# Search Frictions, Network Effects and Spatial Competition: Taxi versus Uber

Bo Bian

Department of Economics, Penn State

April 15, 2018

# Introduction

- Taxi industry has matching friction due to fixed pricing structure (Buchholz (2016))
  - Locations have heterogeneous profitability.
  - Drivers oversupply locations with high conditional expected profit (on being matched).
  - In equilibrium, excess supply/demand coexist across locations.
- Uber has more flexible prices than taxis
  - If demand $>$ supply of a location, Uber applies surge pricing to motivate drivers' supply.
  - Uber is more efficient to match drivers and riders within a location.
- Traffic problems and congestion in Manhattan.
- Rapid growth of ridesharing industry v.s. regulatory policy.

# Introduction

Research Questions:

- How is the performance of Uber's flexible pricing on matching efficiency?
- To what extent, worse traffic condition affects matching efficiency?
- Does competition between Uber and taxis improve or reduce matching between drivers and riders?
- Policy: What's the effects of NYC government's proposal to cap Uber vehicles in 2015? Competition $\downarrow$, matching $?$
- Is there other non-price factor that affects matching? Network effects?

# Introduction

Network Effects

- Define a location-time pair as a market.
- Consider a firm-market pair as a two-sided platform.
  - Drivers of fixed firm choose platforms (among markets) to search.
  - Riders of fixed market choose platforms (among firms) to hail a ride.
- Do (direct/indirect) network effects exist (i.e. waiting time, matching probability)?
- What is the impact of network effects on matching efficiency? (better or worse)

# Introduction

To answer the questions:

- I model Uber and taxi drivers' dynamic spatial search decisions.
- I model riders' static discrete choice demand.
- To incorporate network effects:
  - Add matching probability of supply side to drivers' search decisions.
  - Add demand and supply into mean utility as proxy for waiting time, and matching probability of demand side.
- Assume explicit matching within market: random matching for taxis and perfect matching for Uber.
- Simulate three counterfactuals:
  - Eliminate Uber's surge multiplier
  - Improve traffic conditions.
  - Restrict the number of Uber vehicles.
  - Compare results with/without network effects.

# Data

- NYC Taxi&Limousine Commission Taxi trips records:
  - Taxis: pickup/dropoff datetime, location, trip distance, time, fare.
  - Uber: pickup date time, rough location.
- UBER API: real-time surge multipliers are collected from 79 location spots every 10 minutes from Nov 2015-June 2016.
- NYC Metropolitan Transportation Authority: weekly aggregate MetroCard swipes of subway stations.
- **Limitation:**
  - Supply and demand are not directly observed.
  - Dropoffs of Uber are not available.

# Data

- Sample Construction:
  1. Pickups of taxis/Uber in a representative weekday from 6am-4pm in April 2016(taxi 173,000 trips and Uber 49,000 trips).
  2. Define 10 minutes interval as a period (totally 60).
  3. Choose 40 geographic locations of NYC (2400 markets).
  4. Calculate average trip time and distance between markets using taxi pickup/dropoffs.
  5. Calculate prices using trip time, distance and surge multiplier.
  6. Calculate subway ridership paying full fare as outside option.
  7. Calculate distribution of taxis' dropoffs in 2010 and 2016. Use 2010's dropoffs as travelling patterns of population.

