Learning from consumer travel distance: Firm expansion, contraction, and exit

Jeongwon Jang [1] Daisoon Kim [2] Yoonsoo Lee [1]

[1]Sogang University

[2]NC State University

VEAMS - Virtual East Asia Macroeconomics Seminar Series, June 3, 2022

Motivation

Important role of **customer acquisition** in firm's life cycle growth dynamics

- See e.g. Arkolakis (2010, 2016), Foster et. al. (2016), Eaton et. al. (2021), Einav et. al. (2021)
- Understand the firm growth in the context of customer-firm interaction.

Market size and consumer travel distance

Which market traits are relevant to firm growth?

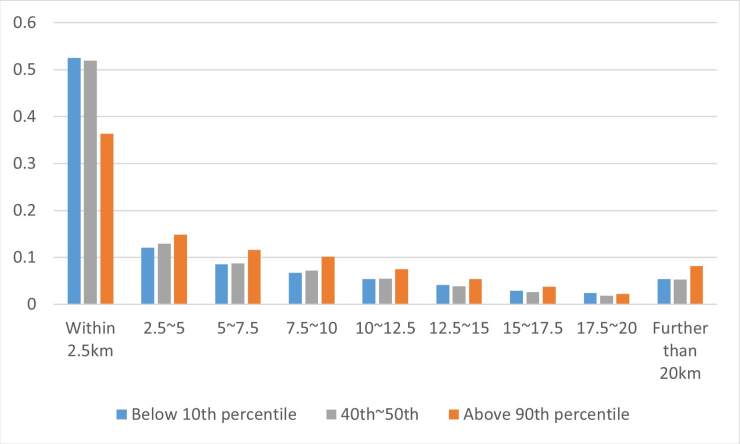
- Traditional market size measures: population, income, total sales, etc.
- more recently: floating population, card transactions, etc.

Travel distance of consumers could be a important margin of **market size** that affects firm's customer acquisition process (i.e. growth)

In non-tradable service sectors, **"trade"** is generated by **shopping trips** of customers

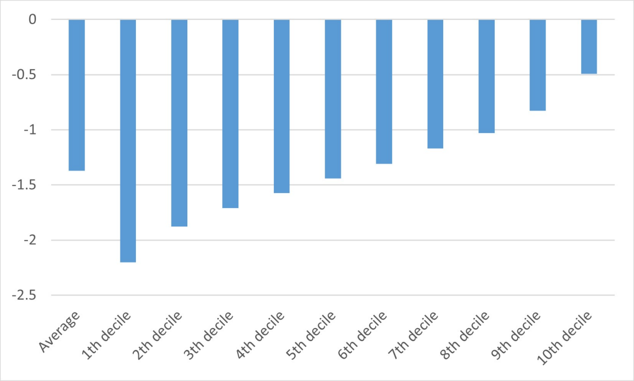
- In urban areas, a large amount of transaction comes from non-residents who may live far away.
- **The size and the composition of customer pools** are strongly related to **how far consumers are willing to travel to the market**

Larger markets have higher sales share from remote customers.



Share of consumption for each shopping distance group by market size measured by total sales.

Strong gravity in large markets



Point estimation on distance coefficients

Notation, notes, and remarks:

- All coefficients are from estimations of the gravity equation where dependent variable is the sales from each customer address and independent variables are Euclidean distance between the market location and customer address, total sales for the market, and total consumption for the customer address. All point estimates all statistically significant at the 1% level.

What we do

Provide a new perspective on market size using credit card transaction data

- Sales weighted average of **travel distance** for each market

  Unit of market: pair of 4digit KSIC-8digit township

  e.g., Convenience stores in town A, Korean restaurants in town A, Korean restaurants in town B

Empirically investigate the relationship between consumer travel distance and firm growth

Build a theoretical framework to understand the empirical findings

Key findings

We find that the **growth rates** of non-tradable service firms are positively correlated with **consumer travel distance**, after controlling various variables including **total sales** in the market

- Longer travel distance $\Rightarrow$ rapid growth of expanding firms
- Longer travel distance $\Rightarrow$ rapid shrink of contracting firms
- Young firms show more rapid growth and contraction in markets with longer travel distance.
    - but the travel distance does not matter for old firms
- Exit probability is higher in markets with longer travel distance.

These findings can be rationalized by firm learning from customers with different travel distances.

- Lower uncertainty on a shop's product makes consumers willing to travel longer distance.
- A firm learns better from long-distance consumers (i.e., low uncertainty or high precision of signals).

Literature review

Importance of demand side factors in firm growth

- Understanding firm growth as customer acquisition process [Arkolakis ‘10, ’16, Gourio Rudanko, 14; Foster et. al. ‘16, Einav et. al. ‘21, Eaton et. al. ‘21]
- Emphasizing demand appeal in firm growth [Foster et. al. ‘08, Hottman et. al. ‘16, Eslava Haltiwanger ‘20, Ignaszak Sedlácek ‘21, Kehrig Vincent ‘21]

Efficiency gain of large market size.

