Delivery in the city: evidence on monopolistic competition from New York restaurants

Nathan Schiff[1] Jacob Cosman[2] Tianran Dai[1]

[1]Shanghai University of Finance and Economics
[2]Zillow

December 2019

Product Differentiation and the Response to Competition

Research question: how do firms respond to new competition in markets with many firms and differentiated products?

Theoretical prediction depends on model:

- Monopolistic competition (Dixit Stiglitz) implies no strategic competition, thus no direct response
- Spatial competition (Hotelling, Salop) suggests strategic response from close competitors

This paper: examine response of existing restaurants to competition from new entrants by measuring *menu changes*

Overview and Preview of Results

New longitudinal dataset from ∼11,700 restaurants in New York City over 68 consecutive weeks (∼550,000 menus)

DiD Matching strategy: compare treated restaurant's menu before and after new competition with matched control restaurant's menu

Examine competition in both geographic and product space, and measure response across large set of outcomes (menu changes, restaurant characteristics)

Intended contribution: provide first evidence on causal effects of local competition in large differentiated markets

Results:

- Treated and control restaurants change prices and products frequently but *no evidence* of competitive response (se small)
- However, restaurants in top decile of entry frequency (23 median entrants) 5% more likely to exit market after one year than lowest decile (0 median entrants)

Local vs Global Competition

"In markets characterized by monopolistic competition, market power is accompanied by a low level of strategic interaction, in that the strategies of any particular firm do not affect the payoff of any other firm...In contrast, in spatial models, even in the limit of a continuum of firms, strategic interaction remains. In that case, firms interact locally, and neighbors count, no matter how large the economy is." MWG 1995

Monopolistic competition serves as a theoretical framework in many fields (Trade, IO, Labor, Urban) and used in analysis of important industries (services, retail)

Further, local vs global competition distinction is important to understand response to input cost shocks.

Restaurant industry is especially relevant as a large employer of minimum wage workers; monopolistic competition implies perfect price pass-through but spatial competition suggests lower profitability (Aaronson and French 2007; Draca, Machin, Van Reenen 2011)

Existing Literature

Most empirical work on differentiated markets focuses on market size effects (Syverson 2004, Campbell and Hopenhayn 2005, Campbell 2011, Hottman 2016)

Smaller literature examines local competition (Netz and Taylor 2002; Pinkse, Slade, and Brett 2002; Pinkse and Slade 2004; Sweeting 2010)

Existing work mostly cross-sectional data, few firms, structural assumptions.

Advantages of our dataset and approach:

1. Panel allows better control for heterogeneity, large number of firms gives much greater variation in treatment and allows for isolation of treated firms (minimize confounders)
2. Menus allow for comprehensive measurement of response to competition (not limited to price) and granular measurement of competition (useful for treatment assignment and matching)

Menu Data

Text of menus from over 11,700 restaurants in NYC (with entry, exit), every week from 11/27/2016 to 3/11/2018 (68 periods)

Data source is large online food delivery service (Grubhub); consumers browse menus and then order and pay using this service

Important feature: when order is placed restaurant is obligated to provide service at prices; ensures prices and items are accurate, allowing us to observe high frequency changes

Observe location, cuisines, item names, item descriptions, prices, promotions, review counts and ratings

Data has significant noise (oscillators/time-of-day effects, extreme values in prices and item counts); missing prices for weeks 23-26, missing item names 44-48

To address noise we use a number of alternative specifications

Example Menu: China King Express (327 items)

[u'55. Chicken Mei Fun', u'Thin noodles.', u'$5.50', u'+', u'Mei Fun'],
[u'56. Beef Mei Fun', u'Thin noodles.', u'$5.75', u'+', u'Mei Fun'],
[u'57. Shrimp Mei Fun', u'Thin noodles.', u'$5.75', u'+', u'Mei Fun'],
[u'58. House Special Mei Fun', u'Thin noodles.', u'$5.75', u'+', u'Mei Fun'],
[u'59. Singapore Chow Mei Fun',
u'Thin noodles. Spicy.',
u'$5.75',
u'+',
u'Mei Fun'],
[u'60. Mixed Vegetables', u'With white rice.', u'$4.95', u'+', u'Vegetables'],
[u'62. Eggplant with Garlic Sauce',
u'With white rice. Spicy.',
u'$4.95',
u'+',
u'Vegetables'],
[u'63. Bean Curd Homestyle',
u'With white rice. ',
u'$4.95',
u'+',
u'Vegetables'],
[u'64. Broccoli with Brown Sauce',
u'With white rice.',
u'$4.95',
u'+',
u'Vegetables'],

Restaurant Characteristics: Levels

	mean	median	sd	min	p1	p99	max	N
item count	124.44	100.00	88.66	10.00	15.0	399.0	500	419782
median item price	8.62	8.00	3.35	2.50	3.0	18.5	25	419782
mean item price	9.40	8.82	3.88	2.28	3.9	22.9	49	419782
min item price	1.59	1.25	1.42	0.00	0.0	8.0	25	419782
max item price	32.52	22.50	49.29	2.99	7.5	190.0	2199	419782
cuisines	4.05	4.00	3.11	0.00	0.0	14.0	35	423214
reviews	380.63	206.00	509.99	1.00	4.0	2326.0	10064	370764
stars	3.72	4.00	1.19	1.00	1.0	5.0	5	395984
food rating	85.30	88.00	9.62	0.00	50.0	100.0	100	406096
order rating	89.61	92.00	9.01	0.00	56.0	100.0	100	406093
delivery rating	86.09	89.00	11.09	0.00	46.0	100.0	100	406079

Stats averaged across restaurant-periods; excludes outliers, oscillators, missing item name periods, missing price periods