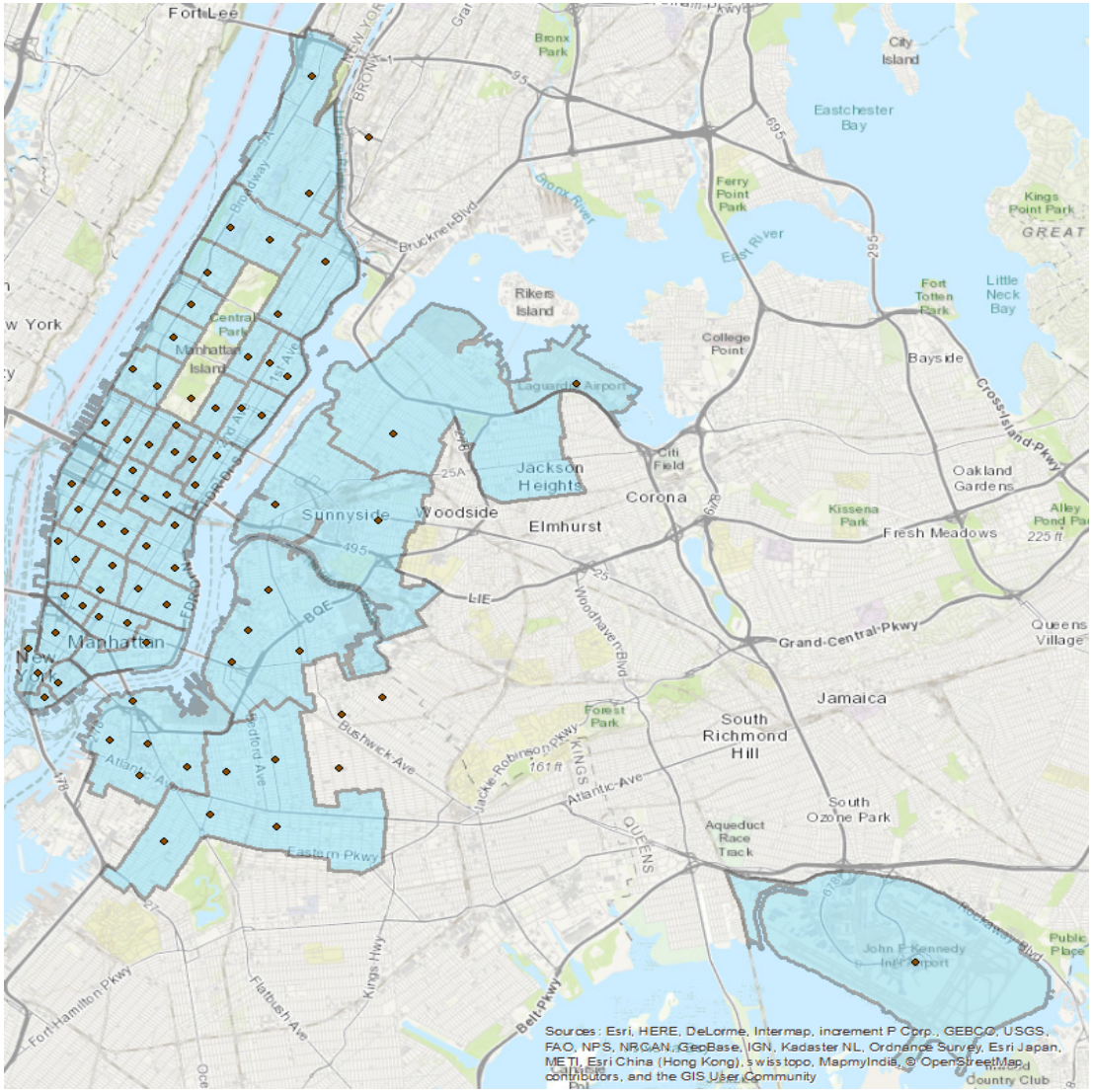



Figure: 40 select markets in the sample

# Empirical Model: Rider's nested logit demand

Passengers $c$ in location $i$ at time $t$ chooses from $f = y, x, o$ with utility:

$$\begin{aligned}\ln(U^{ij}_{cft,pre}) &= \ln(\tau^i_{ft}) + \ln(U^{ij}_{cft,post}) \\ &= \underbrace{\theta_1 \ln(v^i_{ft}) + \theta_2 \ln(u^i_{ft}) + d_x + d_i + t + \xi^i_{ft}}_{\delta^i_{ft}} \\ &\quad + \alpha^{ij} \ln(p^{ij}_{ft}) + \underbrace{\zeta^{ij}_{cgt} + (1-\beta)\nu^{ij}_{cft}}_{\varepsilon^{ij}_{cft}}\end{aligned} \tag{1}$$

- $\tau^i_{ft}$: demand side's matching probability.
- $v^i_{ft}$: supply of firm $f$.
- $\xi^i_{ft}$ unobserved demand shock.
- $d_x, d_i$: dummies.
- $\varepsilon^{ij}_{cft}, \nu^{ij}_{cft}$ follows type I extreme value distribution.