- Stronger selection [Syverson ‘04, Campbell Hopenhayn ‘05, Melitz Ottaviano ‘08]
- Innovation and productivity enhancement[Atkeson Burstein‘10, Lileeva Trefler‘10, Bustos ‘11]

Data and Measurements

Credit card transaction data

Basic information

- Monthly credit card transaction data compiled by Statistics Korea
- The card company covers 35 million individual users, accounting for about two-thirds of the total population
- 1.4 billion observations from Oct 18 to Dec 19
- We focus on transactions in retail and restaurants located in Seoul brought by customers in Seoul metropolitan area (Seoul city and Gyeonggi province)

Data structure

- Transactions are grouped as disaggregated cells defined by combination of supply-side info (Industry, location) + Customer-side info (Address, sex, age, etc.)
- Each cell contains transaction amount, transaction count, number of cards.
- We define market as a pair of Industry i (4digit KSIC) – Location l (8digit township).
- For each market, we grouped transactions by customer's address j (8digit township).

Measuring customer travel distance

We measure customer travel distance as sales weighted average of Euclidean distance between market location l and customer address j.

Travel distance measurement: For each market (pair of i and l)

$$ConsDist_{i,l} = \sum_{j \in L} s_{i,l,j} Distance_{l,j}, \quad \text{where} \quad s_{i,l,j} = \frac{Sales_{i,l,j}}{\sum_{j \in L} Sales_{i,l,j}}$$

- We average 15 months of *ConsDist*

Summary statistics

	Variables	Mean	Sd	P25	P50	P75	obs
	ConsDist (km)	5.33	2.10	3.89	4.95	6.38	5245
Market-level	TotalS (Million won)	1636.8	4258.8	217.41	505.11	1363.2	5245
	Avg # of Establishments in the market	28.40	49.25	6	12.52	30.90	5245
	Growth	0.11	0.45	-0.09	0.02	0.19	117222
Firm-level	AvgRevenue (Millon won)	395.20	1989.01	43.50	126.50	331	117222
	Age	7.75	7.49	2	5	11	117222
	POP (Thousand people)	23.36	9.07	17.46	22.66	28.70	424
Location-level	FloatPop (Thousand people)	201.65	150.27	99.57	154.57	249.74	424
	AvgIncome (Millon won)	35.88	9.44	30.11	32.70	37.82	424

Does Consumer Travel Distance matter?

Property of ConsDist

Table 1: Relationship between ConsDist and shopping behaviors

VARIABLES	(1) log(ConsDist)	(2) log(ConsDist)	(3) log(ConsDist)	(4) log(ConsDist)
log(# of cards)	0.117***	0.127***	0.129***	0.0883***
	(0.004)	(0.004)	(0.004)	(0.004)
log(Avg Volume)	0.183***		0.171***	0.128***
	(0.010)		(0.010)	(0.008)
log(Avg visit per card)		-0.521***	-0.416***	-0.329***
		(0.046)	(0.046)	(0.034)
Observations	5,335	5,335	5,335	5,335
R-squared	0.372	0.35	0.382	0.739
Fixed effect	Industry	Industry	Industry	Industry Township

Relationship between consumer travel distance and total sales

We now focus on the relationship between the consumer travel distance measure and total sales

- Sales variation across markets: Average over 15 months

$$\log(\mathit{TotalS}_{i,l}) = \beta_1 + \beta_2 \log(\mathit{ConsDist}_{i,l}) + \beta_3 \log(\mathit{POP}_l) + \beta_4 \log(\mathit{FloatPop}_l) + \beta_4 \log(\mathit{AvgIncome}_l) + \psi_i + \varepsilon_{i,l}$$

- Sales variation within a market over time

$$\log(\mathit{TotalS}_{i,l,t}) = \beta_1 + \beta_2 \log(\mathit{ConsDist}_{i,l,t}) + \beta_3 \log(\mathit{FloatPop}_{l,t}) + \psi_i \times \phi_l + \varepsilon_{i,l,t}$$

Distance matters for total sales variation

Table 2: Relationship between ConsDist and total sales in the market

VARIABLES	(1) log(Total sales)	(2) log(Total sales)	(3) log(Total sales)	(4) log(Total sales)
log(ConDist)	0.501***	0.728***	0.176***	0.177***
	-0.0552	-0.0417	-0.00458	-0.00457
log(POP)	0.101***	0.122***		
	-0.0378	-0.0248		
log(FloatPop)	0.571***	0.674***		0.0744***
	-0.0317	-0.0213		-0.00744
log(AvgIncome)	-0.0281	-0.182***		
	-0.0905	-0.0591		
Observations	5,777	5,777	91,605	91,605
R-squared	0.118	0.627	0.971	0.971
Fixed effects	No	Industry	Market	Market

Firm Growth and Consumer Travel Distance

Firm level growth

Firm growth measure:

- For establishment s in year t,

$$Growth_{s,t} = \frac{Revenue_{s,t} - Revenue_{s,t-1}}{\frac{1}{2}(Revenue_{s,t} + Revenue_{s,t-1})}$$

Firm level data from SBR (Statistical Business Register) from Statistics Korea

- Tax data and Census on Establishments from 2018-2019
- Use observations that also appear in Census on Establishments (census + revenue data)
- Merge with credit card transaction data by industry and location
- We restrict the sample into continuers.

