

14.273: Advanced Topics in IO

Search - Spatial Search Frictions

Tobias Salz

*Graduate IO. Parts of notes gratuitously borrowed from John Asker, Nick Buchholz, Allan Collard-Wexler, Chris Conlon, Kei Kawai, Robin Lee, Ariel Pakes, Paul Scott, and Matt Shum.

Introduction

There has been a recent interest in IO to understand how spatial search frictions give rise to market outcomes. This is particular relevant in the transportation sector.

Spatial search frictions

- The need to coordinate trade in a physical space is a major source of search frictions.
- In equilibrium, market organization will reflect spatial frictions.

Research questions

- How are supply & demand determined in spatial search markets?
- How does the extend of frictions vary across equilibria?
- What are the welfare consequences of those frictions?
- What can be done to reduce those frictions?

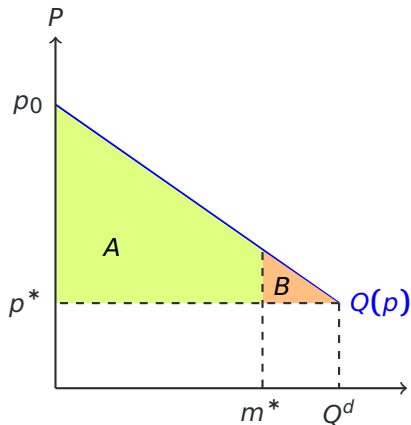
Spatial Search Frictions in IO



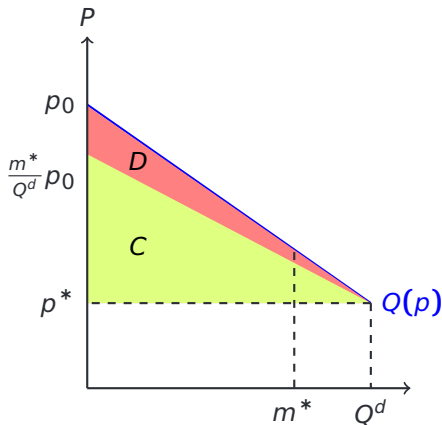
Welfare Implications

Frictions reduce surplus, more so when matches are random.

Frictions with Sorting



Frictions with Random Matching



Conceptually important papers:

- Lagos (2000) formulates a model of spatial search frictions.
- Lagos (2003) calibrates Lagos (2002) with NYC Taxis data.

Here, we will discuss three papers that combine those ideas with concepts from industry dynamics:

1. Buchholz (2018): Endogenous location choice model with search frictions, NYC taxi industry. (build on Lagos (2003))
2. Brancaccio, Kalouptsi, and Papageorgiou (2018): Endogenous location choice model with search frictions, world shipping industry. (build on Lagos (2003))
3. Frechette, Lizzeri, and Salz (2018): Entry/exit model with search frictions, NYC taxi industry.

NYC Taxis, Setting

Industry Background

Institutional Details:

- Medallion limit: 13,237 total of taxis in the system.
- Fare structure as of September 4: \$2.50 + \$2.50/mi.
- Waiting/idling fare exists.
- Yellow cabs not permitted to pre-arrange rides.
- Drivers lease medallions by the shift.
 - Most medallions require two shifts per day.
 - Drivers keep income net of lease and fuel costs.
- Data is from a period before Uber entered NYC.

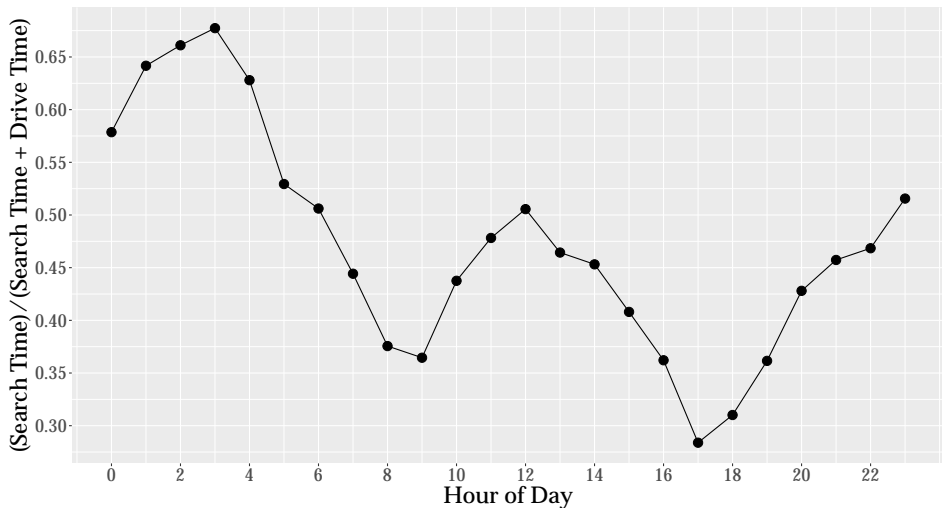
Data:

- TPEP (taxi passenger enhancement project).
- Lat/lon trip start/end point, fare, driver + medallion id.