Menu Changes and Durations

Restaurants *are* changing online menus

	mean	median	mean dur	med dur	N
item count	8.91	3.00	3.90	1	141666
median price	0.84	0.50	7.67	2	72001
mean price	0.28	0.09	3.69	1	149781
min price	0.96	0.50	30.16	23	18307
p25 price	0.54	0.26	7.54	2	73193
p75 price	0.98	0.50	7.85	2	70363
max price	14.07	3.05	20.86	10	26471

Stats calculated for unique changes specific to each var, cols 1-2 use absolute changes, duration is number continuous periods with no var change

Within Restaurant Changes

$$Y_{rt} = \beta * Weeks_{rt} + \eta_r + \epsilon_{rt}$$

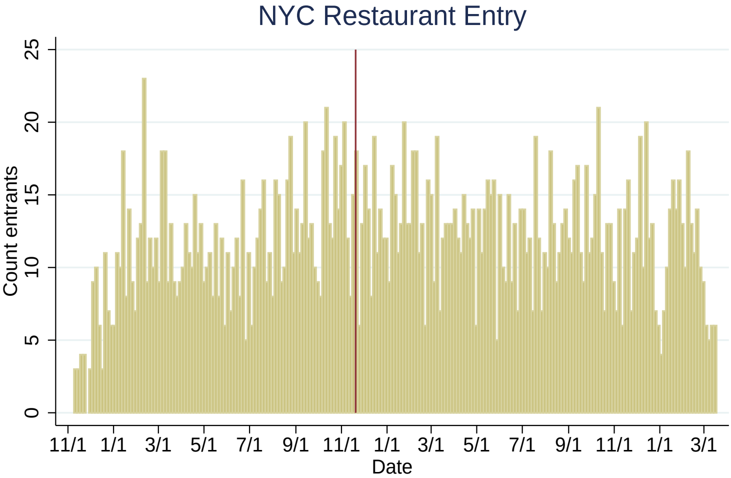
	(1) item ct	(2) p50 item prc	(3) mean item prc	(4) min item prc	(5) max item prc	(6) reviews ct	(7) food rtng
weeks observed	0.0886***	0.0068***	0.0088***	0.0001	0.1142***	5.2740***	-0.0113***
	(0.0048)	(0.0002)	(0.0004)	(0.0001)	(0.0139)	(0.0835)	(0.0009)
Observations	456153	456153	456153	456153	456153	404211	441055
Clusters	11302	11302	11302	11302	11302	10403	10576

Average restaurant increases median item price by \$0.3536 over one year, about 4.1% of average median item price (\$8.62)

Larger changes for more expensive items

Reviews increasing each week (5.3), menu length increases by 4.68 items per year (3.7% of median length)

Entry (data from inspections and Yelp)



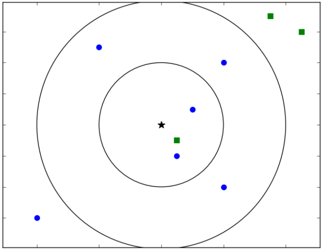
Entry date calculated as earliest of inspection and Yelp date, 11/1/2015-3/17/2018.
Bin width is 4 days; there are 2,585 entrants over the entire period.
Menu data period, 11/27/2016-3/17/2018, right of vertical line.

Defining Treatment: Geography and Timing

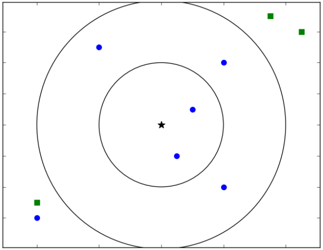
We identify restaurants with *only* one new competitor in a geographic area, within a specified duration; control restaurants have no changes

- Set treatment radius ρ_T, exclusion radius ρ_X, duration d
- Treated at t if exactly one entrant in ρ_T between period $t-1$ and t, nothing else in $\rho_X > \rho_T$, nothing else between $t-2d$ and $t+2d$
- Control at t if no entrants between $t-2d$ and $t+2d$ within ρ_X
- Neither treatment nor control at t if
    - Any entrants at other period between $t-2d$ and $t+2d$
    - Any entrant within ρ_X but outside ρ_T
    - Multiple entrants within ρ_T at t

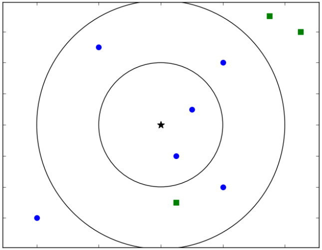
Geographic assignment of treatment and control



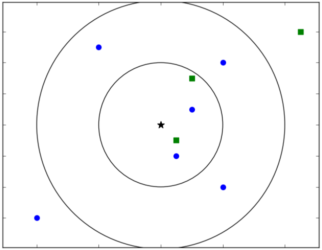
(a) Treatment



(b) Control

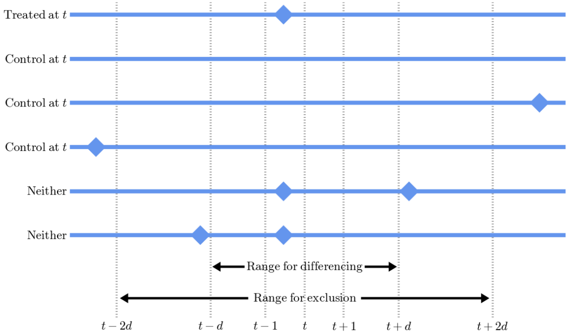


(c) Neither

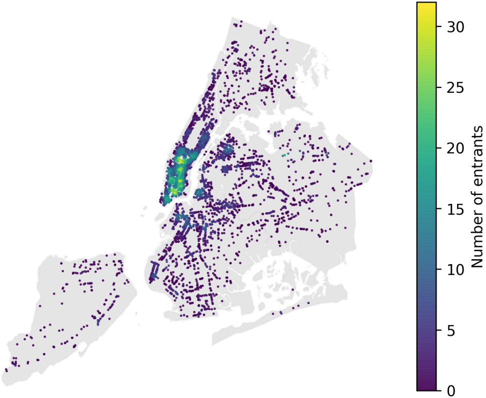


(d) Neither

Treatment Timing



Problem: Entry Far More Likely in Certain Neighborhoods



Demographics *and* restaurant characteristics differ by treatment status (for d=4,6,8) restaurants, locations