Relationship between consumer travel distance and firm growth

$$
\begin{aligned}
Growth_{s,i,l} =& \beta_1 + \beta_2 \log(ConsDist_{i,l}) + \beta_3 \log(TotalS_{i,l}) + \beta_4 \mathbb{1}\{Growth > 0\} \times \log(ConsDist_{i,l}) \\
&+ \beta_5 \mathbb{1}\{Growth < 0\} \times \log(ConsDist_{i,l}) + \beta_6 \mathbb{1}\{Growth > 0\} \times \log(TotalS_{i,l}) \\
&+ \beta_7 \mathbb{1}\{Growth < 0\} \times \log(TotalS_{i,l}) + X_{s,i,l}\gamma + Age_{s,i,l} + \psi_i + \phi_l + \varepsilon_{s,i,l},
\end{aligned}
$$

s, i, l refer to firm, industry, and location respectively

Dummy for expanding or contracting firms

Control variables

- Average number of firms in the market in 2018 and 2019
- Firm's average revenue of 2018 and 2019
- Floating population of the market location
- Population of the market location
- Average income of the market location
- $Age_{s,i,l}$=Categorized as 4 bins: {1, 2-5, 6-10, 11-}

Firm growth and ConsDist

Table 3: Firm growth and consumer travel distance

VARIABLES	(1) Growth	(2) Growth	(3) Growth	(4) Growth	(5) Growth	(6) Growth
log(ConsDist)	0.0127	0.00578	-0.00716	-0.0258***	0.00333	-0.0242***
	(0.016)	(0.005)	(0.008)	(0.005)	(0.012)	(0.008)
log(TotalS)	0.0289***	0.0122**	0.00430*	0.00584**	0.00307	0.0042
	(0.005)	(0.005)	(0.002)	(0.003)	(0.003)	(0.003)
$\mathbb{1}\{\text{Growth}>0\}\times$ $log(ConsDist)$		0.107***		0.109***		0.109***
		(0.017)		(0.017)		(0.017)
$\mathbb{1}\{\text{Growth}<0\}\times$ $log(ConsDist)$		-0.0923***		-0.0755***		-0.0753***
		(0.025)		(0.025)		(0.026)
$\mathbb{1}\{\text{Growth}>0\}\times$ $log(TotalS)$		-0.0134**		-0.0156**		-0.0157**
		(0.006)		(0.006)		(0.006)
$\mathbb{1}\{\text{Growth}<0\}\times$ $log(TotalS)$		0.0210***		0.0188***		0.0189***
		(0.003)		(0.003)		(0.003)
Observations	117,222	117,222	117,222	117,222	117,222	117,222
R-squared	0.004	0.378	0.274	0.509	0.278	0.511
Fixed effects	No	No	Age Industry	Age Industry	Age Industry Township	Age Industry Township

Shopping trip vs Stopping by

Customers may travel far to certain markets because of trip-chaining (not for the purpose of shopping only)

- Case1: Long distance commuters stop by the coffee shop near the work place.
- Case2: Long distance shoppers (for jewelry) stop by coffee shop nearby.
- In this case, customers travel distance may be overstated.

Our approach to tackle this

- Use Relative ConsDist compared to ConsDist of convenience stores (CVS) located in the same location.

$$Rel\_ConsDist_{i,l} = \frac{ConDist_{i,l}}{ConDist\_ConvStore_l}$$

- CVS's ConsDists is a good measure for trip chaining because customers barely travel far to visit certain CVS.

Shopping trip still matters

Table 4: Relationship between firm growth and relative travel distance (compared to CVS's)

VARIABLES	(1) Growth	(2) Growth	(3) Growth	(4) Growth
log(Rel_ConsDist)	0.00329	-0.0370***		
	(0.011)	(0.010)		
log(ConsDist)			0.000474	-0.0261***
			(0.011)	(0.006)
log(TotalS)	0.00306	0.00560*	0.00372	0.00589**
	(0.003)	(0.003)	(0.002)	(0.003)
$\mathbb{1}\{\text{Growth}>0\}\times$ log(Rel_ConsDist)		0.144***		
		(0.017)		
$\mathbb{1}\{\text{Growth}<0\}\times$ log(Rel_ConsDist)		-0.0907***		
		(0.025)		
$\mathbb{1}\{\text{Growth}>0\}\times$ log(ConsDIst)				0.110***
				(0.017)
$\mathbb{1}\{\text{Growth}<0\}\times$ log(ConsDIst)				-0.0753***
				(0.025)
$\mathbb{1}\{\text{Growth}>0\}\times$ log(TotalS)		-0.0112*		-0.0158**
		(0.006)		(0.006)
$\mathbb{1}\{\text{Growth}<0\}\times$ log(TotalS)		0.0148***		0.0188***
		(0.003)		(0.003)
Observations	117,095	117,095	117,095	117,095
R-squared	0.274	0.508	0.274	0.509
Fixed effects	Age Industry	Age Industry	Age Industry	Age Industry

Robustness check: using consumption share from remote customers

Table 5: Relationship between consumption share from remote customers and firm growth

VARIABLES	(1) Growth	(2) Growth	(3) Growth	(4) Growth
Sales_further_5km	-0.0116	-0.0528***		
	(0.020)	(0.016)		
Sales_further_10km			-0.0155	-0.0875***
			(0.022)	(0.024)
log(TotalS)	0.00398*	0.00473*	0.00387	0.00486*
	(0.002)	(0.003)	(0.002)	(0.003)
$\mathbb{1}\{\text{Growth}>0\}\times$ Sales_further_5km		0.240***		
		(0.045)		
$\mathbb{1}\{\text{Growth}<0\}\times$ Sales_further_5km		-0.179**		
		(0.064)		
$\mathbb{1}\{\text{Growth}>0\}\times$ Sales_further_10km				0.320***
				(0.0597)
$\mathbb{1}\{\text{Growth}<0\}\times$ Sales_further_10km				-0.194**
				()
$\mathbb{1}\{\text{Growth}>0\}\times$ log(TotalS)		-0.0130**		-0.0110*
		(0.006)		(0.006)
$\mathbb{1}\{\text{Growth}<0\}\times$ log(TotalS)		0.0180***		0.0155***
		(0.003)		(0.003)
Observations	117,222	117,222	117,222	117,222
R-squared	0.274	0.508	0.274	0.508
Fixed effects	Age Industry	Age Industry	Age Industry	Age Industry

