Graduate IO: Price Dispersion and Search

November 6, 2016

Agenda

- price dispersion and search
  - theory: Varian (1980)
  - empirical: Sorensen (2000)
- structural estimation of search models
  - Hong and Shum (2006), Hortacsu and Syverson (2004), De Los Santos, Hortacsu, and Wildenbeest (2012)

Search

- an important assumption of the models we have developed in the previous lectures is that consumers know the prices that sellers are charging

Search

- an important assumption of the models we have developed in the previous lectures is that consumers know the prices that sellers are charging
- however, in most retail markets, consumers do not know the prices charged by different retailers and have to learn them
  - this activity is costly, so consumers are unlikely to be fully informed
  - one of the main uses of the Internet is to provide price information cheaply

Search

- an important assumption of the models we have developed in the previous lectures is that consumers know the prices that sellers are charging
- however, in most retail markets, consumers do not know the prices charged by different retailers and have to learn them
  - this activity is costly, so consumers are unlikely to be fully informed
  - one of the main uses of the Internet is to provide price information cheaply
- question: how does the consumer's lack of information about prices affect competition among firms?

A Simple Model: Varian (1980)

- supply
  - N firms supply a homogenous good
  - product costs are zero

A Simple Model: Varian (1980)

- supply
  - N firms supply a homogenous good
  - product costs are zero
- demand
  - consumers have unit demands and willingness to pay of r
  - two types of consumers: I informed and M uninformed consumers
    - an informed consumer knows the prices charged by all N firms
    - an uninformed consumer does not know the prices and randomly selects one firm to shop

The Game

- n firms post prices simultaneously
  - informed consumers buy from the firms setting the lowest price if it does not exceed r
  - uninformed consumers randomly select a firm and buy if the firm's price does not exceed r

The Game

- n firms post prices simultaneously
  - informed consumers buy from the firms setting the lowest price if it does not exceed r
  - uninformed consumers randomly select a firm and buy if the firm's price does not exceed r

- let $U = \frac{M}{n}$ be the number of uninformed consumers each firm gets

The Game

- n firms post prices simultaneously
  - informed consumers buy from the firms setting the lowest price if it does not exceed r
  - uninformed consumers randomly select a firm and buy if the firm's price does not exceed r
- let $U = \frac{M}{n}$ be the number of uninformed consumers each firm gets
- payoffs

$$\pi_i(p_i, p_{-i}) = \begin{cases} p_i(U + I) & \text{if } p_i < p_j,\ j \neq i \\ p_i\left(U + \frac{I}{m}\right) & \text{if } p_i = p_j,\ j = 1, ..., m-1 \\ p_i U & \text{if } r \geq p_i > p_j,\ j \neq i \end{cases}$$

Equilibirium Analysis

- claim 1: there is no equilibrium in which all firms charge the same price

Equilibirium Analysis

- claim 1: there is no equilibrium in which all firms charge the same price
- proof: without loss of generality, assume that $n = 2$
  - $p = 0$ is not an equilibrium because each firm can raise price to r and earn rU
  - if they charge $r \geq p > 0$, then a slight cut in price gets all the informed consumers, which increases profits.

Equilibirium Analysis

- claim 1: there is no equilibrium in which all firms charge the same price
- proof: without loss of generality, assume that $n = 2$
  - $p = 0$ is not an equilibrium because each firm can raise price to r and earn rU
  - if they charge $r \geq p > 0$, then a slight cut in price gets all the informed consumers, which increases profits.
- thus, the *Law of One Price* does not hold!

Equilibirium Analysis (Cont.)

- claim 2: there is no equilibrium in pure strategies

Equilibirium Analysis (Cont.)

- claim 2: there is no equilibrium in pure strategies
- proof: by previous claim, one firm sets a higher price than the other, let $p^L < p^H$
  - clearly, $p^H = r$ since firm H's profits are pU which is strictly increasing in p
  - but then firm L should raise price p^L until it is slightly below r
  - but then firm H wants to undercut p^L

Equilibirium Analysis (Cont.)

