The Struggle of Taxi Industry in the Age of Sharing Economy

Unveiling the optimal driving pattern and operation strategy for cabdrivers in post-2013 Chicago area

Dongping Zhang

M.A. in Computational Social Science
The University of Chicago

April 5, 2017

The Taxi-Uber War

Table of Contents

1. Background
   - Why post-2013?
2. Research Question
3. Motivation
   - Why is this question interesting?
   - Some Facts in 2015
   - What is the contribution?
4. Data
   - Source of Data
   - What are my variables of interest?
5. Model
   - Types of Analyses
   - Computational Tools
6. Summary

Background

Why post-2013?

- Uber launched in Chicago on September 22, 2011
- Lyft launched in Chicago on May 11, 2013

Table of Contents

1. Background
   - Why post-2013?
2. Research Question
3. Motivation
   - Why is this question interesting?
   - Some Facts in 2015
   - What is the contribution?
4. Data
   - Source of Data
   - What are my variables of interest?
5. Model
   - Types of Analyses
   - Computational Tools
6. Summary

Research Question

What am I studying?

What does spatially and temporally quantified daily digital traces of the most productive cabdrivers imply the optimal operation pattern and driving strategy in post-2013 Chicago area?

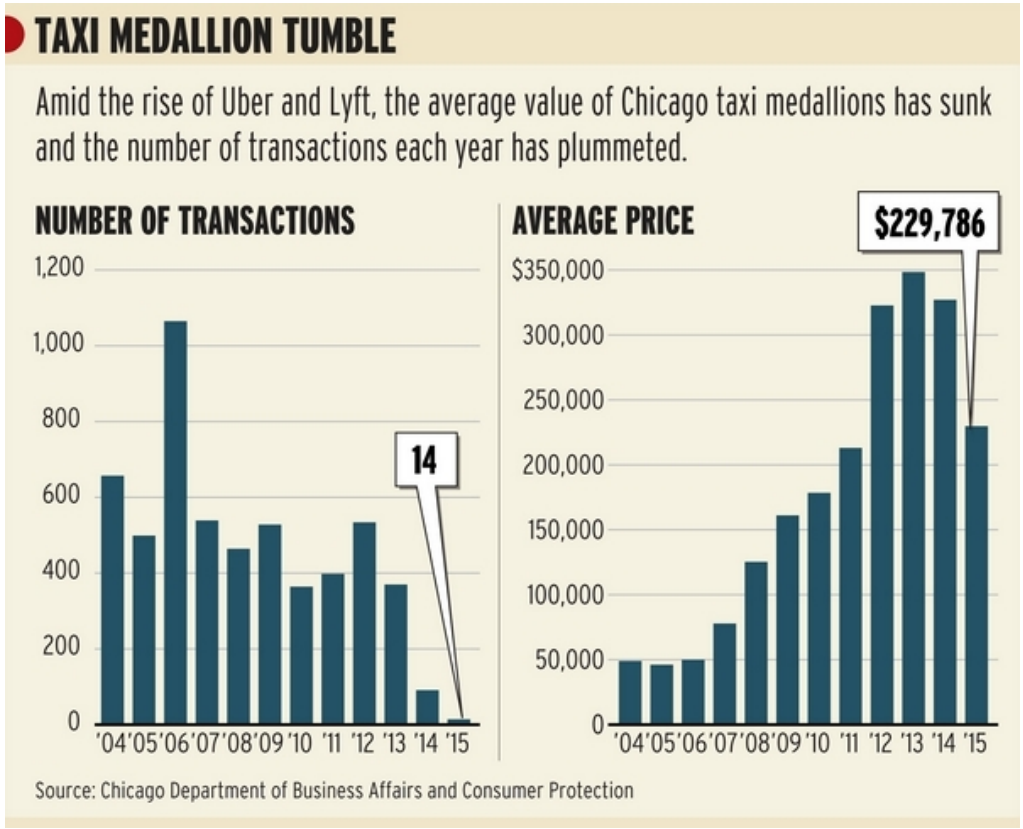
Lit Review

- Liu, L., Andris, C., & Ratti, C. (2010). Uncovering cabdrivers' behavior patterns from their digital traces. *Computers, Environment and Urban Systems, 34*(6), 541-548.

