Inferring causal effects of work zone on crashes: A case study in Pennsylvania

Graduate Student: **Zhuoran Zhang**, PhD Candidate, CEE, CMU

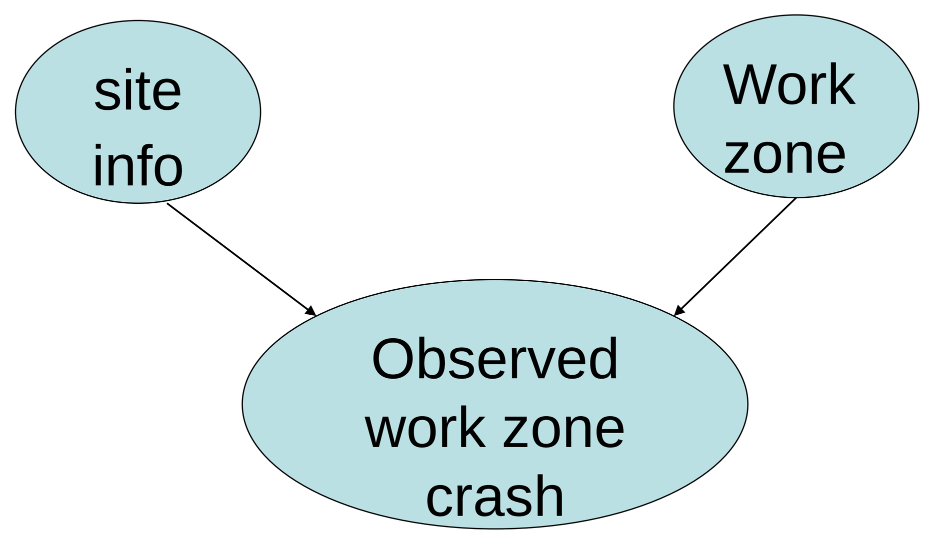
Research Advisor: **Burcu Akinci**, Prof., CEE, CMU. Sean Qian**.**, Prof., CEE, CMU.

Motivation and research objective

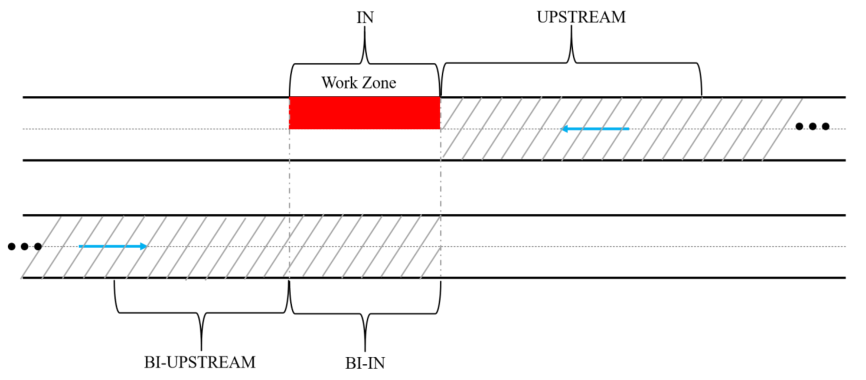
- Federal Highway Administration (FHWA) estimated more than 20 percent of U.S. highways are under construction during peak construction season, resulting in more than 3,000 work zone (FHWA, 2009). The amount of work zones performed on roads is increasing due to deteriorating roadway system and constant demand of transportation capacities (Yang et al., 2014).
- The work zone crash number is increased by 41% in 2015 compared with 2013 (FHWA, 2017). Fatal crashes in work zones increased by 3 percent while fatal crashes outside of work zones decreased by 1.5 percent from 2016 to 2017 in the U.S.(FHWA, 2019).
- The impact of work zones on crash rate is still not well understood and this impact is associated with work zone deployment strategies and traffic conditions (Ozturk et al., 2014; Tsyganov et al., 2003; Yang et al., 2015).
- To mitigate the safety issue and make informed decisions on work zone deployment strategy and traffic control policy, this research aims to inferring causal effects of different types of work zone on crashes.