- claim 2: there is no equilibrium in pure strategies
- proof: by previous claim, one firm sets a higher price than the other, let $p^L < p^H$
  - clearly, $p^H = r$ since firm H's profits are pU which is strictly increasing in p
  - but then firm L should raise price p^L until it is slightly below r
  - but then firm H wants to undercut p^L
- we need to look for an equilibrium in which firms cannot forecast the prices of their rivals and undercut them
  - firms are perceived as choosing prices randomly

Mixed-Strategy

- let $F(p)$ denote the probability that a rival posts a price less than p, then the expected profit to a firm when it charges p is

$$\pi(p) = [1 - F(p)]^{n-1} p(U + I) + \left\{1 - [1 - F(p)]^{n-1}\right\} pU$$

  - the first term on the RHS is the probability that p is the lowest price, in which case firm demand is $U + I$
  - the second term on the RHS is the probability that p is not the lowest price, in which case firm demand is only U

Mixed-Strategy

- let $F(p)$ denote the probability that a rival posts a price less than p, then the expected profit to a firm when it charges p is

$$\pi(p) = [1 - F(p)]^{n-1} p (U + I) + \left\{1 - [1 - F(p)]^{n-1}\right\} pU$$

  - the first term on the RHS is the probability that p is the lowest price, in which case firm demand is $U + I$
  - the second term on the RHS is the probability that p is not the lowest price, in which case firm demand is only U

- in a mixed strategy equilibrium, the firm's profits at every price p has to be constant: $\pi(p) = k$
  - solving the equilibrium equation yields

$$1 - F(p) = \left[\frac{k - pU}{pI}\right]^{\frac{1}{n-1}}$$

Mixed-Strategy (Cont.)

- clearly, the upper bound on prices is r, therefore, the unknown constant k satisfies

$$F(r) = 1 \Rightarrow k = rU$$

  - highest price, demand is only U

Mixed-Strategy (Cont.)

- clearly, the upper bound on prices is r, therefore, the unknown constant k satisfies

$$F(r) = 1 \Rightarrow k = rU$$

  - highest price, demand is only U

- substituting, we obtain

$$1 - F(p) = \left[\frac{(r-p)\,U}{pI}\right]^{\frac{1}{n-1}}$$

Mixed-Strategy (Cont.)

- clearly, the upper bound on prices is r, therefore, the unknown constant k satisfies

$$F(r) = 1 \Rightarrow k = rU$$

  - highest price, demand is only U

- substituting, we obtain

$$1 - F(p) = \left[\frac{(r-p)\,U}{pI}\right]^{\frac{1}{n-1}}$$

- the lower bound of the set of prices that firms will charge is obtained by setting $F(p) = 0$

$$\underline{p} = \frac{rU}{U+I}$$

Mixed-Strategy (Cont.)

- the lower bound $\underline{p}$ is strictly positive
  - firms earn positive profits: escape the Bertrand trap
  - the equilibrium density of prices is U-shaped: firms will tend to either price near r or near the lower bound

Mixed-Strategy (Cont.)

- the lower bound $\underline{p}$ is strictly positive
  - firms earn positive profits: escape the Bertrand trap
  - the equilibrium density of prices is U-shaped: firms will tend to either price near r or near the lower bound
- intuition: firm either goes for the informed consumers or is content exploiting the uninformed consumers

Mixed-Strategy (Cont.)

- the lower bound $\underline{p}$ is strictly positive
  - firms earn positive profits: escape the Bertrand trap
  - the equilibrium density of prices is U-shaped: firms will tend to either price near r or near the lower bound
- intuition: firm either goes for the informed consumers or is content exploiting the uninformed consumers
- remarks: as U falls, market becomes more competitive, prices fall and lower prices more likely

Search Cost and Internet

- prior to the Internet, search costs were attributed to dispersion in the geographical location of stores
  - it was costly to visit every store and determine who was offering the lowest price

Search Cost and Internet

- prior to the Internet, search costs were attributed to dispersion in the geographical location of stores
  - it was costly to visit every store and determine who was offering the lowest price
- however, price search engines like *Pricewatch.com* provide consumers with lost of prices at very low cost
  - online travel agents like *Travelocity* and *Expedia*

Search Cost and Internet

- prior to the Internet, search costs were attributed to dispersion in the geographical location of stores
  - it was costly to visit every store and determine who was offering the lowest price
- however, price search engines like *Pricewatch.com* provide consumers with lost of prices at very low cost
  - online travel agents like *Travelocity* and *Expedia*
- this led many researchers to predict less price dispersion and lower margins in online markets than in brick and mortar markets

