Transportation Research with Cloud Resources: Ridesourcing Use and Safety

Ria Kontou, PhD

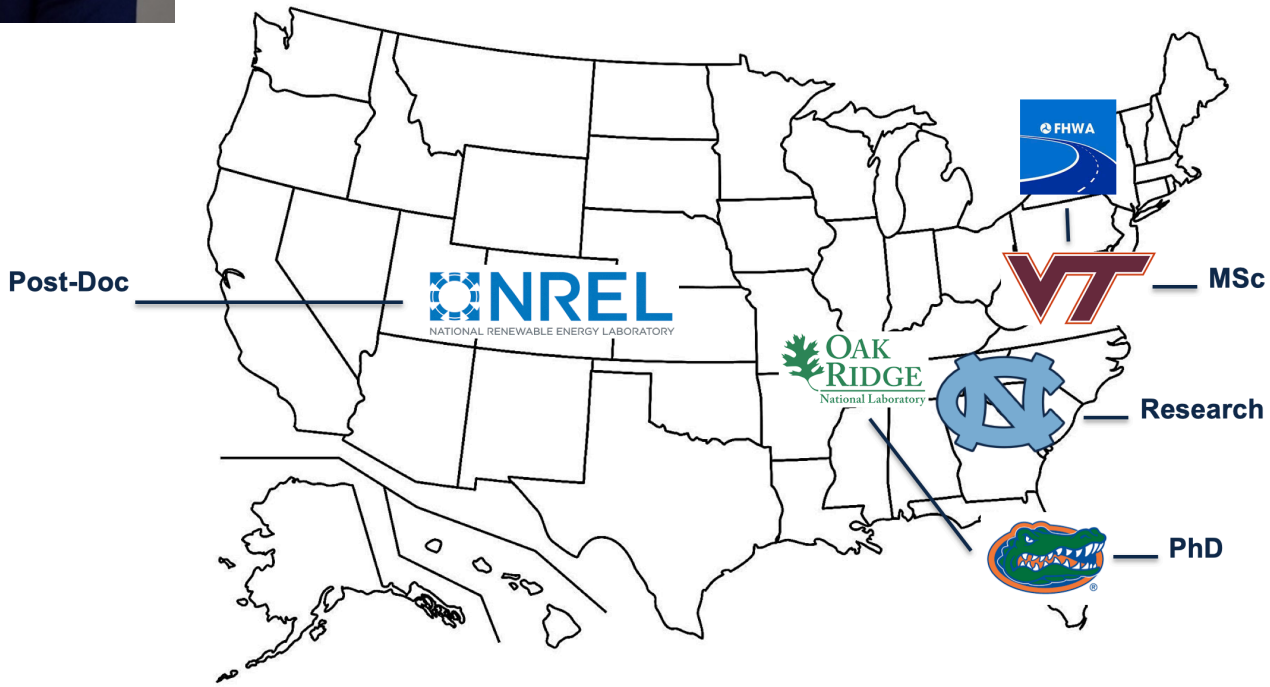
Joint work with Dr. Noreen McDonald
Department of City and Regional Planning
University of North Carolina at Chapel Hill

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

Ria Kontou: Brief Introduction

Present Dept. of City and Regional Planning UNC

Focus Transportation



Transportation Engineering Focus

GOALS

Safe and Sustainable Transportation Futures

- minimize externalities
- conserve resources
- equitable mobility

MEANS

Emerging Transportation Technologies & Services

- shared mobility
- electrification
- connectivity & automation
- multimodality

METHODS

Modeling

- optimization & heuristics
- machine learning
- simulation

Introduction: Ridesourcing & Safety

Why would ridesourcing use be associated with road crashes?

