SHRP 2 S08: Analysis of the SHRP 2 Naturalistic Driving Study Data — Relationship between Driver Behavior and Safety on Rural 2-Lane Curves

Shauna Hallmark, Nicole Oneyear, and Dan McGehee
July 2013
Objective

- Assess the relationship between driver behavior and characteristics, roadway factors, environmental factors, and likelihood of lane departures using NDS and roadway data
  - Develop models to quantify the relationship between driver behavior and the roadway environment
  - Focus on curves on rural 2-lane paved roadways (lane tracker does not function on unpaved roadways)
Data Needs

- Conducted literature review to identify factors of interest
- Reviewed available data from NDS and S04
  - Developed data collection/reduction methodology for Phase I
    - Some roadway data available but mobile mapping data not yet available
    - Unconsented drivers not yet identified
  - Revised with updated data for Phase II
  - Need time series data and data reduced at the event level
IRB

- Obtained IRB from ISU
- Completed data sharing agreement with VTTI
- Fairly straightforward
- Listed specific variables and methodology for data request
- Time consuming for ISU to approve
  - Phase I lesson learned: start early
Sampling Plan for Phase II

- Balanced:
  - Need for statistically representative sample based on potential number of covariates
  - Cost to procure data
  - Time to reduce and analyze data
  - Time constraints for Phase II

- 800 to 1000 traces

- 5 states appeared to have large rural component
Data Collection and Reduction

- NDS and RID not yet merged
- Overlaid both VTTI trip maps and RID where mobile mapping was available
- Evaluation criteria
  - No turning or passing lanes in curve
  - No stop controlled intersection on the major approach and in curves and tangents
  - No signal controlled intersections in curves or tangents
  - No railroad crossings within curve or tangents
  - No sites within 0.5 miles of town
  - No construction (need to check against forward view)
Curve Identification

- Queried RID for rural 2-lane curves of interest
  - radius, presence of RS, etc.
- When mobile mapping not available
  - used other RID data, Google Earth
    - radius using chord-offset method
- Overlaid with weighted trip maps
  - Number of trips and drivers
Curve Identification

- Identified additional roadway characteristics with Google
- Extracted other characteristics necessary for sample selection for roadways where mobile mapping not

Roadway Data
- Distance between curves
- Curve type (isolated, S-curve)
- Radius
- Presence of countermeasures:
  - Chevrons
  - Edgelines
  - Paved shoulders
  - rumble strips
  - Other curve signing
  - RPMs

Data sources: ESRI, Florida DOT
Data Request – Phase II

• Step 1: Narrow to 400 to 500 curves based on desired curve characteristics
  • Curve radius
  • Shoulder width
  • Other geometry of interest
  • Countermeasures of interest
Data Request – Phase II

• Step 2: request driver characteristics for initial curves from VTTI
  – Matrix of driver ID, age, vehicle type
  – Narrow to 350 curves based on desired driver/roadway/vehicle characteristics
  – Request one forward view for each curve
    • Confirm roadway condition and identify sites with construction/other issues
    • Phase I lesson learned
      – Number of traces had new traffic control/roadway conditions or were in work zone
Data Request – Phase II

• Step 3: develop final data request
  – Select sites to represent range of characteristics considering sample size
  – Exclude traces
    • Lane tracker not functioning
    • Poor video quality (if known)
    • No forward/side acceleration
    • Need sufficient samples with steering wheel and pedal position
  • Phase I lesson learned: decide critical DAS variables and then ensure they are present in samples requested
    – some variables not collected on certain vehicles
    – some variables not reporting for some time periods
Data Request – Phase II

• Step 3: develop final data request
  – Ideally receive in batches so request could be fine tuned if necessary
  – Traces will include curve and tangent sections
  – Need to ensure that near-crash/severe events are included
    • Indicate thresholds (i.e. include any trace where lateral acceleration $\geq 0.5\ g$)
Reduction of Kinematic Driver Factors

- Driver distraction – forward and over shoulder video at secure data enclave
  - type
  - duration -- coded by video time
  - Attempting to include fatigue using steering wheel reversal – Phase II
- Driver response – DAS variables
  - speed
  - forward acceleration
  - steering angle input
  - braking
  - gas pedal input
- Driver forward attention - manual reduction from forward video at secure data enclave
  - Head position
  - Scan (glance) location
- Head position available (not used in Phase I), evaluating for Phase II

Eye Glance Locations
1. Forward
2. Rear-view Mirror
3. Up
4. Steering Wheel
5. Center Console
6. Down
7. Right
8. Left

Not shown: Over the Shoulder, Missing and Other
Reduction of Roadway Factors

• Phase I: manually reduced variables using aerial imagery, Google and forward video view - currently updating with mobile mapping

• Phase II: using RID
  – Mobile mapping is most accurate when variables are available
  – Some state roadway information
  – Aerial imagery
  – Forward video view from DAS