Search Cost and Internet

- it is much easier to test this prediction in online markets
  - even if products are identical (e.g., "Da Vinci Code" book) when they are sold in brick and mortar stores, they are differentiated by location
  - products in online markets are not differentiated by location

Search Cost and Internet

- it is much easier to test this prediction in online markets
  - even if products are identical (e.g., "Da Vinci Code" book) when they are sold in brick and mortar stores, they are differentiated by location
  - products in online markets are not differentiated by location
- nevertheless, a long list of papers have found that price dispersion among E-retailers is similar to that of brick and mortar retailers and margins are not extremely low (e.g., Amazon reports average markups of 15%)
  - one reason may be obfuscation, online retailers try to make it difficult for consumers to determine the true price, e.g., shipping cost, taxes, etc.
  - See Ellison and Ellison: "Search, Obfuscation, and Price Elasticities on the Internet", *Econometrica*, 2009.

Sorensen (2000)

- objective: test a comparative static prediction of Varian's theory of price dispersion
  - consumers have stronger incentives to search for lowest price when they have to purchase drug frequently

Sorensen (2000)

- objective: test a comparative static prediction of Varian's theory of price dispersion
  - consumers have stronger incentives to search for lowest price when they have to purchase drug frequently
- hypothesis: lower markups, less dispersion for drugs that are purchased more frequently

Sorensen (2000)

- objective: test a comparative static prediction of Varian's theory of price dispersion
  - consumers have stronger incentives to search for lowest price when they have to purchase drug frequently
- hypothesis: lower markups, less dispersion for drugs that are purchased more frequently
- data: prices of 152 top-selling prescription drugs
  - 10 pharmacies in Middletown, NY
  - 11 pharmacies in Newburgh, NY

Data Overview

- in New York, pharmacies are required to post prices on a large poster provided by State Board of Pharmacy
- the two towns are geographically isolated so consumers in these towns have to buy drugs at these pharmacies

Data Overview

- in New York, pharmacies are required to post prices on a large poster provided by State Board of Pharmacy
- the two towns are geographically isolated so consumers in these towns have to buy drugs at these pharmacies

- drug characteristics were collected from Moseby's Genrx: primary use, dosage, duration of therapy, wholesale prices

Data Overview

- in New York, pharmacies are required to post prices on a large poster provided by State Board of Pharmacy
- the two towns are geographically isolated so consumers in these towns have to buy drugs at these pharmacies

- drug characteristics were collected from Moseby's Genrx: primary use, dosage, duration of therapy, wholesale prices

- drug dosage/therapy duration used to compute expected number of purchase times per year

Descriptive Statistics

- price variation: posted prices averaged \$13.17, with 10th and 90th percentiles of \$4.91 and \$25.36

Descriptive Statistics

- price variation: posted prices averaged \$13.17, with 10th and 90th percentiles of \$4.91 and \$25.36

- pharmacies could not be easily sorted into low-price and high-price categories

Descriptive Statistics

- price variation: posted prices averaged \$13.17, with 10th and 90th percentiles of \$4.91 and \$25.36
- pharmacies could not be easily sorted into low-price and high-price categories

Price Rankings by Pharmacy
A. Middletown

	Price Group		
Pharmacy	Lowest 3	Middle 4	Highest 3
Eckerd	45	103	10
Eckerd	29	102	27
Immediate	43	54	61
K-Mart	56	57	45
Medicine Shoppe	99	49	10
Price Chopper	80	67	11
Rite-Aid	3	11	144
Rite-Aid	2	18	138
Rx Place	38	104	16
Wal-Mart	79	67	12

Descriptive Statistics

B. Newburgh

	Price Group		
Pharmacy	Lowest 3	Middle 3	Highest 3
Ace	26	112	30
Hudson	33	106	29
Medical Arts	73	65	30
Price Chopper	134	27	7
Rite-Aid	4	23	141
Rite-Aid	10	45	113
Rite-Aid	18	34	116
Rx Place	64	70	34
Wal-Mart	142	22	4

Descriptive Statistics

B. Newburgh

Pharmacy	Price Group		
	Lowest 3	Middle 3	Highest 3
Ace	26	112	30
Hudson	33	106	29
Medical Arts	73	65	30
Price Chopper	134	27	7
Rite-Aid	4	23	141
Rite-Aid	10	45	113
Rite-Aid	18	34	116
Rx Place	64	70	34
Wal-Mart	142	22	4