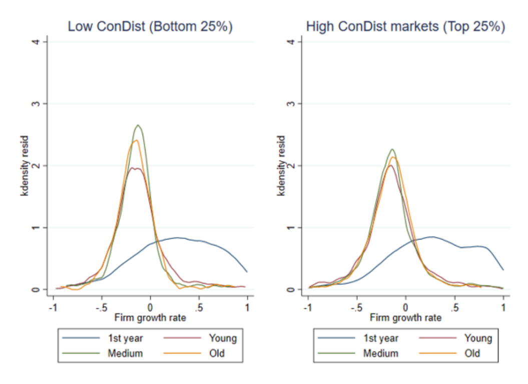
Additional robustness checks

Table 6: Additional robustness checks

	Dependent variable: Firm growth			
VARIABLES	(1)	(2)	(3)	(4)
	Restricting sample to Single-establishment firms		Winsorize growth rate	
log(ConsDist)	-0.00322	-0.0251***	-0.00623	-0.0285***
	(0.008)	(0.005)	(0.007)	(0.004)
log(TotalS)	0.00236	0.00392	0.00505**	0.00712**
	(0.002)	(0.003)	(0.002)	(0.003)
$\mathbb{1}\{\text{Growth}>0\}\times$ log(ConsDist)		0.103***		0.109***
		(0.018)		(0.017)
$\mathbb{1}\{\text{Growth}<0\}\times$ log(ConsDist)		-0.0638**		-0.0672***
		(0.026)		(0.022)
$\mathbb{1}\{\text{Growth}>0\}\times$ log(TotalS)		-0.0157**		-0.0154**
		(0.006)		(0.006)
$\mathbb{1}\{\text{Growth}<0\}\times$ log(TotalS)		0.0200***		0.0173***
		(0.003)		(0.002)
Observations	112,213	112,213	117,222	117,222
R-squared	0.299	0.517	0.29	0.531
Fixed effects	Age Industry	Age Industry	Age Industry	Age Industry

Firm Learning:
Young vs Old Firms

Faster learning in markets with longer travel distance



The evolution of residual firm growth rate distribution by age in western restaurant industry

---

Notation, notes, and remarks:

- We first regress the sales growth of firm on the total sales in the market and all control variables and obtain the residual growth rate. We classify firm age into four categories: 2-5 (Young), 6-10 (Medium), and 11- (Old)

Learning from customer acquisition

Hypothesis: Long ConsDist facilitates firm's learning about their potential, leading to rapid adjustment for both expansion and contraction.

- Remote customers may have better signals about firm's type
- They would travel long distance when they have better information or strong preference on the market, otherwise they do not want to pay higher travel cost
- If ConsDist is related to learning, correlation b/w ConsDist and firm growth would diminish in age

$$
\begin{aligned}
Growth_{s,l,i} = \beta_1 &+ \beta_2 \log(ConsDist_{i,l}) + \beta_3 \log(Age_{s,i,l}) \times \log(ConsDist_{i,l}) + \beta_4 \log(TotalS_{i,l}) \\
&+ \beta_5 \log(Age_{s,i,l}) \times \log(TotalS_{i,l}) + \log(Age_{s,i,l}) + X_{s,i,l}\gamma + \psi_i + \phi_l + \varepsilon_{s,i,l},
\end{aligned}
$$

Age and learning from ConsDist

Table 7: Age heterogeneity in relationship between ConsDist and firm growth

VARIABLES	(1) Growth	(2) Growth
log(ConsDist)	0.0509**	-0.0905***
	(0.021)	(0.024)
log(TotalS)	0.0710***	-0.0270***
	(0.009)	(0.007)
log(Age)	0.0551	0.129**
	(0.089)	(0.060)
log(ConsDist)$\times$ log(Age)	-0.0264**	0.0319**
	(0.012)	(0.011)
log(TotalS)$\times$ log(Age)	-0.0292***	0.00612**
	(0.005)	(0.002)
Observations	62,970	47,350
R-squared	0.296	0.11
Fixed effects	Industry	Industry
Sample	Positive growth firms	Negative growth firms

Exit probability

Faster learning increases exit probability conditional on control variables.

Table 8: Exit probability

	Dependent variable: Exit probability					
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Logit	Logit	Logit	OLS	OLS	OLS
log(ConsDist)	0.582***	0.374***	0.206***	0.0952***	0.0584***	0.0295***
	(0.021)	(0.026)	(0.029)	(0.003)	(0.004)	(0.004)
log(TotalS)	-0.0377***	-0.0152**	0.00257	-0.00639***	-0.00155	-0.00024
	(0.004)	(0.007)	(0.011)	(0.001)	(0.001)	(0.002)
Negative previous year growth dummy		0.520***	0.489***		0.0761***	0.0709***
		(0.016)	(0.016)		(0.002)	(0.002)
Previous year entry dummy		0.375***	0.361***		0.0675***	0.0635***
		(0.031)	(0.032)		(0.005)	(0.005)
Observations	147,583	147,574	147,519	147,583	147,574	147,574
R-squared				0.005	0.044	0.053
Fixed effect	No	Age	Age Industry	No	Age	Age Industry

Model with
Learning & Shopping trips

Firm learning from consumers vs. experience

In standard learning models, a firm learns from its experience of operations.