- Ridesourcing associated with reduction of alcohol-involved crashes and driving under the influence arrests

Supported by Morrison et al. 2017 &
Dills and Mulholland 2018


Miami Herald

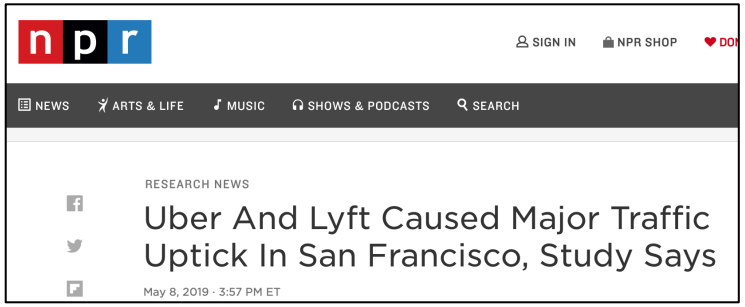
SOUTH FLORIDA

DUI arrests in South Florida plummet. Uber, Lyft, millennials among the reasons why

BY CHARLES RABIN JULY 31, 2018 01:07 PM

- Drivers cruising contributes to congested city centers and associated with increase in crashes

Supported by Barrios et al. 2018


npr

SIGN IN NPR SHOP

NEWS ARTS & LIFE MUSIC SHOWS & PODCASTS SEARCH

RESEARCH NEWS

Uber And Lyft Caused Major Traffic Uptick In San Francisco, Study Says

May 8, 2019 · 3:57 PM ET

Research Objective

Research Goal

Uncover effects of ridesourcing use on: road crashes,
injuries,
fatalities,
DWI arrests rates.

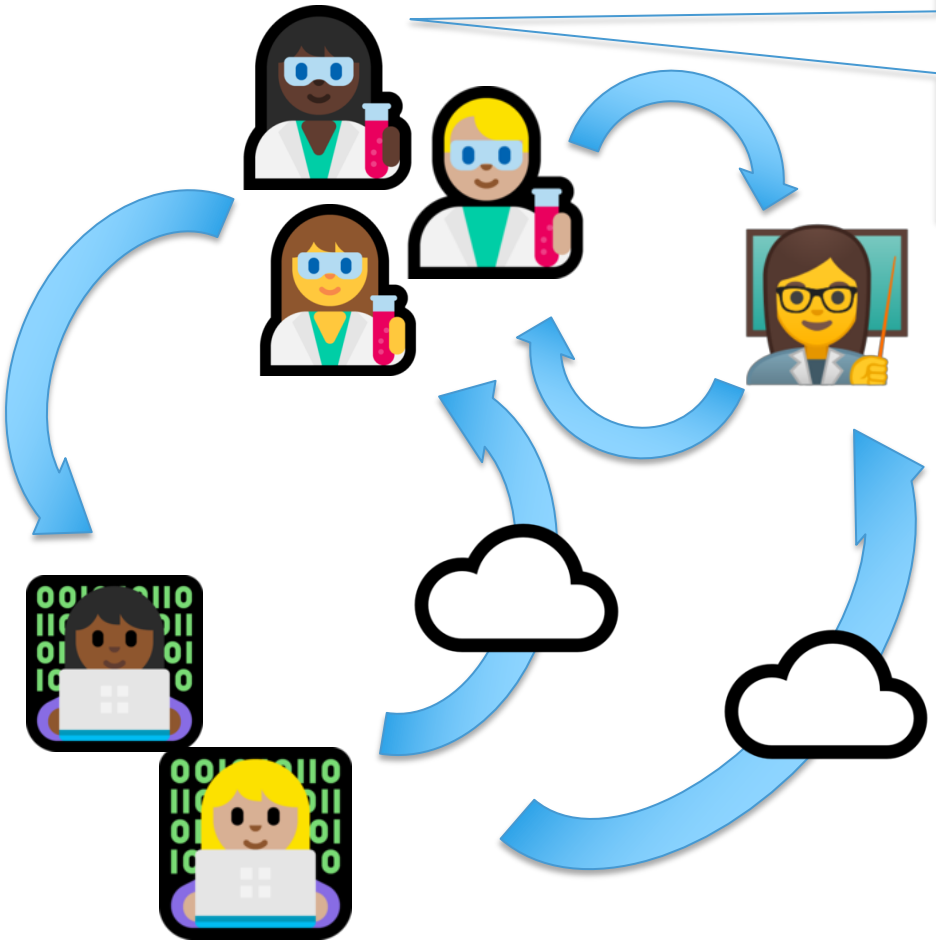
Contributions

First empirical study that uncovers such effects while accounting for intensity of ridesourcing demand using a **data-rich**, **reproducible** approach.