- (kind of) consistent with randomization on individual drug prices (a lot of uncertainty in rivals' prices)

Regression I

- regression equation

$$RANGE_{ij} = \beta_0 + \beta_1 PFREQ_i + other\ controls + \epsilon_{ij}$$

where $RANGE_{ij}$ is price range of drug i in city j, $PFREQ_i$ is purchase frequency of drug i

Regression I

- regression equation

$$RANGE_{ij} = \beta_0 + \beta_1 PFREQ_i + other\ controls + \epsilon_{ij}$$

where $RANGE_{ij}$ is price range of drug i in city j, $PFREQ_i$ is purchase frequency of drug i

PRICE DISPERSION AND PURCHASE FREQUENCY

	DISPERSION MEASURE			
	Range (1)	Standard Deviation (2)	Residual Range (3)	Residual Standard Deviation (4)
Purchase frequency	−.336	−.173	−.266	−.102
	(.123)	(.076)	(.061)	(.016)
Wholesale cost	.280	.180	.215	.069
	(.033)	(.020)	(.043)	(.014)
Branded with generic competition	−.803	−1.480	−1.842	−.362
	(1.037)	(.641)	(.861)	(.248)
Branded without generic competition	−1.505	−2.010	−1.967	−.772
	(2.108)	(1.303)	(1.060)	(.339)
Newburgh dummy	−2.686	−3.172	−1.493	−.916
	(.633)	(.314)	(.791)	(.271)
Constant	20.070	7.321	14.570	5.283
	(4.343)	(2.563)	(1.062)	(.448)
R^2	.371	.447	.258	.253
$\hat{\rho}$	.338	.585	.149	.648

Regression II

AVERAGE MARGINS AND PURCHASE FREQUENCY

	DEPENDENT VARIABLE		
	Average Margin (1)	Average Price (2)	Average Relative Margin (3)
Purchase frequency	−.262 (.102)	−.137 (.105)	.001 (.003)
Wholesale cost	...	.994 (.032)	...
Wholesale cost × generic dummy	...	−.208 (.059)	...
Branded with generic competition	2.101 (.720)	−.668 (1.056)	−.235 (.020)
Branded without generic competition	3.415 (1.660)	−.123 (1.891)	−.255 (.046)
Newburgh dummy	1.681 (.174)	1.648 (.140)	.047 (.005)
Constant	12.69 (2.435)	11.86 (2.581)	.463 (.068)
R^2	.229	.895	.510
$\hat{\rho}$	.915	.936	.898

- margins are negatively correlated with drug frequency: 37% lower for drugs that are purchased monthly versus drugs purchased only once

Remarks

- alternative explanations
  - pharmacy heterogeneity: fixed effects account for 33% of the dispersion in prices
  - cost heterogeneity: differences in drug acquisition costs across pharmacies are too small

Remarks

- alternative explanations
  - pharmacy heterogeneity: fixed effects account for 33% of the dispersion in prices
  - cost heterogeneity: differences in drug acquisition costs across pharmacies are too small
- main conclusion: price dispersion is substantial, it is positively correlated with drug purchase frequency

Consumer Search Models

- Varian (1980) isn't technically a model of consumer search
- various subsequent authors proposed models in which consumers' "informedness" is endogenous
  - consumers have search costs
  - the search for lower prices if expected benefit of search is greater than search cost
- two main modeling approaches
  - fixed sample size search
  - sequential search

Nonsequential Search: Burdett and Judd (1983)

- consumers know price distribution $F(p)$

Nonsequential Search: Burdett and Judd (1983)

- consumers know price distribution $F(p)$
- choose in advance how many price quotes to obtain
  - one rational: maybe price quotes come after a delay

Nonsequential Search: Burdett and Judd (1983)

- consumers know price distribution $F(p)$
- choose in advance how many price quotes to obtain
  - one rational: maybe price quotes come after a delay
- if cost of each price quote is c, then getting n price quotes gives an expected total purchase cost equal to

$$cn + \int_0^\infty np\,[1 - F(p)]^{n-1}\, dF(p)$$

  - this is a convex function of n, so there exists a unique integer that minimizes total purchase cost (or two adjacent integers that tie)

Sequential Search: Stahl (1989)

- consumers know price distribution $F(p)$
  - fraction μ of zero-search-cost "shoppers"
  - fraction $1 - \mu$ have common search cost $c > 0$ per price quote

