Spatial matching game and e-hail programs in NYC taxicab industry

Kin-ping Jeremy Wong
Georgetown University
kjw65@georgetown.edu

Abstract

The taxicab in NYC has been well-known for its search friction, as illustrated in the theory paper by Lagos (2000). Even when matching is perfect within an area, profit maximizing drivers overcrowd at the profitable zones, leading to frictions at equilibrium. Buchholz (2017) and Frachett, Lizzeri & Saez (2016) estimates a dynamic general equilibrium model to quantify the magnitude of the effects of entry restrictions and matching frictions. I present a model featuring a dynamic spatial matching game to studying the effect of using e-hail apps to improve matching efficiency. Counterfactual results indicate that substituting 30% street hail demand with e-hail technology generates 22% earnings and 23% utilization gains for e-hail drivers.

1. Background on the taxicab market

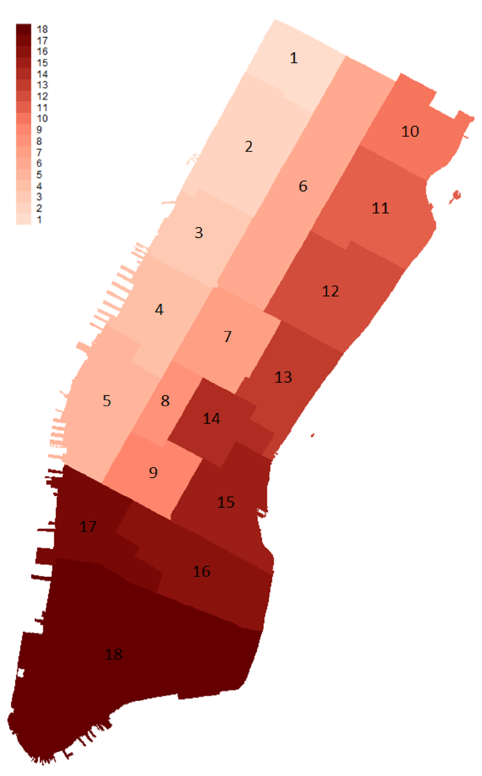
Cab and minifleet owners, to operate, must own a medallion licensed by TLC (Taxicab and Limousine Commission). The number of medallions is capped at 13,587; the fare is fixed. From 2013, TLC also licensed green cabs, but these are not allowed to pick up passengers in Manhattan and the airports. In 2013 June, TLC launched the e-hail (Pilot) Program and allowed app developers to dispatch cabs and match passengers' requests which have contributed 0.39% yellow cab rides, with a completion rate of 53%. Most of the requests

Outcome	2014	2015-2016
Cancelled by driver	154,758	391,183
Cancelled by passenger	507,729	1,251,154
Completed by green cabs	1,232,676	3,540,561
Completed by yellow cabs	706,669	combined
No response	799,867	1,014,333
Total requests	3,415,592	6,197,232

I study the data in 2016 June. By that time, e-hail apps have been launched for three years, so we can assume e-hail drivers were using it optimally and making search decisions optimally. Waiting time data of Uber Taxi was collected from its application programming interface for the month.

2. Model

The New York City area is divided into 20 zones, among which 18 are Manhattan zones as shown above. The day is discretized into 144 10-minute periods. Cab drivers can be equipped with an e-hail app or not. At the beginning of each period t, vacant cabs at each zone z decide where to search for passengers. Whatever their searching decisions, the drivers will travel in the same zone for x minutes before reaching their target zones. From each drivers the point of view, matches follow a Poisson arrival process such that the probability of picking up a passenger within a period s is:



$$1 - \exp -\lambda_z^t s \tag{1}$$

With τ periods left in his shift and given his belief $\tilde{\lambda}$, the value function for going to next zone z_2 is given by the following Bellman equation:

$$\begin{aligned} V\left(z_2|z,t,\tau;\tilde{\lambda}\right) &= P\left(ride|z,t;\tilde{\lambda}\right) EV(ride|z,t) \\ &+ \left(1 - P\left(ride|z,t;\tilde{\lambda}\right)\right)\left[P\left(ride|z_2,t;\tilde{\lambda}\right)\right. \\ &\left.\times EV(ride|z_2,t) + \left(1 - P\left(ride|z_2,t;\tilde{\lambda}\right)\right) V(z_2,t+1,\tau-1)\right] \end{aligned}$$

The Bellman equation for e-hail drivers is defined similarly, except that now the drivers consider also the probability of getting an e-hail request. Both e-hail and regular green cab drivers are restricted to pick up passengers in non-Manhattan zones only. Under the current zone grouping green cabs only pick up in one zone, so they are accounted in the aggregate matching processes, but their behaviors are not modeled.