Integrated Cloud Computing

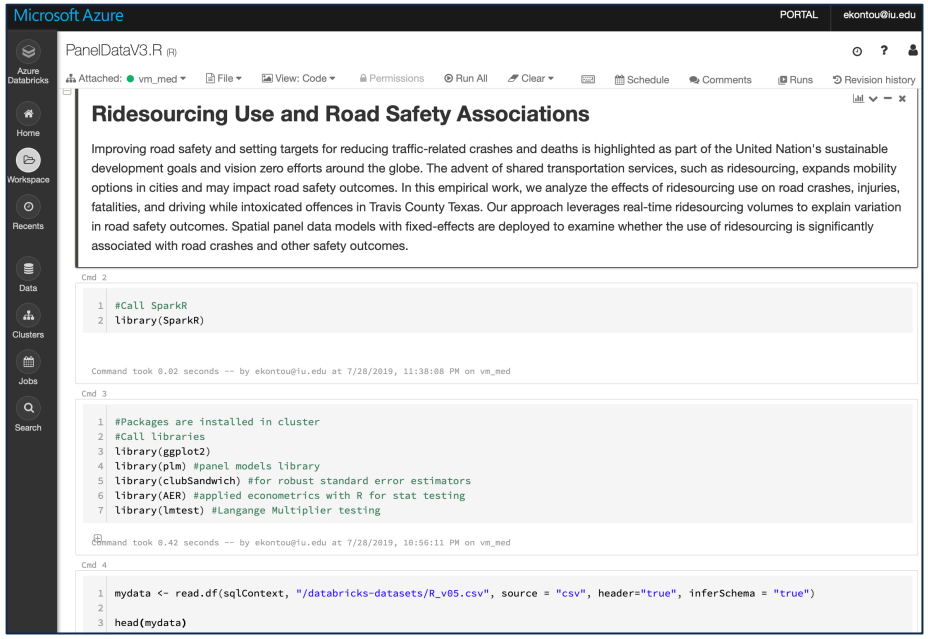
- **Reproducibility across platforms**
  - Azure Databricks
  - Jetstream
- **Collaboration tools**
  - GitHub
- **Containers**
  - Docker
  - Singularity
- **Workflows**

Research Networks & Humanware



- Computing Resources?
- Reproducibility?
- Workflow Automation?

Knowledge Gained



- Azure Databricks Notebooks and SparkR (credits)
- Microsoft Azure Machine Learning Workshop at UNC
- Indiana University PTI Community/Connections

Statistical Analysis

Modeling Efforts – Inference

- Fixed-effects OLS
- Poisson and Negative Binomial count panel data models
- Spatial lag panel data models
- Spatial error panel data models
- Spatial autoregressive model with autoregressive and heteroskedastic error terms
    - Fixed-effects and random effects testing
    - Different techniques (max log likelihood, Lasso, Monte Carlo likelihood)

Empirical Datasets (1)

- **Safety Outcome**

  Source: TXDOT Crash Records Information System
  Austin Police Department Crime Reports

