024	0.362	9.250	<0.001	(2.392, 3.823)
Functional Class 2: Nonlimited access	1.60	4.935	0.487	16.160	<0.001	(4.066, 5.988)
Functional Class 3: Nonlimited access	1.00	2.712	0.251	10.780	<0.001	(2.262, 3.251)
Functional Class 4: Nonlimited access	0.99	2.694	0.251	10.650	<0.001	(2.245, 3.233)
Functional Class 5: Nonlimited access	1.56	4.772	0.498	14.980	<0.001	(3.890, 5.855)
Ramp	2.22	9.180	0.517	39.350	<0.001	(8.220, 10.251)
Daytime	0.21	1.239	0.063	4.200	<0.001	(1.121, 1.370)
Urban	–0.13	0.882	0.062	–1.790	0.074	(0.769, 1.012)
Windshield wiper use (on/off)	–0.06	0.940	0.023	–2.530	0.011	(0.896, 0.986)
RDCW disabled	–0.16	0.855	0.040	–3.320	0.001	(0.780, 0.938)
Maximum segment speed	0.19	1.205	0.004	50.080	<0.001	(1.196, 1.214)
Minimum segment speed	–0.11	0.894	0.002	–44.880	<0.001	(0.890, 0.898)
No. of brake applications on segment	0.00	1.003	0.003	1.000	0.317	(0.997, 1.008)
Segment distance (mi)	–0.03	0.973	0.005	–5.460	<0.001	(0.964, 0.983)

Number of observations = 331,641
log likelihood = –10922.3
LR chi-squared (16) = 8595.98
probability > chi-squared = 0.001
pseudo R^2 = 0.2824

Table 3.34. Single-Level Cohort-Based Logit Model, CSW with Driver Attributes

Variable	Coefficient	OR	SE for OR	z	p > z	95% CI for OR
Functional Class 1: Limited access	baseline	1.00	NA	NA	NA	NA
Functional Class 2: Limited access	0.26	1.295	0.204	1.640	0.100	(0.951, 1.764)
Functional Class 3: Limited access	0.27	1.305	0.766	0.450	0.650	(0.413, 4.122)
Functional Class 1: Nonlimited access	1.02	2.770	0.369	7.650	<0.001	(2.134, 3.596)
Functional Class 2: Nonlimited access	1.64	5.174	0.565	15.040	<0.001	(4.176, 6.410)
Functional Class 3: Nonlimited access	1.07	2.926	0.298	10.520	<0.001	(2.395, 3.573)
Functional Class 4: Nonlimited access	1.02	2.776	0.283	10.000	<0.001	(2.272, 3.390)
Functional Class 5: Nonlimited access	1.70	5.458	0.627	14.770	<0.001	(4.357, 6.836)
Rural nonramp	baseline	1.00	NA	NA	NA	NA
Urban nonramp	–0.05	0.953	0.081	–0.570	0.572	(0.807, 1.126)
Urban ramp	2.13	8.410	0.775	23.100	<0.001	(7.020, 10.075)
Rural ramp	2.28	9.797	1.657	13.490	<0.001	(7.033, 13.648)
Daytime	0.28	1.321	0.078	4.730	<0.001	(1.177, 1.482)
Wet conditions	–0.31	0.737	0.062	–3.620	<0.001	(0.625, 0.870)
RDCW disabled	–0.15	0.863	0.044	–2.860	0.004	(0.780, 0.955)
Maximum segment speed	0.19	1.205	0.005	44.550	<0.001	(1.195, 1.215)
Minimum segment speed	–0.11	0.893	0.003	–40.160	<0.001	(0.888, 0.898)
No. of brake applications on segment	0.01	1.006	0.004	1.700	0.089	(0.999, 1.013)
Segment distance (mi)	0.13	1.137	0.069	2.100	0.036	(1.008, 1.281)
Female with bachelor's or above	–0.32	0.727	0.064	–3.620	<0.001	(0.612, 0.864)
Male with bachelor's or above	–0.01	0.994	0.002	–3.260	0.001	(0.990, 0.997)
Male years of driving experience	–0.01	0.989	0.002	–5.580	<0.001	(0.985, 0.993)
Female years of driving experience	0.0004	1.0004	0.0057	0.730	0.466	(1.000, 1.000014)
Female last year's mileage driven	0.99	1.000	0.0034	2.930	0.003	(1.000, 1.000)
Male last year's mileage driven	–0.02	0.978	0.006	–3.370	0.001	(0.965, 0.991)

Number of observations = 279,166
log likelihood = –9231.2975
pseudo R^2 = 0.2889
LR chi-squared (23) = 7501.19

Table 3.34 shows a model that incorporates driver attributes as predictors. As in the first CSW logit model, all functional classes increase CSW alert odds compared with the baseline, but the two limited access functional classes are insignificant. Compared with rural nonramp locations, urban nonramps decrease alert odds but are insignificant. All ramp locations increase CSW alert odds and are significant (this is expected since most CSW alerts occurred on ramps). Daytime conditions slightly increase odds, while wet conditions slightly decrease odds. Driving with the alert system disabled decreases the odds of a CSW alert but is insignificant. The effects of maximum and minimum speed are similar to their effects in the first CSW logit model, but the OR for brake applications is slightly higher (but is still insignificant). Segment distance increases alert odds but is insignificant. Gender interactions with education, experience, and mileage driven in the previous year mostly have odds ratios (ORs) close to 1.0, meaning they do not have much of an effect on CSW alert odds. Gender–education interactions are significant, as are males interacted with last year's mileage (all other driver attributes are at least marginally significant). The pseudo R^2 is similar to that for the first CSW logit model, indicating decent model fit.

CHAPTER 4

Conclusions, Implications for SHRP 2 Safety Program, and Suggested Research

This report contains many models (although only a portion of those estimated) and many findings. To provide structure to these findings and their implications for the SHRP 2 Safety program, the chapter is organized according to the five original research questions. The chapter concludes with suggestions for future research.

Research Question 1

What is the relationship between events (e.g., crashes, near crashes, and incidents) and pre-event maneuvers? What are the contributing driver factors, environmental factors, and other factors?

This broad question encompassed many different models yielding a variety of findings. The general structure of the event-based models was to use predictor variables representing driver, context (i.e., roadway and environment), and event attributes. A set of tests was conducted to specifically explore changes in parameter estimates if variables from only one or two of these components were included in the model. Specifically, models were estimated with context-only, driver-only, and event-only variables (and combinations of only two of the components). Resulting parameter estimates changed substantially depending on how many of the three components were represented in the model; importantly, the exclusion of any of the components led to major changes in estimated parameters (see Chapter 3).

