Project Title: Mitigating crash risks in work zones: causal inference and Crash Modification Factors

Recipient/Grant (Contract) Number: Carnegie Mellon University, Grant #: 69A3552344811

Center Name: Safety21 National University Transportation Center for Promoting Safety

Research Priority: Promoting Safety

Principal Investigator(s): Sean Qian Carnegie Mellon University

PI Contact: seanqian@cmu.edu

Project Partners:

- California Department of Transportation
- toXcel Inc

Research Project Funding: $202,042.00

Project Start and End Date: 07-01-2023 to 06-30-2024

Project Description:

In the U.S., a work zone crash occurred every five minutes during 2015-2019. It is still unclear the causes of those crashes, such as work zone configurations, weather conditions, work zone duration, and roadway characteristics. The causal impact of those potential factors to work zone crashes may vary substantially cross different types of roads and traffic flow characteristics. Agencies have been working on mitigating work zone crash risk by implementing work zone countermeasures, such as increasing work zone duration, left-hand merge, downstream lane shift, increasing the inside shoulder width, and two-way two-lane operations. The effectiveness of such countermeasures is typically evaluated by a Crash Modification Factor (referred to as CMF throughout this proposal), and more generally, Crash Modification Functions. To this end, crash modification factors (CMF) are typically unknown for work zones. For example, the Manual on Uniform Traffic Control De- vices (MUTCD) provides qualitative recommendations of signs, flaggers, or closure settings for work zones with different characteristics, but no quantitative CMFunctions. FHWA CMF clearinghouse only provides the CMFunctions for implementing left-hand merge and downstream lane shift in rural areas and modifying shoulder width in urban areas. It is critical to understand the root causes of work zone crashes and propose effective strategies to reduce crash occurrences. This project will continue the work zone safety analysis models developed by CMU Mobility Data Analytics Center, by extending to a full set of all potential causal factors and deploying the models to a number of state agencies (e.g., PA, MD, CA). Based on the models, the team will also establish a systematic approach to estimate the CMF for work zones under various roadway and work zone characteristics. In addition, an online web-based traffic safety analysis tool for selected deployment partners will be developed. Up-to-date safety data from various data providers can be acquired, archived and analyzed to enhance the web application over time. The safety data providers include State and local agencies, Police Department, Waze, and other private data sources (such as INRIX and TomTom). The team will integrate and analyze large-scale crash and incident data and developed an online tool to visualize and forecast crash types, frequencies and severity for an actual or hypothetical work zone deployments on each road segment, along with mitigating strategies for agencies' decision making. This research is based upon funded research "Mobility Data Analytics Center" in the years of 2016-2023, with the focus on data-driven safety analysis and improvement. In the past two years, we have started building a data engine and a prototype web application to demonstrate the feasibility of multi-source data-driven decision making for state DOTs. We started from the PA where we have close partnerships with many local entities and have successfully applied our data analytics tools in several case studies. Our intension in this project is update the safety analysis models for both PA and MD (possible for CA too) and develop Crash Modification Factors specifically for work zones on state owned roads, which fills the gap that CMFs were rigorously developed for roads without active construction projects. First, we will implement a rigorous causal inference model (from MAC's prior research studies) to infer the causal effect of work zones on crash risk across different work zone configurations, roadway functional classifications, weather conditions, and traffic conditions. The causal forest model avoids potential spurious heterogeneous treatment effects

(HTE) by systematically identifying the heterogeneity of treatment effects. In addition, the developed method incorporates the causal forest method with the fixed-effect variable representing road segments to mitigate the unobserved confounding bias in work zone safety studies. The proposed method will be implemented using multi-source data sets of thousands of work zones in PA and MD between 2018 and mid-2022 to control for the complex built and natural environments and reduce the bias of the estimated HTE. Second, safety analytics of work zones for the City of Pittsburgh and two selected counties in MD (Montgomery and Howard County) for developing initial Crash Modification Factors (CMFs). We will implement our rigorously developed econometrics models, Regression Discontinuity, to relate the work zone configurations to the safety risks, e.g., work zone length, buffer zone, roadway characteristics, traffic volumes, speed limits, etc. Crash Modification Factors can then be estimated from those causal models that encapsulate the impact of one factor to the work zone safety, holding everything else the same. This will be compared to the Highway Safety Manual to ensure CMFs are consistent with other road conditions and characteristics. We will build a web-GIS application to provide a user interface to visualize all work zones by crash risk, potential conflicts with traffic flow, and buffer zone, roadway characteristics, traffic volumes, speed limits, identify safety hot spots, highlight potential crash causes. In addition, suggestions on how to improve work zones safety for each work zone will be provided as part of the web-GIS application. Users can visualize the current cause of work zone related crashes and be advised with measures that can potentially reduce work zone safety risks. This will also be streamlined with WZDx, work zone data exchange, to ensure the output data format is consistent with WZDx, which can be used by all stakeholders following this national standard. Third, establishment of an online traffic safety analysis tool for PA and MD using the safety data in 2022-2023. We continue to collect and archive up-to-date safety data from various data providers in both states and enhance the web application. The safety data providers include PennDOT, MDOT SHA, Waze, and other private data sources (such as INRIX and TomTom). We will integrate and analyze large-scale crash and incident data, and developed an online tool to visualize and forecast crash types, frequencies and severity for each road segment in PA and MD. The web application allows travelers and agencies access historical, real-time, and forecasted traffic safety metrics on state owned roads. The servers hosting the web application will be optimized for load balancing. We will continue to interview various data resource providers in both states to enhance the quality and quantity of massive data, including governmental agencies, consulting firms and private data providers.

Outputs:

Modeling scripts: The scripts for work zone safety analysis and CMF calculations, with a description of different causal variables used in the model.

June 30, 2024

Will be shared with state DOT officials for establishing CMF of work zones, and safety analysis dashboard.

Final report: A technical report summarizing all data sets, innovations, technical details of proposed econometric models, solution algorithms, CMFs validated case studies, and findings.

June 30, 2024

The report will be fully edited and ready for publication in academic journals.

Outcomes/Impacts:

The expected outcome of this research is a novel framework of work zone safety models and tools that are generally applicable for any regional transportation networks. It also estimates CMFs for work zones that quantifys the safety risks of work zones and optimizes safety metrics of upcoming work zone designs. This framework will be delivered with a set of open-source codes shared online, followed by a prototype web application that implements it using multi-modal data collected over many years in the state of PA and MD. The application also provides user interfaces to manage various scenarios of work zone configurations, weather conditions and driving behaviors, and visualize the resultant system metrics for any road segments of interest. One case study will be conducted for assisting public agencies on setting guidelines for planning and operating work zones for PA. We will actively engage state DOTs, PennDOT, MDOT SHA and Caltrans, to gauge their interest and deploy those tools for their day-to-day operations.