



## Insights from big and small data

Which trips and travelers are captured by location-based services data?

**Innovations in Travel Modeling**

June 27, 2018

Leah Flake, RSG  
Vince Bernardin, RSG  
Elizabeth Greene, RSG  
Rebekah Anderson, ODOT  
Jeffrey Dumont, RSG

# Case Study Background



**Evaluated similarities & differences between smartphone data collected for a household travel survey using RSG’s rMove™ app and location-based services (LBS) data collected passively from smartphone apps for FHWA’s TMIP**

- Data characteristics for the same day of data between LBS and rMove:

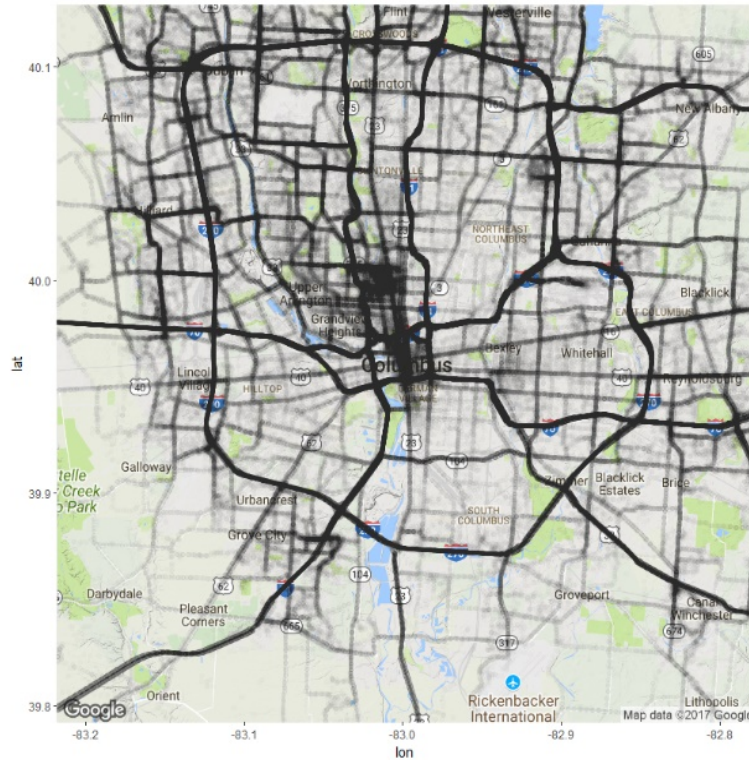
	LBS	rMove
Total devices	95,697	222
Points	12,128,310	124,681
Median time between points	147 seconds (2.5 minutes)	4 seconds (0.06 minutes)
Standard deviation between points	2.2 hours (132 minutes)	0.72 hours (43 minutes)

- High variance in LBS data collection among devices (many with high point frequency, many with sparse point frequency)

