Evaluating Spatial Pricing in Ride-Sourcing Systems

Natalia Zuniga-Garcia

Mauricio Tec
James G. Scott
Natalia Ruiz-Juri
Randy B. Machemehl

The University of Texas at Austin

INFORMS Annual Meeting
Phoenix, AZ

November 5^{th}, 2018

Outline

Introduction

Ride-sourcing companies or TNCs, provide pre-arranged or on-demand transportation service for compensation.

Background

From **drivers** perspective, trips may be mispriced relative to other trip opportunities, leading to inefficiencies on a network level:

- Loss of service reliability
- Limit long-term driver participation

Recent research efforts have addressed ride-sourcings spatial mispricing problem with the objective of reducing *search frictions*[1] using:

- Spatial surge pricing models
- Spatio-temporal pricing mechanisms
- Search and matching models
- Non-linear pricing models

[1] Imbalance between driver supply and passenger demand across geographic areas that causes the presence of high matching and reaching times.

Introduction

Motivation

- Methods focused on the optimization of the platform revenue and do not evaluate the driver perspective
- Limited evidence on the driver opportunity cost of the trip destination
- Lack of understanding of the spatial structure of driver productivity[2]
- Limited empirical evaluations

Objective

Analyze the spatial structure of ride-sourcing operational and driver performance variables to support the need for new pricing strategies.

We explore the spatial structure of ride-sourcing search frictions and driver performance variables as a function of the trip destination.

[2] We define the driver productivity in terms of profit per unit time.

Introduction

Analyzing operational and performance variables at a high-definition spatial level requires additional data analytics methods. We propose the use of a spatial smoothing or denoising technique that:

- Allows fine resolution analysis
- Compensates for inherent sampling noise
- Enhances interpretability

Contributions

1. Empirical evidence of spatial and temporal variation of driver productivity variables as a function of trip destination.
2. Temporal and spatial evaluation of different ride-sourcing operational measures and search frictions in Austin.
3. Implementation of a spatial denoising methodology to analyze high-definition spatial variables.

Methodology

Ride-Sourcing Data

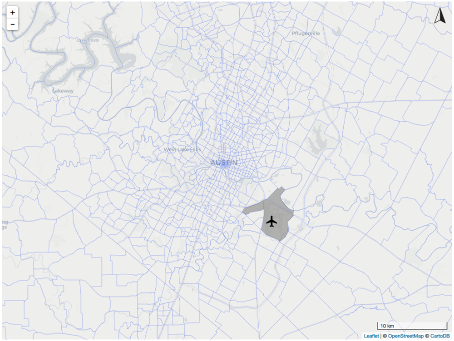
Austin-based TNC (Ride Austin) trips[3] during the period that Uber and Lyft were out of the city.

- Space discretization
  - Data is summarized over 1,305 traffic analysis zones (TAZs)
  - TAZ areas vary from 0.01 km^2 in the Central Business District (CBD) to 30 km^2 in the rural area, with an average of 2 km^2
- Time discretization
  - Weekday AM-peak
  - Weekday PM-peak
  - Weekday off-peak
  - Weekend

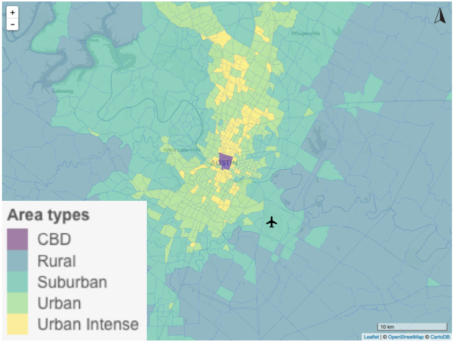
[3] https://data.world/ride-austin

Methodology

Ride-Sourcing Data



(a) TAZs in Austin (airport shaded)



(b) TAZ area types

Figure: Description of TAZs

Methodology

Ride-Sourcing Data

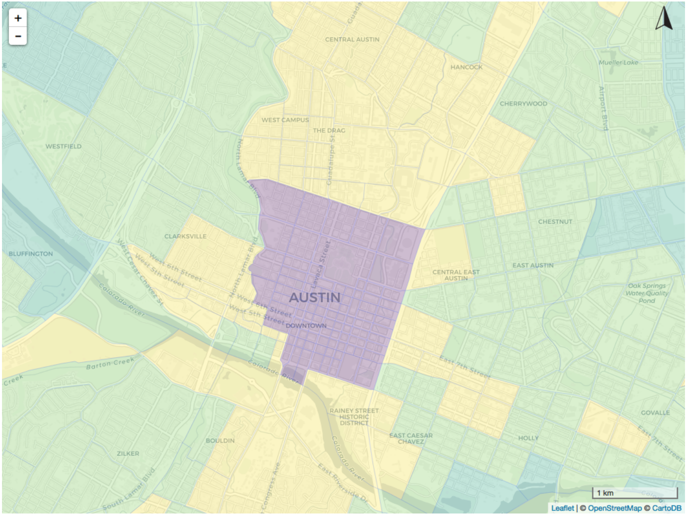


Figure: TAZ area types (downtown)

