Lecture 12: Search and price dispersion

EC 105. Industrial Organization.

Matt Shum
HSS, California Institute of Technology

Information costs

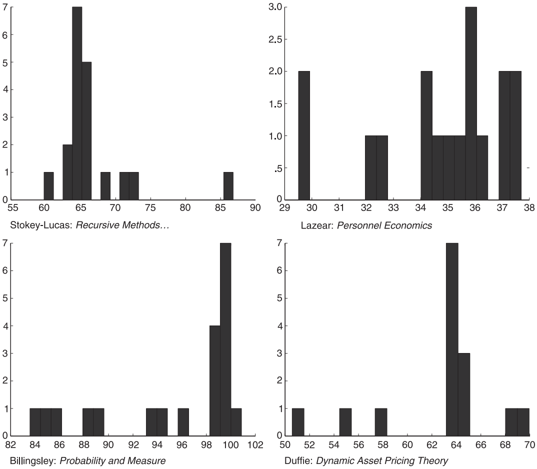
- In many markets, consumers or firms may lack important information about each other
- Two types of information costs:

1. **Search costs**:
   - Consumers are not fully aware of firms' prices in the market
   - gasoline, furniture
2. **Verification costs**:
   - Consumers are not fully aware of product *quality*: how reliable the product is
     - automobiles, apartments, electronics
   - Firms are not fully aware of consumers' characteristics, actions
     - insurance: are people driving safely, do they have unreported health risks?

We focus on search costs in this lecture.

Why are prices for the same item so different across stores?

FIGURE 1
RAW HISTOGRAMS OF ONLINE PRICES



A puzzle considering basic economic theory: review this.

Consider the benchmark of *perfect competition*

Leaving the PC world

- One important implicit assumption of PC paradigm is that consumers are aware of prices at all stores. This implies an infinitely elastic demand curve facing firms. (ie. if one firm raises prices slightly, he will lose all demand).
- Obviously, this assumption is not realistic. Here we consider what happens, if we relax just this assumption, but maintain other assumptions of PC paradigm: large #firms, perfect substitutes, etc.

Search model

- Each consumer demands *one* unit;
- Starts out at one store, incurs cost $c > 0$ to search at any other store.
- Consumer only knows prices at stores that she has been to, and buys from the canvassed store with the lowest price. "free recall"
- Utility u from purchasing product: demand function is

$$\begin{cases} \text{purchase if } p \leq u \\ \text{don't purchase otherwise} \end{cases} \tag{1}$$

- What is equilibrium in this market?

Diamond paradox

- Claim: a nonzero search cost $c > 0$ leads to equilibrium price equal to u ("monopoly price")
- Assume that marginal cost=0, so that under PC, $p = 0$
- n firms, with n large. Consumers equally distributed initially among all firms.
- Start out with all firms at PC outcome. What happens if one firm deviates, and charges some p^1 such that $0 < p^1 < c$?
  - Consumers at this store?
  - Consumers at other stores?
  - How will other stores respond?
  - By iterating this reasoning ....

Now start at “monopoly outcome”, where all firms are charging u.

- What are consumers’ purchase rules?
- Do firms want to undercut? Given consumer behavior, what do they gain?
- Role for advertising?
- P. Diamond (1971), “A Theory of Price Adjustment”, *Journal of Economic Theory*

Remarks

- Diamond result quite astounding, since it suggests PC result is "knife-edge" case.
- But still doesn't explain price *dispersion*
- Assume consumers differ in search costs
- Two types of consumers: "natives" are perfectly informed about prices, but "tourists" are not.

Tourist-natives model

- Tourists and natives, in proportions $1-\alpha$ and α. L total consumers: αL natives, and $(1-\alpha)L$ tourists.
- Tourists buy one unit as long as $p \leq u$, but natives always shop at the cheapest store.
- Each of n identical firms has U-shaped AC curve
- Each firm gets equal number of tourists $\left(\frac{(1-\alpha)L}{n}\right)$; natives always go to cheapest store.
- Consider world in which all firms start by setting $p^c = \min_q AC(q)$.
- Note that deviant store always wants to price *higher*. Demand curve for a deviant firm is kinked (graph). Deviant firm sells exclusively to tourists.

Deviant firm will always charge u. Only tourists shop at this store. If charge above u, no demand. If below u, then profits increase by charing u.