- **Control for Ridesourcing Exposure**

  Ride-sourcing Sum of Origins & Destinations

  Source: RideAustin Hackathon Data

- **Control for Travel Exposure**

  StreetLight data O-D per Travis County TX Census Tract

  Source: StreetLight Data Monthly Analysis

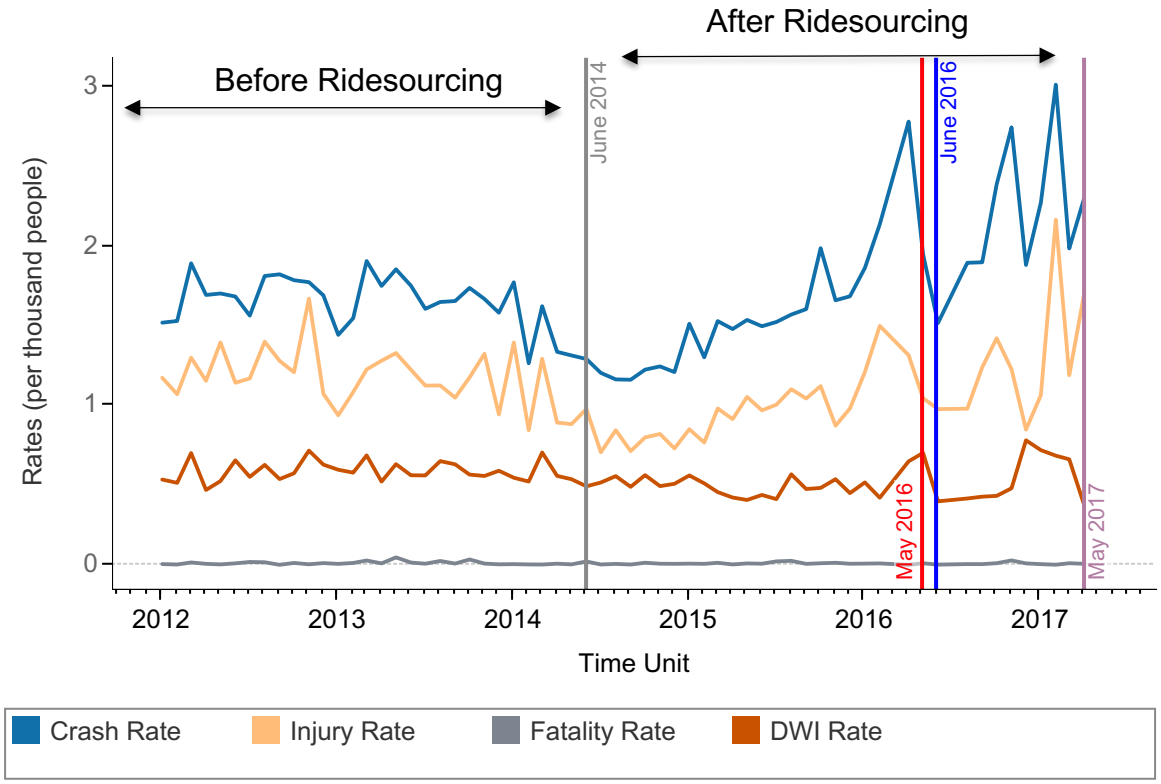
Empirical Datasets (2)

- **Additional Controls**

  census tract & year
  - Median Household Income
  - Population Density
  - Employment Density
  - Percent of Zero Vehicle Ownership

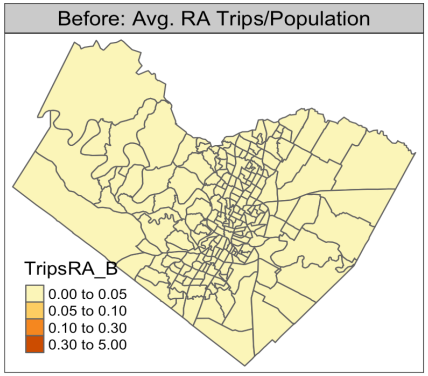
  Source: American Community Survey 2012-2017

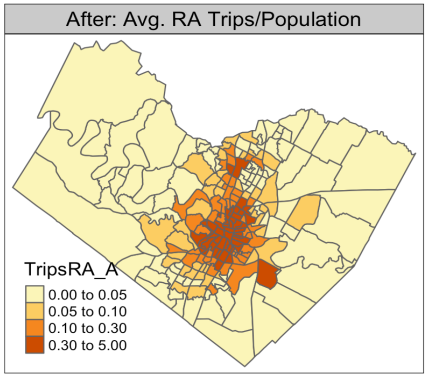
Safety Outcome Rates in Travis County TX