Search Frictions in Taxi Markets

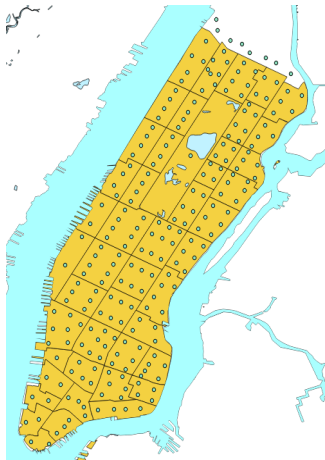


Market clears on search-time and wait-time.



Drivers spend 220% of the time it would take under **optimal route** (Google) to get to next passenger.

Buchholz (2018)



- Divide map into 48 locations
 - Joining census tracts (blue dot centroids) to locations (yellow).
 - Captures 94% of all rides.
- Divide time into 120 periods
 - 5-minute periods, 6a-4p
- Each loc./time, extract data:
 - Number of matches.
 - Customer transitions.
 - Travel time and distance.

Demand arrives in every location every 5-minutes

- Poisson arrival rates λ_i^t

Taxis choose search locations:

- Accounts for fare and continuation value.

Agents/actions/payoffs:

- Vacant drivers choose which locations to search for passengers
- Drivers want to maximize daily profits earned by giving rides

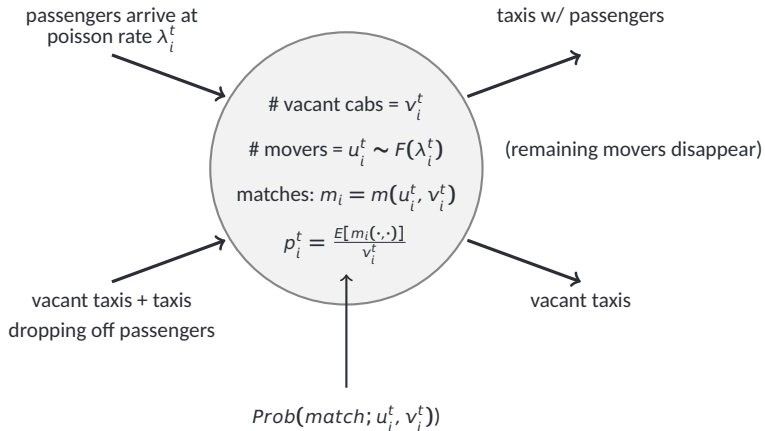
State space, \mathcal{S} : distribution of all NY taxis across city over the day

- 48 locations \times 120 periods = 5,760-element matrix.
- Taxis take demand functions, prices as exogenous.

Key Parameters:

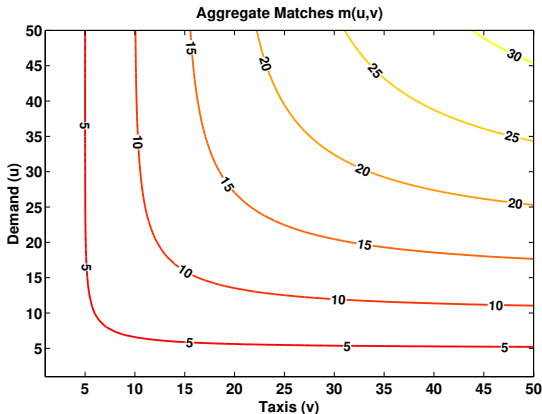
- Distribution of demand arrivals in each location-time (5,760 rates)

Arrivals and matches in a location form probabilities:



Matching function (Butters '77, Burdett, Shi, Wright '01)

- $m(u_i, v_i) = v_i \cdot \left(1 - \left(1 - \frac{1}{v_i}\right)^{u_i}\right)$
- Microfounded on single-good, random matching market
- Bounded, approximately CRS outside of lower limit



“Micro-Foundation” for:

$$m(u_i, v_i) = v_i \cdot \left(1 - \left(1 - \frac{1}{v_i} \right)^{u_i} \right)$$

- Workers send out applications to vacancies v with equal probability.
- Each worker sends out one application.
- The probability distribution for a particular vacancy is then given by:
binomial $(u, \frac{1}{v})$.
- A firm receives at least one application with probability: $1 - (1 - \frac{1}{v})^u$.
- The total number of matches is therefore: $m(u, v) = v \cdot (1 - (1 - \frac{1}{v})^u)$