Implications for SHRP 2 Safety Program: Failure to test the inclusion of context-, driver-, and event-based variables runs the risk of producing a model with biased parameters. Although the data were limited and this is only one realization of an experiment, the results of this test showed substantial parameter changes in the tests of parameter inclusion. Both the magnitudes and SEs of the parameters changed substantially. This is clear evidence that the exclusion of any of the set of variables (i.e., driver, context, and event) is very likely to result in biased parameter estimates, obscuring the effect of any one variable on event occurrence. Based on this result, future analyses of SHRP 2 event-based data (such as in proposed research for the S08 project) should be required to include variables representing driver, context, and event attributes. In addition, thorough tests should be conducted to explore changes in parameter values and significance. The Penn State team is concerned that parameter estimates may exhibit the same characteristics, even in data sets with large sample sizes.

One is left to ponder the reasons for this apparent interaction. One possibility relates to the nature of the variables used to predict the outcome. Among the strongest variables (i.e., those showing the greatest association with crashes or near crashes) were the driver distraction variables. These variables, which were derived from the driver face camera, included distractions such as those attributed to a portable electronic device, internal distractions (such as a pet or other creature within the vehicle), or vehicle-related distractions such as adjusting the climate or audio controls. Some may view these driver actions as endogenous to the event process (i.e., the conditions that led to the event also led to the distraction). While this may not be true in all cases, it is likely true for some. While distraction was used as a predictor variable, the team now understands, after further deliberation, that some distractions may be endogenous and may not be suitable as event predictors. A range of statistical methods to address endogeneity should be considered in these circumstances. In addition, exploring measurement periods beyond the 5-s-before-event criterion used in the VTTI database may be necessary.

So, while the modeling seems feasible, a caution is in order: carefully consider event model specification. Special care should be exercised and perhaps specific models formulated to explore the nature of the endogeneity between distractions and other event-related measures. Although distractions have been used in the modeling (and by others) as predictor variables, the model tests indicate their use may not be valid.

An additional issue of interest is to reach conclusions, however tentative, concerning the efficacy of using categorical-outcome models (such as logit or binary hierarchical models) to compare crash and noncrash events. The Penn State team explored this issue within the limits of the data by comparing crash and near-crash events (combined) with critical incidents. The series of models estimated by the team yielded generally consistent results concerning the effects of particular parameters when using a complete model specification as described above.

Implications for SHRP 2 Safety Program: Given a set of data that is event-based, such as the VTTI data file, it is feasible to apply well-established categorical data analysis techniques to estimate factors that differentiate between the categorical outcomes. In this case, the team differentiated between crashes and near crashes (combined) and critical incidents. This implies that such a differentiation appears feasible for crashes (or other adverse events) and a sample of comparable, similarly described nonevents. Such a comparison was expected to occur in this research, but the data for nonevents in the VTTI file did not contain predictor variables consistent with the events; as a result, the VTTI data did not permit such analysis.

Several strong gender-related differences in factors contribute to crash or near-crash and critical incident occurrence. Gender was important in both driver- and event-based models, hardly a surprise given the extensive literature on gender-related safety differences. Many gender-related factors were revealed as main effects, but they were particularly apparent as interaction terms, especially in driver-based models.

Implications for SHRP 2 Safety Program: Analyses that are directly or indirectly influenced by gender should include tests of a range of main effects and interaction terms. Variables with significant promise in future modeling include level of education and years of driving experience. There were associations between number of previous crashes and traffic violations that varied with gender; these associations were not consistent, but they may warrant attention from researchers on gender issues.

A limited number of vehicle factors rose to significance. There was some indication that older vehicles driven by women were associated with a reduction in event frequency. The team did not interpret this as a direct safety effect. Despite many attempts to replace this variable with others of greater intuitive appeal, this result persisted.

Implications for SHRP 2 Safety Program: Clearly, additional research is needed concerning the analysis of vehicle factors, particularly considering the gender of the driver. The small sample size limited the ability to make inferences concerning vehicle type.

Research Question 2

What hierarchical structure (statistically speaking), if any, exists in the manner in which these relationships need to be explored?

Figure 4.1 shows one hierarchy successfully applied to the analysis of event data. The sketch is intended to convey that individual drivers may have any number of events; they must have at least one, but they may have more. If one were to model this with a count regression approach, each event would enter the model as if it were independent and from a different driver.

Using a hierarchical approach, driver attributes enter at the driver level, once for each driver. Event characteristics are entered as predictors for each event in which they occur. This hierarchical approach (described in Chapter 2, with findings in Chapter 3) provides a conceptually justifiable approach to the modeling of complex events.

Implications for SHRP 2 Safety Program: There are many hierarchical approaches that may be taken with a data set such as those presented in naturalistic driving studies. Much attention has been focused on the analysis of events; the driver-based approach presents one way to analyze drivers at a separate level from the events of interest, providing a much better depiction of the physical process being investigated.

A second hierarchical model was used in the driver-based analysis of the VTTI data. That data structure is shown in Figure 4.2.

In this structure, males and females are accounted for separately, including separate parameter estimates for each gender category. In a single-level model, there would typically be a dummy variable representing the difference between males and females, but not an indication of the actual parameter value for each gender specifically. The hierarchical approach provides this additional information; a model of this type is developed in Chapter 2 and described as applied to VTTI data in Chapter 3.

Implications for SHRP 2 Safety Program: This presents another example of how hierarchical approaches can be applied

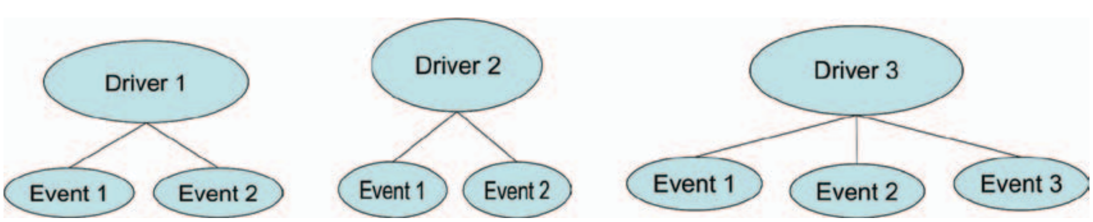

Figure 4.1. Hierarchy analysis of event data.