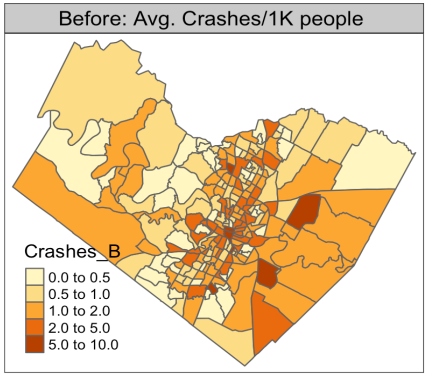
Ridesourcing Exposure in Travis County TX

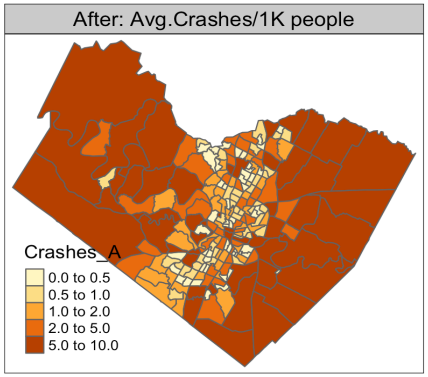
Ridesourcing Exposure





Road Crash Rates





Descriptive Statistics

	Tract-Month-Year Unit without Ridesourcing Mean (Std. Deviation)	Tract-Month-Year Unit with Ridesourcing Mean (Std. Deviation)
Crashes (per 1,000 people)	1.21	1.32
DWI Arrests (per 1,000 people)	0.45	0.36
Fatalities (per 1,000 people)	0.008	0.0091
Ridesourcing Rates (per 1,000 people)	0	0.23
Median Household Income ($)	63,744 (32930)	71,068 (34002)
OD Trips (StreetLight Data)	28,950 (25,160)	31,384 (23,827)
Gas Price ($/gallon)	3.36 (0.19)	2.05 (0.08)
Population	4,847 (2,620)	5,314 (3,029)
Percent of Zero Vehicle Ownership	0.03 (0.04)	0.03 (0.03)
Percent of employment	0.72 (0.12)	0.72 (0.11)
Records	72.50%	27.50%

Preliminary Findings

Results of indicative fixed-effects panel data models

	Crashes	Injuries	DWI Arrests	Fatalities
γ coefficient	-0.0133	-0.0208	-0.0347	-0.00117
P-value	0.229 (n.s.)	0.0544 (<0.1)	<0.001	0.636 (n.s.)
Adjusted std. errors for clustering at the census tract level	0.011	0.0106	0.00804	0.00245
Census tract and month-year fixed effects	Yes	Yes	Yes	Yes
Controls	No	No	No	No
Observations	8720	8720	8720	8720
γ coefficient	-0.013048	-0.023875	-0.033447	-0.00148
P-value	0.2325 (n.s.)	0.0284 (<0.05)	<0.001	0.55 (n.s.)
Adjusted std. errors for clustering at the census tract level	0.0112	0.0106	0.00818	0.00246
Census tract and month-year fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	8720	8720	8720	8720

Preliminary Findings Discussion

- RideAustin use 1% increase is found associated with a 0.3% and 0.2% decrease in DWI arrests and road injuries, respectively.
- No statistically significant association at conventional levels between RideAustin use and crashes or fatalities.
- Potential benefits of implementation of ridesourcing, saving lives and reducing court costs, insurance rate increases, loss of income due to DWI arrests reduction.

Future Directions

- Goals
  - Reproduce the analysis with 2 additional geographic information systems datasets with cloud resources
  - Develop similar workflows for our other projects

- Future work
  - Image processing for urban morphology
  - Leverage GPUs

Thank you and
Thank you Brian Voss, Craig Stewart and IU PTI!
Questions?
ekontou@email.unc.edu