# Empirical Model: Passengers' decision

- The conditional market share is $s^{ij}_{ft} = s^{ij}_{f|gt} * s^{ij}_{gt}$
- Unconditional market share is $s^{i}_{ft} = \sum_j a^{ij}_t s^{ij}_{ft}$
- Demand is $u^{i}_{ft} = \lambda^{i}_t s^{i}_{ft}$ with market size $\lambda^{i}_t$
- Dropoff distribution of firm $f$ is:

$$\tilde{a}^{ij}_{ft} = \frac{a^{ij}_t s^{ij}_{ft}}{s^{i}_{ft}} \tag{2}$$

# Empirical Model: Drivers' decision

- Unmatched drivers in location $i$ make search decision:

$$j^* = \arg\max_j \{V^j_{ft+\chi^{ij}_t} - c^{ij}_t + \rho_f \underbrace{(V^j_{ft+\chi^{ij}_t} - \min_l \{V^l_{ft+\chi^{il}_t}\})}_{\Delta^j_{ft}} \mathbb{1}_{\chi^{ij}_t=1} + \epsilon^j_f\} \tag{3}$$

with ex-ante value in location $\forall i$:

$$V^j_{ft} = \phi^j_{ft}\Big(\sum_l \tilde{a}^{jl}_{ft}(p^{jl}_{ft} - c^{jl}_t + V^l_{ft+\chi^{jl}_t})\Big) + (1-\phi^j_{ft})\mathbb{E}_\epsilon\big[\max_l\{V^l_{ft+\chi^{jl}_t} - c^{jl}_t + \rho_f \Delta^j_{ft}\mathbb{1}_{\chi^{ij}_t=1} + \epsilon^l_f\}\big]. \tag{4}$$

where $\epsilon_f$ is TIEV with scale $\sigma_f$, $\phi^i_{ft}$ is matching probability:

# Empirical Model: Drivers' decision

- Employed drivers travel based on transition $\tilde{a}^{ij}_{ft}$
- Unemployed drivers travel based on search policy:

$$\pi^{ij}_{ft} = \frac{\exp((V^{j}_{ft+\chi^{ij}_t} - c^{ij}_t + \rho_f \Delta^{j}_{ft} \mathbb{1}_{\chi^{ij}_t=1})/\sigma_f)}{\Sigma_l \exp((V^{l}_{ft+\chi^{il}_t} - c^{il}_t + \rho_f \Delta^{l}_{ft} \mathbb{1}_{\chi^{il}_t=1})/\sigma_f)} \tag{5}$$

- State of vehicles updates:

$$\tilde{v}^{i}_{ft+1,k} = \tilde{v}^{i}_{ft,k+1} + \Sigma_j m^{j}_{ft} \tilde{a}^{ji}_{ft} \mathbb{1}_{\chi^{ji}_t=k} + \Sigma_j (v^{j}_{ft} - m^{j}_{ft}) \pi^{ji}_{ft} \mathbb{1}_{\chi^{ji}_t=k}, \forall f, i, k \tag{6}$$

where $\tilde{v}^{i}_{ft+1,k}$ indicates the number of firm $f$'s cars arriving at $i$ in $k$ periods conditional on current period t.