First case: many informed consumers (α large)

- Number $q^u \equiv \frac{(1-\alpha)L}{n}$ of tourists at each store so small that $u < AC(q^u)$.
- In free-entry equilibrium, then, all firms charge p^c, and produce the same quantity L/n.
- If enough informed consumers, competitive equilibrium can obtain (not surprising)

Second case: few informed consumers (α small)

- Assume enough tourists so that $u > AC(q^u)$.
- But now: hi-price firms making positive profits, while lo-price firms making (at most) zero profits. Not stable.
- In order to have equilibrium: ensure that given a set of high-price firms (charing u) and low-price firms (charging p^c), no individual firm wants to deviate. Free entry ensures this.
- Let β denote proportion of lo-price firms.
- Each high-price firm charges u and sells an amount

$$q^u = \frac{(1-\alpha)L(1-\beta)}{n(1-\beta)} = \frac{(1-\alpha)L}{n} \tag{2}$$

- Each low-price firm charges p^c and sells

$$q^c = \frac{\alpha L + (1-\alpha)L\beta}{n\beta} \tag{3}$$

- In equilibrium, enough firms of each type enter such that each firm makes zero profits. Define quantities q^a, q^c such that (graph):

$$AC(q^a) = u; \quad AC(q^c) = p^c.$$

(Quantities at which both hi- and lo-price firms make zero profits.)

- With free entry, n and β must satisfy

$$q^a = \frac{(1-\alpha)L}{n}; \quad q^c = \frac{\alpha L + (1-\alpha)L\beta}{n\beta} \tag{4}$$

- Solving the two equations for n and β yields

$$n = \frac{(1-\alpha)L}{q^a}; \quad \beta = \frac{\alpha q^a}{(1-\alpha)(q^c - q^a)} \tag{5}$$

- N.B: arbitrary which firms become high or low price. Doesn't specify process whereby price dispersion develops.
- As $\alpha \to 0$, then $\beta \to 0$ (Diamond result)

Two search models:

Consider two search models:

1. **Nonsequential search model:** consumer commits to searching n stores before buying (from lowest-cost store). "Batch" search strategy.
2. **Sequential search model:** consumer decides after each search whether to buy at current store, or continue searching.

Question:

from observed prices (as above), can we infer what consumers' search costs in a market are?

Consider data for the four textbooks.

Main assumptions:

- Infinite number ("continuum") of firms and consumers
- Observed price distribution F_p is *equilibrium mixed strategy* on the part of firms, with bounds $\underline{p}, \bar{p}$.
- r: constant per-unit cost (wholesale cost), identical across firms
- Firms sell homogeneous products
- Each consumer buys one unit of the good
- Consumer i incurs cost c_i to search one store; drawn independently from search cost distribution F_c
- First store is "free"
- $\tilde{q}_k$: probability that consumer searches k stores before buying

Consumers in nonsequential model

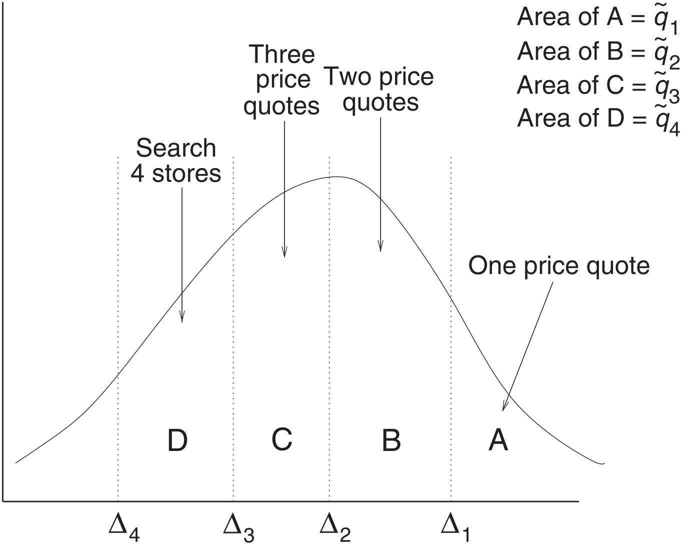
- Consumer with search cost c who searches n stores incurs total cost

$$\begin{aligned} & c*(n-1)+E[\min(p_1,\ldots,p_n)] \\ =& c*(n-1)+\int_{\underline{p}}^{\bar{p}} p\cdot n(1-F_p(p))^{n-1}f_p(p)dp. \end{aligned} \tag{6}$$

- This is decreasing in c. Search strategies characterized by cutoff-points, where consumer indifferent between n and $n+1$ must have cost

$$\Delta_n = E[\min(p_1,\ldots,p_n)] - E[\min(p_1,\ldots,p_{n+1})].$$

  and $\Delta_1 > \Delta_2 > \Delta_3 > \cdots$.
- Similarly, define $\tilde{q}_n = F_c(\Delta_{n-1}) - F_c(\Delta_n)$ (fraction of consumers searching n stores). Graph.