Identification strategy: DiD matching

Core idea:

- find very similar pairs of restaurants where one gets new competitor ("treated") nearby and other does not ("control")
- compare pre/post differences in menus for treated and control

$$Y_{rt} = \beta * D_{rt} + \mu_r + \mu_{L_r} + \xi_{rt} + \xi_{L_r t} + \epsilon_{rt}$$

Intuition:

- Restaurant outcomes (price, menu feature) depend on time-invariant (cuisine, neighborhood) and time-varying factors (trends in tastes, neighborhood change)
- We difference to remove time-invariant effects
- We use matching to address time-varying factors: our assumption is that restaurants in similar areas, selling similar items will experience the same trends

Reduced Form Model of Restaurant Competition

Implementing the Matching

Use two-stage "calliper matching" process (Rubin and Thomas, 2000)

Three steps for our implementation:

1. Define treated and control (geography, duration)
2. Matching stage 1: match treated and control based on *predicted entrants*
    - Use a poisson model to predict count of entrants in each location over entire period
    - Demographics (HH characteristics), house price trends, public transit, number of restaurants in area (long before treatment)
    - Define callipers (bins) as 0.25 of sd of log predicted entrants, based on optimizing demographic balance; ensures common support
3. Matching stage 2: Among subset of matched controls, use menu distance to choose restaurant closest to treated

Menu Distance: Measuring Product Differentiation

Second stage of our matching process requires pairing restaurants with similar menus

Matching by cuisine is not sufficient (ex: "American" cuisine)

We convert text of menu (item names, item descriptions) into distribution over groups of letters; we compare menus from two restaurants by comparing their distributions over these letter groups

Allows us to measure how similar two restaurants, as well as how they change

We also use our method to define competitors in *product space* as an extension to main results

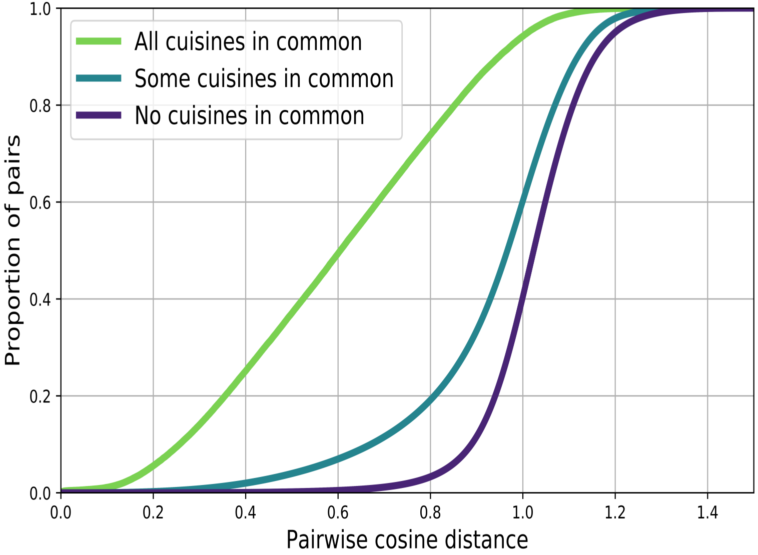
Using text as measure of product differentiation

- Break text down into **n-grams** — consecutive overlapping three-letter strings
  - Top n-grams in data include "chi", "hic", "ick", "cke", "ken"
  - $\approx 20,000$ n-grams appear at least once
- Represent a menu as **vector of n-gram frequency shares**
  - Demean to get deviations from average n-gram frequencies

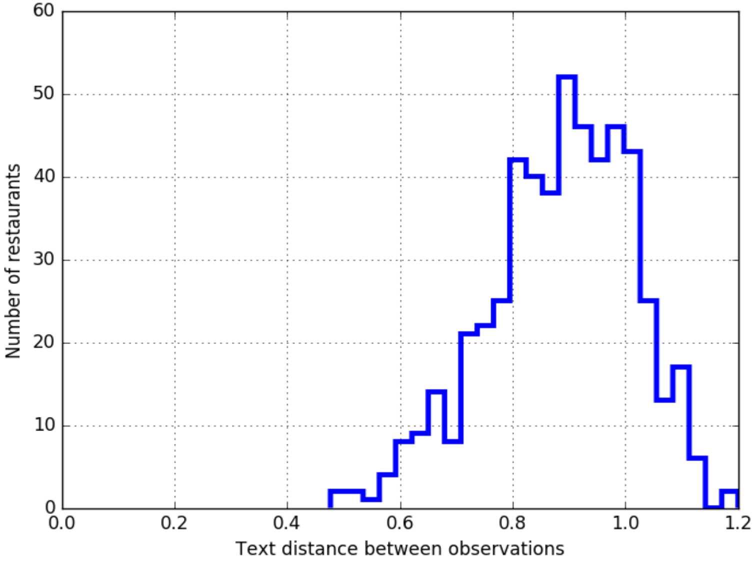
$$S_{ab}^{c} = \frac{\sum_{j=1}^{J}(x_{aj}-\mu_j)(x_{bj}-\mu_j)}{\left(\sum_{j=1}^{J}(x_{aj}-\mu_j)^2 \sum_{j=1}^{J}(x_{bj}-\mu_j)^2\right)^{1/2}} = \cos(\theta)$$