Model: Match and Vacancy values

$$\begin{aligned}
 v_i^t(\mathcal{S}) = & \mathbb{E}_{p_i^t | \lambda_i, \mathcal{S}} \left[p_i^t(\mathcal{S}) \overbrace{\left(\sum_j M_{ij}^t \cdot (\Pi_{ij} + v_j^{t+\tau_{ijt}}) \right)}^{\text{Exp. Value of Fare}} + \right. \\
 & \left. (1 - p_i^t(\mathcal{S})) \cdot \underbrace{\mathbb{E}_{\varepsilon_{a,j}^{t+1}} \left[\max_{j \in \mathcal{L}} \{ v_j^{t+\tau_{ijt}} - c_{ij} + \varepsilon_{a,j} \} \right] \middle| \lambda_i^t}_{\text{Exp. Value of Vacancy}} \right]
 \end{aligned}$$

- M_{ij}^t is the empirical probability of customer travel from $i \rightarrow j$
- Fare received: $\Pi_{ij} = \text{flag fare} + \text{dist. fare} \cdot d_{ij} - c_{ij}$
- Travel time in periods: τ_{ijt}
- $\varepsilon_{a,j}$ logit with $\text{var}(\varepsilon) = \sigma_\varepsilon$
- CCPs between each location pair define policy functions.

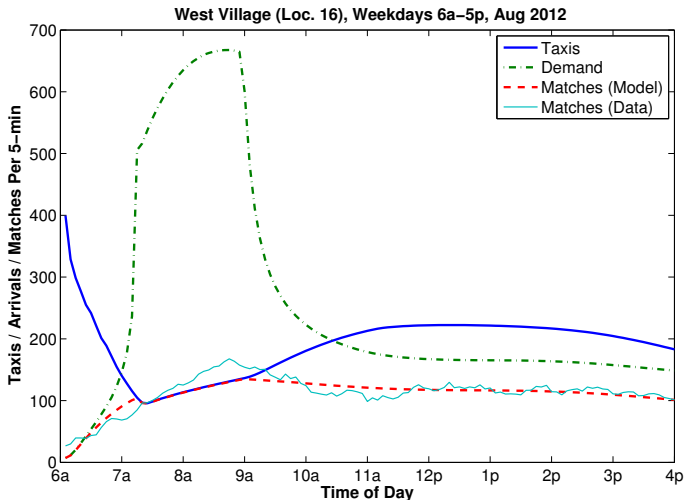
Dimensionality: 48 locations \times 13k taxis \times 120 periods

- Stochastic state transitions: employed and vacant taxis.
- Number of possible (anonymous) taxi distributions $\approx 10^{100}$.
- Infeasible to solve via full integration method. Instead, exploit the large number of agents, transitions deterministic (as if continuum of firms)
 - No need to integrate over improbable events.
 - Permits direct computation of transitions/V. fcts.

Non-Stationary Oblivious Equilibrium:

- Taxis condition on own-state, don't observe competitors' states.
- Expectations are formed through experience (repeated observation of locations at each time of day).

Excess demand in some locations:

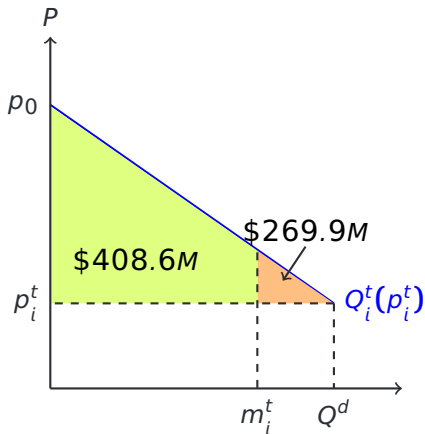


Taxi drivers face constant elasticity demand of the form:

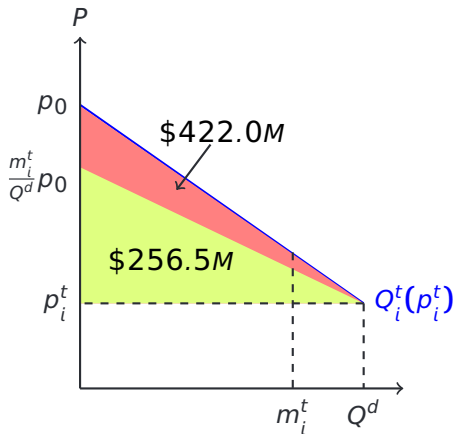
$$\underbrace{\ln(\lambda_{ij}^t)}_{\text{Estimates}} = \sum_{k \in \{2,4,6\}} \mathbf{I}_{d_{ij} \in (k,k')} \left(\alpha_{0,k} + \sum_s \mathbf{I}_{si} \left(\alpha_{1,ks} \underbrace{\ln(P_{ij}^t)}_{\text{Fares}} + \alpha_{2,kst} X_{it} \right) \right) + \delta^t$$

- Price Variation: across origin-destinations ij and time t .
- Elasticities for ride lengths k , # adjacent subway stations s
- X includes rich spatial variation:
 - demographics (income, commute-time)
 - point-to-point transit times
- Time-of-day fixed effects
- Since the fare structure (b, π) is fixed, no simultaneity bias

Costs of frictions w/ sorting



Costs of frictions w/ random matching



Frechette, Lizzeri, and Salz (2018)

Research Question

Frictions:

- Restrictions on extensive margin (entry) and intensive margin (ownership rules) affect capital/labor utilization.
- Shift frictions.
- Prices are fixed.
- Search frictions.