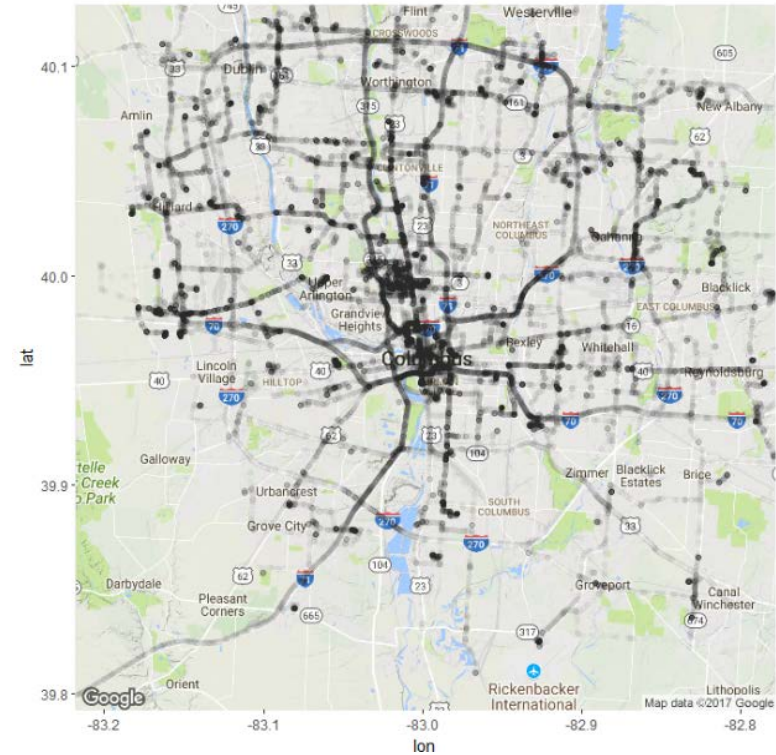


# LBS vs. rMove Spatial Coverage

**DURING THE SAME DAY:**



LBS locations



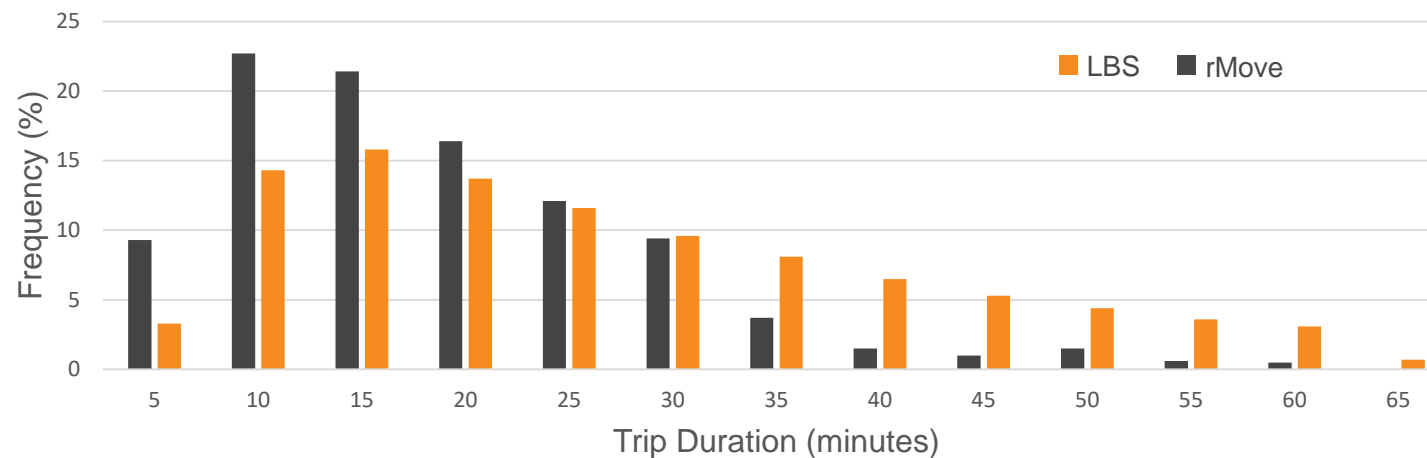
rMove locations



# Trip Inference in LBS Data

- To understand how comprehensively travel patterns are collected from the LBS source, we developed an algorithm to infer trips from LBS data
- Resulting trip characteristics compared to rMove trips:

	LBS	rMove
<b>Number of Trips</b>	378,953	1,187
<b>Median Distance (straight line)</b>	1.56 mi	1.64 mi
<b>Median Travel Time</b>	34 minutes	12 minutes
<b>Number of Trips per Device</b>	4.73	5.34

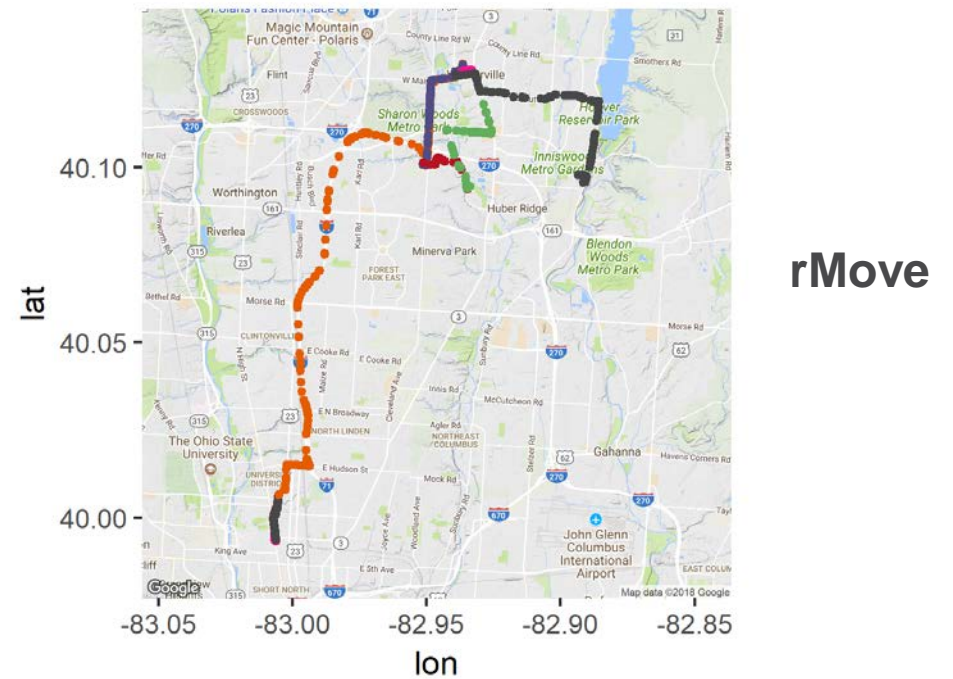


# Matching LBS and rMove users

- Identified common users between rMove and the LBS source to have a better idea of LBS data collection gaps & to help adjust the trip inference algorithm
- Resulted in 26 likely matches (of 222 devices) between rMove and LBS for given day