- Angle between frequency vectors for menus a and b is θ_{ij}
  - $\cos\theta_{ij} = 1$ means menus identical, -1 negatively correlated
  - Use cosine distance: $1-\cos\theta_{ij}$, metric over menus
  - Larger $1-\cos\theta_{ij}$ means menus further apart $[0,2]$

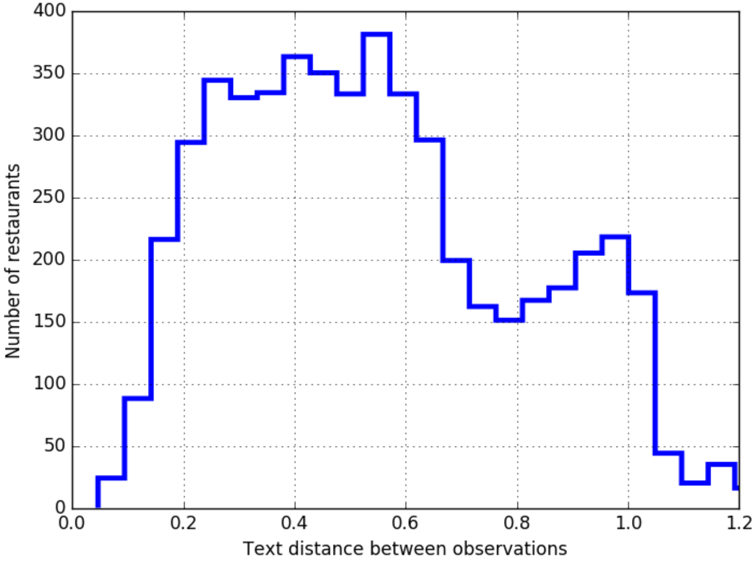
Menu Distance Distributions (CDFs)



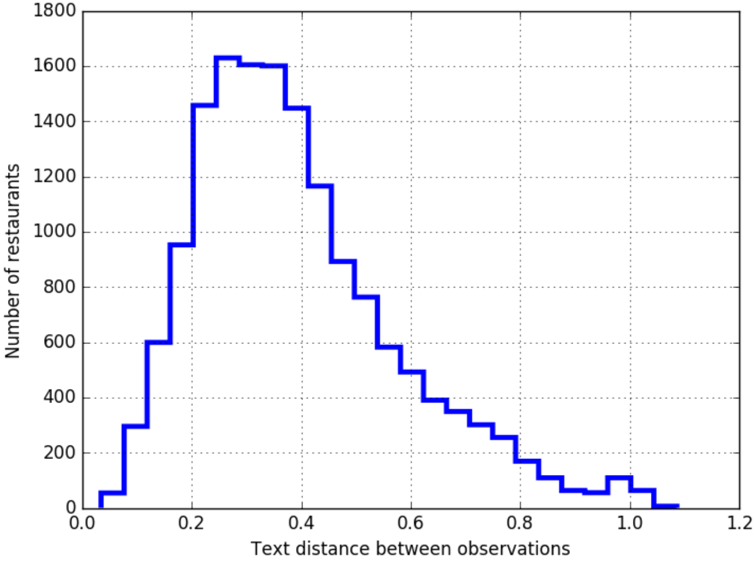
Cuisines: American



Cuisines: Asian, Chinese



Cuisines: Asian, Chinese, Lunch Specials



Post-matching Estimates of Competitive Response

Specifications: unit of obs at 1) restaurant level 2) item level

$$Y_{r,t} = \beta_1 * post_{rt} + \beta_2 * (post_{rt} \times D_{rt}) + \beta_3 * open_{rt} + \eta_h + \eta_r + \varepsilon_{r,t}$$

$$ItemPrice_{i,r,t} = \delta_1 * post_{rt} + \delta_2 * (post_{rt} \times D_{rt}) + \eta_r + \varepsilon_{r,t}$$

No time fixed effects because match exact periods (but included in event-study figures)

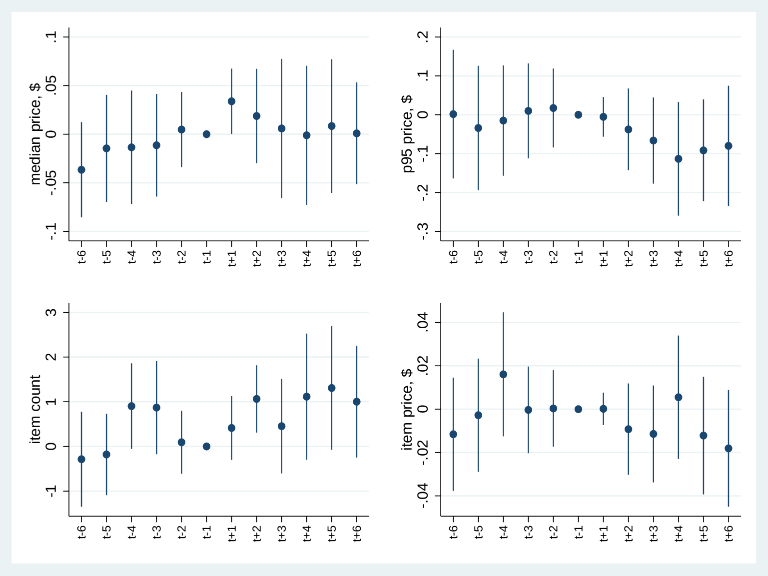
Dealing with noise: drop outliers and missing periods (conservative), include menu-hour and open status fixed effects, run constant item specification