→ Paves the way for entrants such as **Uber** and **Lyft**.

Question: What is the welfare loss from these frictions?

- Counterfactuals that introduce dispatch technology and change entry restrictions.

Structure: General equilibrium model: Search + dynamic labor supply.

- Supply: Drivers make entry and stopping decisions.
- Total number of entrants limited by number of medallions.
- Almost no fare variation.
- Drivers earnings depend on arrival rate of passengers and number of competing taxis.
- Number of passengers function of waiting time.
- Wait-time (for passengers) and search-time (for cabs) clear market.

Demand: and **wait-time** unobserved.

Idea: Taxi search time reveals information about # of passengers.

- $s_1 = g(C_1, D_1)$, $s_2 = g(C_2, D_2)$. If $C_2 = C_1$ and $s_1 > s_2 \Rightarrow D_2 > D_1$
- Problem: $g()$ is not known.

Solution:

- Use geographical structure of Manhattan and simulate $g()$.

Idea: Taxi search time informative about number of people waiting.

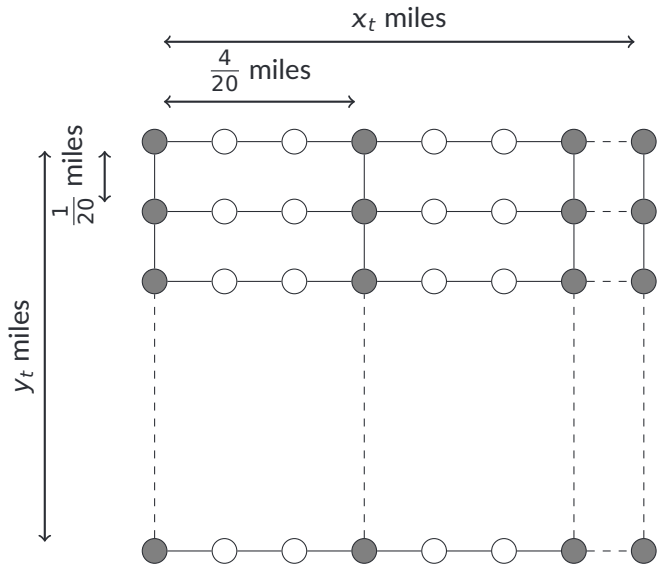
Use **geographical structure** of Manhattan to obtain:

$$\begin{pmatrix} s_t \\ w_t \end{pmatrix} = g(d_t, c_t, \phi_t)$$

- Simulate w_t and s_t for many $(d_t, c_t, mph_t, distance_t)$.
- Invert $g(\cdot)$ to back out d_t and w_t .

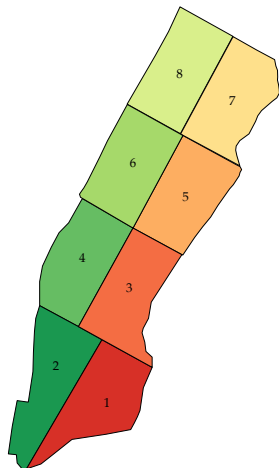
Matching Simulation

A stylized representation of Manhattan



with $y_t/x_t = 4$ and $y_t \cdot x_t = a_t$

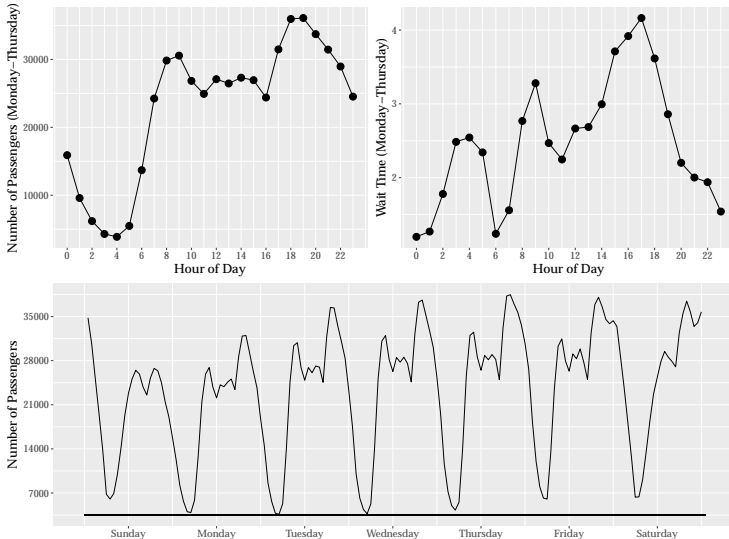
Homogeneity across Areas



Area	Travel Distance	Search time
	Mean	Mean
1	2.67	11.64
2	2.08	9.15
3	1.88	8.6
4	2.06	9.21
5	1.74	8.34
6	1.85	8.79
7	1.94	9.68
8	2.41	11.42