- the firm receives one signal per period, thus the number of signals equals its age.

In our model, a firm learns from a consumer's visit.

- the number of signals equals the number of consumers since its birth.
- the number of consumers matters for the learning.
- moreover, consumers with high-quality signals also matter.
    - In particular, long travel distance consumers tend to have a better quality (high precision) of signal.

Sketch of the model with firm learning and endogenous shopping trips

Why does distance matter? $\Longleftarrow$ Better learning from long distance travelers (precise signals)

Consumer $\omega \in \Omega$ faces uncertainty for attributes on product of firm s

- $a_{s,t}(\omega)$: attributes (appeals), = underlying demand ζ_s + i.i.d. demand shock $\varepsilon_{s,t}(\omega)$
- heterogenous uncertainty: $\varepsilon_{s,t}(\omega) \sim \mathcal{N}(0, \sigma^2_{\varepsilon,s}(\omega))$

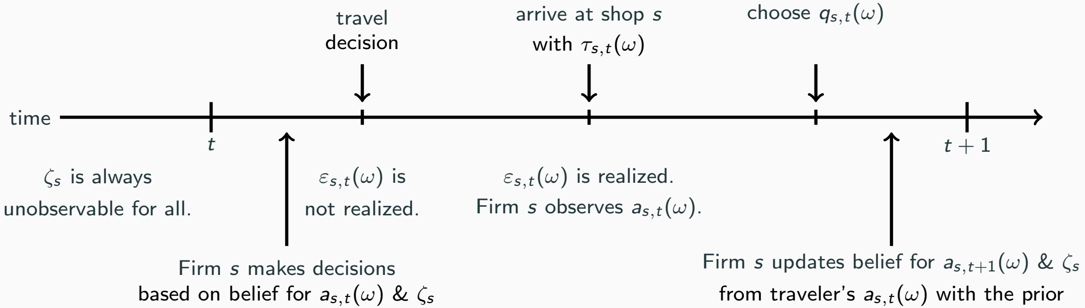
Each type consumer ω with different shopping travel distance (costs)

- $\tau_{s,t}(\omega)$: costs when consumer ω visit shop (firm) s
- Low uncertainty (high precision) consumers are more willing to travel long distance.

Firm $s \in S$ learns from visiting consumers

- Its product's underlying demand ζ_s is unobservable.
- From history of signals of visitors $a_{s,t}(\omega)$, the firm learns ζ_s.
- More consumers and high fraction of consumers with low $\sigma^2_{\varepsilon,s}(\omega)$ in firm learning about its type ζ_s leads to fast convergence of its price, quantity, and revenue.

Timeline: Consumer ω's shopping travel to shop (firm) s



- consumer ω visits shop s if s/he belives travel benefits $v^e_{s,t}(p_{s,t};\omega) >$ costs $\tau_{s,t}(\omega)$.
- consumer ω does not purchase if realized $a_{s,t}(\omega) = \zeta_s + \varepsilon_{s,t}(\omega) \le 0$
- firm s learns if consumer ω visits it regardless of purchase, $q_{s,t}(\omega) > 0$ or $q_{s,t}(\omega) = 0$.

Firm: Environment

Firm s' maximization with many consumers and their demand uncertainty

$$\max \mathbb{E}_{s,t}\left[p_{s,t}q_{s,t} - w_{s,t} \times q_{s,t}\right] \qquad \text{s.t.} \quad q_{s,t} = \sum_{\omega \in \Omega} q^*_{s,t}(p_{s,t};\omega)$$

- consumer ω's demand for firm s' product

$$q^*_{j,t}(a_{j,t}(\omega), p_t(\omega)) = a_{j,t}(\omega)[p_t(\omega)]^{-\theta} \times \underbrace{\mathbb{1}[a_{s,t}(\omega) > 0]}_{\substack{\text{consumer } \omega \text{ purchases when} \\ a_{s,t}(\omega) \text{ is large enough}}} \times \underbrace{\mathbb{1}[v^e_{s,t}(p_{s,t};\omega) > \tau_{s,t}(\omega)]}_{\substack{\text{consumer } \omega \text{ visits when} \\ \text{expected benefits}>\text{travel costs}}}$$

Notation, notes, and remarks:

$q_{s,t}$ & $p_{s,t}$	firm s's quantity and price
$a_{s,t}(\omega)$	attributes (appeals), = underlying demand ζ_s + i.i.d. demand shock $\varepsilon_{s,t}(\omega)$
•	$\zeta_s \sim \mathcal{N}(\mu_\zeta, \sigma^2_\zeta)$ and $\varepsilon_{s,t}(\omega) \sim \mathcal{N}(0, \sigma^2_{\varepsilon,s}(\omega))$
$v^e_{s,t}(p_{s,t};\omega)$	expected benefits of travel
$\tau_{s,t}(\omega)$	costs of travel
$\Omega_{s,t}$	setof visitors, $\Omega_{s,t} = \{\omega \in \Omega : \mathbb{1}[v^e_{s,t}(p_{s,t};\omega) > \tau_{s,t}(\omega)] = 1\} \subseteq \Omega$
$N_{s,t}$	number of visitors, $N_{s,t} = \sum_{\omega\in\Omega} \mathbb{1}[v^e_{s,t}(p_{s,t};\omega) > \tau_{s,t}(\omega)]$