# Estimation Algorithm Overview

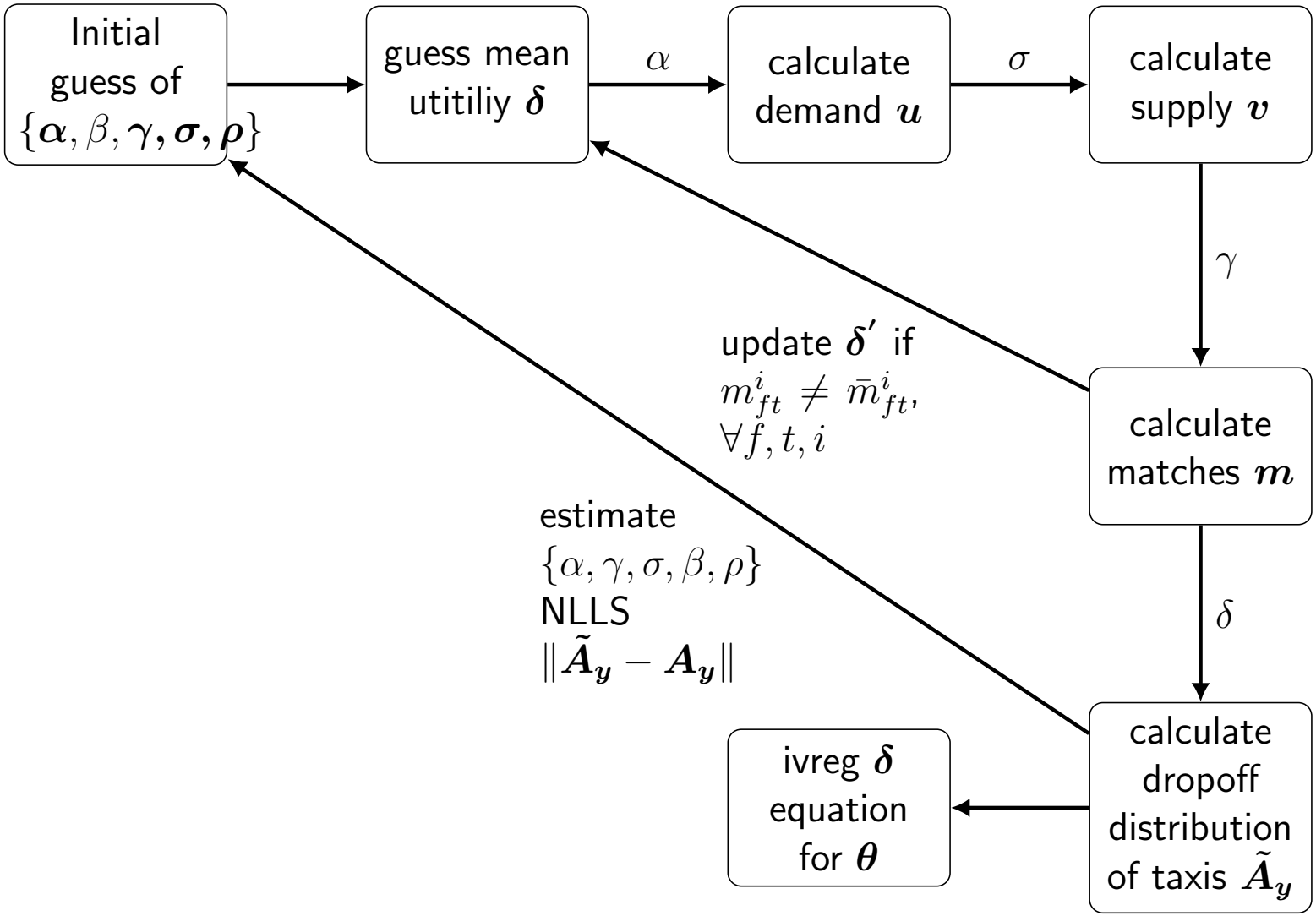



Figure: overview of the estimation process

# Results

Table: Estimates of nonlinear parameters

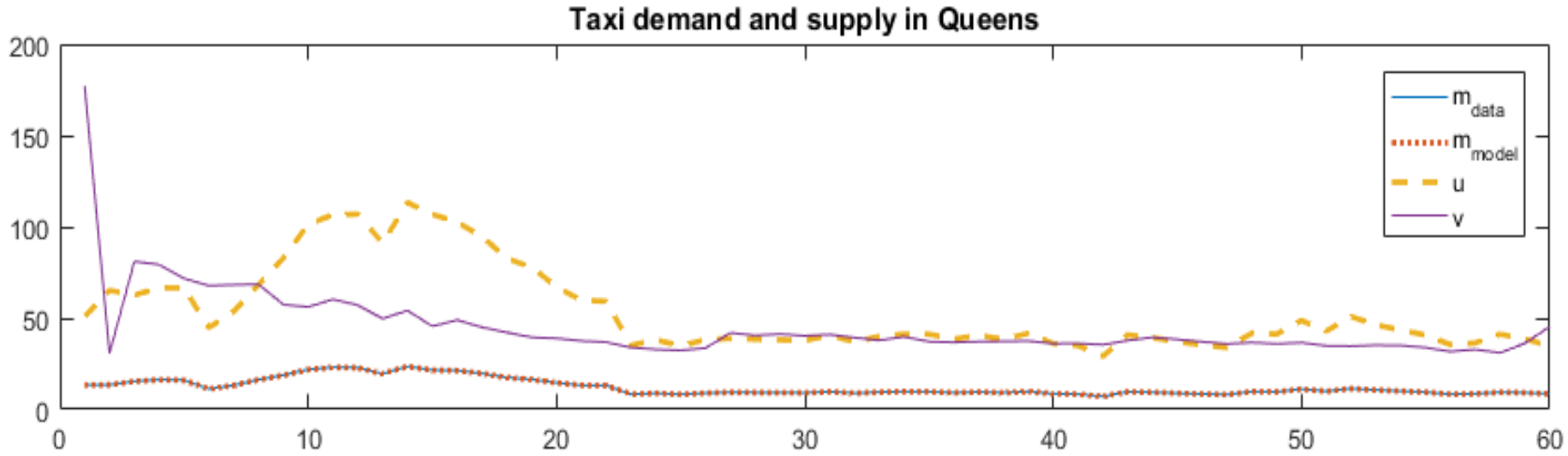
| *panel 1: nonlinear parameter* | | *estiamtes* |
|---|---|---|
| demand side parameters | | |
| $\hat{\alpha}_1$ | | -0.81 |
| $\hat{\alpha}_2$ | | -0.58 |
| $\hat{\alpha}_3$ | | -0.41 |
| $\hat{\alpha}_4$ | | -0.26 |
| $\hat{\beta}$ | | 0.38 |
| supply side parameters | | |
| $\hat{\sigma}_y$ | | 7.67 |
| $\hat{\sigma}_x$ | | 12.65 |
| $\hat{\rho}_y$ | | 0.38 |
| $\hat{\rho}_x$ | | 0.27 |
| matching function | | |
| $\hat{\gamma}_1$ | | 1.11 |
| $\hat{\gamma}_2$ | | 3.67 |
| *mean utilities* | *mean* | *min/max* |
| $\hat{\delta}^i_{yt}$ | 1.21 | -1.79/4.90 |
| $\hat{\delta}^i_{xt}$ | 0.30 | -1.62/4.02 |