- Speed limit
- Advisory speed limit
- Roadway pavement type
- Shoulder type
- Type and quality of lane markings
- Distance to next nearest upstream/downstream curve
- Presence of chevrons
- Number of chevrons
- Presence of curve advisory signs
- Other signing
- Begin curve (PC)
- End curve (PT)
- Presence of rumble strips
- Delineation
- Driving direction
- Estimate of lane width from DAS
- Pavement condition (visible deformities)
- Number of access points within curve and tangent section
Reduction of Vehicle Factors

- Spatially correlated vehicle trace to curve
  - Location of vehicle upstream, within, downstream of curve
  - Identified location of countermeasures for RQ 3

- Calculated lane position from lane tracking variables
  - Needed for lane crossings
Crash Surrogates

- Will not be sufficient crashes
- Need lane departure crash surrogates
- Surveyed literature and considered against available data
  - Time to collision (vehicle location and spatial location of objects not sufficiently accurate; forward radar did not appear to be adequate)
  - Lane deviation (common measure for lane departure and data are available)
    - Offset from lane center
    - Distance from right or left lane line
Research Question 1

• Defines curve area of influence and normal driving
  – Necessary to set incident boundaries for Research Questions 2 and 3
  – Applications for design, signing, countermeasure placement, setting speed limits
• Build on models of theoretical curve driving behavior
• Data sampling: time series (continuous) data at level provided in DAS
  – One row = 0.1 sec
• Evaluated kinematic variables that indicate response

Image source: Campbell et al. (2006)
Research Question 1

- **Initial Results**
  - Relationship between forward acceleration, side acceleration, speed, and gas pedal position and position along curve
  - Influence area within 200 m upstream
  - Define normal offset based on curve geometry
Research Question 1

• Initial Results
  – Evaluated relationship between forward acceleration, side acceleration, pitch, roll, and offset and curve negotiation
  – Highly correlated to driver
Research Question 2

• What is the relationship between driver distraction, other driver, roadway, and environmental characteristics and risk of lane departure?

• Provides insight into which driver, roadway, and environmental factors are safety critical

• Response variables: probability of left or right side lane departure of a certain magnitude (separate models)
  – Also considering lane departure of a certain magnitude (i.e. > 0.3 feet)
Research Question 2

- **Data sampling**
  - Aggregate by epoch – one driver trip through curve or tangent section
  - Curve and tangent sections
  - Included upstream (driver distraction and focus area prior to negotiating curve)
  - 90 traces from Phase 1

- **Model Form**
  - Multivariate logistic regression
  - Odds of a lane departure of a certain magnitude given driver, roadway, environmental variables
  - Stepwise approach
  - Consider correlation among variables

\[
\log \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \epsilon_{ij}
\]
Research Question 2

- Develop simple odds ratio using threshold of \( \leq -0.3 \)
  - Amount of time driver was engaged in a distraction immediately upstream was also statistically significant
  - Drivers engaged in distraction > 25% of time upstream were 4.3 times (2.6, 5.98) more likely to have a right lane departure
RQ 2 – Consideration for Phase II

• Initial results not as good as expected given 90 samples
  • However roadways were similar and few drivers available
    – 70% of lane departure for 3 drivers
    – Driving and probability of lane departure highly correlated to driver

• Indicates need to ensure large sample of drivers over range of conditions

• Better define lane departure ($\leq -0.3$ m + some kinematic indicator)

• Evaluated sample size
  – Approximately 10 samples for each category, around 790 to 1,000 observations
  – Reduction of sample size is feasible given time frame and proposed staffing
Research Question 3

- What roadway cues and countermeasures are the most effective in getting a driver’s attention and how do they affect driver response to horizontal curves?
- Identifies driver response to roadways cues and relationship to curve negotiation
- Analytical Approach
  - Dynamic linear model (DLM) -- time series model
  - Allows for the inclusion of explanatory variables and a stochastic time components
  - Can model several time series at once
Research Question 3

- Data sampling approach
  - time series data (continuous) data at level provided in DAS
    - One row = 0.1 second
- Response variable: deviation within lane or likelihood of a lane departure
- Covariates: similar to research question 2
  - Roadway (with focus on countermeasures)
  - Driver
  - Environmental
Research Question 3

• Initial Results
  – Developed initial model for one curve with multiple traces for 1 driver (same direction of travel)
  – Included driver glance position, driver distraction, and position in relationship to location along the curve
Lessons Learned

- Review existing available datasets (VTTI warehouse and CTRE)
- Do own legwork (providing exact locations expedited request)
- Review past documents by SHRP vendors who have utilized data (S01, S02, etc.)
- Site visit if feasible
- Communicate early and often with RID and NDS providers
- Make and review small initial data request
  - Request 1 vehicle trace for each curve to check for unusual situations
  - Filter data request for critical DAS variables
- Patience!! (have a back-up plan)