Sequential Search: Stahl (1989)

- consumers know price distribution $F(p)$
  - fraction μ of zero-search-cost "shoppers"
  - fraction $1-\mu$ have common search cost $c > 0$ per price quote

- if best price found so far is z and $z < WTP$, expected benefit from an additional search is

$$\int_{\underline{p}}^{z} (z-p)\, dF(p)$$

  - search again if expected benefit is greater than cost c

Sequential Search: Stahl (1989)

- consumers know price distribution $F(p)$
  - fraction μ of zero-search-cost "shoppers"
  - fraction $1-\mu$ have common search cost $c > 0$ per price quote
- if best price found so far is z and $z < WTP$, expected benefit from an additional search is

$$\int_{\underline{p}}^{z} (z-p)\, dF(p)$$

  - search again if expected benefit is greater than cost c
- optimal search rule is a reservation price rule: if you find a price smaller than r^*, then buy; otherwise searching

Remarks

- most models assume consumers know the underlying price distribution

Remarks

- most models assume consumers know the underlying price distribution
- what if instead they learn about the distribution through search?

Remarks

- most models assume consumers know the underlying price distribution
- what if instead they learn about the distribution through search?
- Rothschild (1974)
  - many properties of optimal search behavior look similar
  - but no longer possible to characterize equilibrium price distribution

Structural Estimation

- some papers test the predictions of search models, e.g., Sorensen (2000)

Structural Estimation

- some papers test the predictions of search models, e.g., Sorensen (2000)
- but what about estimating search models directly?

Structural Estimation

- some papers test the predictions of search models, e.g., Sorensen (2000)
- but what about estimating search models directly?
- can we use price distribution to estimate search costs?
  - Hong and Shum (2006), Hortacsu and Syverson (2004), etc.

Hong and Shum (2000): Nonsequential Search Case

- idea: if we impose the equilibrium restrictions of a search model, we can recover the unobserved distribution of search cost from the observed distribution of prices

Hong and Shum (2000): Nonsequential Search Case

- idea: if we impose the equilibrium restrictions of a search model, we can recover the unobserved distribution of search cost from the observed distribution of prices
- cutoff-points: expected price savings from the k^{th} price quote

$$\Delta_k = E\left[\min\left(p_1, ..., p_k\right)\right] - E\left[\min\left(p_1, ..., p_{k-1}\right)\right]$$

Hong and Shum (2000): Nonsequential Search Case

- idea: if we impose the equilibrium restrictions of a search model, we can recover the unobserved distribution of search cost from the observed distribution of prices
- cutoff-points: expected price savings from the k^{th} price quote

$$\Delta_k = E\left[\min\left(p_1, ..., p_k\right)\right] - E\left[\min\left(p_1, ..., p_{k-1}\right)\right]$$

- consumers who obtain k price quotes must have search costs between Δ_k and Δ_{k+1}

Hong and Shum (2000): Nonsequential Search Case

- how many consumers are obtaining k price quotes? use supply-side mixed-strategy equilibrium condition: expected profits must equal at all prices

$$(\bar{p} - c)\,\hat{q}_1 = (p_i - c)\left[\sum_{k=1}^{K} \hat{q}_k k \left(1 - \hat{F}_p(p_i)\right)^{k-1}\right]$$

Hong and Shum (2000): Nonsequential Search Case

- how many consumers are obtaining k price quotes? use supply-side mixed-strategy equilibrium condition: expected profits must equal at all prices

$$(\bar{p} - c)\,\hat{q}_1 = (p_i - c)\left[\sum_{k=1}^{K} \hat{q}_k k \left(1 - \hat{F}_p(p_i)\right)^{k-1}\right]$$

- since $\sum_{k=1}^{K} \hat{q}_k = 1$, above gives $n-1$ equations to solve for K unknowns $\{c, \hat{q}_1, ..., \hat{q}_{K-1}\}$

Hong and Shum (2000): Nonsequential Search Case

- how many consumers are obtaining k price quotes? use supply-side mixed-strategy equilibrium condition: expected profits must equal at all prices

$$(\bar{p} - c)\,\hat{q}_1 = (p_i - c)\left[\sum_{k=1}^{K} \hat{q}_k k \left(1 - \hat{F}_p(p_i)\right)^{k-1}\right]$$

- since $\sum_{k=1}^{K} \hat{q}_k = 1$, above gives $n-1$ equations to solve for K unknowns $\{c, \hat{q}_1, ..., \hat{q}_{K-1}\}$

- use estimates $\hat{q}_k$ to solve for $F_c(\Delta_1), ..., F_c(\Delta_{K-1})$

Hortacsu and Syverson (2004)

- question: what explains the diffuse prices of seemingly similar mutual funds? specifically, what level of search costs would rationalize the observed price dispersion?