Restrict sample to treatment-control pairs where menu distance is less than 5th percentile; cluster standard errors at level of entrant

Run analysis for 4, 6, and 8 week durations

Location Balance Menu Balance

Event study plot for 6 week duration



FE results for 6 week duration

	(1) Med Prc	(2) p95 Prc	(3) Itm Ct	(4) Itm Prc
treated X post	0.025	-0.065	0.703	-0.009
	(0.024)	(0.081)	(0.491)	(0.007)
post	0.046***	0.140*	0.196	0.048***
	(0.016)	(0.073)	(0.261)	(0.006)
open	-0.031**	-0.014	1.824***	
	(0.015)	(0.028)	(0.317)	
Observations	12815	12815	12815	2328974
Clusters	222	222	222	224
Treated	922	922	922	933
DepVarMean	8.19	17.80	154.06	8.58

Similar results for 4 and 8 weeks: significant changes for both treated and control (post), but no differential effects

Additional Results, Robustness

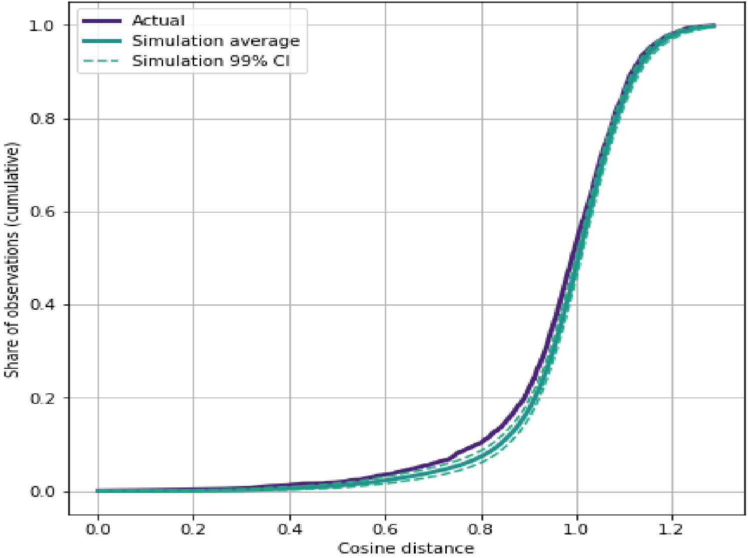
Robustness:

- Additional outcomes: service quality, hours, review count, listed cuisines; no meaningful effects (decrease in review growth for d=4, only)
- Duration issues: long differences, shifted analysis where pre is $[0, d-1]$ and post is $[d+1, 2d]$. Again, no meaningful effects and similar post coefficients
- Heterogeneity: include interactions with count of local competitors, but find no differential response

Analysis of entrant locations: use monte-carlo simulation to compare menu distance of actual entrant locations to random permutations

Interestingly, restaurants of more similar cuisines locate slightly *closer* than expected

Entrant Location Choice: Menu-distance CDF



Extension: Competition in Product Space

Competition in product space (cuisine) may be more important than geographic space

We re-run analysis using menu distance to define treatment:

- Define entry as entry onto the Grubhub delivery site
- Treatment defined as entrant within 2nd percentile of pairwise menu distance distribution, restricted to restaurants within 1.5k (rough delivery radius)
- Match first using control set within 2nd menu distance percentile, then choose restaurant with closest predicted entrant count

Product Space: FE results for 6 week duration

	(1) Med Prc	(2) p95 Prc	(3) Itm Ct	(4) Itm Prc
treated X post	0.033	0.078	-0.098	0.024
	(0.024)	(0.157)	(0.433)	(0.016)
post	0.044***	0.168	0.386	0.030***
	(0.016)	(0.126)	(0.283)	(0.005)
open	0.009	-0.055	2.965***	
	(0.017)	(0.052)	(0.576)	
Observations	8485	8485	8485	1759042
Clusters	347	347	347	348
Treated	679	679	679	682
DepVarMean	8.52	18.83	166.95	9.01

Post coefficients fairly similar to geographic specifications

Restaurants Exiting Online Site

While restaurants don't appear to be strategically responding to entry, it's still possible entry affects likelihood of exit

We don't observe actual exit from the market, but we can proxy for this with exit from the delivery site

However, we can't use our treatment-control methodology to look at exit because this is a permanent outcome, and restaurants change treatment status over time

Instead, we look at how the entrant intensity of an area affects the likelihood of exit: are restaurants locating in areas with high entrant intensity more likely to exit?

To avoid reverse correlation (entrant replaces exiter) we use pre-period entrants to calculate entrant intensity

Identification Strategy: Generalized Propensity Score

Endogeneity: we already know that restaurants in areas with high entry intensity look different

We therefore want to match restaurants by entrant intensity, *but* also allow treatment to be continuous: the number of entrants in area

Generalized Propensity Score (Hirano and Imbens, 2004): matches observations by likelihood of receiving a particular value of the continuous treatment

With this technique we can estimate a "dose-response function": how do different entrant intensities (doses) affect the likelihood of exit?