Table of Contents

1. Background
   - Why post-2013?
2. Research Question
3. Motivation
   - Why is this question interesting?
   - Some Facts in 2015
   - What is the contribution?
4. Data
   - Source of Data
   - What are my variables of interest?
5. Model
   - Types of Analyses
   - Computational Tools
6. Summary

Taxi Medallion Sales Trend



Motivation

Taxi Medallion Price

- Taxi Medallion and related assets were worth $2.5 billion in Chicago and is regarded as one of the "cash-cows"
- Average price for a medallion in Chicago was less than $230,000 in 2015
- 30% drop of medallion sales price

Some Facts in 2015

Number of Drivers

- 156,661 active drivers in ride-hailing industry in Chicago
- 12,955 active taxi chauffeur licenses in Chicago

Number of Taxi Trips

Annual Taxi trips plummeted by 35% since 2014

- 2014: 31,013,591 trips
- 2015: 27,395,382 trips
- 2016: 19,874,714 trips

Motivation

What is the contribution?

- A systematic study of large scale cabdrivers' behaviors in a real and complex city context
- Recognizing high-level human behaviors and decision-makings from their daily digital traces
- The invasion of Uber, Lyft, and other ride-hailing service companies is gradually driving the traditional taxi industry out of business
- Policy implication to alleviate the burden of the traditional taxi industry
- Aiming to uncover some new insight of urban transportation and human mobility

Table of Contents

1 Background
- Why post-2013?

2 Research Question

3 Motivation
- Why is this question interesting?
- Some Facts in 2015
- What is the contribution?

4 Data
- Source of Data
- What are my variables of interest?

5 Model
- Types of Analyses
- Computational Tools

6 Summary

Data

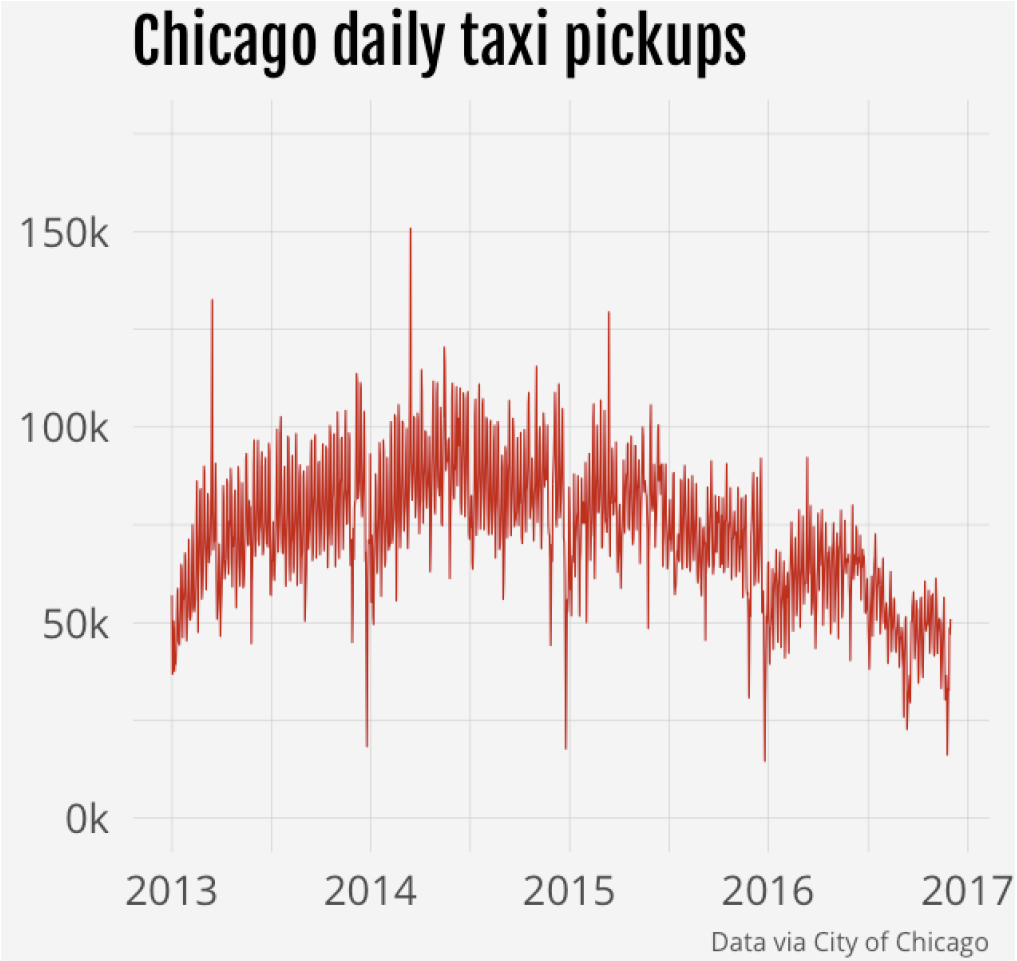
Where did I acquired the data?

