Use of SHRP2 NDS Data to Evaluate Roadway Departure Characteristics

Phase I Final Report and Phase II Proposal

SHRP 2 Implementation Assistance Program

Submitted by: Shauna Hallmark (*shallmar@iastate.edu*) and Omar Smadi (*smadi@iastate.edu*),

Center for Transportation Research and Education at Iowa State University's Institute for Transportation, 2711 South Loop Drive, Suite 4700, Ames, Iowa 50010 (515) 294-8103

September 30, 2015

Table of Contents

1. Introduction and Need for Research 1
1.1 Background 1
1.2 Project Objectives 1
3. Data Sources and Data Reduction 1
3.1 Data Sources 1
3.2 Data Reduction 2
3.3 Overview of Roadway Crash/Near-Crash Events 4
4. Method of Analysis and Research Results 5
4.1 Analysis of Roadway Departure Crash Risk 5
4.2 Evaluating Relationship between Speed Choice and Roadway/Driver Characteristics 6
4.3 Evaluating Relationship between Lateral Position and Roadway/Driver Characteristics ... 7
5. Future Direction and Phase 2 Proposal 10
5.1 Objectives 10
5.2 Benefits of the Proposed Research 11
Task 1: Update IRB 12
Task 2: Data Requests and Data Reduction 12
Task 3: Correlate Roadway Departure Risk to Driver and Roadway Factors 15
Task 4: Evaluation of Driver Speed Choice based on Roadway Characteristics 16
Task 5: Evaluate Driver Response Based on Lateral Position 18
Task 6: Final Report and Outreach 19
Appendix a .... a
A.1 Schedule .... a
A.2. Budget .... a
A.3 References .... b

1. INTRODUCTION AND NEED FOR RESEARCH

1.1 Background

The Federal Highway Administration (FHWA) (1) estimates that 51% of roadway fatalities are roadway departures, while 40% of fatalities are single-vehicle run-off-road (SVROR) crashes. Roadway departures make up 56% of roadway fatalities and 40% are speeding related as noted in a survey of FARS data (2010—2012). The majority of fatal roadway departure crashes (79.5%) occur in rural settings with 38.7% of fatal crashes on rural undivided two lane roads. Curves account for 37.8% of fatal crashes, 50.5% are at nighttime, 75.3% are male, 11.5% are 15 to 20 year olds, and 16.6% occur on wet/icy/snowy roads (2010 to 2012 FARS data) (2).

Since an overwhelming number of fatal crashes are roadway departures, this is an emphasis area for the 45 US states in their Strategic Highway Safety Plans (3). Most highway agencies are pro-active in implementing a range of countermeasures to reduce road departures. However, agencies have only limited information about the effectiveness of some countermeasure. Analyses typically utilize crash studies which are not able to capture the effect of speed and distraction.

1.2 Project Objectives

The SHRP 2 Naturalistic Driving Study (NDS) data provides a unique opportunity to evaluate the relationship between driver and roadway characteristics in a manner not previously possible. These data provide a detailed record of driver and roadway characteristics during actual crashes/near-crashes as well as providing a snapshot of normal driving behavior.

The objective of Phase I was to utilize the SHPR 2 NDS and RID data to investigate the relationship between driver, roadway, and environmental characteristics and roadway departure risk. In particular the role that speeding and driver characteristics such as engagement in secondary tasks can be investigated in a manner not possible with traditional studies.

Phase I looked at roadway departure risk from three different perspectives. First, a logistic regression model was used to assess relevant the relationship between roadway and driver characteristics and crashes/near-crashes. Even with the full-SHRP 2 dataset, crashes and near crashes were still rare events. As a result, the study also used lateral position as a crash surrogate. Driver speed choice was also modeled as a function of driver and roadway characteristics. Each provided different information about roadway departure risk.

3. DATA SOURCES AND DATA REDUCTION

3.1 Data Sources

Several datasets were utilized as described in the following sections.

Roadway Information Database (RID): The RID contains detailed roadway data for around 12,500 centerline miles in the SHRP 2 NDS study states collected with mobile data collection. Roadway attributes include items such as: curve radius and length, presence of rumble strips, lane width, grade, number of lanes, speed limit, etc. The RID also combined data from several sources including state DOTs, HPMS, and other supplemental data which covered most roadways for each study state. Time series traces can be linked to the RID using GPS position.

Trip Density Maps: Trip density maps were created by VTTI to show number of trips and drivers by roadway link. A geographic file was provided by VTTI which could be used in-house to identify roadways with a certain number of trips.

Events: VTTI reduced a set of crashes and near-crashes from the SHRP 2 NDS data (4,246 total) which are available in the Event Detail Table on the SHPP2 InSight website. Over 76 variables are provided including crash type, severity, driver actions, presence of passengers, environmental factors, event narrative, etc. A brief video clip of the forward roadway is included along with graphical display of select vehicle kinematics (i.e. speed, acceleration, distance into trip, wiper status). High level roadway and traffic characteristics are also included such as intersection type, traffic control, alignment, and level of service.

A total of 1,631 roadway departure crashes and near-crashes were identified. Hereafter we will refer to crashes and near-crashes as "*safety critical events*". Events were selected using incident type (i.e. road departure left) or precipitating event (i.e. subject over left lane line) which indicated a roadway departure.

Around 32,586 baseline events are also available in the Event Detail Table. Similar information is provided on the website as for safety critical events except that a forward video clip is not provided. At least one baseline event was selected for each driver. Each baseline event lasts about 21 seconds. Since there are a large number of baseline events which could be mapped to roadway departure events, the team focused on those which included drivers involved in roadway departure safety critical events yielding 3,083 baseline events.

Once relevant safety critical and baseline events were identified, a download of attributes from the Event Detail Table was requested along with GPS coordinates, raw time series data (i.e. speed, acceleration, pedal position), and a video clip of the forward roadway. However, GPS coordinates for crashes were not provided due to privacy concerns.

Existing time series data: The team already had access to a set of time series data which were utilized for a previous project (SHRP 2 S08). The focus of that study was rural 2-lane curves. A set of over 4,000 time series traces were available for over 100 curves in the FL, NC, IN, NY, and PA study areas. This includes tangent and curve driving for over 200 different drivers. Due to time constraints in S08, only a fraction of that data had been utilized.

3.2 Data Reduction

The following describes data reduction that was common to all three analyses. When additional data elements were needed for a particular analysis they are described in the corresponding write-up. Table 1 summarizes the general characteristics.

Roadway Characteristics: When GPS data were available, times series data for each event was geocoded into ArcMap and overlain with the RID and aerial imagery. Roadway variables were extracted using the RID data when available. In some cases a variable was not collected, and in other cases the RID was not available for the study segment because the RID did not cover all roads in the NDS. Google Earth was used to extract the roadway features not included in the

RID such as RPMs or number of chevrons. When radius was not available in the RID, it was measured using aerial imagery and the chord-offset method. NDS forward video was used to determine delineation, pavement condition, roadway lighting, etc.