Research gap in current approaches

- Missing control of confounding variables the inference of work zone effect on crash made the inference biased (Elias and Herbsman, 2000)
  - Site-specific characteristics confounded the observed crash rates in work zone. Thus analysis with only work zone crashes is not reliable. (Chen and Tarko, 2014; Yang et al., 2015b, 2013)
  - Before-after comparison makes a strong assumption that one site keeps the same crash rate in different years, which is not usually held. (Jin et al., 2008; Ozturk et al., 2014; Ullman et al., 2008)
- Control variables like traffic speed change frequently and varies per site, posted speed limit may not reflect the actual speed. (Theofilatos et al., 2019)
- Although crashes occurred in the upstream of work zones is recognized as correlated with work zone, they are usually ignored in previous analysis. (Chen and Tarko, 2014; Venugopal and Tarko, 2000)
- Small sample size (usually <100) of work zone sites made the statistical inference insignificant and hard to analyze work zone deployment strategies. (Elias and Herbsman, 2000)

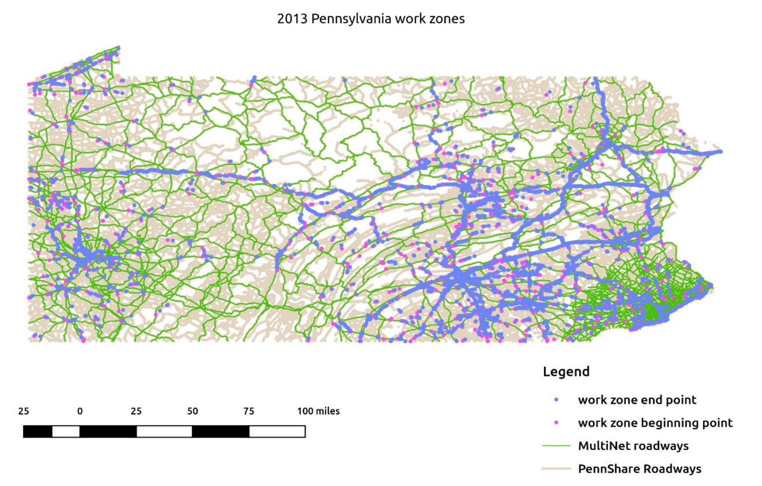


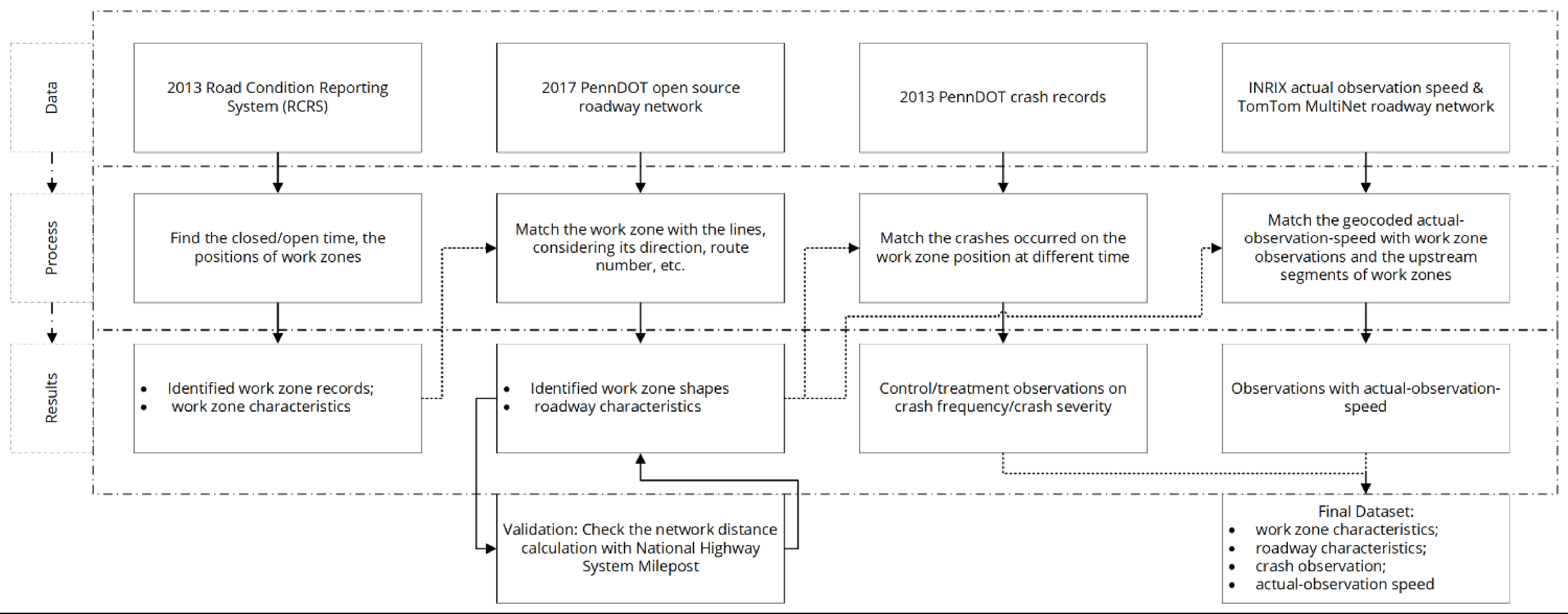
$$P(crash\,|work\ zone) = P(crash|site\ info)P(crash|do(work\ zone))$$