- Use *observed distribution* of request and pickup locations in simulation.
- Demand estimation robust to smaller and bigger divisions.

Demand: Result



Regress recovered demand and wait-time on **rainfall dummy**.

	(1)	(3)	(4)	(6)
	$\log(d_t)$	$\log(d_t)$	$\log(w_t)$	$\log(w_t)$
Rainfall	0.261**	0.247**	0.355**	0.463**
	(0.014)	(0.0083)	(0.011)	(0.0083)
Hour FE	No	Yes	No	Yes
Day of W. FE	No	Yes	No	Yes

Constant elasticity, log-linear demand function:

$$d_t = \exp(a + \sum_{h_t} \beta_{h_t} \cdot 1\{h_t\}) \cdot w_t^\eta \cdot \exp(\xi_t) \Leftrightarrow$$

$$\log(d_t) = a + \sum_{h_t} \beta_{h_t} \cdot 1\{h_t\} + \eta \cdot \log(w_t) + \xi_t$$

Simultaneity:

- Drivers and passengers respond to unobserved shocks.
- $\text{cov}(\xi_t, w_t) \neq 0$

Supply rationed: hard to find supply side instruments.

Dependent Variable:	$\log(w_t)$	$\log(s_t)$	$\log(w_t)$	$\log(s_t)$
Observations	1428	1428	1428	1428
Hour FE	Yes	Yes	No	No
Hour x Day of Week FE	No	No	Yes	Yes
R^2	0.838	0.890	0.895	0.931

- Hour controls account for most variation.

Shift transition:

- Driven by supply side constraints.
- Long run demand adjusts.
- Possibly too coarse.

Trip Distance:

- Longer taxi trips lead to less supply.
- Use trip distance outside of Manhattan.

Number of active cabs in previous periods:

- Shift indivisibilities.
- \rightarrow yields elasticity of ≈ -1.2 .

Shift value of driver i at time t :

$$V(\mathbf{x}_{it}, \epsilon_{it}) = \max \{ \epsilon_{it0}, \pi_t - C_{z_i, h_t}(l_{it}) - f(h_t, k_i) + \epsilon_{it1} \\ + \beta \cdot E_{\epsilon} [V(\mathbf{x}_{i,t+1}, \epsilon_{i,t+1}) | \mathbf{x}_{i,t+1}] \}$$

With:

- $\mathbf{x}_{it} = (l_t, h_t), h_t \in \{0, 23\}$.
- z_j : medallion type, (owner-operated/fleet).
- $C_{z_i, h_t}(l_{it})$: cost of driving (polynomial of shift length).
- $\epsilon_{it0}, \epsilon_{it1} \sim T1EV \text{ iid}, f_{z_i}(h_t, k_i) = f_{z_i} \cdot 1\{h_t \in h_{k_i}\}$: late fine.
- $\pi_t = \pi_0 \cdot \frac{E[DELIVERYTIME_t | h_t]}{E[DELIVERYTIME_t | h_t] + E[SEARCHTIME_t | h_t]}$: hourly earnings.

Definition

A competitive equilibrium in the taxi market at time t is a set: $\{s_{h_t}, w_{h_t}, c_{h_t}, d_{h_t}, \pi_{h_t} : h_t \in H_t\}$, such that:

1. d_{h_t} results from the demand function $d(\cdot)|_{h_t}$ under the waiting time w_{h_t} .
2. s_{h_t} and w_{h_t} result from d_{h_t} and c_{h_t} under the matching function $g(\cdot)|_{h_t}$.
3. c_{h_t} results from optimal starting and stopping under π_{h_t} .
4. π_{h_t} results from s_{h_t} .