Methodology

Description of variables

- Operational (based on trip origin)

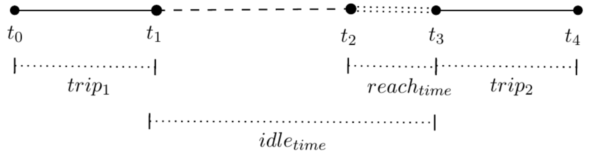


Figure: Driver time diagram

- Productivity, CBD-origin trips only (based on trip destination)

$$Productivity\ \mathbf{A} = \frac{fare_{trip_1}}{t_1} \tag{1}$$

$$Productivity\ \mathbf{B} = \frac{fare_{trip_1}}{t_3} \tag{2}$$

$$Productivity\ \mathbf{C} = \frac{fare_{trip_1} + fare_{trip_2}}{t_4} \tag{3}$$

Spatial Smoothing Approach

Background

- Typically used for a wide range of applications:
  - Predicting crime hotspots
  - Detecting crash hotspots
  - Special event detection
- Approaches types:
  - **Local**
    - Smooth only a local window around each point.
    - *Gaussian* smoothing, average a point over its neighboring values, thus removing noise by blurring.
  - **Global**
    - Define an objective function over the entire graph and simultaneously optimize the whole set of points.
    - *Total variation denoising or fused lasso*, removes noise by emphasizing edges.

Spatial Smoothing Approach

Smoothing

Assume that we have observations y_i, each associated with a vertex $s_i \in \mathcal{V}$ in an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with node set $\mathcal{V}$ and edge set $\mathcal{E}$.

$$y_i = x_i + \epsilon_i \,, \quad i = 1, \ldots, n, \tag{4}$$

where, x_i is the "true" denoised signal and ϵ_i is mean-zero error.

The goal of the smoothing techniques is to estimate x_i in a way that leverages the assumption of spatial smoothness over the underlying graph.

Spatial Smoothing Approach

Graph-Fused Lasso

One way to estimate x is by using the GFL, defined by a convex optimization problem that penalizes the first differences of the signal across edges.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \ell(\mathbf{y}, \mathbf{x}) + \lambda \sum_{(r,s) \in \mathcal{E}} |x_r - x_s|, \tag{5}$$

where, ℓ is the loss function, r is the start node and s the end node, $n = |\mathcal{V}|$, and $\lambda > 0$ is the regularization parameter.

- Equation 5 does not have a closed-form solution. Convex optimization approaches are required.
- We implemented the method developed by Tansey and Scott (2015)[4], which leads to an efficient approach that presents a fast and scalable solution.

[4] Tansey, W., & Scott, J. G. (2015). A fast and flexible algorithm for the graph-fused lasso. arXiv preprint arXiv:1505.06475.

Spatial Smoothing Approach

Loss Function

Penalized weighted least squared-error loss function to take into account the differences in the number of observations within each zone.

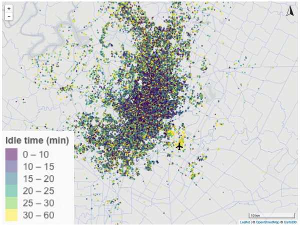
$$\underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad \sum_{i=1}^{n} \frac{\eta_i}{2}(y_i - x_i)^2 + \lambda \sum_{(r,s)\in\mathcal{E}} |x_r - x_s| \tag{6}$$

Where, η_i is the count of trips observed within the i-th TAZ.