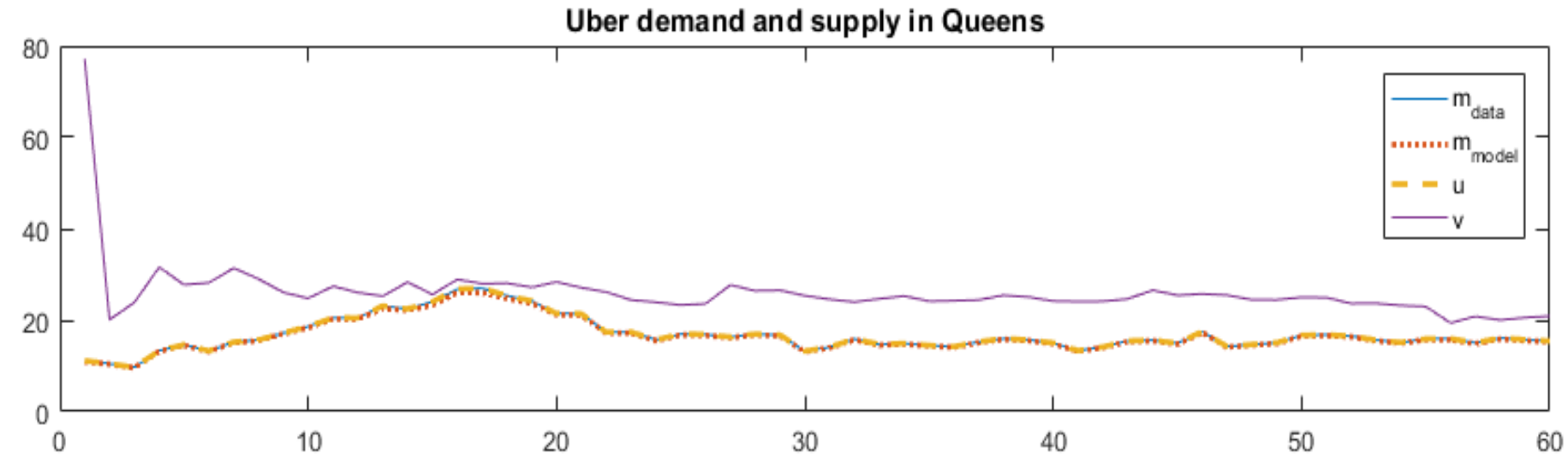
# Results

Table: Linear regression of mean utility

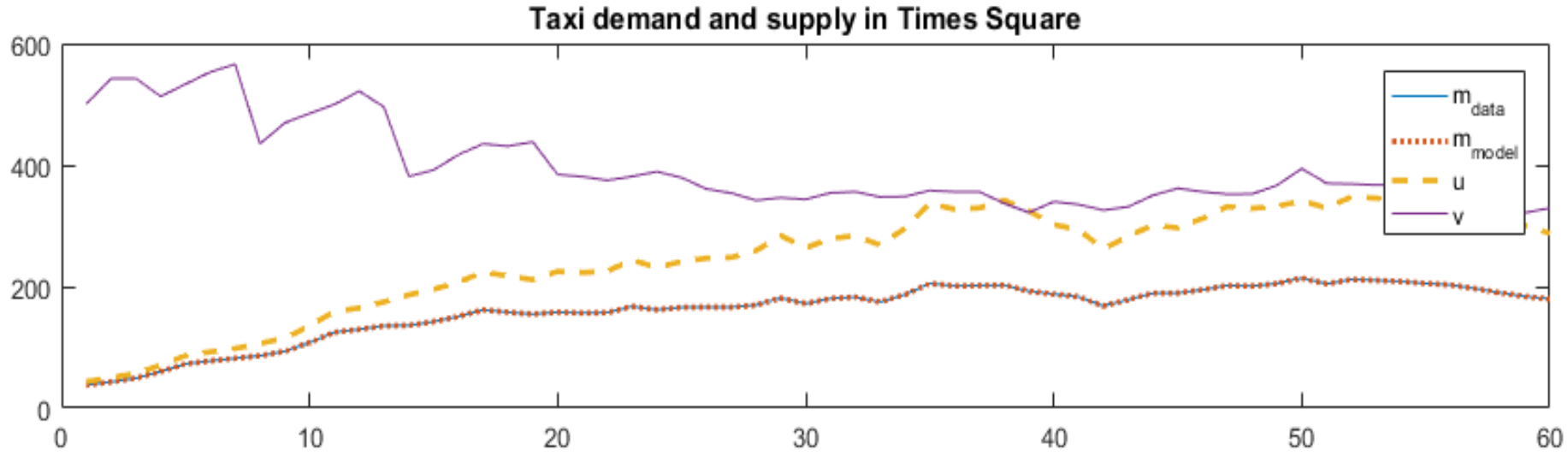
| Dependent variable $\delta^i_{ft}$ | | |
|---|---|---|
| | *OLS* | *2SLSIV* |
| $\ln v$ | 0.054 | 0.496 |
| | (0.019) | (0.07) |
| $\ln v \times d_x$ | -0.053 | -0.224 |
| | (0.039) | (0.11) |
| $\ln u$ | 0.53 | 0.249 |
| | (0.021) | (0.06) |
| $\ln u \times d_x$ | -0.014 | -0.086 |
| | (0.038) | (0.08) |
| Uber dummy $d_x$ | 0.19 | 1.21 |
| | (0.054) | (0.18) |
| constant | -2.226 | -3.71 |
| location fixed effects | YES | YES |
| time fixed effects | YES | YES |