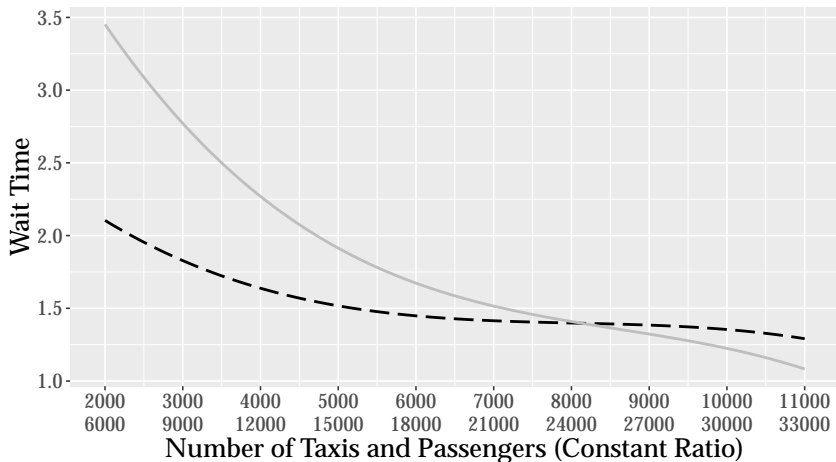
- Shift definition same as in Farber (2008).
- Estimation via **MPEC**.
- Different parameters for owner-operators and mini-fleet.
- Estimate for “average” weekday, Monday-Thursday.
- Match starting and stopping probabilities.

Counter-factuals

Comparing matching functions



Matching Function — Dispatch — Search

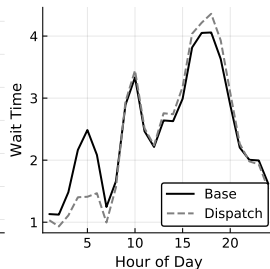
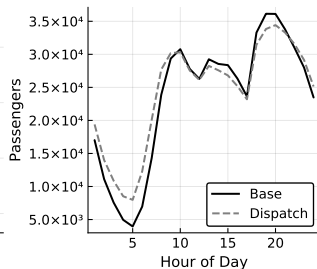
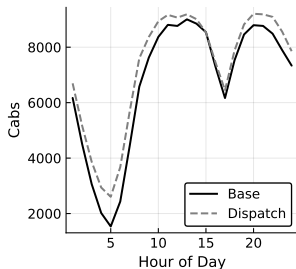


Counterfactuals



Dispatch: Assign empty cab to closest passenger.

	Baseline	Dispatch	$\Delta\%$
Consumer Surplus (million minutes / day)	2.00	2.12	6.11
Driver Revenue (hourly income)	\$39.54	\$41.63	5.27
Medallion Revenue (PV in millions)	\$2.54	\$2.68	5.42
Wait time (average in minutes)	2.7	2.66	-1.47

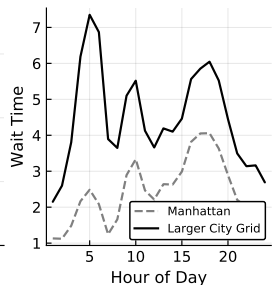
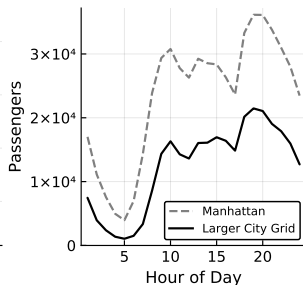
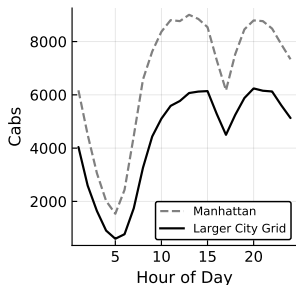


Counterfactuals



Dispatch comparison, Brooklyn

	Baseline	Brooklyn	$\Delta\%$	Brooklyn Dispatch	$\Delta\%$
Consumer Surplus (million minutes / day)	2.0	1.43	-28.3		
Driver Revenue (hourly income)	39.54	\$32.61	-17.54		
Medallion Revenue (PV in millions)	2.54	\$1.97	-22.2		
Wait time (average in minutes)	2.7	4.33	60.4		

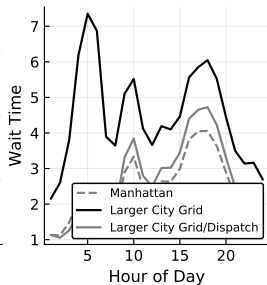
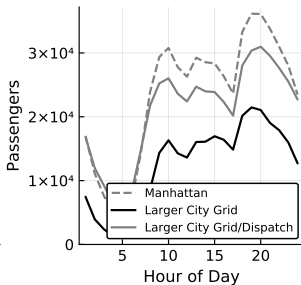
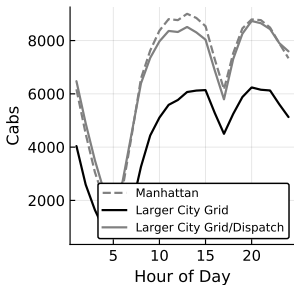


Counterfactuals



Dispatch vs Search on the size of Brooklyn

	Baseline	Brooklyn	$\Delta\%$	Brooklyn Dispatch	$\Delta\%$
Consumer Surplus (million minutes / day)	2.0	1.43	-28.3	1.97	-1.59
Driver Revenue (hourly income)	39.54	\$32.61	-17.54	\$39.12	-1.06
Medallion Revenue (PV in millions)	2.54	\$1.97	-22.2	\$2.52	-1.59
Wait time (average in minutes)	2.7	4.33	60.4	3.02	11.87



Segmentation counterfactual: 50% dispatch, 50% search.