Aggregate matching between C searching cabs and D through street hail follows a Cobb Douglas matching function:

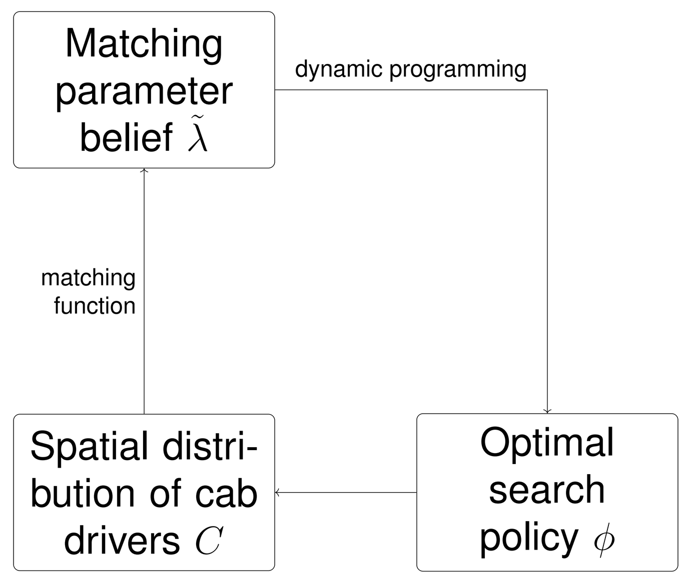
$$m(C,D) = C^\alpha D^\beta \tag{2}$$

While e-hail matching function with C^e e-hail cabs and D^e requests with overall ehail app network factor κ and meet probability p^e follows:

$$m^e(C^e,D^e) = \min\{C^e,D^e\}(1-p^e)^\kappa \tag{3}$$

Definition 1 *A competitive equilibrium in this model is* $\{S_0,\theta,V,\phi,V^e,\phi^e,\tilde{\lambda}\}$*: an initial state distribution* S_0*, a vector of matching function parameters* θ*, value function and decision rule pairs* $V(z,t,\tau;\tilde{\lambda})$ *and* $\phi(a|z,t,\tau;\tilde{\lambda})$*,* $V^e(z,t,\tau;\tilde{\lambda})$ *and* $\phi^e(a|z,t,\tau;\tilde{\lambda})$ *and belief* $\{\tilde{\lambda}_z^t\}_{z,t}$ *such that:*

- V*,* V^e *and* ϕ*,* ϕ^e *solve the drivers' optimization problem.*
- *Drivers' belief* $\tilde{\lambda}$ *are consistent with the law of motion* Φ *of vacant cabs induced by optimal strategy* ϕ*.*



The interaction between search behavior and actual spatial distribution of cabs forms beliefs as follows: (At equilibrium the belief drivers hold $\tilde{\lambda}$ and the realized matching parameter λ will be consistent.)

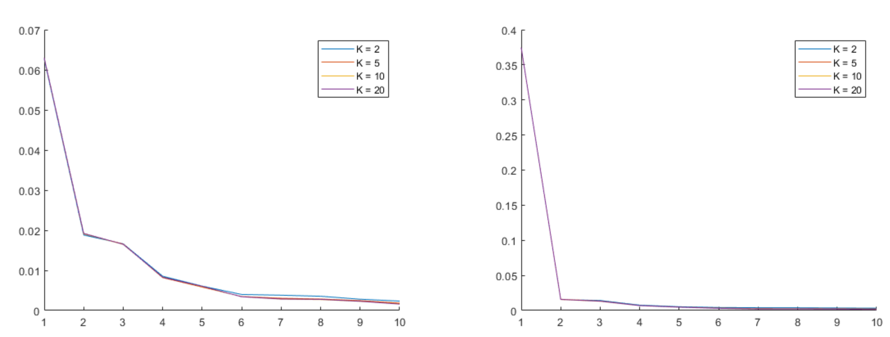
3. Estimation

From the data one can directly infer the cab shift starting times and lengths (initial states S_0), payoffs conditioned on endpoints and trip duration ($\pi(s,s')$), and state transitions due to trips ($Q(s,s')$). Matching parameters θ is estimated using the method of simulated moments. Given an initial guess of θ the following contraction property of the matching game:

- Given belief $\lambda(\theta)$ solve for drivers' search decisions using Bellman equation. Simulate the spatial distributions.
- Update drivers' beliefs using the implied matching parameters. Repeat until convergence.

Algorithm 1: *Algorithm for computation of competitive equilibrium given* θ*.*

We rely numerically on the existence of an equilibrium over a large set of simulations. Beliefs converge:



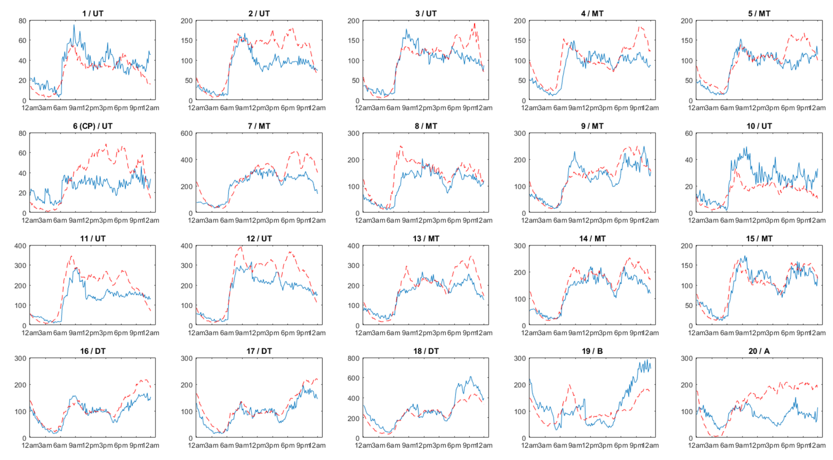
(a) (b)