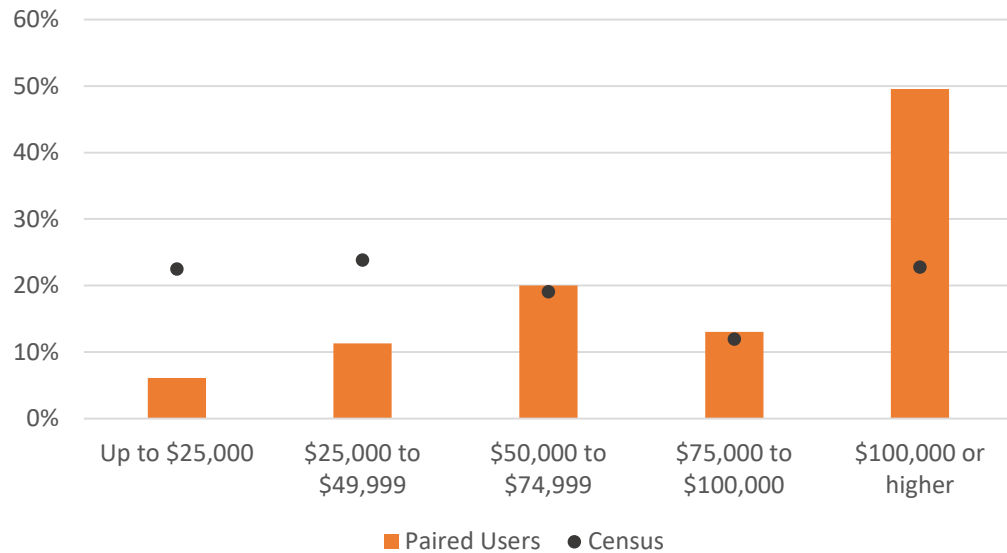


10:48  
15:32  
15:57  
17:36  
18:04  
18:07  
19:24  
19:27

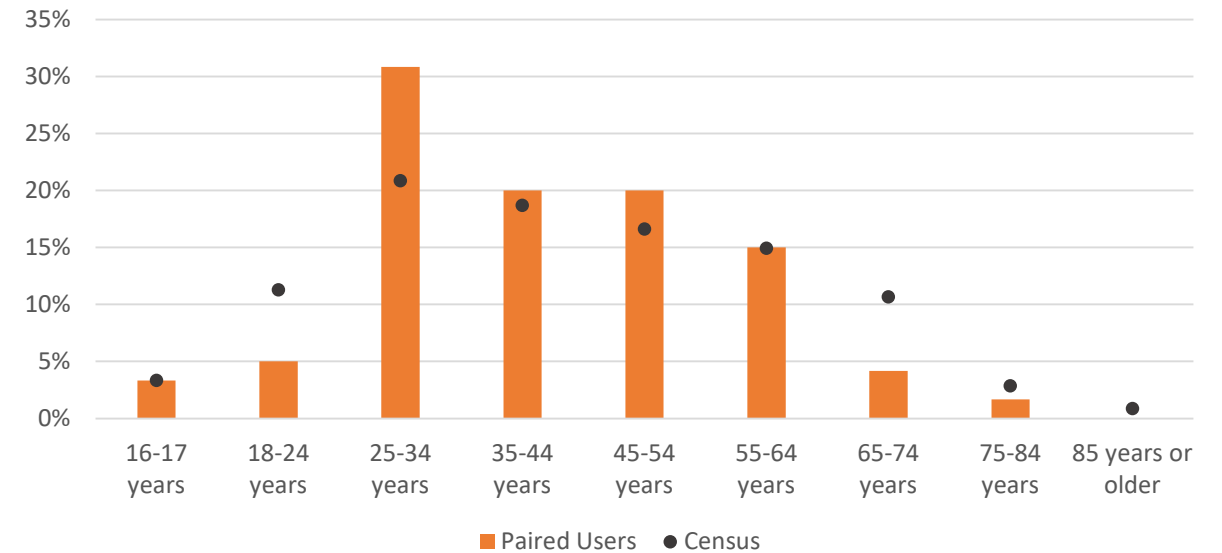


# Demographics of LBS Users Matched to rMove Survey

## INCOME



## AGE



- The subset of users in the LBS dataset is younger than in the Census
- LBS subset has relatively fewer people in the low-income ranges (less than \$50k) compared to Census



# Complementarity of Passive LBS and Travel Surveys

Passive LBS and rMove data together allow us to understand travel better than either alone

## PASSIVE LBS

- **Spatial coverage** from LBS data is complete
  - to a degree that is impossible to achieve with survey data
- **Device matching** with targeted spatial data is feasible
- **User persistence** appears to be very good
- **Longitudinal** observation may be possible (e.g., for trend monitoring)

## TRAVEL SURVEYS (rMove)

- **Temporal coverage** of trip data is complete
  - Identifies missing short trips in passive LBS data
- **Traveler characteristics** (and mode and purpose) are observed
  - Identifies LBS bias towards young adults, more affluent

- If we can measure bias, we can correct for it -

Potential for blended datasets with representativeness of surveys and volume of passive data





## Contacts

[www.rsginc.com](http://www.rsginc.com)

**Leah Flake**  
Consultant  
[flake.leah@rsginc.com](mailto:flake.leah@rsginc.com)

**Vince Bernardin**  
Director

**Elizabeth Greene**  
Director