Firm learning: Beliefs and updates

Firm s's belief for ζ_s: normal with mean $b_{s,t}(\mu[\zeta_s])$ and variance $b_{s,t}(\sigma^2[\zeta_s])$

$$b_{s,t}(\mu[\zeta_s]) = \left[\frac{\bar{\bar{\sigma}}^2_{\varepsilon,s,t-1}}{\bar{\bar{\sigma}}^2_{\varepsilon,s,t-1} + N_s^{t-1}\sigma^2_\zeta}\right]\mu_\zeta + \left[\frac{N_s^{t-1}\sigma^2_\zeta}{\bar{\bar{\sigma}}^2_{\varepsilon,s,t-1} + N_s^{t-1}\sigma^2_\zeta}\right]\bar{\bar{a}}_{s,t-1}$$

$$b_{s,t}(\sigma^2[\zeta_s]) = \frac{\bar{\bar{\sigma}}^2_{\varepsilon,s,t-1}\sigma^2_\zeta}{\bar{\bar{\sigma}}^2_{\varepsilon,s,t-1} + N_s^{t-1}\sigma^2_\zeta}$$

- (In the standard model, $\bar{\bar{\sigma}}^2_{\varepsilon,s,t-1} = \sigma^2_{\varepsilon,s}$ and $N_s^{t-1} = \text{age}_{s,t-1} N_s$ where $N_s = 1$.)

Firm s's belief for $a_{s,t}(\omega)$: normal with mean $b_{s,t}(\mu[a_{s,t}(\omega)])$ and variance $b_{s,t}(\sigma^2[a_{s,t}(\omega)])$

$$b_{s,t}(\mu[a_{s,t}(\omega)]) = b_{s,t}(\mu[\zeta_s])$$

$$b_{s,t}(\sigma^2[a_{s,t}(\omega)]) = b_{s,t}(\sigma^2[\zeta_s]) + \sigma^2_{\varepsilon,s}(\omega)$$

Notation, notes, and remarks:

N_s^{t-1} number of of customers from $t - \text{age}_{s,t}$ to $t-1$, $\quad N_s^{t-1} = \sum_{i=t-\text{age}_{s,t}}^{t} N_{s,i}$

$\bar{\bar{a}}_{s,t-1}$ firm s' average of signals, i.e., average of all visitors' signals from $t - \text{age}_{s,t}$ to $t-1$

$\bar{\bar{\sigma}}_{\varepsilon,s,t-1}$ firm s' average of visitors' variance, i.e., average of all visitors' $\sigma^2_{\varepsilon,s}(\omega)$ from $t - \text{age}_{s,t}$ to $t-1$

Consumer: Preference

Utility with travel costs

$$U\left(x_t(\omega) + \sum_{\omega \in \Omega} \left\{ g\left(q_{s,t}(\omega); a_{s,t}(\omega)\right) - \tau_{s,t}(\omega) \right\}\right)$$

$$= 1 - \frac{1}{\kappa} \exp\left(- \kappa x_t(\omega) - \kappa \sum_{s \in S} \left\{ \frac{\left[a_{s,t}(\omega)\right]^{\frac{1}{\theta}} \left[q_{s,t}(\omega)\right]^{\frac{\theta-1}{\theta}}}{(\theta-1)/\theta} \times \mathbb{1}\left[a_{s,t}(\omega) > 0\right] - \tau_{s,t}(\omega) \right\}\right)$$

Consumer heterogeneity in locations and attributes

- To purchase product ω, consumer j should visit shop ω with cost $\tau_{s,t}(\omega)$.
- attributes (appeals): $a_{s,t}(\omega) = \zeta_s + \varepsilon_{s,t}(\omega)$ where $\zeta_s \sim \mathcal{N}(\mu_\zeta, \sigma_\zeta^2)$ and $\varepsilon_{s,t}(\omega) \sim \mathcal{N}(0, \sigma_{\varepsilon,s}^2(\omega))$

All consumers can purchase numeraire $x_t(\omega)$ without any travel costs.

Consumers without learning (memoryless)

Consumer: Optimal purchase (quantity) decision

Backward induction: For given travel decision $d_{s,t}(\omega) = \mathbb{1}\left[v^e_{s,t}(p_{s,t};\omega) > \tau_{s,t}(\omega)\right]$,

Purchase (quantity) decision after observing $a_{s,t}(\omega)$

$$\max_{x_{s,t},\{q_{s,t}(\omega)\}_{\omega\in\Omega}} U\left(x_t(\omega) + \sum_{s\in S}\left[g\left(q_{s,t}(\omega); a_{s,t}(\omega)\right) - \tau_{s,t}(\omega)\right] d_{s,t}(\omega)\right) \quad \text{s.t. } x_{s,t} + \sum_{s\in S} p_{s,t}(\omega) q_{s,t}(\omega) \le m_t(\omega)$$

Purchase (quantity) decision after observing $a_{s,t}(\omega)$

$$q_{j,t}(\omega) = a_{s,t}(\omega) p_{s,t}^{-\theta} \times \mathbb{1}\left[a_{s,t}(\omega) > 0\right] \times d_{s,t}(\omega)$$

Indirect utility

$$U\left(m_t(\omega) + \sum_{s\in S}\left[\frac{a_{s,t}(\omega)}{\chi_{s,t}} \times \mathbb{1}\left[a_{s,t}(\omega) > 0\right] - \tau_{s,t}(\omega)\right] d_{s,t}(\omega)\right)$$

Notation, notes, and remarks:

