Price dispersion and search

Paul Schrimpf

UBC
Economics 565

November 25, 2015

References

- Theory and overview of reduced form empirical work: Baye, Morgan, and Scholten (2006)
- Structural empirical papers:
    - Hong and Shum (2006), Moraga-González and Wildenbeest (2008), De los Santos, Hortaçsu, and Wildenbeest (2012), Wildenbeest (2011)

Section 1

Introduction

- Homogenous product and competitive market:
    - Theory $\Rightarrow$ one price
    - Reality $\Rightarrow$ price dispersion
- Explanations:
    - Unobserved product heterogeneity
        - Likely part of explanation, but it is largely tautological
    - **Imperfect information and search costs**
        - Stigler (1961)

Table 1a: Measures of Price Dispersion Reported in the Literature in Offline Markets[1]

Study	Data Period	Product Market	Intervals of Estimated Price Dispersion Measures	Dispersion Measure
Bailey (1998)	1997	Books	13.2%	Standard Deviation
		Books	10.4%	Standard Deviation
		Compact Discs	17.6%	Standard Deviation
		Compact Discs	11.0%	Standard Deviation
		Software	7.1%	Standard Deviation
		Software	8.1%	Standard Deviation
Borenstein and Rose (1994)	1986	U.S. Airline	0.018 - 0.416	Gini coefficient
Carlson and Pescatrice (1980)	1976	Consumer Sundries	3.3% - 41.4%	Coefficient of Variation
Eckard (2004)	1901 - 2001	Baking Powder, Sugar, Salt -- 1901	3.1% - 10.1%	Coefficient of Variation
		Baking Powder, Sugar, Salt -- 2001	0.0% - 13.4%	Coefficient of Variation
Friberg, Ganslandt and Sandstrom (2001)	1999	Books	\$54.00 - \$122.00	Range
		Books	\$21.94 - \$76.20	Standard Deviation
		Compact Discs	\$20.00 - \$40.00	Range
		Compact Discs	\$12.91 - \$23.86	Standard Deviation
		Books (Sweden)	\$19.00 - \$58.00	Range
		Compact Discs (Sweden)	\$21.00 - \$46.00	Range
Lach (2002)	1993 - 1996	Refrigerator (Israel)	4.9%	Coefficient of Variation
		Chicken, Flour, Coffee (Israel)	11.4% - 19.7%	Coefficient of Variation
Marvel (1976)	1964 - 1971	Regular Gasoline	\$0.048	Range
		Regular Gasoline	\$0.015	Standard Deviation
		Premium Gasoline	\$0.048	Range
		Premium Gasoline	\$0.017	Standard Deviation
Pratt, Wise and Zeckhauser (1979)	1975	Various Products and Services	4.4% - 71.4%	Coefficient of Variation
		Various Products and Services	11.0% - 567.0%	Range
		Various Products and Services	7.2% - 200.0%	Value of Information
Roberts and Supina (2000)	1963 - 1987	Wood Products	13.8% - 90.2%	Coefficient of Variation
		Fabrics	18.8% - 78.1%	Coefficient of Variation
		Coffee	14.3% - 25.1%	Coefficient of Variation
		Ready-Mixed Concrete	13.2% - 37.2%	Coefficient of Variation
		Newsprint	4.5% - 8.2%	Coefficient of Variation
		Gasoline	6.2% - 11.8%	Coefficient of Variation
		Tinplate Steel Cans	25.0% - 31.0%	Coefficient of Variation
		Pan Bread	26.0% - 49.6%	Coefficient of Variation
		Corrugated Shipping Containers	21.8% - 39.6%	Coefficient of Variation
Scholten and Smith (2002)	1976 - 2000	Consumer Sundries -- 1976	3.3% - 41.4%	Coefficient of Variation
		Consumer Sundries -- 2000	1.6% - 42.0%	Coefficient of Variation
		Consumer Sundries -- 2000	5.7% - 28.4%	Coefficient of Variation
Sorensen (2000)	1998	Prescription Drugs	\$13.17	Range
		Prescription Drugs	22.0%	Coefficient of Variation
Stigler (1961)	1953	Anthracite Coal	\$3.46	Range
		Anthracite Coal	\$1.15	Standard Deviation
	1959	Identical Automobiles	\$165.00	Range
		Identical Automobiles	\$42.00	Standard Deviation
Villas-Boas (1995)	1985 - 1987	Coffee	21.5%	Coefficient of Variation

[1] Table 1a includes studies comparing offline and online price dispersion.