Table 1: Characteristics Included in Analyses and Data Sources

Feature	RID	Google Earth	forward video	driver video	event detail table/	provided by VTTI
roadway characteristics						
curve radius	✓	✓				
rumble strips, paved shoulders	✓	✓				
chevrons, advance curve signing	✓	✓				
RPMs, guard rail,		✓				
speed limit	✓	✓				
pavement condition			✓			
lane line condition			✓			
driver characteristics						
glance duration, secondary tasks				◙	▲	
impairment, seat belt use				◙	▲	
number of passengers				◙	▲	
hands on wheel				◙	▲	
speeding, too fast for conditions					▲	
age, gender, years driving						✓
number of violations, crashes						✓
following another vehicle			✓			
environmental/other characteristics						
time of day, lighting condition			✓			
ambient & surface condition			✓			
LOS					✓	

◙ for safety critical events ▲ for time series traces ✓ all data

Driver Characteristics: Static driver characteristics were collected with the NDS data and provided by VTTI. They include factors such as age, gender, annual mileage, etc. and were available for all sources of data. Vehicle type was also provided.

Driver Distraction: Several driver characteristics were provided in the event detail table for safety critical and baseline events. They include: impairment, maneuver judgement, distraction type, impairment, engagement in secondary tasks, etc. Visually distracting tasks and longer eyes-off-road glance duration have been noted as contributing crash factor (4; 5; 6; 7) but did not appear to be coded consistently in the Event Detail Table. Secondary tasks were only coded in safety critical events if they occurred within a 5 seconds window prior to start of the event. As a result, distractions that occurred upstream of the conflict were not included. For baseline events, secondary tasks were only coded for the last 6 seconds of the baseline epoch. Secondary tasks were coded when they involved non-driving related glances away from the driving task. As a result, duration of glances away from the forward view should correspond to length of secondary tasks but this does not appear to be the case since a number of tasks were recorded > 6 seconds and it is unlikely drivers were looking away for this entire period.

Driver glance location and distraction were coded for 583 time series traces for the SHRP 2 S08 study (8). Attention was measured by location inside and outside the vehicle (i.e. forward roadway) where a driver was focused for each sampling interval. Scan position, or eye movement, has been used by several researchers to gather and process information about how drivers negotiate roadways (9). Because eye tracking is not possible with NDS data, glance location was used as a proxy. Glance locations, represent practical areas of glance locations for manual eye glance data reduction. Secondary tasks were coded when they involved a glance away from the driving task. Another 261 additional traces were reduced for the Phase I study.

Environmental Characteristics: Time of day, pavement condition (i.e. wet, dry), lighting conditions (i.e., day, dawn, dusk, night/no lighting, night w/lighting) were provided in the Event Detail Table for safety critical and baseline events or were reduced from the forward view.

3.3 Overview of Roadway Crash/Near-Crash Events

Characteristics of safety critical events were summarized to provide a snapshot of relevant factors. Events were first categorized by type of roadway departure as shown in Figure 1. As noted, the majority of events (28%) were drivers who unintentionally drifted from their lane or left their lane during a turning maneuver and 2% lost control during a turn. Presumably these events occurred at an intersection. Almost 20% occurred in a parking lot. Intersection and parking lots accounted for over half of the safety critical events but are not what is typically considered a roadway departure.

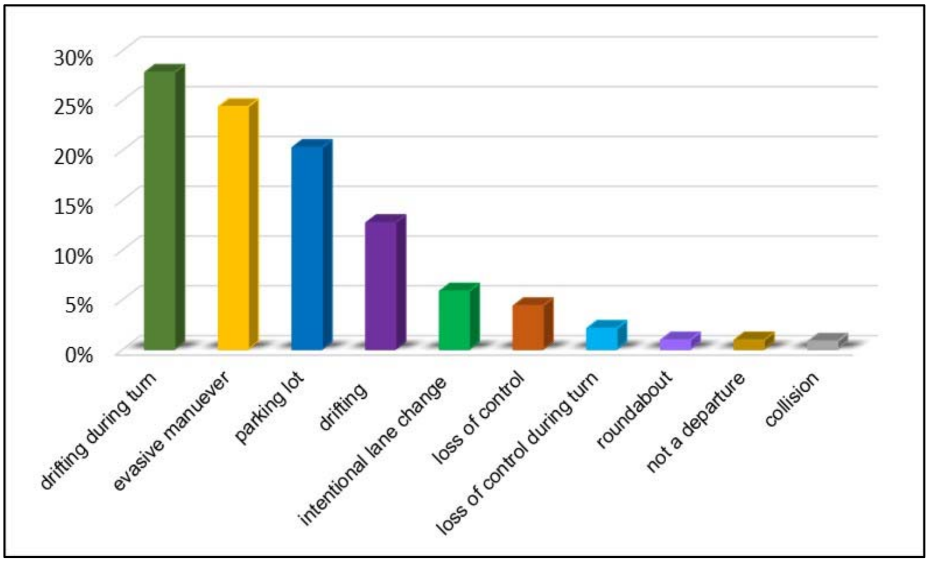

Figure 1: Types of Roadway Departures

Around 24% departed their lane during an evasive maneuver, 13% drifted from their lane, 6% occurred during an intentional lane change, 4% lost control, and 1% were departed their lane as a secondary event.

Drifting and loss of control are the types of events that roadway countermeasures are most likely to address. Actions by other drivers which resulted in collision or evasive maneuvers may also be addressed by roadway departure countermeasures since these drivers were likely to have left their lane resulting in the action by the subject driver. Evasive maneuvers and loss of control may be attributed to speed which suggests the need for speed management.

Since intersection and parking lot events are not traditional roadway departures they were excluded from the analyses. When these were excluded the 53% of the remaining roadway departures occurred on 2-lane roadways, 26% occurred on 4-lane roadways, and 15% occurred on 6- or more lane roadways.

Speed and distraction for safety critical events were compared against baseline events. Speeding (over speed limit or too fast for conditions) was noted for 17.6% of safety critical vs. 4.1% of baseline events. Drivers in crashes/near-crashes were slightly more likely to be noted as drowsy (2.3%) than baseline drivers (1.9%) and were almost twice as likely to be involved in a cellphone related task than baseline drivers (12.6% versus 8.4%).

4. METHOD OF ANALYSIS AND RESEARCH RESULTS

The SHRP 2 NDS data were utilized to address roadway departure risk from several different perspectives. The most important objective was to assess crash risk. To accomplish this, safety critical events were analyzed using baseline events as a measure of exposure to identify the impact of driver and roadway characteristics on roadway departure crash risk. Although all roadway departures are of interest, due to time and resource constraints the study focused on rural 2-lane roadways to demonstrate proof of concept.

Even with the full SHRP 2 dataset, crashes/near-crashes are still rare events. Only 783 traditional roadway departure events were available resulting in a limited sample size when spread across different roadway types. As a result, the team investigated various other surrogate measures which could be used to assess roadway departure risk. Speed and lateral position were both used as a surrogate safety measure. A speed prediction model was developed to relate speed to roadway and driver characteristics and another model was used to predict lateral offset as a function of roadway and driver characteristics.