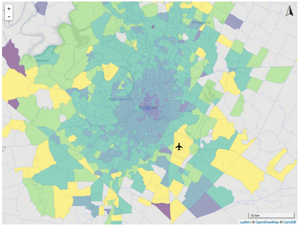
Choosing the Regularization Parameter

- Split the data into a training and a test set
- Estimate the out-of-sample prediction error using the root mean square error (RMSE) criterion

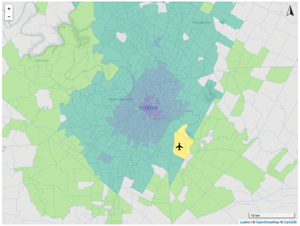
Spatial Smoothing Approach



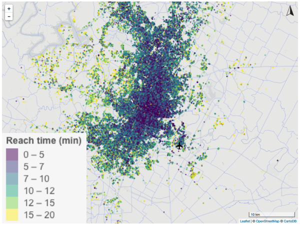
(a) Idle time data



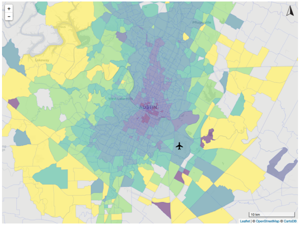
(b) Non-smoothed



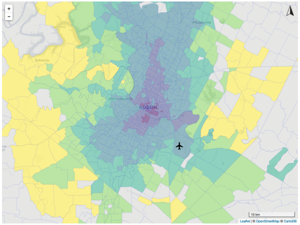
(c) Smoothed



(d) Reach time data



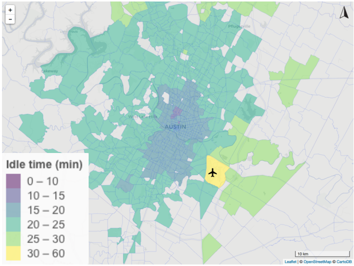
(e) Non-smoothed



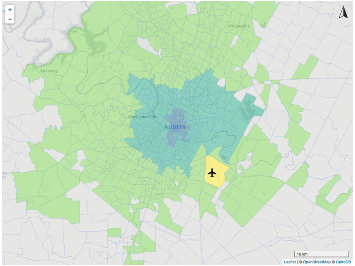
(f) Smoothed

Figure: GFL denoising example

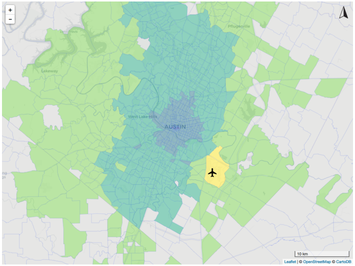
Results - Operational Variables



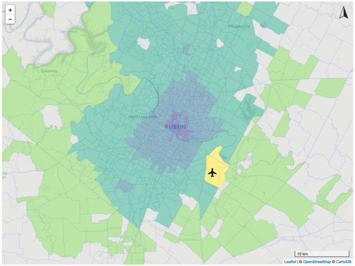
(a) Idle time AM-peak



(b) Idle time PM-peak

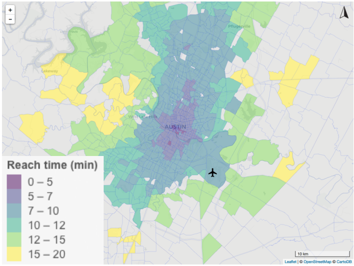


(c) Idle time off-peak

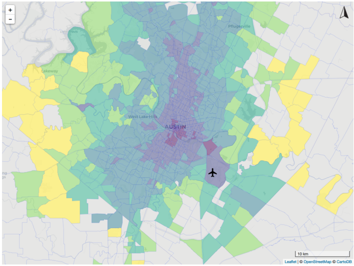


(d) Idle time weekend

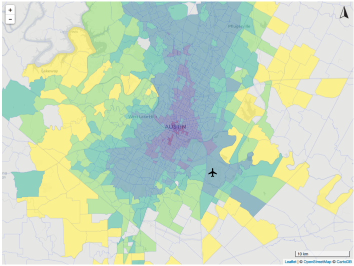
Results - Operational Variables



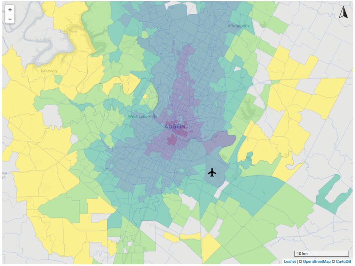
(a) Reach time AM-peak



(b) Reach time PM-peak

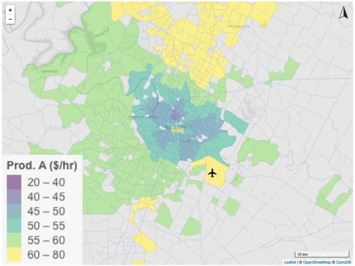


(c) Reach time off-peak

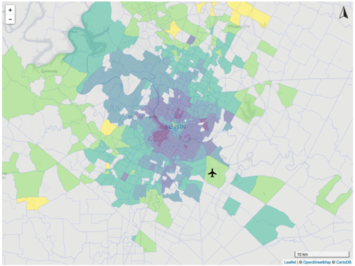


(d) Reach time weekend

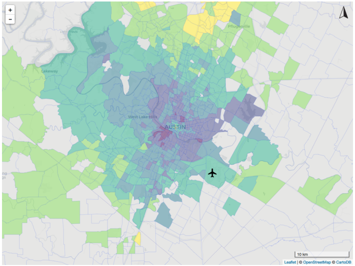
Results - Productivity Variables