Chicago Open Data Portal powered by Socrata

What are my variables of interest that are also included in the raw dataset?

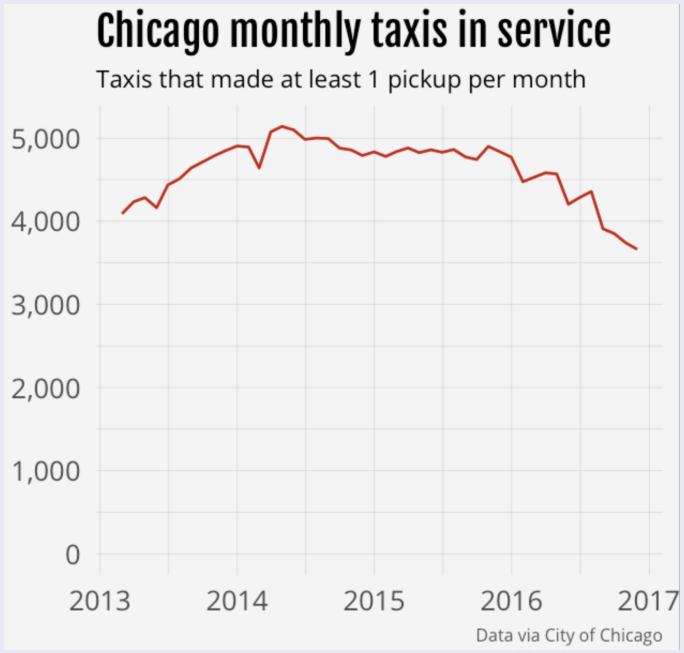
- Trip ID
- Taxi ID
- Trip Start Timestamp
- Trip End Timestamp
- Trip Miles
- Fare, Tips, Tolls, Extras
- Pickup Coordinates
- Dropoff Coordinates