4.1 Analysis of Roadway Departure Crash Risk

The objective of this analysis was to understand how driver distraction, roadway characteristics, and environmental factors affect the likelihood of roadway departure events.

Model Description: Data were modeled at the event level (i.e. one row per safety critical event or baseline). A total of 108 safety critical events on rural two-lane roadways were identified. Baseline events for the same drivers were extracted and used as measure of exposure. Safety critical events only included those where the driver drifted off the roadway or lost control. Secondary roadway departures resulting from avoidance maneuvers or crashes were not included in this analysis but will considered in Phase II. The odds of a roadway departure was the dependent variable and co-variates listed in Table 1 were tested in the model.

In an earlier study, Hallmark et al. (8) evaluated several statistical methods to predict roadway departure events from the SHRP 2 NDS including generalized linear, Bayesian, and regression tree models. Logistic regression was found to be the most appropriate statistical method to analyze the event level data.

The best fit model was selected by checking goodness-of-fit using the likelihood ratio chi-square test. The test compared the model with predictors to the model with intercept parameter only. Model statistics such as the likelihood ratio test, the Hosmer & Lemeshow test, Akaike information criterion (AIC), and R-square values were used to select the best fit model

Results and Conclusions: The final logistic regression model and model diagnostics predicted the likelihood of roadway departure event with 8 explanatory variables as shown in Table 2.

Table 2: Logistic Regression Analysis of Roadway Departure Event

Parameter	Estimate	SE	p-value	Odds ratio (95% confidence interval)
β_0 (intercept)	-3.1520	0.4408	0.000	0.04 (0.02, 0.10)
β_1 (Curb)	1.2924	0.4146	0.002	3.64 (1.63, 8.36)
β_2 (Paved Shoulder)	-2.1506	1.4972	0.000	0.12 (0.04, 0.29)
β_3 (Delineation)	1.1889	0.3919	0.002	3.28 (1.53, 7.18)
β_4 (Curve)	2.1318	0.3933	0.000	8.43 (4.00, 18.83)
β_5 (Icy/Snowy Surface)	2.5801	1.1431	0.024	13.20 (2.01, 168.78)
β_6 (Wet Surface)	2.4046	0.5510	0.000	11.07 (3.95, 34.78)
β_7 (Visual Distraction)	1.5721	0.4095	0.000	4.82 (2.20, 11.04)
β_8 (Younger Driver)	1.3062	0.3559	0.000	3.69 (1.86, 7.56)
Log Likelihood = -105.9974			*Likelihood ratio test = 235.39 (p value << 0)*	
Pseudo R-squared = 0.572			*AIC = 229.99*	
Hosmer & Lemeshow = 3.5016 (p value = 0.8991)				

As noted, safety critical events were 3.6 times more likely to occur on a roadway with curb and gutter. Tire-strike events are reasonably low risk events and were over-represented in the dataset since they trigger the 0.5 g force threshold used to detect safety critical events. The team debated whether tire strikes should be excluded from the analysis but it was decided that curb strikes still indicate a lane line crossings.

The analysis also indicated that safety critical events were 0.12 times less likely when paved shoulders were present and 3.3 times more likely when lane lines were either not present or obscured. Roadway departures were 8.4 times more likely to be on a curve than tangent section. Roadway departures were correlated to roadway surface condition (13 times more likely on icy/snowy roads and 11 times more likely on wet roads). Roadway departure events were 4.8 times more likely when a distraction was present and 3.7 time more likely to be younger drivers.

Conclusions and Limitations: Preliminary results indicated that relationships between safety critical events roadway and driver characteristics could be investigated. The major limitation was sample size. As noted in the logistic regression, several variables had very large confidence intervals (i.e. icy/snowy surface) which we believe is a result of the small number of observations in the baseline events. Inconsistencies in glance duration and secondary task coding were also limitations. Both limitations are addressed for Phase II.

4.2 Evaluating Relationship between Speed Choice and Roadway/Driver Characteristics

The objective of this analysis was to predict speed as a function of driver and roadway characteristics. Drivers involved in safety critical events were more likely to be coded as speeding suggesting speed is a contributing crash factor. Speed has also been used as a surrogate safety measure by other researchers (10; 11; 12). Speed was collected in all vehicles in the SHRP 2 NDS and in most cases speed appears to be available and reliable.

Data Utilized in Model: Baseline events on 2-lane tangent roadway sections were identified. Events where vehicles were slowed significantly due to congested traffic were excluded or if an intersection was present in the tangent section. This resulted in 240 baseline events which were included in the final analysis. VTTI only reduced driver distraction for the last 6 seconds of baseline events. As a result, mean speed was calculated for this period for each baseline.

Model Description: A number of researchers have used linear regression models to relate vehicle speed to explanatory variables (13; 14; 15). Since there were more than one observation for some drivers, correlation between observations needed to be accounted for. As a result, a mixed effect model was used to account for the within-driver interdependency. Mean speed was the predictor variable. The linear mixed model is similar to a linear regression model, but includes both fixed effect and random effect in the same model. The relationships of the independent variables are often assumed to be additive. The lme4 package in the statistical package R was used to estimate the model. Co-variates tested are shown in Table 1.

Results and Conclusions: The final model was chosen by using model diagnostics such AIC, Bayesian information criterion (BIC), and log likelihood test. The quasi R-squared was 0.766 for the mixed linear model and was found to fit the data appropriately. As discussed above, the mixed linear model added a random effect term repeated measures for the same driver.

Table 3: Fixed Effects of the Speed Prediction Model

parameters	estimate	s. e.	t values	p-value
intercept	19.05	2.5064	7.601	0.000
speed limit	0.66	0.0539	12.261	0.000
guardrail	3.10	1.3157	2.360	0.0252
younger driver	2.95	1.1111	2.652	0.0164
wet surface	-3.84	1.2133	-3.163	0.0079
street parking	-2.86	1.1589	-2.464	0.0216
following car	-3.07	0.5397	-5.696	0.0004
secondary task duration	-0.45	0.1825	-2.492	0.0207

The explanation of the fixed effect is similar to the linear regression model. Eight parameters were statistically significant and were included in the final model as shown in Table 3.

As expected, speed is higher as speed limit increases. Presence of guardrail increases drivers' speed by 3.10 mph which may be due to better delineation of the roadway. Younger driver (16 to 24) travel 2.95 mph higher than other age groups. Presence of a wet surface decreases mean speed by 3.84 mph. The presence of street parking decreases vehicle speed by 2.86 mph. Following another vehicle decreases mean speed by 3.07 mph. For each additional second that a driver engages in a secondary tasks, the mean speed was decreased by 0.45 mph.

Conclusions and Limitations: Overall the model successfully predicted vehicle mean speeds on two-lane tangent roadways as a function of driver, roadway, and environmental factors.

The main limitation was sample size. As a result, it was not possible to include a variety of roadway characteristics. Coding of driver distraction was also a limitation as described in Task 2. Both can be addressed in Phase II.