- $\chi_{s,t} = (\theta - 1) p_{s,t}^{\theta-1}$

Consumer: Optimal travel decision (1/2)

Travel decision with unobserved $a_{s,t}(\omega)$

$$\max_{\{d_{s,t}(\omega)\}_{s\in S}} \mathbb{E}_{\omega,t}\left[U\left(m_t(\omega)+\sum_{s\in S}\left[\frac{a_{s,t}(\omega)}{\chi_{s,t}}\times \mathbb{1}\left[a_{s,t}(\omega)>0\right]-\tau_{s,t}(\omega)\right]d_{s,t}(\omega)\right)\right]$$

Optimal decision rule: consumer ω visit shop s if

$$1>\mathbb{E}_{\omega,t}\left[\exp\left(-\kappa\left[\frac{a_{s,t}(\omega)\times \mathbb{1}\left[a_{s,t}(\omega)>0\right]}{(\theta-1)p_{s,t}^{\theta-1}}\right]+\kappa\tau_{s,t}(\omega)\right)\right]$$

Notation, notes, and remarks:

- $\chi_{s,t}=(\theta-1)p_{s,t}^{\theta-1}$

Consumer: Optimal travel decision (2/2)

We obtain $d_{s,t}(\omega) = \mathbb{1}\left[v^e_{s,t}(p_{s,t};\omega) > \tau_{s,t}(\omega)\right]$ as follows:

Optimal travel decision rule: consumer j visit shop ω if

$$v^e_{s,t}(p_{s,t};\omega) = \left\{\left(\frac{1}{\chi_{s,t}}\right)\mu_\zeta - \kappa\left(\frac{1}{\chi_{s,t}}\right)^2\left[\frac{\sigma^2_\zeta + \sigma^2_\varepsilon(\omega)}{2}\right]\right\} - \frac{1}{\kappa}\ln\Theta_{s,t}(\omega) > \tau_{s,t}(\omega)$$

- high uncertainty $(\sigma^2_\zeta + \sigma^2_\varepsilon(\omega))$ and travel costs $(\tau_{s,t}(\omega))$ discourage, but high expected appeal (μ_ζ) encourage shopping travel to s.
- $\Theta_{s,t}(\omega)$ are negligible with high μ_ζ, i.e., $\Theta_{s,t}(\omega) \to 0$ and $\frac{\partial\Theta_{s,t}(\omega)}{\partial\sqrt{\sigma^2_\zeta+\sigma^2_\varepsilon(\omega)}} \to 0$ as $\mu_\zeta \to \infty$

Notation, notes, and remarks:

- $\chi_{s,t} = (\theta - 1)p^{\theta-1}_{s,t}$
- $\Theta_{s,t}(\omega) = \Phi\left(\{\mu_\zeta - (\kappa/\chi_{s,t})[\sigma^2_\zeta + \sigma^2_\varepsilon(\omega)]\}/\sqrt{\sigma^2_\zeta + \sigma^2_\varepsilon(\omega)}\right)$ where $\Phi(\cdot)$ is the standard normal distribution function.

Conclusion

Conclusion

Travel distance of consumers is important in firm's customer acquisition process and thus its expansion, contraction, and exit.

In particular, we find that

- Longer distance $\Rightarrow$ rapid growth of expanding firms
- Longer distance $\Rightarrow$ rapid shrink of contracting firms & high exit probability
- Such results are strong among young firms but not in old firms.

These findings are consistent with our model of firm learning and endogenous shopping trips of consumers with different travel distances.

Appendix

of cards variation

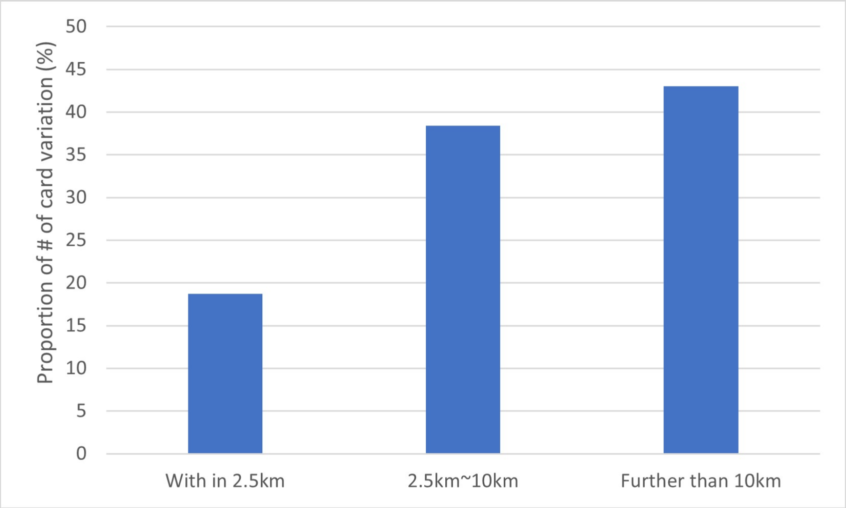
The new margin of # of customers

- Einav et. al. (2021) states that **# of cards** accounts for majority of variation of firm sales variation
- We quantify the relative importance of **# of cards** variation from each **travel distance group** in total # of cards variation of markets

Decomposition

- Three travel distance group: A-less than 2.5km, B-further than 2.5km less than 10km, C-further than 10km
- # of cards=# of cards from A + # of cards from B + # of cards from C