Area of A = $\tilde{q}_1$
Area of B = $\tilde{q}_2$
Area of C = $\tilde{q}_3$
Area of D = $\tilde{q}_4$
Three price quotes
Two price quotes
Search 4 stores
One price quote
D
C
B
A
Δ_4
Δ_3
Δ_2
Δ_1

Firms in nonsequential model

- Firm's profit from charging p is:

$$\Pi(p) = (p - r)\left[\sum_{k=1}^{\infty} \tilde{q}_k \cdot k \cdot (1 - F_p(p))^{k-1}\right], \quad \forall p \in [\underline{p}, \bar{p}]$$

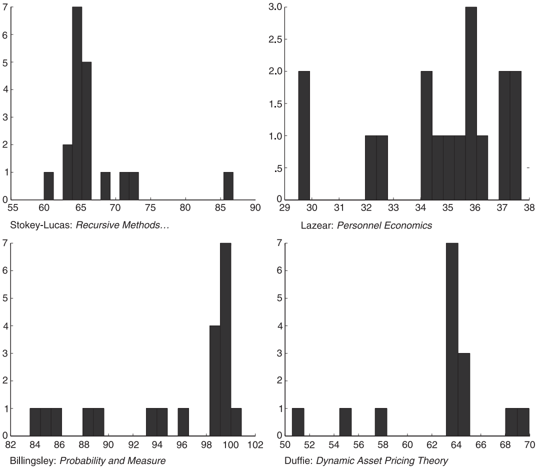
- Since firms use mixed streategy, they must be indifferent btw all p:

$$(\bar{p} - r)\tilde{q}_1 = (p - r)\left[\sum_{k=1}^{\infty} \tilde{q}_k \cdot k \cdot (1 - F_p(p))^{k-1}\right], \quad \forall p \in [\underline{p}, \bar{p}] \qquad (7)$$

Reconsider this data

FIGURE 1

RAW HISTOGRAMS OF ONLINE PRICES



Estimating search costs corresponding to graphs

- We observe data $P_n \equiv (p_1, \dots p_n)$. Sorted in increasing order.
- Take $\underline{p} = p_1$ and $\bar{p} = p_n$
- Given picture above, we want to estimate consumer cutpoints Δ_i, $i = 1, 2, 3, \ldots$ and also the $\tilde{q}_k$, probabilities that consumers search k stores.
- Consumer cutpoints $\Delta_1, \Delta_2, \dots$ can be estimated by computing

$$\Delta_n = E[\min(p_1, \dots, p_n)] - E[\min(p_1, \dots, p_{n+1})].$$

  using the observed prices P_n.
- Assume that consumers search at most K $(< N - 1)$ stores. Then can solve for $\tilde{q}_1, \ \dots, \ \tilde{q}_K$ from

$$(\bar{p} - r)\tilde{q}_1 = (p_i - r)\left[\sum_{k=1}^{K-1} \tilde{q}_k \cdot k \cdot (1 - \hat{F}_p(p_i))^{k-1}\right], \quad \forall p_i, i = 1, \dots, n-1.$$

  where $\hat{F}_p = Freq(p \leq \tilde{p}) = \frac{1}{n}\sum_i \mathbf{1}(p_i \leq \tilde{p})$ is the *cumulative distribution function* of the observed prices.
    - $n - 1$ equations with K unknowns.

Nonsequential model: results

TABLE 2 Search-Cost Distribution Estimates for Nonsequential-Search Model

Product	K[a]	M[b]	$\tilde{q}_1^c$	$\tilde{q}_2$	$\tilde{q}_3$	Selling Cost r	MEL Value
Parameter estimates and standard errors: nonsequential-search model							
Stokey-Lucas	3	5	.480 (.170)	.288 (.433)		49.52 (12.45)	102.62
Lazear	4	5	.364 (.926)	.351 (.660)	.135 (.692)	27.76 (8.50)	84.70
Billingsley	3	5	.633 (.944)	.309 (.310)		69.73 (68.12)	199.70
Duffie	3	5	.627 (1.248)	.314 (.195)		35.48 (96.30)	109.13

Search-cost distribution estimates

	Δ_1	$F_c(\Delta_1)$	Δ_2	$F_c(\Delta_2)$	Δ_3	$F_c(\Delta_3)$
Stokey-Lucas	2.32	.520	.68	.232		
Lazear	1.31	.636	.83	.285	.57	.150
Billingsley	2.90	.367	2.00	.058		
Duffie	2.41	.373	1.42	.059		

[a] Number of quantiles of search cost F_c that are estimated (see equation (5)). In practice, we set K and M to the largest possible values for which the parameter estimates converge. All combinations of larger K and/or larger M resulted in estimates that either did not converge or did not move from their starting values (suggesting that the parameters were badly identified).