4.3 Evaluating Relationship between Lateral Position and Roadway/Driver Characteristics

Lane keeping measures such as lateral position and standard deviation of lateral position are widely used as a safety surrogate measure (15; 16; 17; 18). Lateral position was collected in the

SHPR 2 study with sensors that detect lane line markings or contrast in road surface. However, reliable lateral position was only available in a subset of data. In some cases sensors were not functioning. Additionally the lane tracking system is much less reliable when lane line discontinuities and obscured lane lines are present.

A preliminary analysis was conducted to determine whether relationship between lateral position and driver or roadway characteristics could be determined. Lateral position was reliable for 748 baseline events where distractions of different lengths were present. Standard deviation of lateral position (SDLP) was calculated for each event and was compared by glance duration as shown in Figure 2. As noted, SDLP increased with longer distraction durations. This preliminary analysis indicated that lateral position could be used to detect differences in driving.

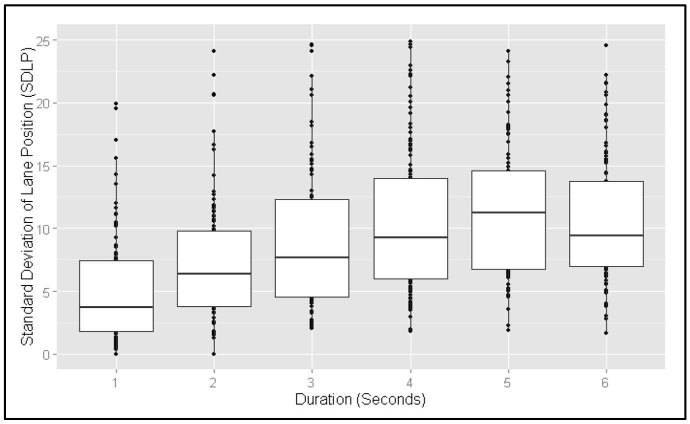

Figure 2: Relationship of Lane Position and Glance Duration

Next a model was developed which related lateral position to driver and roadway characteristics at several points before and within a curve.

Data Utilized in Model: Time series traces were available for rural 2-lane curves as noted in Section 3.1. Vehicle offset (distance between the vehicle centerline and lane centerline) was used as the lateral position metric. Offset data were not always reliable and accuracy was critical since small discrepancies could drastically skew the results of the model so the feasibility of using a particular time series trace was assessed using the lane markings probability variables. A total of 323 traces across 98 unique curves with 68 unique drivers were utilized.

Lateral position, speed, glance location, and distraction were extracted at 100 meters upstream of the point of curvature (PC) and at 7 equidistant points within the corresponding curve. Roadway, environmental, and driver characteristics were linked to the corresponding curve.

Model Description: A linear mixed effects (LME) model was utilized to predict offset within the curve. The LME model was chosen as it allows one to account for random effects due to repeated measures from including multiple traces by the same driver in the same curve. The LME function in the NLME package of R was used to develop the model. Correlation between the variables was examined to determine which variables should be included in the model. Due to the data being of a time series nature, a correction for the autocorrelation was required. Other model statistics, such as AIC and BIC, were evaluated to ensure selection of the best fit model.

Results and Conclusions: The best fit model included the variables shown in Table 4. Lateral placement for each point within the curve was also included in the model but not shown due to space limitations. The model suggests an association that as drivers tend to the inside direction of the curve in the upstream, the offset in the curve also shifts to the inside. It also found a correlation between curves with a radius less than 460 meters shifting 0.07 meters towards the inside of the curve. A driver glancing down at a particular point in the curve is associated with the driver's lane position shifting towards the inside of the curve by approximately 0.08 meters. A similar correlation was found if the driver was distracted at a prior point in the curve.

Table 4: Best fit model

variable	estimate	p value	95% lower	95% upper
intercept	-0.039	0.049	-0.079	-0.0002
Offset at 100 m upstream	0.438	<0.001	0.374	0.502
Small Radius (R<460 m)	0.067	0.050	0.000	0.134
Glancing down	0.080	0.016	0.015	0.146
Distracted in prior section	0.045	0.035	0.003	0.087
σ Driver random effect	0.026			
σ Curve random effect	0.096			
σ Residual	0.3382			
Phi 1	0.770		0.758	0.775
Phi 2	-0.197		-0.246	-0.147

The model was also used to plot vehicle path through a right hand curve (Figure 3). As noted drivers drift the most at the center of curve and drivers may be most vulnerable to a roadway departure at that point.

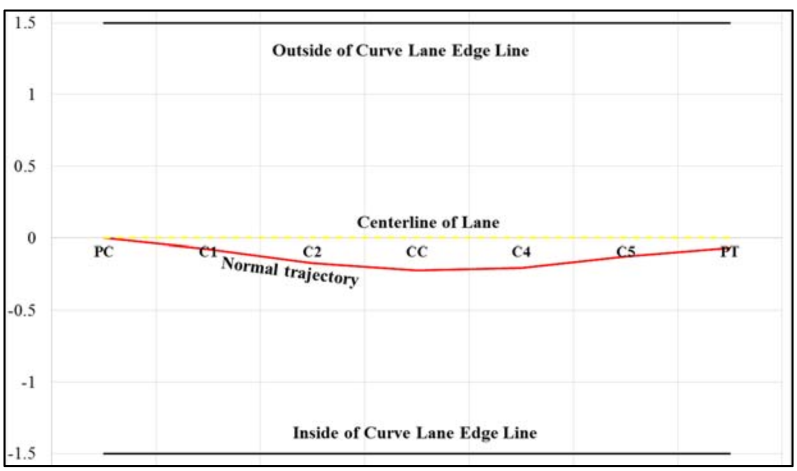

Figure 3: Parameter estimates of vehicle trajectories

Conclusions and Limitations:

The preliminary results indicated the feasibility of the analysis. The model is able to show a direct correlation between glances away from the driving task, presence of a distraction, and curve radius. Although 98 curves were represented, there was not a sufficient sample of the various countermeasures to detect an impact. As a result it is expected that including a sufficient samples of countermeasures desired or roadway characteristics of interest, the impact of roadway factors on lateral offset can be quantified.

The main limitation of this analysis was sample size. Reliable offset data were only available in a subset of the times series traces that were available. Additionally, reduction of driver glance and secondary tasks is time consuming and the number of driver types and roadway features that could be modeled was limited due to project constraints. Both can be addressed in Phase II.

5. FUTURE DIRECTION AND PHASE 2 PROPOSAL

This section outlines the research proposed for Phase II, summarizes the intended outcomes, benefits, and expected products.

5.1 Objectives

The goal of the research proposed for Phase II is to better understand the relationship between driver, roadway, and environmental characteristics and roadway departures. The research will focus on assessing the impacts of specific roadway factors and countermeasures in order to provide agencies with better information about which countermeasures are effective and why.

The SHRP NDS and RID data offer a rich source of data to explore these relationships. A first-hand account of the scenario preceding safety critical events including presence of distractions and speed can be directly investigated. Analyses of these safety critical events provides the most valuable insight into safety risk.