	Baseline	Segmented	$\Delta\%$	Segmented Arbitrage	$\Delta\%$
Consumer Surplus (million minutes / day)	2.00	1.8	-9.88		
Driver Revenue (hourly income)	\$39.54	\$37.39	-5.43%		
Medallion Revenue (present value)	\$2.54	\$2.68	-8.71		
Wait time (minutes average)	2.7	3.32	23.1		

Segmentation counterfactual: 50% dispatch, 50% search.

	Baseline	Segmented	$\Delta\%$	Segmented Arbitrage	$\Delta\%$
Consumer Surplus (million minutes / day)	2.00	1.8	-9.88	2.02	0.9
Driver Revenue (hourly income)	\$39.54	\$37.39	-5.43%	37.19	-5.96
Medallion Revenue (present value)	\$2.54	\$2.68	-8.71	2.27	-10.66
Wait time (minutes average)	2.7	3.32	23.1	2.96	7.8

Driver earnings difference over eight hour shift:

- Without passenger choice: \$35.2.
- With passenger choice: \$70.2.

**Brancaccio, Kalouptsidi, and
Papageorgiou (2018)**

Introduction

- 80% of global trade by volume is carried out by ships
- Large price differentials across space, e.g.
 - Shipping price from Australia to China about 30% more expensive than vice versa.
- 42% of ships currently in transit are without cargo (ballast).
- Spatial equilibrium determines world trade. Trade costs are endogenous and determined *jointly* with trade flows. Standard trade models predict trade costs → trade flows

Bulk Shipping

Industry:

- Homogeneous unpacked dry cargo, for individual shippers on non-scheduled routes.
- Transports raw materials (iron ore, grain, coal, steel, etc.).
- Operate like “taxi drivers, not buses”.
- Contracts through brokers.
- Unconcentrated industry, homogeneous good.

Data:

- Shipping contracts.
- Ship movements (including draft).
- Daily wind from oceanic station.

In labor markets, evidence for **search frictions**:

- Wage dispersion.
- Coexistence of unemployed workers and vacancies.

Here:

- Substantial price dispersion within time/origin/destination (coeff of variation 30%).
- Evidence of unrealized matches.
 - Matches $< \min \{ \text{ships, exporters} \}$.
 - Simultaneous arrivals and departures of empty ships.

Model Overview

- Dynamic **spatial search model**.
- There are I regions in the world.
 - different trip durations between regions.
- **Agents:** in the model:
 - Exporters (freights).
 - Ships.
- In each region i there are e_i exporters/freights awaiting transportation:
- Freights are heterogeneous in:
 1. Revenue (valuation) from delivery, r .
 2. Destination, j .

Ships

In every period a ship is either:

- **Sailing** toward a destination j , either full or empty at sailing cost c^s .
 - Ship traveling from i to j arrives with prob d_{ij} (avg trip duration $1/d_{ij}$).
- **Waiting** in port i at a cost c_i^w .
 - Randomly matches with an exporter and begins trip.
 - If unmatched choose where to search (either wait at port again, or ballast to another region).

Environment: Matching Process

- Exporters and ships search for each other
- $m_i(e_i, s_i)$ new matches
 - s_i unmatched ships and e_i unmatched exporters in region i
 - probability of ship finding an exporter is λ_i
 - probability of exporter finding a ship is λ_i^e
- Search frictions generate rents to be split
 - price τ_{ijr} determined by **Nash bargaining**

Value Functions

- Traveling ship:

$$V_{ij} = -c^s + d_{ij} \cdot \beta \cdot V_j + (1 - d_{ij}) \cdot \beta \cdot V_{ij}$$

- Ship at port start of period:

$$V_i = -c_i^w + \lambda_i \cdot \underbrace{\mathbb{E}_{j,r} (\tau_{ijr} + V_{ij})}_{\text{matched ship value}} + (1 - \lambda_i) \underbrace{U_i}_{\text{unmatched ship value}}$$

- Ship that remained unmatched:

$$U_i = \max \left\{ \beta V_i + \sigma \epsilon_{ii}, \max_{j \neq i} V_{ij} + \sigma \epsilon_{ij} \right\}$$

- Value of unmatched exporter:

$$U_{ijr}^e = \lambda_i^e \underbrace{(r - \tau_{ijr})}_{\text{matched exp. value}} + (1 - \lambda_i^e) \cdot \delta \cdot \beta \cdot U_{ijr}^e$$

Surplus sharing condition (Nash Bargaining):

$$\gamma \cdot \left(\underbrace{\tau_{ijr} + V_{ij}}_{\text{ship inside-opt.}} - \underbrace{E_{\epsilon}(U_i)}_{\text{ship outside opt.}} \right) = (1 - \gamma) \cdot \left(\underbrace{r - \tau_{ijr}}_{\text{ex. inside-opt.}} - \underbrace{U_{ijr}^e}_{\text{ex. outside-opt.}} \right)$$

where γ is the exporter's bargaining power.