[b] Number of moment conditions used in the empirical likelihood estimation procedure (see equation (17)).

[c] For each product, only estimates for $\tilde{q}_1, \ldots, \tilde{q}_{K-1}$ are reported; $\tilde{q}_K = 1 - \sum_{k=1}^{K-1} \tilde{q}_k$.

[d] Indifferent points Δ_k computed as $Ep_{(1:k)} - Ep_{(1:k+1)}$ (the expected price difference from having k versus $k+1$ price quotes), using the empirical price distribution. Including shipping and handling charges.

Sequential model

- Consumer decides after each search whether to accept lowest price to date, or continue searching.
- Optimal "reservation price" policy: accept first price which falls below some optimally chosen reservation price.
- NB: optimal strategy features no recall.

Consumers in sequential model

- Heterogeneity in search costs leads to heterogeneity in reservation prices
- For consumer with search cost c_i, let $z^*(c_i)$ denote price z which satisfies the following indifference condition

$$c_i = \int_0^z (z-p)f(p)dp = \int_0^z F(p)dp.$$

Now, for consumer i, her reservation price is:

$$p_i^* = \min(z^*(c_i), \bar{p}).$$

Results: sequential search model

TABLE 3 **Estimates of Sequential-Search Model**

Product	δ_1	δ_2	Median[a] Search Cost	Selling Cost r	α[b]	$F_c^{-1}(1-\alpha;\boldsymbol{\theta})$	Log-L Value
Stokey-Lucas	.46 (.02)	1.55 (.03)	29.40 (1.45)	22.90 (1.31)	.58	19.19	31.13
Lazear	.40 (.01)	1.15 (.01)	16.37 (1.00)	11.31 (.79)	.69	4.56	34.35
Billingsley	.25 (.01)	2.01 (.04)	9.22 (.94)	65.37 (.83)	.51	8.43	23.73
Duffie	.21 (.02)	4.57 (.29)	10.57 (2.01)	28.24 (1.63)	.54	7.00	18.93

Note: Including shipping and handling charges. Standard errors in parentheses. δ_1 and δ_2 are parameters of the gamma distribution; see equation (13).

[a] As implied by estimates of the parameters of the gamma search-cost distribution.

[b] Proportion of consumers with reservation price equal to $\overline{p}$, implied by estimate of r (see equation (11)).

- Generally, nonsequential (batch) search model generates smaller (more reasonable) search cost estimates
  - Batch search is feature of online search engines?

Test: sequential or non-sequential?

- Look for evidence from online book markets.
  - Do consumer *recall*? (Return to earlier searched stores)
- Nonseq search: recall is possible.
- Seq search: optimal behhavior has no recall
- Hortacsu, de los Santos, Wildenbeest paper

Bookstore	Transactions Number	Transactions %	Visits Number	Visits %
Amazon	10,206	65.5%	249,593	76.3%
Barnes and Noble	3,046	19.6%	25,758	7.9%
Book Clubs				
christianbook.com	615	3.9%	3968	1.2%
doubledaybookclub.com	468	3.0%	4001	1.2%
eharlequin.com	61	0.4%	3647	1.1%
literaryguild.com	326	2.1%	3500	1.1%
mysteryguild.com	188	1.2%	2095	0.6%
Other Bookstore				
1bookstreet.com	10	0.1%	120	0.0%
allbooks4less.com	5	0.0%	199	0.1%
alldirect.com	27	0.2%	490	0.1%
ecampus.com	114	0.7%	1206	0.4%
powells.com	68	0.4%	1326	0.4%
varsitybooks.com	16	0.1%	218	0.1%
walmart.com	183	1.2%	28663	8.8%
booksamillion.com	245	1.6%	2290	0.7%
Total	15,578	100.0%	327,074	100.0%

Table 1: Transactions and Visits by Bookstore

Search window	No. of visited visited		If 2 or more firms, bought from:		Exhausted search?
7 Days	One	76%			
	2 or more	24%	Last firm sampled	65%	
			Recalled	35%	55%
6 Days	One	77%			
	2 or more	23%	Last firm sampled	64%	
			Recalled	36%	55%
5 Days	One	79%			
	2 or more	21%	Last firm sampled	63%	
			Recalled	37%	55%
4 Days	One	80%			
	2 or more	20%	Last firm sampled	61%	
			Recalled	39%	55%
3 Days	One	82%			
	2 or more	18%	Last firm sampled	61%	
			Recalled	39%	56%
2 Days	One	84%			
	2 or more	16%	Last firm sampled	61%	
			Recalled	39%	56%
1 Day	One	86%			
	2 or more	14%	Last firm sampled	61%	
			Recalled	39%	56%
Same day	One	90%			
	2 or more	10%	Last firm sampled	62%	
			Recalled	38%	58%