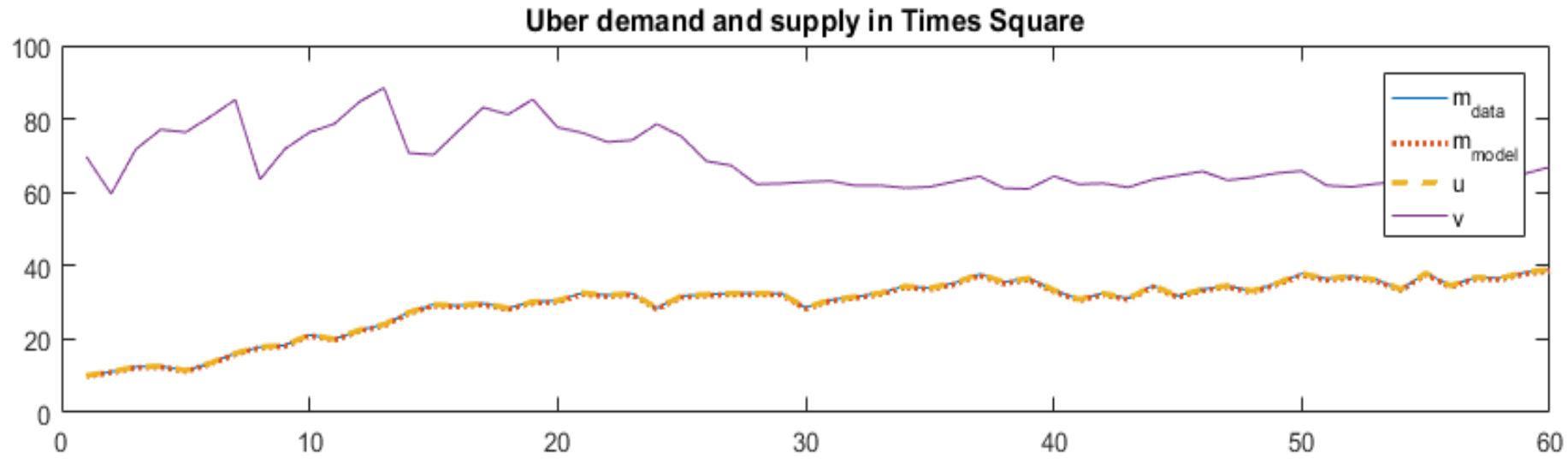
# Results


Taxi demand and supply in Queens
200
150
100
50
0
0
10
20
30
40
50
60
m data
m model
u
v


Uber demand and supply in Queens
80
60
40
20
0
0
10
20
30
40
50
60
m data
m model
u
v

# Results


Taxi demand and supply in Times Square
600
400
200
0
0
10
20
30
40
50
60
m data
m model
u
v


Uber demand and supply in Times Square
100
80
60
40
20
0
0
10
20
30
40
50
60
m data
m model
u
v

# Counterfactuals: Eliminate surge multiplier

| *supply, demand, match* | *Benchmark* | *w/o network* | *with network* |
|---|---|---|---|
| $\Sigma_{i,t} v^i_{yt}$ | 418,100 | 417,230 | 414,860 |
| $\Sigma_{i,t} v^i_{xt}$ | 82,707 | 83,375 | 84,219 |
| $\Sigma_{i,t} u^i_{yt}$ | 283,510 | 278,000 | 273,730 |
| $\Sigma_{i,t} u^i_{xt}$ | 50,224 | 58,374 | 62,065 |
| $\Sigma_{i,t} m^i_{yt}$ | 173,230 | 170,810 | 168,470 |
| $\Sigma_{i,t} m^i_{xt}$ | 47,738 | 52,088 | 54,371 |
| *two type friction* | | | |
| $within friction_y$ | 95,547 | 94,343 | 93,092 |
| | $ 1,286,400 | $ 1,272,300 | $ 1,256,800 |
| $within friction_x$ | 635 | 758 | 786 |
| | $ 10,517 | $ 10,516 | $ 10,804 |
| $cross friction_y$ | 14,738 | 12,844 | 12,165 |
| | $ 203,530 | $ 178,120 | $ 169,460 |
| $cross friction_x$ | 1,850 | 4,661 | 5,002 |
| | $ 34,450 | $ 64,905 | $ 66,603 |
| *welfare* | | | |
| $\$taxiprofit$ | $2,510,400 | $2,480,100 | $2,452,900 |
| $\$Uberprofit$ | $779,350 | $721,360 | $744,330 |
| $consumer\ welfare$ | 505,210 | 511,140 | 510,020 |
| $\Delta\$consumer\ welfare$ | NA | $120,400 | $ 96,977 |
| $\Delta\$social\ welfare$ | NA | $32,110 | $4,457 |