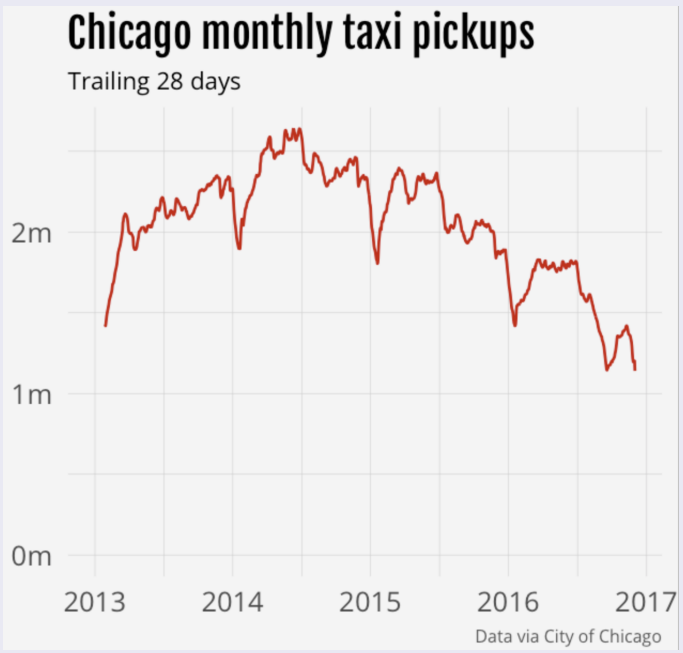
Exploratory Data Analysis 1



Exploratory Data Analysis 2

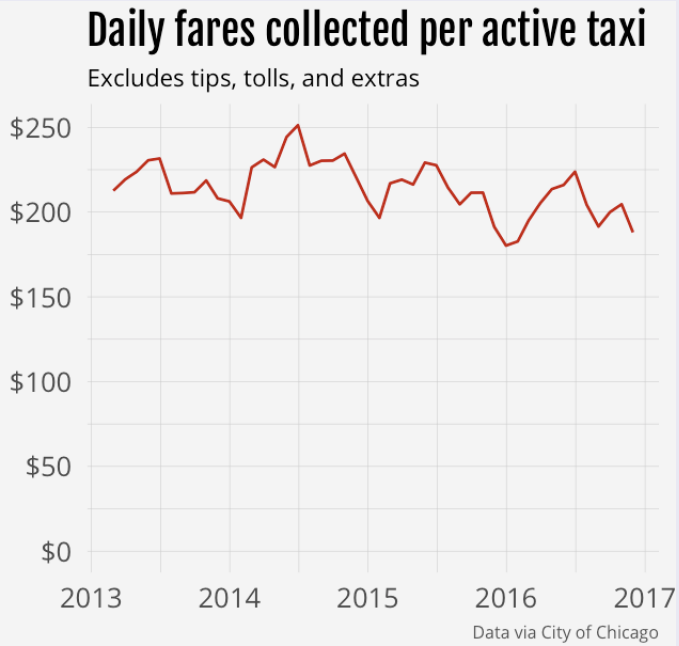
Monthly Taxis in Service & Monthly Taxi Pickups





Exploratory Data Analysis 3

Taxis Daily Fares & Trips per Day



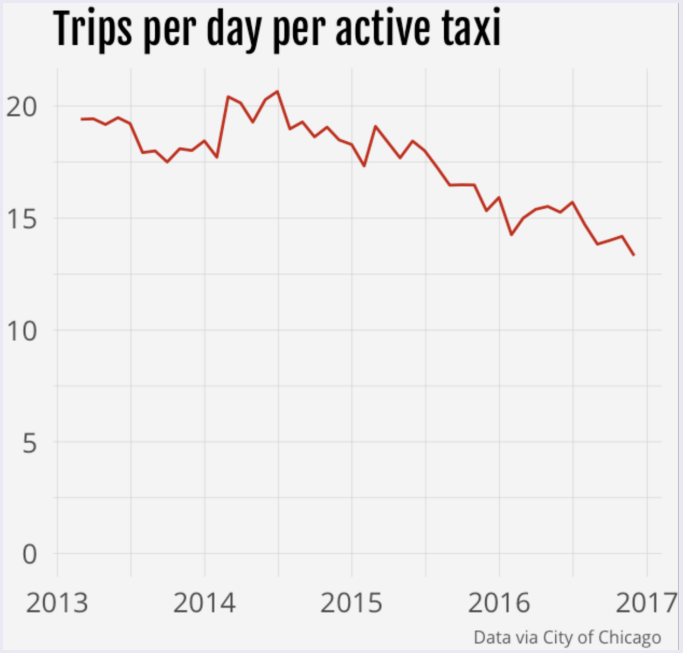


Table of Contents

1. Background
   - Why post-2013?
2. Research Question
3. Motivation
   - Why is this question interesting?
   - Some Facts in 2015
   - What is the contribution?
4. Data
   - Source of Data
   - What are my variables of interest?
5. Model
   - Types of Analyses
   - Computational Tools
6. Summary

Types of Analyses

What analyses would I conduct?

- ESDA (points vs. aggregating pickups and drop-offs by areal unit)
- Spatial Regression Analysis to capture spatial dependency
- 3-Dimensional Flow Analysis
  - pickup points, direction, drop-off points
- Ratio of real path length over shortest path length
- Ratio of real path travel time over shortest path travel time

Computational Tools

What computational tools will I use?

- SQL – RDBMS
- Python
- R
- GeoDa – ESDA
- GeoDaSpace – SRA

Table of Contents

1. Background
   - Why post-2013?
2. Research Question
3. Motivation
   - Why is this question interesting?
   - Some Facts in 2015
   - What is the contribution?
4. Data
   - Source of Data
   - What are my variables of interest?
5. Model
   - Types of Analyses
   - Computational Tools
6. Summary

Summary

- To identify a subset of most productive cabdrivers who have consistently higher daily income from 78,283,687 trip history
- To spatially and temporally quantify the daily digital traces of those most productive cabdrivers
- To uncover potential driving patterns and operation strategies of those most productive cabdrivers

Reference

- Liu, L., Andris, C., & Ratti, C. (2010). Uncovering cabdrivers' behavior patterns from their digital traces. *Computers, Environment and Urban Systems, 34*(6), 541-548.
- Huang, H., Zhang, D., Zhu, Y., Li, M., & Wu, M. Y. (2012). A Metropolitan Taxi Mobility Model from Real GPS Traces. *J. UCS, 18*(9), 1072-1092.
- Li, B., Zhang, D., Sun, L., Chen, C., Li, S., Qi, G., & Yang, Q. (2011, March). Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on* (pp. 63-68). IEEE.