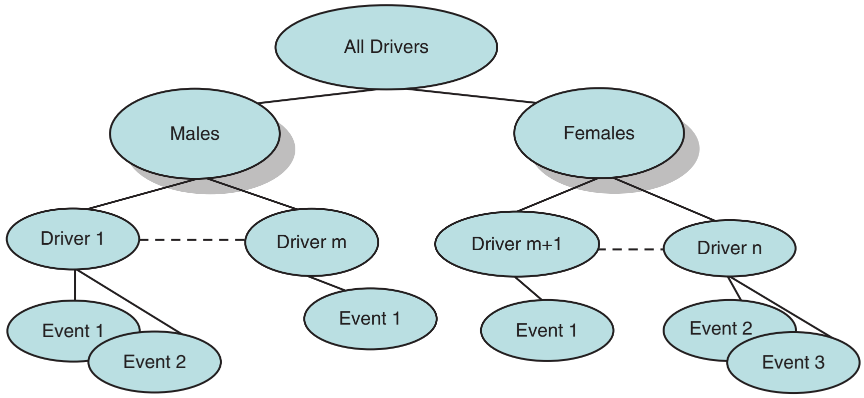

Figure 4.2. Driver-based hierarchical model.

to naturalistic data. The benefits of obtaining gender-specific estimates of factors contributing to the risk of events are clear.

Research Question 3

What kind of elucidative evidence emerges from the analysis of roadway departure crashes in terms of Questions 1 and 2? Is the illustrative hierarchy of relationships generalizable to other nonintersection crash types such as leading vehicle crashes?

In its proposal to SHRP 2, the Penn State team described the desirability of comparing hierarchical modeling structures and models for road departure and lead vehicle collisions. This could not be done because of a lack of available lead vehicle event data in the supplied VTTI database.

The Penn State team considers the notion of elucidative evidence to include surrogate measures and their testing, as well as exposure-based models developed during the study. There is also evidence that several types of predictor variables, such as precipitating event information in the VTTI models, have particularly important roles in the models.

One useful definition of crash surrogates was articulated by Hauer and Gårder (1986) in their focused discussion of the traffic conflicts technique as a surrogate measure: "one should be able to make inferences about the safety of an entity on the basis of a short duration 'conflict count' instead of having to wait a long time for a large number of accidents to materialize." Additional attributes of surrogates as having a time dimension and being responsive to countermeasures (as a crash would be) have been proposed in the Phase 1 report and by Shankar and colleagues (2008). More generally, surrogates can be considered as measures that can be substituted for crashes in a safety analysis: in the data for this project, they are typically vehicle kinematic—and event-related measures that offer some description of vehicle movement and/or position relative to the roadway.

The first example of surrogate testing is contained in the event-based analyses conducted with the VTTI data. Potential surrogates include the precipitating events of subject over lane or road edge and lost control. These two variables were derived by the VTTI data coders as part of the original 100-car data set. In most event models they were strong indicators of crash or near-crash events in the categorical models; in hierarchical models subject over lane or road edge was the second strongest predictor associated with the prediction of a crash or near-crash event. While there is strong association with crash events, this measure does not have a time dimension, so it does not directly meet the desirable criterion suggested in the Phase 1 report and by Shankar and colleagues (2008). Further, Hauer and Gårder's rule could not be applied because the team did not have access to the comparable set of subject behavior for noncrashes. Were such data available, the hierarchical model could be formulated to test the association between this measure and crashes. One is also left to ponder exactly how this measure could be broadly applied outside of SHRP 2's instrumented vehicles, but it is clear that this measure has some potential as a surrogate.

Implications for SHRP 2 Safety Program: The categorical models explored in this study appear to be a useful paradigm to explore surrogates when provided with event-based data. While kinematic measures or combinations of kinematic and roadway position measures were not directly tested with VTTI data, the Penn State team believes they are possible measures for future testing. The subject over lane or road edge variable contained position-only information and was very strongly associated with crash-related events; the team believes that including longitudinal or lateral velocity and lateral position information would enhance this variable's predictive ability.

A limitation of the categorical models deserves mention. Initial event-based models, both bivariate logistic and hierarchical, used improper speed as an event-based predictor. Successful model fit was obtained, but improvement was sought. Driver Impairment 1 (drowsy, sleepy, fatigued) was substituted as a predictor and a much better fit occurred overall, including reduced SEs for several variables. While the team was pleased by the improved fit, there was concern about the apparent model instability. This apparent instability may be the result of the small sample size, but it may also reflect endogeneity among the predictors. As a recommendation to future SHRP 2 analysis contractors, the team suggests that care be exercised in surrogate analyses; additional empirical testing in several other sites and with other drivers should reveal more about this issue.

A method for validating events containing possible surrogates for crashes is proposed and discussed in Chapter 3. The statistical predictions from the event-based model were compared to text descriptions of the event etiology derived from video and kinematic data (using the original VTTI data coding); the comparison showed that events originally coded as critical incidents were statistically estimated to be crashes. It was posited that these events could be used to supplement crash data observed directly. The manipulation of the event-based models is proposed to provide useful information about whether a particular critical incident or near-crash event really was similar, statistically, to a crash event in a similar context. Such a comparison is, of course, dependent on the model being correct.

Implications for SHRP 2 Safety Program: The team offers this method as a way for future SHRP 2 researchers to supplement their crash data. The method was developed along the way to working with event-based data. It may be used by others as needed.

The UMTRI analyses tested several kinematic measures, particularly longitudinal velocity entering curves, as a potential surrogate of event risk, in this case a CSW alert (instead of a crash event). Initial tests of piecewise linear models applied to the data as a whole showed that the measure has some merit, but the models were weakened statistically by the presence of serial correlation in the observations (data were collected at 10 Hz). The team next explored tracking individual drivers through the same location multiple times to see if repeated behaviors or learning occurred and to explore individual variability. The models showed different results from the aggregate. While the results were not stunning, they showed potential and are recommended over aggregate approaches.

Part of the analysis of individual drivers tied the kinematic measures to specific road segments using Google maps. The next set of analysis contracts should have detailed roadway data available for at least a few of the study sites through SHRP 2 S04 contracts. The ability to explore context for this modeling should greatly enhance the findings. The cohort-based modeling also shows promise in quantifying context effects; this method is described further in the discussion of Research Question 5.

Implications for SHRP 2 Safety Program: After estimating a great many models with aggregate data, the team believes an approach that tracks individual drivers repeatedly over the same route has potential for additional insight. Again, detailed and accurate roadway data from project S04 will be essential to these tests. Specifically, a range of kinematic variables can be measured at specific points of documented high crash frequency; these can be compared with a set of individual drivers' kinematic signatures through the same roads. Kinematic measures at crash locations can be compared with similar measures at low-frequency crash locations and tested for their predictive capability.