However, the number of crashes/near-crashes is limited. As a result, full exploitation of the dataset requires evaluating the data from different perspectives to glean additional insight into driver behavior and response to different roadway factors. In order to accomplish this the team proposes to analyze the data from three different perspectives. The major tasks are:

1) Task 3 will develop roadway departure risk factors using crash/near-crashes. This analysis will include all roadway types where these safety critical events have occurred. Phase I indicated that relationships such as distraction and certain roadway characteristics could be derived.

   Risk factors will be presented in the form of odds ratios which can be directly understood and used by roadway agencies to select countermeasures which are appropriate to a particular type of roadway departure.

2) Task 4 will develop a speed prediction model will be developed to assess the relationship between speed and driver and roadway characteristics. Speed will be used as a crash surrogate in this analysis. The relationship between speed and roadway characteristics is important since speed plays a significant role in roadway departure crashes. Additionally, agencies are regularly tasked with addressing speed management and need information to focus limited resources.

   Output from the speed prediction model will be estimates of change in speed due to roadway, driver, or environmental characteristics. This information can be directly utilized by agencies to assess the impact of different countermeasures or policies on speed reduction. This assumes that safety is improved by better managing speeds.

   One of the advantages of a speed prediction model is that speed is likely to be reliable for the majority of SHRP 2 NDS and as a result, selection of a large number of samples to meet task objectives is feasible. Drivers are expected to modulate speed in response to some changing roadway conditions and specific countermeasures such as advance signing. The disadvantage to use of speed as a predictor of driver behavior is that drivers do not necessarily adjust their speed for countermeasures whose intent is to keep the driver on the roadway, such as rumble strips. As a result, use of lateral position as a measure of safety will also be utilized.

3) Task II-5 will develop a lateral position model which will relate position within the lane to roadway and driver characteristics. Lateral position is an ideal safety surrogate measure since roadway departure is the consequence of failure to properly lane keep. Lateral position, however, is less available than speed so sample size and the number of roadway characteristics that can be modeled is limited.

   Output from the lateral position model will be an estimate of how vehicles lane keep when specific roadway or driver characteristics are present. This also provides a surrogate measure of safety risk.

The research addresses 3 of the Focus Areas for the Safety Task Force: Speed, Roadway Features and Driver Performance; and Preceding Contributory Events.

Dr. Linda Boyle, is a professor at the Department of Industrial and Systems Engineering at University of Washington. She is a human factors expert who has conducted a number of instrumented vehicle studies. She also has significant expertise with NDS data. Boyle's research centers on driving behavior, crash countermeasures, crash and safety analysis, and statistical modeling.

Budget and a time schedule are provided in Appendix A. Steve Gent (Project monitor from the Iowa DOT) was not able to attend the IAP project presentation. He provided a letter in support of the project (see Appendix A).

5.2 Benefits of the Proposed Research

The results of this research will provide more information about which specific roadway features are correlated to increased risk of road departure. It will also provide valuable information about how drivers interact with roadway features and the impact that has on the effectiveness of countermeasures. This will allow agencies to make better decision about countermeasure selection.

Understanding driver behavior will also provide invaluable information about why certain countermeasures work. The research has implications for roadway design, selection of sign type and placement, sight distance, and selection and application of countermeasures.

Model output will include estimates of safety risk for a given roadway factor or countermeasures, expected increases or decreases in speed, and expected improvements or degradation in lateral position. This information can be used to assess the expected benefit of incorporating a particular countermeasure or making a particular roadway improvement. This can be used directly to assess the impact of investment decisions with scarce resources.

Additionally, the research will provide results that will aid agencies in understanding the relationship between driver distraction and road departure risk which can be used by policy makers.

The proposed tasks to accomplish the research objectives are briefly described below. In order to focus resources, objectives 2 and 3 will focus on **rural 2- and 4-lane roadways** which represent the bulk of roadway departure crashes.

Task 1: Update IRB

It will be necessary to update IRB documents as required by Iowa State University. Dr. Boyle was not included in Phase I. She has completed the necessary IRB training and will complete her own IRB if needed.

Task 2: Data Requests and Data Reduction

Several datasets are already available from Phase I as shown in Table 5. One of the major limitation noted in Phase I was sample size. Additional data will be needed to ensure that roadway characteristics of interest can be included and a sufficient sample size is available. Since the three research objectives overlap, much of the data can be carefully selected and utilized for more than one objective as noted.

Once all data needs have been determined, data sharing agreements (DSA) will be developed in conjunction with VTTI. DSA will also include requesting permission to continue using data that are already in-house.

Task 3 will utilize the crash/near-crash data which was already reduced Phase I. Time series traces of normal driving will also be necessary on the same or similar roadways as a measure of exposure as described in Task 2. Tasks 4 and 5 will evaluate speed and lateral position as a function of driver and roadway characteristics using time series data.

Table 5: Summary of In-house Data

	event detail table data	time series data	linked to RID	forward video	distraction/ glance coded
crashes/near-crashes*	778	778	256	778	NA
baseline	3,083	2,186	2,186	2,186	NA
rural 2-lane trip traces	NA	4,106 over 219 curves	4,106	1013	515

*excludes intersection and parking lot events

Addressing Limitations Noted in Phase I: Several limitation were noted in Phase I which will be addressed in Phase II as noted in the corresponding section. Limitations include:

- sample size
- GPS not provided for crashes due to privacy concerns so roadway factors not reduced
- baseline incidents are not representative of safety critical event characteristics
- glance duration not indicated in safety critical and baseline events
- secondary tasks only coded for sections of safety critical and baseline events
- roadway characteristics and countermeasures included were limited

Crash/near-crash/baseline event data: 738 roadway departure safety critical events were identified in Phase I and roadway data reduced as noted in Section 3.2 for near-crashes. GPS coordinates are needed for crashes so that corresponding roadway characteristics can be reduced. We have offered to code roadway characteristic for VTTI using an interim ID which we could not use to link to event IDs. We would reduce all relevant roadway factors and provide this information back to VTTI. They would be able to match the interim ID to the event ID and make this data available to all researchers. As an alternative, roadway factors will be coded during a planned trip to reduce driver glance and distraction since the crash data can be accessed in the secure data enclave.

Time Series Data: A set of time series data are already available as noted in Table 5. Other data will be requested as needed in the form of traces along roadway segments of interest (1 trace is 1 driving trip across an individual roadway segment). Roadway segments where near-crashes occurred have already been identified and locations for crashes will be identified in Phase II. At least 2 traces will be requested for each segment (1556 traces) and will be used as a measure of exposure for analysis of safety risk in Task 3 in lieu of baseline events since it was difficult to select baseline events with similar roadway characteristics. Traces on rural 2 or 4-lane roadways can also be utilized for Task 4 and 5.

Additional time series traces will be requested for Tasks 4 and 5. Roadway segments will be selected to maximize the number of roadway characteristics and countermeasures that can be included. Countermeasures are of particular interest since agencies desire a better understanding of which treatments are the most like to reduce speeds and subsequently roadway departures. Table 1 listed roadway characteristics which were included in analyses in Phase I and will also be included in Phase II.