# Counterfactuals: Improve traffic conditions

| *supply, demand, match* | *Benchmark* | *w/o network* | *with network* |
|---|---|---|---|
| $\Sigma_{i,t} v^i_{yt}$ | 418,100 | 486,710 | 498,920 |
| $\Sigma_{i,t} v^i_{xt}$ | 82,707 | 93,670 | 91,857 |
| $\Sigma_{i,t} u^i_{yt}$ | 283,510 | 283,030 | 300,120 |
| $\Sigma_{i,t} u^i_{xt}$ | 50,224 | 50,182 | 48,391 |
| $\Sigma_{i,t} m^i_{yt}$ | 173,230 | 181,340 | 191,700 |
| $\Sigma_{i,t} m^i_{xt}$ | 47,738 | 48,709 | 46,681 |
| *two type friction* | | | |
| $within friction_y$ | 95,547 | 93,090 | 98,923 |
| | $ 1,286,400 | $ 1,252,900 | $ 1,328,200 |
| $within friction_x$ | 635 | 635 | 611 |
| | $ 10,517 | $ 10,531 | $ 10,406 |
| $cross friction_y$ | 14,738 | 8,596 | 9,501 |
| | $ 203,530 | $ 122,910 | $ 134,920 |
| $cross friction_x$ | 1,850 | 836 | 1,097 |
| | $ 34,450 | $ 15,832 | $ 21,807 |
| *welfare* | | | |
| $\$ taxi profit$ | $2,510,400 | $2,617,800 | $2,741,500 |
| $\$ Uber profit$ | $779,350 | $797,180 | $784,460 |
| $consumer\ welfare$ | 505,210 | 505,210 | 536,070 |
| $\Delta\$ consumer\ welfare$ | NA | $0 | $ 748,900 |
| $\Delta\$ social\ welfare$ | NA | $ 125,230 | $ 985,110 |

# Counterfactuals: Restrict Uber cars

| *supply, demand, match* | *Benchmark* | *w/o network* | *with network* |
|---|---|---|---|
| $\Sigma_{i,t} v^i_{yt}$ | 418,100 | 418,120 | 417,610 |
| $\Sigma_{i,t} v^i_{xt}$ | 82,707 | 65,330 | 64,622 |
| $\Sigma_{i,t} u^i_{yt}$ | 283,510 | 283,030 | 287,300 |
| $\Sigma_{i,t} u^i_{xt}$ | 50,224 | 50,182 | 43,871 |
| $\Sigma_{i,t} m^i_{yt}$ | 173,230 | 173,070 | 174,890 |
| $\Sigma_{i,t} m^i_{xt}$ | 47,738 | 44,351 | 40,734 |
| *two type friction* | | | |
| $within friction_y$ | 95,547 | 95,596 | 96,812 |
| | $ 1,286,400 | $ 1,287,100 | $ 1,304,800 |
| $within friction_x$ | 635 | 700 | 602 |
| | $ 10,517 | $ 11,670 | $ 9,994 |
| $cross friction_y$ | 14,738 | 14,367 | 15,595 |
| | $ 203,530 | $ 198,620 | $ 215,750 |
| $cross friction_x$ | 1,850 | 4,870 | 2,534 |
| | $ 34,450 | $ 91,152 | $ 46,764 |
| *welfare* | | | |
| $\$ taxi profit$ | $2,510,400 | $2,507,800 | $2,544,700 |
| $\$ Uber profit$ | $779,350 | $716,100 | $660,030 |
| $consumer\ welfare$ | 505,210 | 505,210 | 501,540 |
| $\Delta\$ consumer\ welfare$ | NA | $0 | $ -98,829 |
| $\Delta\$ social\ welfare$ | NA | $-65,850 | $-183,849 |

# Conclusion

- I contribute to the literature by considering network effects in the search and matching market and its impact on efficiency.
- I find positive feedback loop between drivers and riders.
- In the first simulation, I find Uber's surge pricing improves matching efficiency between Uber drivers and passengers.
- In the second simulation, I find a strong positive effects of traffic conditions on matching efficiency.
- In the third simulation, less competition makes matching less efficient for both taxis and Uber.
- Network effects are important in both quantities and directions of welfare analysis.