Driver-based models used self-reported annual mileage (provided by drivers during 100-car study interviews) as exposure. While measured miles and time of travel would have been preferred, the team felt that self-reported exposure would be a reasonable start. Measured travel from the UMTRI study was used for exposure in the cohort-based analyses described in the discussion of Research Question 5. The driver-based models using VTTI data showed that exposure (in this case, miles driven per year) is essential to the study of the expected number of events per year for drivers. There was a strong association with the expected number of events, and the inclusion of the variable greatly improved model fit.

Implications for SHRP 2 Safety Program: Not surprisingly, the driver-based analyses indicated that exposure was essential to the modeling of the expected number of events. It is clear that travel by individual drivers should be identified to the extent possible through the face camera or other technologies. Researchers interested in identifying high-risk drivers should explore the hierarchical models formulated in Chapter 2 and empirically tested in Chapter 3. The team developed a model that clearly identified drivers who were outliers with respect to the number of events they experienced. Drivers with exceptionally high, as well as low, numbers of events were identified. This method can be used in the identification of outlier drivers in subsequent SHRP 2 analysis activities.

In consideration of the previous comments about driver distraction and endogeneity, it is of interest to briefly discuss the findings of the analysis with respect to this variable. All distractions are not alike in their effect on event occurrence: virtually all of the event-based models showed substantial differences in the effect of distractions on event occurrence. Most generally, internal distractions (e.g., reading, moving an object in the vehicle, dealing with a pet or insect) were most strongly associated with crash or near-crash event occurrence; passenger-related distractions and occurrences of the driver talking, singing, or daydreaming also had consistent positive correlation. Interestingly, the use of a wireless device was poorly correlated to event occurrence. These findings must be considered in the context that the VTTI data were collected before 2006, when cell phone usage was at a lower level than now, and devices generally had fewer features than in 2010.

Implications for SHRP 2 Safety Program: These findings, taken as a whole, reveal that distractions are an important factor to measure in future SHRP 2 analysis efforts. If a portion of SHRP 2 funds are to be used to preprocess S07 project data to produce event files, then distractions would seem to be a high-priority measure to obtain for each event. The event data would be even more useful if matched with nonevent data collected from all drivers that include comparable distraction measures.

Research Question 4

In terms of elucidative evidence, what types of behavioral correlates emerge? For example, are attitudinal measurements indicative of revealed behavior in terms of headway maintenance and speed reductions?

The principal measure of behavioral correlation was the DDDI collected by VTTI during the original 100-car data collection effort (Dula and Ballard 2003). The DDDI consists of 28 statements to which the driver is asked to respond on a 5-point Likert scale (never, rarely, sometimes, often, and always). Each of the categories of response is assigned an integer from 1 to 5. Example test statements include the following: "I verbally insult drivers who annoy me"; "Passengers in my car/truck tell me to calm down"; and "I will weave in and out of slower traffic." Based on previous research, the responses to the questions are divided into three categories of driving: aggressive driving (AD), negative emotional (NE) driving, and risky driving (RD). Each category is intended to capture a different aspect or component of dangerous driving.

The DDDI was generally associated with an increase in crashes or near crashes in the event-based models and was also positively associated with the number of events in the driver-based models. The results were not always easy to interpret or consistent with intuition. In driver-based models, the AD component was associated with an increase in the number of events, but for females only. In the event-based models, this same component was associated with a reduction in crash or near-crash events (which could be interpreted as an increase in the likelihood of critical incidents). So, while there were associations in the data, and they were generally consistent and statistically significant (within the limits of the data), there is a concern that the findings were not as interpretable as would be desired. The DDDI developers cite validation studies conducted on a simulator (Dula and Ballard 2003), but no additional references to the use of this index were found during a Web search.

Implications for SHRP 2 Safety Program: The Penn State team believes that the testing conducted with the DDDI confirms the importance of including some measure of driver risk propensity in the remaining SHRP 2 data analyses. The current plans for the data collection projects (S07) call for the use of other metrics for estimating risk taking. The team expects that these metrics will be shown to be important in the subsequent modeling and hopes that the results are more consistent than those obtained in the present analyses.

In addition to the DDDI, the Life Stress Index was administered to participating primary drivers in the 100-car study. This tool attempts to measure the amount of stress present in one's life as a whole by using factors such as stress at work, difficulty with personal relationships, and challenges in the family environment. The Life Stress Index was positively associated with crashes and near crashes in some event-based models, but it was not a predictor in the driver-based models. Although the Life Stress Index is another important metric to have, it is not as important as a driving-focused metric such as the DDDI.

Implications for SHRP 2 Safety Program: It would be of interest to obtain a metric for life stress, but this is not as important as driver-based risk-taking measurement. The proposed testing for the S07 projects includes a number of perceptual and cognitive tests; psychological tests include metrics for risk taking, risk perception, driver style and behavior, and thrill and adventure seeking. This should provide more than ample measures of driver predisposition for events.

Research Question 5

If elucidative evidence does in fact emerge in terms of attitudinal correlates and how their interactions vary by context, is it plausible to parse out the marginal effects of various context variables on crash risk by suitable research design?

The response to this research question has two parts. First, the team discussed the effects of the various components of context, specifically roadway-related factors, time of day, and traffic levels, on the probability of crash and near-crash occurrence. These inferences are drawn from the event-based models with the 100-car data. Second, the team considered the cohort-based analyses conducted with the UMTRI data and expanded on their possible role in SHRP 2 projects, particularly S08.

Context was an extremely important factor to consider; several aspects of context were revealed to be associated with crash or near-crash outcomes. Roadway-related factors were important descriptors of context in the series of event-based models. The presence of curves was a significant factor in differentiating critical incidents from crashes and near crashes. While there was some inconsistency in the magnitude of the effect, horizontal curves, in general, indicated a modest increase in risk. Horizontal curve presence does not show the magnitude of influence of driver behavior variables such as distractions, but it is clearly important in defining context.

Implications for SHRP 2 Safety Program: Context, as related to roadway and roadside geometry and features, is planned to be collected through SHRP 2 Safety Project S04. The analyses indicate this is an extremely important activity. The event modeling described in Chapter 3 reveals that failure to include context-related variables will yield a model with substantially biased parameter estimates. Accurately assessing the influence of factors such as distractions and predisposition is impossible without the inclusion of context. In the 100-car study, many of the context variables were obtained from video of the event. In the remaining SHRP 2 projects it seems to be envisioned that much of these data will be obtained from the enhanced GIS data collected as part of the S04 activity. A cost savings will certainly be realized if context data can be gathered in this way, but it is likely that a degree of checking will be necessary to verify roadway and roadside features obtained from the GIS with camera data from the vehicles.