Entry of exporters:

- \mathcal{E}_i ex ante homogeneous potential exporters in market i choose whether and where to export, then draw r .
- Potential entrant exporter (where κ_{ij} is the production/exporting cost):

$$\max \left\{ \epsilon_0^e, \max_{j \neq i} \left\{ E_r U_{ijr}^e - \kappa_{ij} + \epsilon_j^e \right\} \right\}$$

Matching function estimation in the literature:

- Labor Markets: unemployed workers, vacancies, matches observed.
- Taxi Cabs: taxis, matches observed, passengers unobserved.

This paper:

- Imposes strong functional form assumptions (matters for welfare)

No search frictions:

$$\underbrace{m_{it}}_{\text{matches}} = \underbrace{\min(e_{it}, s_{it})}_{\min(\text{exporters}, \text{ships})} \quad (1)$$

Search frictions:

$$\underbrace{m_{it}}_{\text{matches}} = \underbrace{m_i(e_{it}, s_{it})}_{m(\text{exporters}, \text{ships})} \leq \min(e_{it}, s_{it}) \quad (2)$$

How do we distinguish (1) from (2), if one side unobserved/mis-measured?

Test for **search frictions**:

- Consider markets with $\min\{s, e\} = e$
- Then:
 - If $m = \min\{s, e\}$, changing s exogenously doesn't affect m
 - If $m \leq \min\{s, e\}$, changing s exogenously can affect m
 - Weather exogenously changes s - does it affect m ?

Find that matches affected by weather in all markets.

Matching Function

Use lit on **nonparametric identification** (Matzkin 2003)::

$$\underbrace{m_{it}}_{\text{matches}} = \underbrace{m_i(s_{it}, e_{it})}_{m(\text{ships, exporters})}$$

- Independence s_{it}, e_{it} : Correlation between m_{it} and s_{it} is informative about $\frac{\partial m_i}{\partial s}$
- Assume homogeneity of degree 1:

$$\text{knowing } \frac{\partial m_i}{\partial s} \quad \text{we also know } \frac{\partial m_i}{\partial e}$$

- Instrument: **sea weather** (wind speed) exogenously shocks ship arrivals.

Matching Function

We show how to estimate $m_i(e_{it}, s_{it})$ non-parametrically *and* **recover unobserved freights** e_{it}

- Matzkin 2003:

$$Y = m(X, \epsilon)$$

- Can I recover both $m(\cdot)$ and “shock” ϵ ?
- Necessary assumptions
 - $m(X, \epsilon)$ str. increasing in ϵ .
 - $X \perp \epsilon$, or a valid instrument (sea weather).
- Identification needs an additional assumption:
 - Either fix pdf of ϵ
 - Or assume $m(\cdot)$ is homogeneous of degree 1 and a known point

Matching Function

Matzkin argument:

$$\begin{aligned}
 F_{m|s=s_t}(m_t|s=s_t) &= \Pr(m(s, e) \leq m_t | s = s_t) \\
 &= \Pr(e \leq m^{-1}(s, m_t) | s = s_t) \quad (\text{monotonicity}) \\
 &= \Pr(e \leq m^{-1}(s_t, m_t)) \quad (\text{independence}) \\
 &= F_e(e_t)
 \end{aligned}$$

Solution 1: Assume F_e (e.g. uniform) gives us both $m(\cdot)$ and e point-wise.

Solution 2:

- Homogeneity: $m(\alpha s, \alpha e) = \alpha m$
- Suppose we know $m(\alpha s^*, \alpha e^*) = \alpha m^*$, some (m^*, s^*, e^*)
- Then: $F_e(\alpha e^*) = F_{m|s}(\alpha m^* | s = \alpha s^*)$
- And vary α : set $1 = m(1, s^*)$, smallest s^* such that in all markets $m_i \leq e_i$ (conservative w.r.t. search frictions).

Counterfactual

“Counterfactual”: open up Northwest Passage.



Counterfactual

- North America sees exporting increase.
- China/Japan-Korea exports unaffected: ships' outside option higher and exports do not increase.
- Other countries: also affected by *distant and local* shock:
 - Higher outside option of ships: increases price, decreases exports.
- Spatial model: both trade flows and trade costs are equilibrium objects