Distance matters for # of cards variation



The result of variance decomposition across travel distance groups

Averages

Firm ω's information set at the beginning of t

- average attributes (signals) from $t - \text{age}_{s,t}$

$$\bar{\bar{a}}_{s,t-1} = \frac{\sum_{i=t-\text{age}_{s,t}}^{t-1} \sum_{\omega \in \Omega_{s,i}} a_{s,i}(\omega)}{\sum_{i=t-\text{age}_{s,t}}^{t-1} N_{s,i}} = \sum_{i=t-\text{age}_{s,t}}^{t-1} \left(\frac{N_{s,i}}{N_s^{t-1}} \right) \bar{a}_{s,i}$$

- average variances of noises from $t - \text{age}_{s,t}$

$$\bar{\bar{\sigma}}^2_{\varepsilon,s,t-1} = \frac{\sum_{i=t-\text{age}_{s,t}}^{t-1} \sum_{\omega \in \Omega_{s,i}} \sigma^2_{\varepsilon,s}(\omega)}{\sum_{i=t-\text{age}_{s,t}}^{t-1} N_{s,i}} = \sum_{i=t-\text{age}_{s,t}}^{t-1} \left(\frac{N_{s,i}}{N_s^{t-1}} \right) \bar{\sigma}^2_{\varepsilon,s,i}$$

Notation, notes, and remarks:

$\bar{a}_{s,t}$ average attributes of visitors at t, $\bar{a}_{s,t} = \sum_{\omega \in \Omega_{s,t}} a_{s,i}(\omega)/N_{s,t} = \zeta_s + \sum_{\omega \in \Omega_{s,t}} \varepsilon_{s,t}(\omega)/N_{s,t} = \zeta_s + \bar{\varepsilon}_{s,t}$

$N_{s,t}$ number of of visitors, $N_{s,t} = \sum_{\omega \in \Omega} \mathbb{1}\{q^*_{s,t}(p_{s,t};\omega) > 0\} = N_{L,s,t} + N_{H,s,t}$

N_s^{t-1} number of visitors from $t - \text{age}_{s,t}$, $N_s^{t-1} = \sum_{i=t-\text{age}_{s,t}}^{t} N_{s,i}$

Consumer: Optimal travel decision: derivations

For convenience,

- $\exp(-\tilde{a}_{s,t}(\omega)) = \exp(-a_{s,t}(\omega)\kappa/\chi_{s,t}) \sim \mathcal{LN}(-\tilde{\mu}_{s,t}, \tilde{\sigma}_s^2(\omega))$
- $\tilde{\mu}_t(\omega) = \mu_\zeta \kappa/\chi_{s,t}$ and $\tilde{\sigma}_{s,t}^2(\omega) = (\kappa/\chi_{s,t})^2[\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)]$

Optimal decision rule: consumer ω visit shop s if

$$
\begin{aligned}
\exp\left(-\kappa\tau_{s,t}(\omega)\right) > \mathbb{E}_{\omega,t}&\left[\exp\left(-\frac{\kappa}{\chi_{s,t}}a_{s,t}(\omega)\right)\middle| a_{s,t}(\omega) > 0\right]\Pr\left(a_{s,t}(\omega) > 0\right) \\
&= \mathbb{E}_{\omega,t}\left[\exp\left(-\tilde{a}_{s,t}(\omega)\right)\middle|\exp(-\tilde{a}_{s,t}(\omega)) < 1\right]\Pr\left(\exp(-\tilde{a}_{s,t}(\omega)) < 1\right) \\
&= \exp\left(-\tilde{\mu}_t(\omega) + \frac{\tilde{\sigma}_{s,t}^2(\omega)}{2}\right)\Phi\left(\frac{\tilde{\mu}_t(\omega)}{\tilde{\sigma}_{j,t}(\omega)} - \tilde{\sigma}_{j,t}(\omega)\right) \\
&= \exp\left(-\left(\frac{\kappa}{\chi_{s,t}}\right)\mu_\zeta + \left(\frac{\kappa}{\chi_{s,t}}\right)^2\left[\frac{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}{2}\right]\right)\Phi\left(\frac{\mu_\zeta - (\kappa/\chi_{s,t})[\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)]}{\sqrt{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}}\right)
\end{aligned}
$$

Notation, notes, and remarks:

- $\chi_{s,t} = (\theta - 1)p_{s,t}^{\theta-1}$
- $\Phi(\cdot)$ is the standard normal distribution function.

Note that

Distribution change effect Θ dampens our main result

$$\Theta_{s,t}(\omega) = \Phi\left(\frac{\mu_\zeta - (\kappa/\chi_{s,t})[\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)]}{\sqrt{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}}\right)$$

- increases with μ_ζ but decreases with $\sqrt{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}$.

$$\frac{\partial \Theta_{s,t}(\omega)}{\partial \sqrt{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}} = -\left[\frac{\mu_\zeta}{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)} + \frac{\kappa}{\chi_{s,t}}\right]\left[\frac{\phi\left(\{\mu_\zeta - (\kappa/\chi_{s,t})[\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)]\}/\sqrt{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}\right)}{\Phi\left(\{\mu_\zeta - (\kappa/\chi_{s,t})[\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)]\}/\sqrt{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}\right)}\right]$$

which converges to zero when μ_ζ goes to infinity.

Notation, notes, and remarks:

- $z_{j,t}(\omega) = \{\mu_\zeta - (\kappa/\chi_{s,t})[\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)]\}/\sqrt{\sigma_\zeta^2 + \sigma_\varepsilon^2(\omega)}$
- $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal pdf and cdf.