Time of day, specifically dawn or dusk, was a substantial factor increasing risk and again contributed importantly to the definition of context in which crash or near-crash events occur. This variable was consistently significant and positive in all event-based models and had ORs that exceeded some driver distraction and precipitating event factors. These findings are consistent with sleep- and fatigue-related studies of crash risk for private drivers and the motor carrier industry.

Implications for SHRP 2 Safety Program: Future research projects conducted as part of SHRP 2 Safety Project S08 need to seriously consider the identification of dawn and dusk driving. This is an important element of context. As before, a comparison of crashes, near crashes, and critical incidents to a sample of nonevents with comparable attributes would serve to validate these findings.

Run-off-road crashes were consistently and negatively associated with increased traffic levels; this seems like a sensible association, as drivers are more likely to have other types of crashes, near crashes, and critical incidents under more congested traffic conditions. The association was not as strong as other variables previously discussed.

Implications for SHRP 2 Safety Program: It would be advantageous if some measure of traffic level could be collected or available for the S08 analysis projects; measurement other than through vehicle cameras would also be beneficial. The traffic data are important but, in the team's judgment, not as important as the measurement of roadway and roadside features and time of day (dawn or dusk).

In cohort-based models formulated with the UMTRI data, context was generally more strongly associated with event outcome (i.e., CSW alerts) than driver-based variables. This general finding supports the emphasis on context that has stimulated much discussion during research symposia. Nevertheless, the Penn State team wishes to caution that there is no implication that driver actions are unimportant. In fact, the team views context and driver attributes as mutually complementary and closely linked.

The team would like to emphasize that the cohort-based approach allows, for the first time, it is believed, the ability to use naturalistic driving data to examine driver and context factors in a consistent exposure framework that includes both. This ability is only possible with the detailed data available from a naturalistic driving database in which an individual driver is monitored through a series of contexts (such as in the UMTRI RDCW data set).

Implications for SHRP 2 Safety Program: The team believes that cohort analysis represents a breakthrough in analysis paradigms for naturalistic data. The driver is tracked through a roadway network defined as homogeneous based on the needs of the analysis team. Once segments are defined, events (using appropriate screening criteria) can be allocated to the segments. The user or analyst can make the segment designation as fine or coarse as roadway and roadside data allow. This framework provides the measurement of the driver's actions and behavior throughout the driver's travel, not just in the seconds immediately preceding or following a crash. Nevertheless, there is likely to always be a demand to study the details of the crash process in the few seconds before and after a crash or other event. The cohort approach provides a structure for the analyst to flexibly define how the behavior of the driver can be studied.

The team used a range of statistical methods to provide examples of how the cohort-based data structure can be used. These are intended to assist future SHRP 2 safety studies by providing guidance about data manipulation and variable formulation.

Suggested Research

The analyses completed to date offer a number of lessons learned concerning methodological issues in the analysis of naturalistic driving data. Among the more important are

1. Even with the much larger data set available in the SHRP 2 S07 project, there is a need to be rigorous in the application of Poisson, NB, ZIP, and other count regression techniques. As documented in this report, the estimation of literally hundreds of models will be necessary to obtain consistent estimates of model parameters. Models containing main effects may not be sufficient. It was not until interaction-based models were tested that the count-based models started to yield consistent parameter estimates and improved goodness of fit. Model estimation following this suggestion should be considered good practice.
2. The overdispersion parameter (i.e., α) in the NB model is an important indicator of heterogeneity and needs to be thoroughly studied. Including predictors for the parameter yielded much improved fit with this data set.
3. Context is extremely important. The elasticity for time of day (dawn or dusk) in the event-based models was as large as the most important driver distraction variables. The presence of horizontal curves was also marginally significant. These findings reinforce the importance of the

roadway-related data collection activity within SHRP 2 safety projects and the need to thoroughly consider how these data can be included in more precise and accurate event-prediction models. Clearly it will be important to be able to identify noncrash events with common attributes in the data to better estimate crash risk in the larger data set.

4. There is a need to continue to test models that integrate kinematic data with broader data characterizing the event, driver attributes, and context. The event-based models contained in Chapter 3 only scratch the surface of these formulations. There is a particular need to include vehicle kinematics to more closely tie vehicle location (e.g., within the lane) and movement (e.g., longitudinal speed) within event-based model frameworks. Hierarchical models offer particular advantages given their flexibility and relaxation of assumptions concerning variable probability distributions. The cohort formulations discussed in Chapter 3 seem to offer particular promise with respect to kinematic variable inclusion.
5. The technology-intense group of vehicles in the SHRP 2 field study requires the careful consideration of driver adaptation to technology reflected in changes in driver behavior in vehicles with and without warning systems. Adaptation has been a topic of many research papers; the evidence continues to build of its importance in any effectiveness analysis. Extensions of the cohort-based formulations offer promise in exploring linkages between driver, context, and kinematic variables.
6. Future naturalistic driving studies could define homogeneous trip segments to facilitate inclusion of kinematic variables. As the size of the road segment shrinks, there will be an improved ability to capture potentially significant vehicle kinematics linked to geometric features. This suggestion is closely linked to the more general question of improved precision and resolution for road-segment event modeling, and it is the substantial promise of linking S07 and S04 databases.

The Penn State team has explored the provided data sets with methods that the authors believe fit the need. The team hopes it has contributed to an improved understanding of promising methods to analyze naturalistic driving data.

References

Abdel-Aty, M. A., R. Kitamura, and P. P. Jovanis. 1995a. Investigating Effect of Travel Time Variability on Route Choice Using Repeated-Measurement Stated Preference Data. In *Transportation Research Record 1493,* TRB, National Research Council, Washington, D.C., pp. 39–45.

Abdel-Aty, M. A., R. Kitamura, and P. P. Jovanis. 1995b. Exploring Route Choice Behavior Using Geographic Information System–Based Alternative Routes and Hypothetical Travel Time Information Input. In *Transportation Research Record 1493,* TRB, National Research Council, Washington, D.C., pp. 74–80. http://pubsindex.trb.org/view/452596.