Hortacsu and Syverson (2004)

- question: what explains the diffuse prices of seemingly similar mutual funds? specifically, what level of search costs would rationalize the observed price dispersion?
- strategy: use price and quantity data from S&P 500 funds, estimate a search model that allows for vertical differentiation in addition to search frictions

Model: Brief Overview

- consumers know empirical CDF of offered utilities: $u_1 < u_2 < ... < u_N$, but cannot identify any product's position
- optimal search implies cutoff points of the search cost distribution

$$c_j = \sum_{k=j}^{N} \rho_k \left(u_k - u_j\right)$$

where ρ_k is the sampling probability of k

Model: Brief Overview

- consumers know empirical CDF of offered utilities: $u_1 < u_2 < ... < u_N$, but cannot identify any product's position
- optimal search implies cutoff points of the search cost distribution

$$c_j = \sum_{k=j}^{N} \rho_k \left(u_k - u_j\right)$$

where ρ_k is the sampling probability of k

- market shares can be mapped to these cutoffs

$$q_1 = \rho_1 \left[1 - G\left(c_1\right)\right]$$

$$q_2 = \rho_2 \left[1 - G\left(c_1\right)\right] + \frac{\rho_2}{1-\rho_1}\left[G\left(c_1\right) - G\left(c_2\right)\right]$$

$$\vdots$$

  - search cost distribution: recovered from market shares and optimal search condition

De Los Santos, Hortacsu, and Wildenbeest (2012)

- question: which model of search, sequential or non-sequential, is better supported by data on actual consumers' searches

De Los Santos, Hortacsu, and Wildenbeest (2012)

- question: which model of search, sequential or non-sequential, is better supported by data on actual consumers' searches

- strategy: use Comscore data on book purchases (and searches) to test key predictions of the sequential model: no recall (should always buy from the last store visited), and price dependence (decision to search again should depend on price of last store searched)

Comscore Data

- tracks web usage of a large sample of users, including online purchases

Comscore Data

- tracks web usage of a large sample of users, including online purchases
- not a random sample of internet users (because users must agree to be tracked)

Comscore Data

- tracks web usage of a large sample of users, including online purchases
- not a random sample of internet users (because users must agree to be tracked)
- this paper: focus on searches and purchases of books
  - approx 15,500 purchase transactions from 2002 to 2004
  - approx 325,000 site visits to 15 online bookstores

Testing Predictions of Sequential Search

- no recall: under sequential search, buyers should purchase from last store visited (unless their search exhausts all stores: would buy from the lowest price in this case, by assumption)

Testing Predictions of Sequential Search

- no recall: under sequential search, buyers should purchase from last store visited (unless their search exhausts all stores: would buy from the lowest price in this case, by assumption)
- price dependence: under sequential search, the probability of another search depends on the value of the last price quote (only keep searching if last quoted price is high)

Summary Stats: Stores

Bookstore	Transactions Number	Transactions %	Visits Number	Visits %
Amazon	10,197	65.5%	249,593	76.3%
Barnes and Noble	3,042	19.6%	25,758	7.9%
Book Clubs				
christianbook.com	615	3.9%	3,968	1.2%
doubledaybookclub.com	468	3.0%	4,001	1.2%
eharlequin.com	61	0.4%	3,647	1.1%
literaryguild.com	322	2.1%	3,500	1.1%
mysteryguild.com	187	1.2%	2,095	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%	1,206	0.4%
powells.com	68	0.4%	1,326	0.4%
varsitybooks.com	16	0.1%	218	0.1%
walmart.com	183	1.2%	28,663	8.8%
booksamillion.com	246	1.6%	2,290	0.7%
Total	15,561	100.0%	327,074	100.0%