Figure 1: *(a): convergence of beliefs for street hail matching parameters, (b): convergence of beliefs for e-hail parameters*

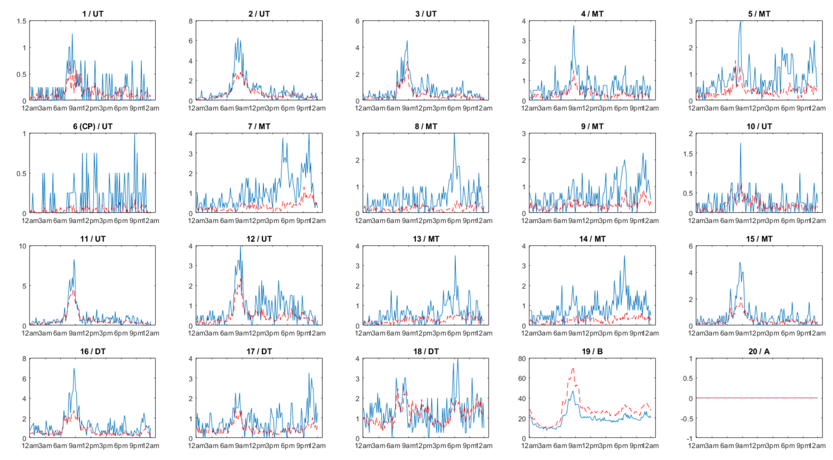
Therefore the estimation strategy follows a nested simulated method of moments algorithm in which the inner loop uses the above fixed point algorithm to solve for the equilibrium moments given an initial guess of matching parameters θ, which is updated in the outer loop. The moments matched are the observed number of street hail and e-hail matches for the 2,880 (zone,period) pairs.

4. Results

The results fit the data well, the following graphs show the simulated and actual moments (matches for all zone-time pairs).



(a)



(b)

Figure 2: *(a): simulated and actual matches from street hail, (b): simulated and actual e-hail matches*

Panel data for individual driver shifts are not available for the 2016 data. However, the estimated model fitted for the 2013 data fits the shift earnings, searching time and utilization rates closely.

5. Counterfactual experiments

The following counterfactual experiments are run: what happens if a portion of passengers currently hailing cabs on street switch to using e-hail apps? In each of the experiments, the street hail demand parameters of every zone are decreased in the given percentage, and the implied number of matches substituted are added to the e-hail requests. The ehail app network strength is kept constant in each of these cases and we assume the percentages for yellow and green ehail drivers grow by the same percentages. We also kept constant the cancellation rates and meeting rates.

	Benchmark	30%	50%	70%
		% of substituted rides		
shift earnings (non-ehail)	\$232.02	\$182.29	\$146.33	\$107.88
shift earnings (ehail)	\$243.47	\$296.37	\$312.45	\$313.11
earnings per hour (non-ehail)	\$26.32	\$20.67	\$16.57	\$12.19
earnings per hour (ehail)	\$27.58	\$33.69	\$35.47	\$35.59
utilization rate (non-ehail)	48.93%	39.82%	33.37%	26.57%
utilization rate (ehail)	49.13%	60.70%	64.38%	65.12%
total street hail rides	317,800	244,984	200,821	153,886
total ehail rides	2,417	76,371	118,478	153,072
total rides in non-yellow zones	20,195	17,300	14,233	11,806

Counterfactual results

We see from above that when passengers substitute to use e-hail apps, the earnings and utilization for e-hail drivers increases, at the expense of the non-ehail drivers. Also, now that Manhattan areas also have high enough ehail request arrival rates, so ehail yellow cab drivers are incentivized to stay searching in Mnahattan, decreasing the overall rides outside Manhattan. Also note that under the current ehail app network factor there is a capacity and limit for earning and utilization improvements.

The counterfactual results are under assumptions that the app network factor remains constant, therefore in real life, one would expect an even higher efficiency gain for the transition from street hail to ehail.

6. Discussion

By slightly changing the ehail matching function, the model can be further extended to study the effects of implementing carpool algorithms. A limitation for the above analyses is that passengers are assumed to behave in an exogenous way, whereas in real life they could respond to changes in waiting times. With availability of UberTaxi data in the time frame, this can be modeled in a future extension. Also, the model provide scopes to model ride sharing apps such as Uber, whose drivers are subject to a similar matching function.