Table 4: Test of “no recall” hypothesis

Search Window	Total	Amazon	Barnes & Noble	Book Clubs	Other bookstores
Recall percentage by firm					
7 Days	35%	50%	20%	18%	23%
6 Days	36%	51%	22%	20%	24%
5 Days	37%	52%	23%	23%	26%
4 Days	39%	54%	24%	27%	27%
3 Days	39%	54%	25%	28%	27%
2 Days	40%	54%	26%	28%	28%
1 Day	40%	54%	26%	31%	27%
Same Day	39%	53%	25%	30%	26%
Distribution of Recall Transactions					
7 Days	1302	70%	18%	6%	6%
6 Days	1283	68%	20%	7%	6%
5 Days	1230	67%	20%	7%	6%
4 Days	1187	66%	20%	8%	6%
3 Days	1103	65%	21%	8%	7%
2 Days	994	64%	22%	7%	7%
1 Day	845	63%	22%	7%	7%
Same Day	588	64%	22%	6%	8%

Table 5: Recall by Firm

When do people search? Lewis/Marvel paper

Table 1: Rankings of GasBuddy.com local price sites

Alexa.com traffic rank for February 22, 2007
(a period of low gas prices
compared to those prevailing in the previous year)

Location	Alexa Traffic Rank
GasBuddy.com	70,456
San Jose	680,565
San Francisco	905,384
Los Angeles	622,956
San Diego	1,552,704
Chicago	2,178,953
Seattle	546,295
Detroit	890,277
Boston	2,180,036
Denver	644,598
Columbus, Ohio	1,103,777
Dallas	1,661,050
Atlanta	910,937
...	
gaspricewatch.com	164,500
fuelgaugereport.com	693,807

When do people search? Lewis/Marvel paper

Table 3: Gasoline Price Movements as a Determinant of Price Search

Variable	Coefficient (Newey-West std. error)	*t*-statistic for difference in coefficients	
α_0 (intercept)	2.56*** (0.118)		
α_1 (date)	0.0016*** (0.0002)		
α_2 (price change +, most recent 4 days)	5.77*** (0.885)	$\alpha_2 \neq -\alpha_3$	2.49
α_3 (price change, -, most recent 4 days)	−0.52 (2.064)		
α_4 (price change, +, days 5-20)	4.48*** (0.814)	$\alpha_4 \neq -\alpha_5$	3.55
α_5 (price change, -, days 5-20)	−1.75** (0.838)		
α_6 (price change, +, days 20-50)	1.59*** (0.335)	$\alpha_6 \neq -\alpha_7$	4.80
α_7 (price change, -, days 20-50)	0.21 (0.297)		
Number of Observations	676		

When do people search? Lewis/Marvel paper

Table 4: Gasoline Price Dispersion as a Function of Price Movements

Specification				
Sample:	49 cities	City/date pairs with 30 prices		full sample
Dependent Variable:	median range	$p^{max} - p^{min}$	median range	$p^{max} - p^{min}$

Variable	Coefficients (Newey-West standard errors in parenthesis)			
β_1 (price change +, most recent 4 days)	0.30 (0.173)	0.11 (0.119)	0.25* (0.147)	0.10 (0.129)
β_2 (price change, -, most recent 4 days)	−2.41*** (0.222)	−1.81*** (0.153)	−2.31*** (0.178)	−1.25*** (0.183)
β_3 (price change, +, days 5-20)	−0.13 (0.101)	−0.27*** (0.089)	−0.13 (0.097)	−0.55*** (0.100)
β_4 (price change, -, days 5-20)	−1.02*** (0.094)	−1.09*** (0.068)	−1.04*** (0.072)	−1.12*** (0.080)
β_5 (price change, +, days 20-50)	−0.01 (0.057)	−0.08* (0.047)	−0.08 (0.053)	−0.14*** (0.052)
β_6 (price change, -, days 20-50)	−0.14*** (0.037)	−0.13*** (0.030)	−0.13*** (0.032)	−0.12*** (0.032)

When do people search?

- Pretty convincing evidence that people search when prices have been rising.
- Which contributes to greater price dispersion.
- (Also to "asymmetric price response"?)