Abdel-Aty, M. A., R. Kitamura, P. P. Jovanis, P. Reddy, and K. M. Vaughn. 1995. New Approach to Route Choice Data Collection: Multiphase, Computer-Aided Telephone Interview Panel Surveys Using Geographical Information Systems Data Base. In *Transportation Research Record 1493,* TRB, National Research Council, Washington, D.C., pp. 159–169. http://pubsindex.trb.org/view/452606.

Aguero-Valverde, J., and P. P. Jovanis. 2006. Spatial Analysis of Fatal and Injury Crashes in Pennsylvania. *Accident Analysis and Prevention,* Vol. 38, No. 3, pp. 618–625.

Aguero-Valverde, J., and P. P. Jovanis. 2007. Identifying Road Segments with High Risk of Weather-Related Crashes Using Full Bayesian Hierarchical Models. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C. http://pubsindex.trb.org/view.aspx?id=801198.

Aguero-Valverde, J., and P. P. Jovanis. 2008. Analysis of Road Crash Frequency with Spatial Models. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2061,* Transportation Research Board of the National Academies, Washington, D.C., pp. 55–63. http://trb.metapress.com/content/fu64443w36487316/.

Aguero-Valverde, J., and P. P. Jovanis. 2010. Bayesian Multivariate Poisson Lognormal Models for Crash Severity Modeling and Site Ranking. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2136,* Transportation Research Board of the National Academies, Washington, D.C., pp. 82–91. http://trb.metapress.com/content/y458112n72mj133w/.

Ayres, G., B. Wilson, and J. LeBlanc. 2004. Method for Identifying Vehicle Movements for Analysis of Field Operational Test Data. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1886,* Transportation Research Board of the National Academies, Washington, D.C., pp. 92–100. http://trb.metapress.com/content/x715q2n5g1n27068/.

Bogard, S., P. Fancher, R. Ervin, M. Hagan, and Z. Bareket. 1998. *Intelligent Cruise Control Field Operational Test.* Final report, Vol. 3. National Highway Traffic Safety Administration, U.S. Department of Transportation. http://ntl.bts.gov/lib/jpodocs/edldocs1/9005/Vol3.pdf.

Bonneson, J. A. 2000. *NCHRP Report 439: Superelevation Distribution Methods and Transition Designs.* Transportation Research Board of the National Academies, Washington, D.C.

Chang, H. L., and P. P. Jovanis. 1990. Formulating Accident Occurrence As a Survival Process. *Accident Analysis and Prevention,* Vol. 22, No. 5, pp. 407–419.

Chayanan, S., M. Nebergall, V. Shankar, N. Juvva, and Y. Ouyang. 2003. *Interaction Between the Roadway and Roadside: An Econometric Analysis of Design and Environmental Factors Affecting Segment Accident Rates.* Research Report WA-RD 562.1. Washington State Department of Transportation, Olympia.

Cleves, M., and I. Cañette. 2009. *Analysis of Multiple Failure-Time Survival Data.* Stata Corp., College Station, Tex.

Congdon, P. 2001. *Bayesian Statistical Modelling.* John Wiley & Sons, Chichester, United Kingdom.

DaSilva, M., and W. G. Najm. 2006. Development of Collision Avoidance Data for Light Vehicles: Near-Crash/Crash Event Data Recorders. Report DOT-VNTSC-NHTSA-06-03. John A. Volpe National Transportation Systems Center, Cambridge, Mass. http://ntl.bts.gov/lib/34000/34700/34724/DOT-VNTSC-NHTSA-06-03.pdf.

Design Quality Assurance Bureau and New York Department of Transportation. 2003. *Recommendations for AASHTO Superelevation Design.* http://downloads.transportation.org/SuperelevationDiscussion9-03.pdf.

Dingus, T. A. 1995. Moving from Measures of Performance (MOPs) to Measures of Effectiveness (MOEs). *Safety Evaluation of Intelligent Transportation Systems Workshop Proceedings,* ITS America and National Highway Traffic Safety Administration, Reston, Va.

Dingus, T. A., S. Hetrick, and M. Mollenhauer. 1999. Empirical Methods in Support of Crash Avoidance Model Building and Benefits Estimation. *ITS Journal,* Vol. 5, No. 2, pp. 93–125.

Dingus, T. A., S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R. R. Knipling. 2006. *The 100-Car Naturalistic Driving Study, Phase II: Results of the 100-Car Field Experiment.* Report DOT HS 810 593. National Highway Traffic Safety Administration, U.S. Department of Transportation. www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/Driver%20Distraction/100CarMain.pdf.

Dingus, T. A., V. L. Neale, S. G. Klauer, A. D. Petersen, and R. J. Carroll. 2006. The Development of a Naturalistic Data Collection System to

Perform Critical Incident Analysis: An Investigation of Safety and Fatigue Issues in Long-Haul Trucking. *Accident Analysis and Prevention,* Vol. 38, No. 6, pp. 1127–1136.

Dula, C. S., and M. E. Ballard. 2003. Development and Evaluation of a Measure of Dangerous, Aggressive, Negative Emotional, and Risky Driving. *Journal of Applied Social Psychology,* Vol. 33, No. 2, pp. 263–282.

Ervin, R., J. Sayer, D. LeBlanc, S. Bogard, M. Mefford, M. Hagan, Z. Bareket, and C. Winkler. 2005. Automotive Collision Avoidance System Field Operational Test Report: Methodology and Results. Report DOT HS-809-900. National Highway Traffic Safety Administration, U.S. Department of Transportation. www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2005/809900All.pdf.

Fitzpatrick, K., M. Wooldridge, O. Tsimoni, J. Collins, P. Green, K. Bauer, K. Parma, R. Koppa, D. Harwood, I. Anderson, R. Krammes, and B. Poggioli. 2000. *Alternative Design Consistency Rating Methods for Two-Lane Rural Highways.* Federal Highway Administration Report FHWA-RD-99-172, Springfield, VA.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2003. *Bayesian Data Analysis,* 2nd ed. Chapman and Hall/CRC, Boca Raton, Fla.

Glennon, J. C. 2006. Calculating Critical Speed: A Motor-Vehicle Crash Reconstruction Method Fraught with Error. Motor Vehicle Accident and Reconstruction Services. CrashForensics.com. www.johncglennon.com/papers.cfm?PaperID=42.

Goldstein, H. 1995. *Multilevel Statistical Models,* 2nd ed. Edward Arnold, London.

Greenberg, R. S., S. R. Daniels, W. D. Flanders, J. W. Eley, and J. R. Boring. 1993. *Medical Epidemiology.* Norwalk, CT: Appleton & Lange.