In addition, we will focus on ensuring that selection of roadway segments of interest includes a wide range of roadway departure countermeasures. Table 6 summarizes 14 roadway countermeasures which are known to be available in the SHRP 2 data and can reasonably be included. Other roadway characteristics, such as varying lane width or shoulder type will also need to be accounted for resulting in approximately 20 distinct roadway characteristics. There are also 10 distinct driver characteristics (i.e. age) and behaviors (i.e. speeding) that can reasonably be obtained and modeled resulting in 30 potential co-variates. A good rule of thumb is 50 observations for each co-variate tested resulting in minimum sample of 1,500 traces.

The team will review the traces requested for Task 3 to determine whether a sufficient sample is available for each desired roadway or driver characteristic. Reliability of the speed and lateral position data for each in-house trace will be assessed. The team will then determine how many of the existing traces are usable and how many additional are needed. For instance, in order to include paved shoulders a minimum of 50 traces with paved shoulders is desirable.

Additional locations will be identified by overlaying the trip map with the RID and then querying locations of interest (i.e. rural 4-lane undivided with rumble strips). However many of the desired roadway characteristics, such as presence of RPMs, were not coded in the RID. As a result, it will be necessary to use the supplemental RID data, video log, and potentially Google Earth to manually identify roadways with the desired characteristics. Each identified segment will be cross-checked with the trip map (described in Task 2) ensuring a sufficient number of drivers are likely to be available.

Table 6: Roadway Factors Known to be Present by Data Category

Countermeasure	crash/near-crash	baseline	full SHRP 2 dataset
RPM*	X	X	X
chevrons	X	X	X
rumble strips (center/edge line)	X	X	X
varying curve radii (and other curve characteristics)	X	X	X
advisory curve signing	X	X	X
lane line marking presence and quality (i.e. faded)	X	X	X
guardrail	X	X	X
paved shoulders	X	X	X
estimate of clear zone	X	X	X
speed limit (to calculate speeding)	X	X	X
post mounted delineators			X
on-pavement curve sign			X
speed feedback sign			X
widened centerline			X
**locations with raised pavement markers will be identified as part of NCHRP 5-21*			

An approximate 1 mile section will be identified for each roadway segment and a buffer created for each. Buffers will be provided to VTTI so the corresponding trips through each buffer can be identified. The team will work with VTTI to specify a filter which can be used to exclude trips where speed and lateral position are not reliable. Speed is more likely to reliable than lane position. As result it is expected that fewer samples with reliable lane position will be available so we will ensure the minimum sample is met for speed. Both tangent and horizontal curves will be included.

Time series data (i.e. speed, acceleration, pedal position, GPS location) for each corresponding trip will be requested from VTTI along with a clip of the forward video, driver characteristics (i.e. age, gender, number of violations),

Reduction of Roadway/Environmental Factors: New time series traces will be overlain with the RID and roadway factors extracted. When not available, the forward video view, Google Earth and other data sources will be utilized as described in Section 3.2.

Many roadway factors will be consistent for the segment (i.e. lane width) but there may be some variations including presence of a horizontal curve or a segment with a passing lane. Since the time series traces will be geocoded, varying roadway characteristics can be related to the appropriate time stamp.

Time of day can be extracted from the time series data. Ambient conditions (i.e. raining) can be obtained from a review of the forward view or inferred from wiper/headlight use. Roadway surface condition will also need to be extracted from the forward view.

Reduction of Driver Factors: Studies have indicated that visually distracting tasks, such as dialing a hand held device, were much riskier than secondary tasks that did not involve glancing away from the driving task (4; 5; 6). Peng (7) found longer eyes-off-road glances were positively correlated with higher roadway departure risks.

Distractions and secondary tasks in the event detail table were described as being associated with a glance away from the driving task. As a result, secondary task length should correlate with glance duration. However, review of the data for crashes/near-crashes suggests that this may not be the case. For instance, a large number of events had distractions from 6 to 24 seconds and is unlikely that drivers glanced away from the driving task for that long. Since driver attention away from the forward roadway had been correlated to crash risk, it is important to ensure glance duration is correctly coded. First we will have a conversation with VTTI to clarify. If the issue cannot be resolved we will explore whether glance location should be coded in a visit to the secure data enclave.

Glance and secondary tasks data are available for some of the times series traces that the team already has access to as indicated in Table 5. Driver glance and engagement in secondary tasks will be reduced for additional traces as needed at the VTT secure data enclave. The team developed a methodology to reduce driver glance location and distraction associated with glances away from the driving task in SHPR 2 S08. We will utilize that a similar methodology (19). **Dr. Linda Boyle** joins the team as a human factor and NDS expert and will assist with defining the final driver data reduction methodology. Glance and secondary tasks will be coded by time stamp so that when data within the segment are sampled, the corresponding driver behaviors can be included.

Data Security Plan: all team members and staff who have access to the data have IRB training and will be included in the DSA. InTrans hosts and manages several servers on-site. For IRB-protected data we have an isolated backup routine on separate media that is always within control of either IT staff (in the locked server room) or the project PI. When not in the locked server room, media is located off-site at a locked location within a locked fireproof safe with only IT staff and the project PI having access to the keys. In addition to IRB training, appropriate use of the data are discussed with all team members. When data with personally identifying information is found in data used in-house, VTTI is alerted and the data are deleted.

Task 3: Correlate Roadway Departure Risk to Driver and Roadway Factors

A preliminary analysis of safety critical events was conducted in Phase I using logistic regression (Section 4.1). The limitations present in the model for rural 2-lane roadway will be addressed and the methodology expanded to the other roadway types.

Data Required: The analysis will include all roadway departures events as noted in Phase I (all roadway types). Data will be modeled at the event level (i.e. one row per crash/near-crash/normal driving event). As a result, data will be aggregated at that level. Average speed, standard deviation of speed, and maximum speed will be calculated and included in the model.

Methodology: Results from Phase I indicated that logistic regression was the appropriate tool to identify the contributing factors to roadway departure events. Logistic regression is ideal since the data are suited to this type of model and explanatory variables can easily be included.

Probability of a roadway departure crash will be modeled as the dependent variable. Roadway, driver, and environmental conditions listed in Table 1 will be included as co-variates. Although sample size is somewhat limited, we will test severity (i.e. probability of a crash versus near-crashes) as a dependent variable as well.

Expected Outcome and Benefits: The expected outcome is risk factors (in the form of odds ratios) that can be used by national/state/local transportation agencies to assess the advantages of certain types of roadway countermeasures. Odds ratios can be calculated for each contributing factor which is helpful since the concept of odds ratio can be easily understood by stakeholders and practitioners.

The odds ratio can be used similar to a crash modification factor (CMF) in that it provides an actual estimate of the effectiveness or risk for a particular roadway or driver characteristic. As an example, the outcome in Phase I showing roadway departures are 3.3 times more likely when lane lines are missing or obscured can be used to calculate the benefit/cost of installing or maintaining edge lines. The impact of speeding, impairment, and distraction on safety risk can also be evaluated.