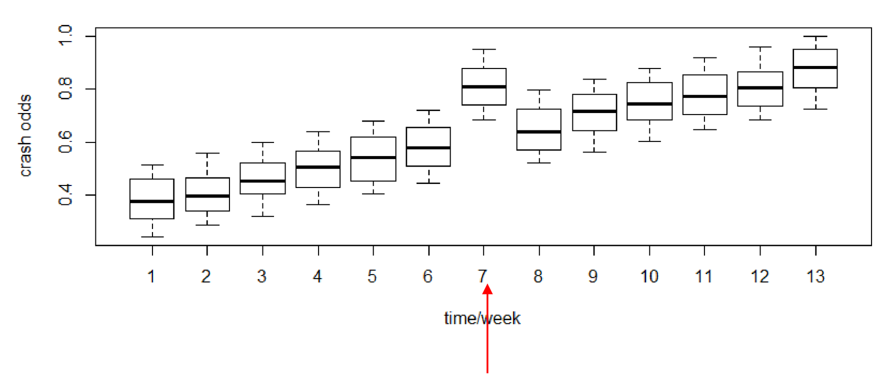


Envisioned Approach

- A regression discontinuity design is applied to elicit the causal effects as a quasi-experimental pretest-posttest design.
- Use actual-observed high-temporal-resolution speed data (5-min) to represent traffic speed before crash occurred
- Use geography analysis to identify crashes occurred in the upstream or downstream of work zone, test whether work zone caused more crash in these area.
- Utilize more than 10,000 work zone records to infer the causal effect on crash and the effectiveness of work zone deployment strategies.







Work zone occurrence

Preliminary results

- Average crash occurrence odds ratio with work zone is related to 1.723 times higher compared with crash occurrence odds ratio without work zone.
- 1% increase in work zone length will bring 26.6% increase on the log crash occurrence odds ratio
- Road segments with fewer lane counts (lane count =1) is related with higher possibility of work zone crash likelihood
- Work zone crash likelihood is increased 83.3% during daytime than nighttime
- The beginning period (first six hours) of work zone has 35% more possibility of work zone crashes than the following periods.