Estimation of Dose-Response with GPS

We predict *pre-period* entrant intensity (λ_r, count in 54 weeks) in every location with another poisson regression, likelihood of any particular count of entrants n_r over the pre-period is:

$$GPS_r(n) = Pr(n|\lambda_r) = \frac{\lambda_r^n e^{-\lambda_r}}{n!}$$

We model the hazard of exit using a Cox model:

$$\phi_r(t|n_r) = \phi_0(t) * exp(\gamma * n_r)$$

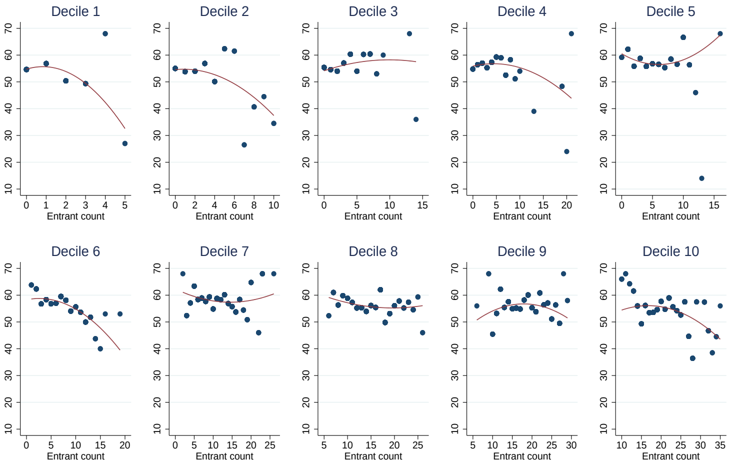
Hazard of exit conditional on the GPS (flexible form):

$$\phi_r(t|n_r) = \phi_0(t) * exp(\gamma_1 * n_r + \gamma_2 * GPS_r(n_r))$$

We estimate the dose-response function using average of predictions for given entrant n (dose) level:

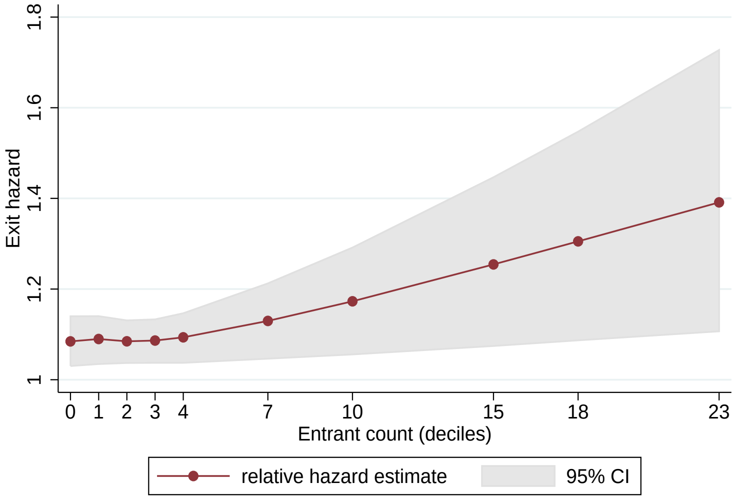
$$E[\phi_r(t|n)/\phi_0(t)] = \frac{1}{R} \sum_r (exp(\hat{\gamma}_1 * n + \hat{\gamma}_2 * GPS_r(n)))$$

Survival vs pre-period entrant count (predicted deciles)



Survival time in weeks graphed against pre-period entrant count, by predicted entrant decile.
Each point represents mean survival time for restaurants with the same entrant count.
Lines show quadratic fit with entrant count bins weighted by number of restaurants.
Sample restricted to restaurants surviving at least 10 weeks and in common support.

Effect of Entrant Intensity on Exit Hazard



Relative hazard plotted at median of entrant count deciles.
95% confidence interval calculated from 1000 bootstrap samples.

Summary of Results

- Restaurants faced with a new entrant (treated) do not change prices or menu characteristics differently from a control group
- This "zero effect" is not the result of a general lack of change; for many measures we observe statistically significant changes over time. However, changes are roughly equal for treated and control.
- Our results are consistent across a large number of outcomes and robustness specifications; they do not seem to be driven by strategic entry choices (monte-carlo exercise)
- We do observe that restaurants locating in areas with higher entrant intensity are more likely to exit. Using the estimated DRF, restaurants in 80.5% of restaurants in first entry decile survive one year but only 75.6% in highest entry decile survive

Concluding Discussion

In this paper we provide some of the first estimates of the response to local competition in a large market with differentiated firms

We use a matching strategy that takes advantage of location characteristics and distance between products, measured using text; this methodology could be useful in other studies where product differentiation is described by text (real estate, investment prospectuses, politica candidates)

Our results are consistent with canonical models of monopolistic competition, as well as recent work showing effect of preference heterogeneity on competition (Gabaix et. al. 2016)

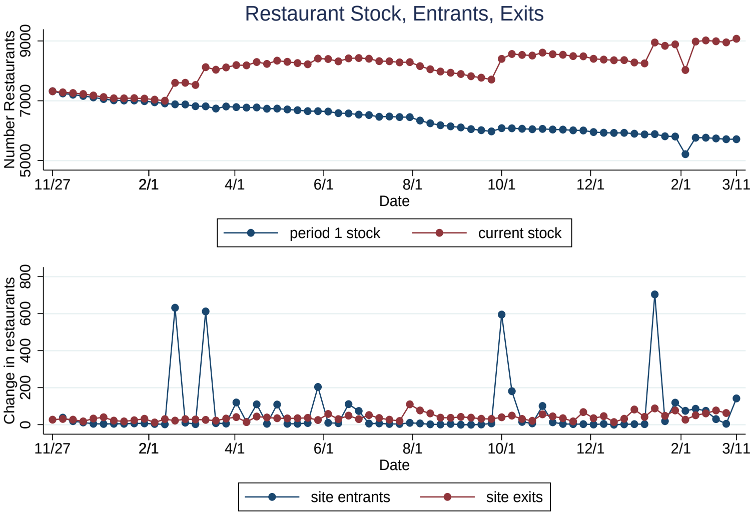
However, results suggest a puzzle: why don't restaurants respond if new competition affects profits (exit)?

Empirical work on endogeneous product differentiation could be useful to better understand this puzzle

END SLIDES

Thank you!