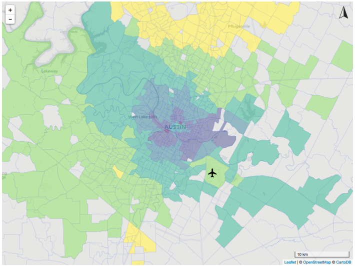
(a) Prod. A AM-peak



(b) Prod. A PM-peak

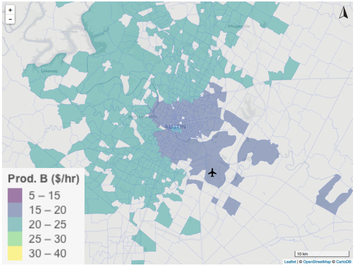


(c) Prod. A off-peak

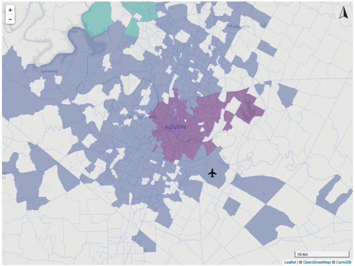


(d) Prod. A weekend

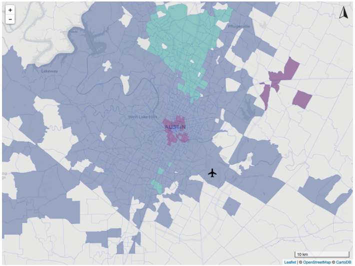
Results - Productivity Variables



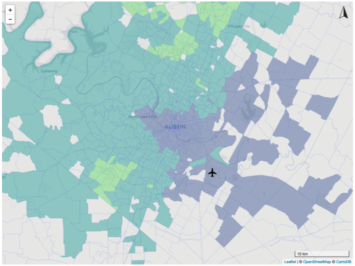
(a) Prod. B AM-peak



(b) Prod. B PM-peak

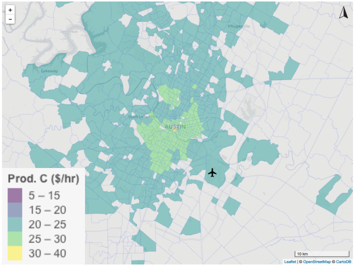


(c) Prod. B off-peak

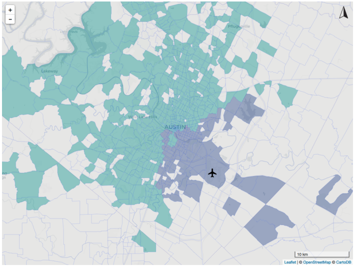


(d) Prod. B weekend

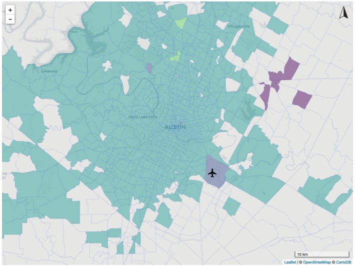
Results - Productivity Variables



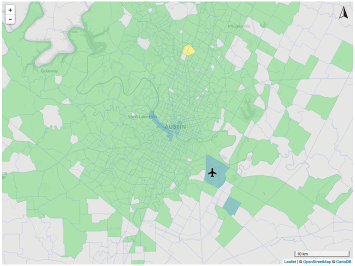
(a) Prod. C AM-peak



(b) Prod. C PM-peak



(c) Prod. C off-peak



(d) Prod. C weekend

Conclusions

Primary findings of this research suggest that there are differences in space and time that can affect ride-sourcing search frictions and driver productivity. Providing spatio-temporal pricing strategies could be one way to balance driver equity across the network.

- **Driver and operator point of view**
  More efficient driver supply method.
- **Planners and engineers perspective**
  Understand the characteristics of of the ride-sourcing service in Austin.
- **Pricing strategies and policies**
  Warranty fair conditions in driver compensation.

Thanks!

Questions or Comments?
nzuniga@utexas.edu