Task 4: Evaluation of Driver Speed Choice based on Roadway Characteristics

The objective of this task is to assess the relationship between speed and driver/roadway characteristics. Speed will be used as a crash surrogate in this analysis. The relationship between speed and roadway characteristics is important since speed plays a significant role in roadway departure crashes. This task will expand the speed models developed in Phase I.

The models will initially include all of the countermeasures indicated in Table 6. It is expected however that some roadway countermeasures, such as better delineation, may actually increase speed but may not necessarily be a safety risk. The lateral position model (Task 5) will bridge this gap by assessing safety risk from a different perspective.

Data Required: Time series traces will be used for this analysis and roadway, driver, and environmental characteristics reduced as described in Task 2. It is expected that approximately 1,500 traces will be available for each roadway type and will represent a range of roadway and driver characteristics. Data will be reduced as described in Task 2. Data will be sampled along each segment at 300 meters as shown in Figure 4 for tangent sections. Speed will be extracted upstream, at the point of curvature (PC) and center of curve (CC) as shown in Figure 5. Each discrete sample point will include the corresponding roadway, driver, and environmental factors. As noted, driver distraction would be noted for Point 2 in Figure 4. Multiple samples allows inclusion of the changing driver state (i.e. distraction) and location (tangent versus curve). Speed will be averaged over a 2-3 second interval around the data point.

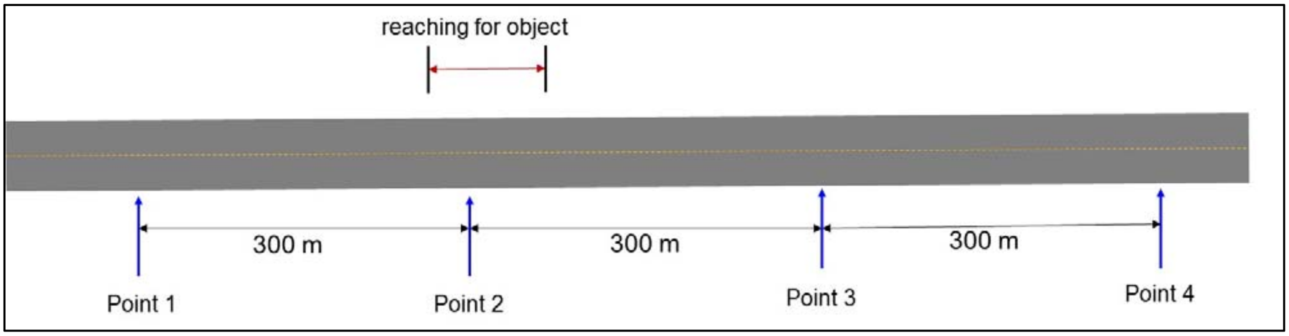

Figure 4: Data Collection Methodology along Tangent Section

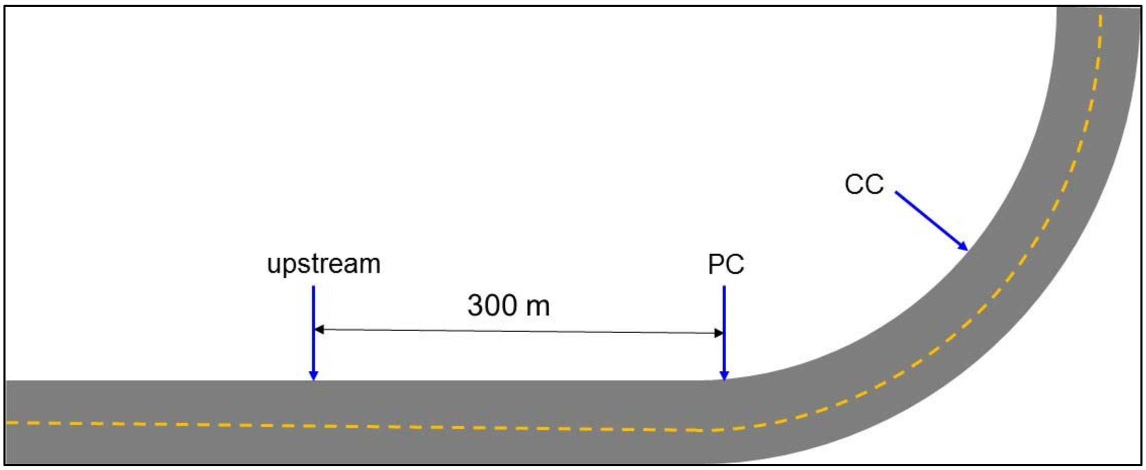

Figure 5: Data Collection Methodology along Curve

Methodology: A linear mixed effects model will be used as described in Section 4.2 for Phase I. Mean speed at a particular point will be modeled as a function of roadway and driver characteristics. Other metrics such as standard deviation of speed, amount over the posted speed limit (or probability of exceeding the posted or advisory speed), and change in speed can also be used as the dependent variable. Characteristics listed in Table 1 and countermeasures listed in Table 14 will be evaluated as co-variates. Models will include curve and tangent sections.

The model will predict speed as a function of roadway/driver characteristics. For instance change in speed from the tangent to within curve can be utilized to assess the impact of curve countermeasures, such as presence of chevrons. The change in driver speed can also be used to determine whether drivers slow down as the encounter a particular countermeasure (i.e. speed differential upstream of a dynamic speed feedback sign can be used to model driver response).

Expected Outcome and Benefits: Model outcome is an estimate of speed due to statistically significant co-variates. Results will indicate expected increase/decrease in mph for a given roadway, driver, or environmental characteristic. For instance, presence of paved shoulders may decrease speed by 'XX' mph. This provides information that easily understood and can be applied by stakeholders.

Most of the roadway countermeasures which can be included in this analysis are not specifically geared towards speed management. Rather their function is typically to keep on the roadway (i.e. RPM) or minimize the impact when vehicles leave the roadway, such as paved shoulders. Countermeasures which improve delineation or create a more forgiving roadside may provide additional safety but may also increase speeds. Consequently results for this task will need to be presented in the overall context of their impact on safety and not just speed reduction.

Task 5: Evaluate Driver Response Based on Lateral Position

The objective of this task is to use lateral position in the form of offset from the lane center as a surrogate measure for roadway departure risk. However, lateral position is not as reliable as speed. For instance, the lane tracking system included in the SHPR 2 DAS had difficulty calculating offset when lane lines were missing or obscured. Additionally discontinuities in lane lines which occur at intersections, turn lanes, etc. also affect the ability to calculate lane position. As a result sample size is not expected to be as large as for the speed model. As noted in Section 4.3, resources will be focused to ensure a sufficient sample of time series traces is available along roadways with countermeasures which are expected to directly impact lateral position to order to conduct a robust statistical analysis of those factors.

A model of lateral position for curve driving was developed in Phase I and demonstrated the feasibility of the method. Models will be developed for to include both tangent and curve segments on rural 2-lane and rural 4-lane roadways.