Sample by Week



Data is 68 wks, 11/27/2016-3/11/2018; entrants not defined in first period, exits not defined in last period.

Determining Entry with Inspections and Yelp

Problem: earliest appearance on delivery site is not entry date

Many restaurants have been in existence long before submitting menu to site (Yelp, inspections)

Unfortunately, inspection dates very noisy, permit applications not linked to restaurant address

Defining entry date:

- Start with all restaurants whose first inspection occurs during sample period (NYC DOHMH)
- For each restaurant, find date of earliest Yelp review
- Take subset where -20 $<$inspect date-yelp date$<$ 100; range is adhoc based on histogram
- Earliest of two dates we define as entry date

Are Treated and Control Restaurants Comparable? Restaurant Characteristics

	Menu stats	
	t-tests	N
item count	-9.88***	126233
mean item price	0.16*	126233
median item price	0.18**	126233
p25 item price	0.10*	126233
p95 item price	0.14	126233
stars	0.04	121925
review count	24.56*	112771
order rating	0.54**	123590
food rating	0.35	123589
delivery rating	0.84***	123590

Tests difference between treated and control.
Calculated using values 4 periods before treatment.
Sample excludes outliers and missing price periods.

Back

Are Treated and Control *Locations* Comparable? Location Characteristics

	Demographics	
	t-tests	N
age.25.29	0.015***	126813
age.30.39	0.017***	126813
age.70.79	-0.001**	126813
race.white	0.063***	126813
race.black	-0.038***	126813
hh.family	-0.058***	126813
hh.married	-0.026***	126799
educ.degree	0.080***	126799
poverty	-0.015***	126799
income.100.150	0.005***	126799
income.150.200	0.005***	126799
unit.detached	-0.043***	125600
competitors 500m	10.694***	116750

Tests difference between treated and control.
All demographics calculated as percent of area.
Competitors calculated 4 periods pre-treatment.
Sample excludes outliers and missing price periods.

Post-Match Location Balance

Variable		Q1	Q2	Q3	Q4	Q5
Competitors within 100 m	Without callipers	1.36	0.74	0.02	0.82	1.13
	With callipers	0.22	0.06	0.05	0.09	0.06
Competitors within 500 m	Without callipers	2.60	0.98	0.11	1.29	1.93
	With callipers	0.75	0.02	0.12	0.05	0.31
Competitors within 1 km	Without callipers	2.44	0.91	0.03	1.29	1.95
	With callipers	0.69	0.07	0.03	0.07	0.24
One-bedroom rent	Without callipers	1.39	1.02	0.11	1.23	1.20
	With callipers	0.07	0.07	0.11	0.08	0.20
Two-bedroom rent	Without callipers	1.38	1.01	0.11	1.23	1.19
	With callipers	0.07	0.06	0.11	0.08	0.20
White	Without callipers	0.80	0.64	0.20	0.91	0.54
	With callipers	0.01	0.05	0.08	0.14	0.37
Black	Without callipers	0.50	0.44	0.02	0.68	0.63
	With callipers	0.02	0.09	0.03	0.26	0.36
Asian	Without callipers	0.14	0.05	0.27	0.07	0.53
	With callipers	0.08	0.02	0.13	0.04	0.37
Latino	Without callipers	0.58	0.57	0.13	0.69	0.92
	With callipers	0.09	0.02	0.05	0.06	0.06
Family household	Without callipers	1.85	0.86	0.11	0.96	1.79
	With callipers	0.43	0.06	0.02	0.08	0.26
Married household	Without callipers	0.92	0.38	0.13	0.40	1.24
	With callipers	0.21	0.06	0.01	0.17	0.39
Enrolled in college	Without callipers	0.32	0.19	0.21	0.16	0.71
	With callipers	0.09	0.03	0.03	0.25	0.45
College graduate	Without callipers	1.71	0.90	0.04	1.19	1.30
	With callipers	0.28	0.01	0.02	0.03	0.15
Poverty	Without callipers	0.60	0.55	0.06	0.81	0.53
	With callipers	0.02	0.04	0.05	0.17	0.32

Post-Match Menu Balance

Variable		Q1	Q2	Q3	Q4	Q5
Median price	Before matching	0.24	0.17	0.13	0.17	0.14
	After matching	0.02	0.32	0.17	0.04	0.09
95th perc price	Before matching	0.24	0.16	0.17	0.19	0.17
	After matching	0.11	0.18	0.01	0.13	0.05
Item count	Before matching	0.17	0.16	0.18	0.24	0.19
	After matching	0.15	0.18	0.27	0.37	0.30
Quality	Before matching	0.22	0.14	0.19	0.14	0.18
	After matching	0.03	0.12	0.17	0.11	0.18
Timeliness	Before matching	0.20	0.16	0.20	0.17	0.18
	After matching	0.13	0.11	0.07	0.01	0.04
Accuracy	Before matching	0.18	0.13	0.21	0.16	0.16
	After matching	0.00	0.09	0.17	0.17	0.13
Cuisines Jaccard	Before matching	0.91	0.92	0.92	0.93	0.93
	After matching	0.62	0.58	0.63	0.66	0.72
Cuisines equal	Before matching	0.01	0.00	0.01	0.01	0.01
	After matching	0.05	0.14	0.10	0.07	0.07
Cuisines subset	Before matching	0.10	0.07	0.06	0.05	0.06
	After matching	0.49	0.52	0.41	0.32	0.28