Hanowski, R. J., J. Hickman, M. C. Fumero, R. L. Olson, and T. A. Dingus. 2007. The Sleep of Commercial Vehicle Drivers Under the 2003 Hours-of-Service Regulations. *Accident Analysis and Prevention,* Vol. 39, No. 6, pp. 1140–1145.

Hanowski, R. J., J. S. Hickman, W. Wierwille, and A. Keisler. 2007. A Descriptive Analysis of Light Vehicle–Heavy Vehicle Interactions Using in Situ Driving Data. *Accident Analysis and Prevention,* Vol. 39, No. 1, pp. 169–179.

Hanowski, R. J., M. A. Perez, and T. A. Dingus. 2005. Driver Distraction in Long-Haul Truck Drivers. *Transportation Research Part F: Traffic Psychology and Behaviour,* Vol. 8, No. 6, pp. 441–458.

Hauer, E., and P. Gårder. 1986. Research into the Validity of the Traffic Conflicts Technique. *Accident Analysis and Prevention,* Vol. 18, No. 6, pp. 471–481.

Heinrich, H. W., D. Petersen, and N. Roos. 1980. *Industrial Accident Prevention.* New York: McGraw-Hill Book Company.

Helai, H., C. H. Chor, and M. M. Haque. 2008. Severity of Driver Injury and Vehicle Damage in Traffic Crashes at Intersections: A Bayesian Hierarchical Analysis. *Accident Analysis and Prevention,* Vol. 40, No. 1, pp. 45–54.

Hilbe, J. M. 2007. *Negative Binomial Regression.* Cambridge University Press, Cambridge, United Kingdom.

Ja Lim, H., J. Liu, and M. Melzer-Lange. 2007. Comparison of Methods for Analyzing Recurrent Events Data: Application to the Emergency Department Visits of Pediatric Firearm Victims. *Accident Analysis and Prevention,* Vol. 39, No. 2, pp. 290–299.

Jones, A. P., and S. H. Jorgensen. 2003. The Use of Multilevel Models for the Prediction of Road Accident Outcomes. *Accident Analysis and Prevention,* Vol. 35, No. 1, pp. 59–69.

Jovanis, P. P., and H.-L. Chang. 1989. Disaggregate Model of Highway Accident Occurrence Using Survival Theory. *Accident Analysis and Prevention,* Vol. 21, No. 5, pp. 445–458.

Jovanis, P. P., and H. Chang. 1986. "Modeling the Relationship of Accidents to Miles Traveled." *Transportation Research Record* 1068, TRB, National Research Council, Washington, D.C., pp. 42–51.

Jovanis, P. P., and F. Gross. 2007. Estimation of Safety Effectiveness of Changes in Shoulder Width with Case Control and Cohort Methods. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2019,* Transportation Research Board of the National Academies, Washington, D.C., pp. 237–245. http://trb.metapress.com/content/aukqg56541163427/.

Kiefer, R. J., M. T. Cassar, C. A. Flannagan, D. J. LeBlanc, M. D. Palmer, R. K. Deering, and M. Shulman. 2003. *Forward Collision Warning Requirements Project Final Report Task 1: Refining the CAMP Crash Alert Timing Approach by Examining "Last-Second" Braking and Lane-Change Maneuvers Under Various Kinematic Conditions.* DOT HS 809 574. National Highway Transportation Safety Administration and Department of Transportation, Washington, D.C.

Kim, D.-G., Y. Lee, S. Washington, and K. Choi. 2007. Modeling Crash Outcome Probabilities at Rural Intersections: Application of Hierarchical Binomial Logistic Models. *Accident Analysis and Prevention,* Vol. 39, No. 1, pp. 125–134.

Klauer, S. G., T. A. Dingus, V. L. Neale, J. D. Sudweeks, and D. J. Ramsey. 2006. *The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data.* Report No. DOT-HS-810-594. National Highway Traffic Safety Administration, U.S. Department of Transportation. www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2006/DriverInattention.pdf.

Lambert, D. 1992. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics,* Vol. 34, No. 1, pp. 1–14.

LeBlanc, D., J. Sayer, C. Winkler, R. Ervin, S. Bogard, J. Devonshire, M. Mefford, M. Hagan, Z. Bareket, R. Goodsell, and T. Gordon. 2006. *Road Departure Crash Warning System Field Operational Test: Methodology and Results. Volume 1: Technical Report.* University of Michigan Transportation Research Institute, Ann Arbor.

Lee, S. E., E. C. B. Olsen, and W. W. Wierwille. 2004. *A Comprehensive Examination of Naturalistic Lane-Changes.* Report DOT HS 809 702. National Highway Traffic Safety Administration, U.S. Department of Transportation. www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2004/Lane%20Change%20Final.pdf.

Lenguerrand, E., J. L. Martin, and B. Laumon. 2006. Modelling the Hierarchical Structure of Road Crash Data: Application to Severity Analysis. *Accident Analysis and Prevention,* Vol. 38, No. 1, pp. 43–53.

Lord, D. 2006. Modeling Motor Vehicle Crashes Using Poisson-Gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter. *Accident Analysis and Prevention,* Vol. 38, No. 4, pp. 751–766.

Lord, D., S. Washington, and J. N. Ivan. 2007. Further Notes on the Application of Zero-Inflated Models in Highway Safety. *Accident Analysis and Prevention,* Vol. 39, No. 1, pp. 53–57.

Lord, D., S. P. Washington, and J. N. Ivan. 2005. Poisson, Poisson-Gamma and Zero-Inflated Regression Models of Motor Vehicle Crashes: Balancing Statistical Fit and Theory. *Accident Analysis and Prevention,* Vol. 37, No. 1, pp. 35–46.

McLaughlin, S. B., J. M. Hankey, and T. A. Dingus. 2008. A Method for Evaluating Collision Avoidance Systems Using Naturalistic Driving Data. *Accident Analysis and Prevention,* Vol. 40, No. 1, pp. 8–16.

Miaou, S. P. 1994. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention,* Vol. 26, No. 4, pp. 471–482.

Milton, J. C., and F. L. Mannering. 1996. *The Relationship Between Highway Geometrics, Traffic Related Elements, and Motor Vehicle*

Accidents. Final Research Report, WA-RD 403.1. Washington State Department of Transportation, Washington.

Mooney, C. Z., and R. D. Duval. 1993. *Bootstrapping: A Nonparametric Approach to Statistical Inference.* Sage University Paper Series on Quantitative Applications in the Social Sciences 07-095. Sage, Newbury Park, Calif.