Data Needs: A subset of the data utilized in Task 4 will be used for Task 5. Since lateral position is less reliable than speed, only a subset of the time series traces will have sufficient fidelity to include in this task. As a result, this reduces the number of co-variates that can be included. Since a lateral position model is more likely to be able to be able to pick up differences in behavior due to countermeasures geared towards keeping vehicles on the roadway (i.e. rumble strips, lane line quality), we will focus on these types of countermeasures for these models.

Driver, roadway, and environmental characteristics will be reduced as noted in Task 2. Factors which are not continuous through the trace, such as distractions or presence of a curve, will be correlated to the appropriate time stamp and spatial position. Tangent segments along each time series trace will be sampled at 300 meter intervals as shown in Figure 4 and will be selected so they are at least 300 meters upstream or downstream of an adjacent curve to ensure curve reaction behavior is not included in tangent samples.

Nine points were sampled in the analysis presented for Phase I which was also used to demonstrate path through a curve. Extraction of the data at multiple points is resource intensive and additional points within the curve does not provide any significant benefit in terms of identifying relevant characteristics that affect lateral position. Data will be sampled at a point 300 meters upstream (represents non-curve driving), at the PC since drivers who enter the curve too fast are more likely to vary lane position at initial curve entry, and at CC. The CC was identified in the Phase I analysis as the point with the most significant lateral offset. The point immediately upstream can be used to assess the change in speed from normal upstream to curve driving.

Each sampling point will be one observation in the lateral position model and the corresponding roadway, driver, and environmental factors extracted. The correlation between multiple points from the same time series trace will be accounted for in the model. Additional multiple observations for the same driver and roadway will handled using repeated measure.

Methodology: A linear mixed effects model was utilized for Phase I and the methodology will be expanded for Phase II. A brief description of the model form was provided in Section 4.2. The LME model was chosen as it allows one to account for random effects due to repeated measures from including multiple traces by the same driver in the same curve. Lane offset, distance from right lane line, or standard deviation of offset will be the predictor variables. Lane offset is the distance between the center of the vehicle and the center of the lane. Distance from the right lane line is calculated from lane offset, lane width, and vehicle track width. They essentially model the same thing but use of position from the right lane line may be easily understood by practitioners since the impact can be visualized.

The best fit model will include those characteristics which are statistically significant. Model fit will be assessed using common statistics such as AIC as described in Section 4.2.

Expected Outcome and Benefits: The expected outcome for this task is a model that predicts the relationship between lane position and driver and roadway countermeasures. The results will be presented in a manner that the implications of lane position within the context of safety risk are explained. For instance, factors which result in vehicles drifting more within their lane can be assumed to be an increased safety risk. Alternatively countermeasures which decrease offset and variation are expected to improve safety.

Task 6: Final Report and Outreach

The main product will be a final report which will include:

- Introduction and summary of literature
- Summary of data request and data reduction protocols
- Results and conclusions for risk analysis
- Results and conclusion for speed prediction model
- Results and conclusion for lateral position model
- Conclusions and lessons learned

Since the intended audience is state, county, and local highway agencies, tech briefs will also be developed which summarize findings in a format that is more easily used by practitioners. For instance, charts, matrices, or clearly defined figures provide quick access to needed information. A tech brief (2 to 4 pages) will be developed to summarize background and results for the risk analysis. The odds of a safety critical event occurring when a particular roadway or driver characteristic is present will be presented along with confidence intervals. This information can be used in a similar manner as CMFs which are easily understood by practioners. As a result, task outcomes can be used to assess the direct impact of the characteristics included in the model on safety risk.

Results of the speed prediction and lateral position models will be combined in one tech brief. Although they will likely pick up the impact of different countermeasures, there will be some overlap in the roadway and driver factors which are included. When speed and lateral position are improved for a given countermeasure, this information will be highlighted. Alternatively, a roadway or driver characteristics that increased speed and decreased lateral position (or vice versa) would also be explained in the context of both pieces of information.

Countermeasures which improve delineation or create a more forgiving roadside may provide additional safety but also result in higher speeds. Consequently results of the speed prediction model will need to be presented in the overall context of their impact on safety and not just speed reduction. Additionally the outcome of the lateral position model, which is shift in lateral position, is not as easily understood as odds ratio or a change in expected speed, the tech brief will need to provide recommendations that can be used by practitioners based on model results. As a result, it will be necessary to present this information along with some guidance on how results can be utilized.

The tech briefs will be geared towards highway agencies who are primarily interested in roadway factors and countermeasures. Project findings related to the relationship between driver distraction/attention and speed will also be summarized for policy makers. This type of knowledge can be used to assess the impact of policies such as texting laws.

Outreach is an important component of the research effort. All products will be posted on the website for the Institute for Transportation. The site is indexed to be searchable by all major search engines. The team are partners in the National Center for Rural Safety (lead by Western Transportation Institute). The team also works closely with the Iowa Local Technical Assistance program (LTAP). The team will work with both groups to ensure products resulting from the research are useful to stakeholders. Additionally, we will work with them to find opportunities to outreach project results.

IOWA DOT
SMARTER | SIMPLER | CUSTOMER DRIVEN

www.iowadot.gov

Office of Traffic & Safety
800 Lincoln Way, Ames, Iowa 50010
Phone: 515-239-1129 | Email:steve.gent@dot.iowa.gov

September 11, 2015

To: The SHRP2 Safety Task Force
From: Steve Gent, P.E., Director of Traffic and Safety
Iowa Department of Transportation
Re: IAP Projects Phase 1 Presentations in Arlington, VA on October 19 – 20, 2015

State travel restrictions have prevented me from attending the October IAP Phase 1 Project Presentation with Iowa State University. As the Iowa Department of Transportation contact on this project, please accept the following summary of my thoughts regarding the Roadway Departure project.

On behalf of the Iowa Department of Transportation, I strongly endorse the need for additional funding to continue to examine and understand roadway departure crashes using the SHRP2 Safety Data (Naturalistic Driving Study database and Roadway Information Database). Roadway departure crashes account for approximately 56% of traffic fatalities in the United States. Further research into roadway departure crashes will give practitioners the knowledge and understanding to significantly improve the safety of our roadways.

The Iowa State team has demonstrated the ability to capture critical data from the SHRP2 databases and to answer specific questions that safety professionals need to know to enhance roadway safety. The areas currently being finalized include:

- the cause of the crash: human factor, roadway factor and/or vehicle factor
- contributing factors: distraction, speeding, roadway surface condition, traffic condition, etc.
- how driver distraction impacts lane keeping/lateral position
- lane keeping/lateral position around curves
- how drivers change their speeds on curves
- lane keeping/lateral position related to roadway characteristics such as lane width, shoulder type, presence of guardrail

The major focus of this research is to evaluate roadway departure crashes, however from the list above, clearly this research also addresses crashes (and near-crashes) related to horizontal curves, speeding, distraction and adverse weather conditions which are emphasis areas for roadway safety professionals. The Iowa State team has extensive experience with the SHRP2 Safety Data and possesses the knowledge and skills necessary to further mine these rich databases.

Again, the Iowa Department of Transportation strongly recommends this Roadway Departure proof-of-concept research be granted Phase 2 status. We believe this expanded research will produce meaningful results that result in safer roadways throughout the nation.