Results

FE Results: d=4 weeks

	(1) Med Prc	(2) p95 Prc	(3) Itm Ct	(4) Itm Prc
treated X post	0.006	-0.025	0.448*	-0.007
	(0.016)	(0.060)	(0.264)	(0.005)
post	0.027**	0.074	-0.004	0.030***
	(0.011)	(0.056)	(0.135)	(0.004)
open	-0.030***	-0.012	1.660***	
	(0.010)	(0.019)	(0.199)	
Observations	19016	19016	19016	3383522
Clusters	285	285	285	311
Treated	1668	1668	1668	1811
DepVarMean	8.40	17.95	148.57	8.69

FE Results: d=8 weeks

	(1) Med Prc	(2) p95 Prc	(3) Itm Ct	(4) Itm Prc
treated X post	0.018	-0.038	0.560	0.000
	(0.023)	(0.098)	(0.578)	(0.010)
post	0.047***	0.095	0.057	0.048***
	(0.016)	(0.075)	(0.287)	(0.008)
open	-0.017	0.044	1.780***	
	(0.015)	(0.048)	(0.424)	
Observations	8116	8116	8116	1462892
Clusters	148	148	148	150
Treated	498	498	498	502
DepVarMean	8.11	17.19	158.69	8.53

Other Results: d=6 weeks

	(1) Food Rtng	(2) Delivery Rtng	(3) Order Rtng	(4) Wkly Hrs	(5) Num Cuisines	F
treat_post	0.040	-0.034	-0.023	0.123	-0.041	
	(0.070)	(0.065)	(0.050)	(0.652)	(0.050)	
Observations	15488	15488	15488	15484	15768	
Clusters	223	223	223	224	224	
Treated	934	934	934	935	935	
DepVarMean	86.62	87.41	90.87	60.12	4.39	

Extended duration FE results: d=6 weeks

	(1) Med Prc	(2) p95 Prc	(3) Itm Ct	(4) Itm Prc
treated X post	-0.016	0.057	0.285	0.012
	(0.019)	(0.068)	(0.349)	(0.019)
post	0.048***	0.079***	0.150	0.033***
	(0.012)	(0.020)	(0.234)	(0.006)
open	-0.045***	-0.049	1.543***	
	(0.013)	(0.053)	(0.331)	
Observations	9208	9208	9208	1648274
Clusters	193	193	193	211
Treated	739	739	739	861
DepVarMean	8.23	17.49	156.90	8.58

Reduced Form Model of Restaurant Competition

Outcomes Y_{it} for restaurant i in neighborhood $j(i)$ at time t:

$Y_{it} = \beta_i * D_{it} + \mu_i + \mu_{j(i)} + \xi_{it} + \xi_{j(i)t} + \epsilon_{it}$

- D_{it} indicates a single new competitor has entered (we focus on cases of single entry over a duration d)
- Time-invariant terms μ_i,μ_j capture restaurant and neighborhood effects (French, or expensive neighborhood)
- Time-varying terms $\xi_{it} + \xi_{j(i)t}$ could be period of poor management, fad for i's cuisine, gentrification

Let $Y_{it}^1 = \beta_i + Y_{it}^0$ and $Y_{it} = D_{it} Y_{it}^1 + (1 - D_{it}) Y_{it}^0$

Then $ATT = E[Y_{it}^1 - Y_{it}^0 | D_{it} = 1] = E[\beta_i | D_{it} = 1] = \beta$

Competition as Selection

$$Y_{it} = \beta_i * \mathbb{I}\{t \geq k\} * D_{it} + \mu_i + \mu_{j(i)} + \xi_{it} + \xi_{j(i)t} + \epsilon_{it}$$

$$D_{it} = \mathbb{I}\left\{\theta_i + \theta_{j(i)} + \psi_{ik} + \psi_{j(i)k} \geq 0\right\} \mathbb{I}\{t \geq k\}$$

If any competition terms are correlated with restaurant behavior terms then it cause a selection problem and biases β

First step is to estimate using differences around treatment period

Let $\triangle Y_{i,k+\tau} = Y_{i,k+\tau} - Y_{i,k-\tau}$

$ATT = E[\triangle Y^1_{i,\tau} - \triangle Y^0_{i,\tau} | \triangle D_{i,\tau} = 1] = \beta$

This removes correlation of D_{it} and μ terms

Matching to Remove Selection on Time Varying Shocks

$$Y_{it} = \beta_i * \mathbb{I}\{t \geq k\} * D_{it} + \mu_i + \mu_{j(i)} + \xi_{it} + \xi_{j(i)t} + \epsilon_{it}$$

Second step is to use matching to address time-varying shocks

We match using a propensity score based on area characteristics, $P(X_{j(i)t})$ and menu distance M_i

The key identifying assuming is conditional mean independence:

$$E[\triangle Y^0_{ik} | P(X_{jt}), M_i, \triangle D_{i\tau} = 1] = E[\triangle Y^0_{ik} | P(X_{jt}), M_i, \triangle D_{i\tau} = 0]$$

In words: conditional on matching, counterfactual change in price for treated (had they not been treated) is equal to change in price for those not treated (in expectation)

Back

Hazard Conditional on GPS

	(1) surv. time	(2) surv. time	(3) exit haz.	(4) exit haz.	(5) exit haz.	(6 exit
observed entrants	-0.0198	-0.1952***	0.0029	0.0108***	0.0140***	-0.0
	(0.0237)	(0.0640)	(0.0031)	(0.0039)	(0.0051)	(0.0
predicted entrants		0.2025***				
		(0.0686)				
GPS				0.6198***	0.6080***	-0.1
				(0.1773)	(0.1768)	(0.7
obs. ents. X GPS					-0.0624	0.0
					(0.0675)	(0.0
obs. ents.^{2}						0.00
						(0.00
GPS^2						0.6
						(0.9
Observations	9310	9310	9310	9310	9310	93
Likelihood	-39855.1	-39850.7	-15718.4	-15712.6	-15712.2	-157