Summary Stats: Searches

	2002		2004	
	Mean	Std. Dev.	Mean	Std. Dev.
Duration of each website visit (in minutes)				
Visits not within 7 days of transaction	8.89	13.03	7.69	12.36
Visits within 7 days, excluding transactions	12.72	15.83	11.02	15.00
Visits within 7 days, including transactions	19.04	18.26	15.74	17.37
Transactions only	28.06	17.69	26.08	17.71
Total duration, excluding transaction visits	32.47	49.80	38.41	78.33
Total duration, including transaction visits	43.88	43.27	47.43	66.11
Number of stores searched	1.27	0.54	1.30	0.56
Number of books per transaction	2.38	2.10	2.20	1.95
Transaction expenditures (books only)	36.67	40.64	32.21	35.68
Number of books purchased	17,956		17,631	
Number of transaction sessions	7,559		8,002	
Number of visits within 7 days	18,350		25,556	
Number of visits not within 7 days	94,011		189,157	

Evidence of Recall

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

Evidence of Price Dependence

- dependent variable: indicator for searching more than 1 store

Variable	(A)	(B)	(C)	(D)	(E)
Panel A. All transactions					
Coefficients					
Intercept	-1.817 (0.093)	-1.796 (0.123)	-0.818 (0.154)	0.295 (0.133)	—
First price lower or equal	-0.071 (0.119)	0.090 (0.152)	0.040 (0.157)	-0.223 (0.165)	-0.073 (0.371)
Loyal	—	—	-1.446 (0.151)	—	—
Average marginal effects					
First price lower or equal	-0.008 (0.014)	0.011 (0.019)	0.005 (0.019)	-0.055 (0.041)	-0.015 (0.078)
Loyal	—	—	-0.171 (0.017)	—	—
Number of observations	2,593	1,504	1,504	649	235

- weak relation between decision to continue searching and observed prices

Outline of Demand Estimation with Search

- basic idea: probability i purchases from j is $P_{iS}P_{ij|S}$
  - $P_{ij|S}$ is the probability of choosing j if set of searched stores was S
  - P_{iS} is probability i chooses to search the set of stores S

Outline of Demand Estimation with Search

- basic idea: probability i purchases from j is $P_{iS} P_{ij|S}$
  - $P_{ij|S}$ is the probability of choosing j if set of searched stores was S
  - P_{iS} is probability i chooses to search the set of stores S

- expected payoff if search set is S

$$m_{iS} = E\left[\max_{j \in S} \{\mu_j + X_j \beta_j + \alpha_i p_j + \epsilon_{ij}\}\right] - kc_i$$

Computing $P_i S$

- if consumers know ϵ and prices have T1EV distribution, then

$$E\left[\max_{j \in S}\{u_{ij}\}\right] = \alpha_i \sigma \log\left(\sum_{j \in S} \exp\left[\frac{\mu_j + X_i \beta_j + \epsilon_{ij} + \alpha_i \gamma_j}{\alpha_i \sigma}\right]\right)$$

  where γ_j and σ are location and scale parameters of stores' price distributions

Computing P_iS

- if consumers know ϵ and prices have T1EV distribution, then

$$E\left[\max_{j\in S}\{u_{ij}\}\right] = \alpha_i\sigma \log\left(\sum_{j\in S}\exp\left[\frac{\mu_j + X_i\beta_j + \epsilon_{ij} + \alpha_i\gamma_j}{\alpha_i\sigma}\right]\right)$$

  where γ_j and σ are location and scale parameters of stores' price distributions

- to smooth the choice probabilities, add a logit shock ς_{iS} to each m_{iS}, so

$$P_{iS} = \frac{\exp\left(m_{iS}/\sigma_\varsigma\right)}{\sum_{S'}\exp\left(m_{iS'}/\sigma_\varsigma\right)}$$

Estimation Details

- $P_{ij|S}$ cannot be calculated analytically
- calculate γ_j's and σ_p from observed price distributions in a first step, then treat as known by consumers
- normalization: variance of ϵ's set to 1, so variance of choice-set error σ_ς is estimated relative to variance of store-specific errors

Main Results

- average search costs around \$1.35
- own-price elasticities around -1 for Amazon, around -2 for B&N
- if demand is estimated assuming full information, then estimated elasticities are much smaller
  - consumers' unresponsiveness to price differences attributed to small α instead of to search frictions

Reading for Next Class

- Ackerberg, D., "Empirically Distinguishing Informative and Prestige Effects of Advertising", *RAND Journal of Economics*, 2001.