Neale, V. L., T. A. Dingus, S. G. Klauer, J. Sudweeks, and M. Goodman. 2005. *An Overview of the 100-Car Naturalistic Study and Findings.* Report VTRC-05-0400. National Highway Traffic Safety Administration, U.S. Department of Transportation.

NHTSA. 2005. *Automotive Collision Avoidance System Field Operational Test (ACAS FOT).* Report DOT HS 809 886. National Highway Traffic Safety Administration, U.S. Department of Transportation, Washington, D.C. www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2005/ACAS%20FOT%20Final%20Program%20Report%20DOT%20HS%20809%20886.pdf.

Poch, M., and F. L. Mannering. 1996. Negative Binomial Analysis of Intersection-Accident Frequencies. *Journal of Transportation Engineering*, Vol. 122, No. 2, pp. 105–113.

Rasbash, J., H. Goldstein, and W. J. Browne. 2000. *The MLwiN Command Interface.* Institute of Education, London.

Rasbash, J., H. Goldstein, and W. J. Browne. 2002. *A User's Guide to MLwiN.* Institute of Education, London.

Sayer, J. R., J. M. Devonshire, and C. A. Flannagan. 2005. The Effects of Secondary Tasks on Naturalistic Driving Performance. Report UMTRI-2005-29. University of Michigan Transportation Research Institute, Ann Arbor.

Shankar, V., S. Chayanan, S. Sittikariya, and M. Shyu. 2004. "Marginal impacts of design, traffic, weather and related interactions on roadside crashes," *Transportation Research Record 1897,* pp. 156–163.

Shankar, V., P. P. Jovanis, J. Aguero-Valverde, and F. Gross. 2008. Analysis of Naturalistic Driving Data: Prospective View on Methodological Paradigms. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2061,* Transportation Research Board of the National Academies, Washington, D.C., pp. 1–8. http://trb.metapress.com/content/l0146u5267476673/.

Shankar, V., F. Mannering, and W. Barfield. 1995. Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies. *Accident Analysis and Prevention,* Vol. 27, No. 3, pp. 371–389.

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

Shankar, V. N., S. Chayanan, S. Sittikariya, M.-B. Shyu, N. K. Juvva, and J. C. Milton. 2004. Marginal Impacts of Design, Traffic, Weather, and Related Interactions on Roadside Crashes. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1897,* Transportation Research Board of the National Academies, Washington, D.C., pp. 156–163. http://trb.metapress.com/content/b6305j14lq348567/.

Shinar, D. 2007. *Traffic Safety and Human Behavior.* Elsevier, Oxford, United Kingdom.

Sittikariya, S., and M.-B. Shyu. 2006. Some Insights on Roadway Infrastructure Design for Safe Elderly Pedestrian Travel. *IATSS Research,* Vol. 30, No. 1, pp. 21–26.

Sittikariya, S., and V. Shankar. 2007. Bayesian Formulations of Heterogeneity in Traffic Accident Count Models. *Proc., 11th World Conference on Transport Research,* Berkeley, Calif.

Spacek, P. 2005. Track Behavior in Curve Areas: Attempt at Typology. *Journal of Transportation Engineering,* Vol. 131, No. 9, pp. 669–676.

Stutts, J., J. Feaganes, D. Reinfurt, E. Rodgman, C. Hamlett, K. Gish, and L. Staplin. 2005. Driver's Exposure to Distractions in Their Natural Driving Environment. *Accident Analysis and Prevention,* Vol. 37, No. 6, pp. 1093–1101.

Sullivan, J. M., O. Tsimhoni, and S. Bogard. 2007. Warning Reliability and Driver Performance in Naturalistic Driving. Report UMTRI-2007-44. University of Michigan Transportation Research Institute, Ann Arbor.

Thomas, A., B. O'Hara, U. Ligges, and S. Sturtz. 2006. Making BUGS Open. *R News,* Vol. 6, No. 1, pp. 12–17.

Tijerina, L., M. Gleckler, D. Stoltzfus, S. Johnston, M. J. Goodman, and W. W. Wierwille. 1999. *A Preliminary Assessment of Algorithms for Drowsy and Inattentive Driver Detection on the Road.* Report DOT HS 808. National Highway Traffic Safety Administration, U.S. Department of Transportation. www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/VRTC/ca/capubs/drowsydr.pdf.

Toledo, T., and D. Zohar. 2007. Modeling the Duration of Lane Changes. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1999,* Transportation Research Board of the National Academies, Washington, D.C., pp. 71–78. http://trb.metapress.com/content/671172564248ll40/.

Vuong, Q. 1989. Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. *Econometrica*, Vol. 57, pp. 307–334.

Washington, S. P., M. G. Karlaftis, and F. L. Mannering. 2003. *Statistical and Econometric Methods for Transportation Analysis.* Chapman and Hall/CRC, Boca Raton, Fla.

Wilson, B. H., M. D. Stearns, J. Koopmann, and C. Y. D. Yang. 2007. *Evaluation of a Road-Departure Crash Warning System.* Report DOT HS 810 854. National Highway Traffic Safety Administration, U.S. Department of Transportation. www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2007/4638-810_854%20RDCW%20EvalCLTest.pdf.

Wolfinger, R., and M. O'Connell. 1993. Generalized Linear Mixed Models: A Pseudo-Likelihood Approach. *Journal of Statistical Computation and Simulation,* Vol. 48, No. 3, pp. 233–243.

Woodward, M. 2005. *Epidemiology: Study Design and Data Analysis,* 2nd ed. Chapman and Hall/CRC, Boca Raton, Fla.

Yannis, G., E. Papadimitriou, C. Antoniou. 2008. Impact of enforcement on traffic accidents and fatalities: a multivariate multilevel analysis. *Safety Science,* Vol. 46, pp. 738–750.

Yannis, G., E. Papadimitriou, and C. Antoniou. 2007. Multilevel Modelling for the Regional Effect of Enforcement on Road Accidents. *Accident Analysis and Prevention,* Vol. 39, No. 4, pp. 818–825.

Related SHRP 2 Research

Integration of Analysis Methods and Development of Analysis Plan (S02)
Roadway Measurement System Evaluation (S03)
Roadway Information Database Developer, Technical Coordination, and Quality Assurance for Mobile Data Collection (S04A)
Mobile Data Collection (S04B)
Design of the In-Vehicle Driving Behavior and Crash Risk Study (S05)
Technical Coordination and Quality Control (S06)
In-Vehicle Driving Behavior Field Study (S07)
Analysis of the SHRP 2 Naturalistic Driving Study Data (S08)