Table 1b: Measures of Price Dispersion Reported in the Literature in Online Markets Only

Study	Data Period	Product Market	Intervals of Estimated Price Dispersion Measures	Dispersion Measure
Ancarani and Shankar (2004)	2002	Books (Italy)	€4.26 - €4.84	Standard Deviation
		Books (Italy)	€20.00 - €22.88	Range
		Compact Discs (Italy)	€2.29 - €2.79	Standard Deviation
		Compact Discs (Italy)	€11.82 - €14.75	Range
Arbatskaya and Baye (Forthcoming)	1998	Mortgage Interest Rates	> 0.25	Range
Arnold and Saliba (2002)	2001	Textbooks	10.7% - 52.6%	Range
		Textbooks	3.5% - 10.0%	Coefficient of Variation
		Textbooks	0.2% - 12.5%	Price gap
Baye, Morgan and Scholten (2003)	2000 - 2001	Consumer Electronics	$123.88 - $143.15	Range
Baye, Morgan and Scholten (2004a)	2000 - 2001	Consumer Electronics	9.1% - 9.7%	Coefficient of Variation
		Consumer Electronics	3.79% - 5.38%	Gap
Baye, Morgan and Scholten (2004b)	1999 - 2001	Consumer Electronics	57.4%	Range
		Consumer Electronics	12.5%	Coefficient of Variation
Baylis and Perloff (2002)	1999	Camera	$342.00	Range
		Scanner	$106.00	Range
Brynjolfsson and Smith (2000)	1998-1999	Books	33.0%	Range
		Compact Discs	25.0%	Range
Chevalier and Goolsbee (2003)	2001	Books	8.1% - 12.3%	Range
Clay, Krishnan, and Wolff (2001)	1999 - 2000	Books	27.7%	Coefficient of Variation
		Books	$7.62	Range
Clay, Krishnan, Wolff and Fernandes (2003)	1999	Books	10.0% - 18.0%	Coefficient of Variation
Clemons, Hann and Hitt (2002)		Travel	$8.03 - $13.40	Range
Ellison and Ellison (2004)	2000 - 2001	Memory Modules	5.9% - 29.0%	Range
Gatti and Kattuman (2003)	2002	Consumer Electronics (France)	3.0% - 15.3%	Coefficient of Variation
		Consumer Electronics (Italy)	4.3% - 14.2%	Coefficient of Variation
		Consumer Electronics (Netherlands)	5.6% - 20.4%	Coefficient of Variation
		Consumer Electronics (Spain)	2.2% - 13.3%	Coefficient of Variation
		Consumer Electronics (Sweden)	6.6% - 14.0%	Coefficient of Variation
		Consumer Electronics (UK)	3.5% - 16.2%	Coefficient of Variation
		Consumer Electronics (Denmark)	6.3% - 20.2%	Coefficient of Variation
		Consumer Electronics (France)	7.8% - 47.4%	Range
		Consumer Electronics (Italy)	9.3% - 27.8%	Range
		Consumer Electronics (Netherlands)	8.9% - 54.6%	Range
		Consumer Electronics (Spain)	3.8% - 32.4%	Range
		Consumer Electronics (Sweden)	16.4% - 50.4%	Range

Section 2

Models of price dispersion

Models of price dispersion

- Key point: relationship between price dispersion and primitives (search cost, market size, number of firms, demand elasticity) depends on modeling assumptions
- Types:
  1. Search
     1. Fixed : gather *n* prices, choose lowest price
     2. Sequential : sequentially gather prices, stop when price low enough
     3. Information clearinghouse : some consumers loyal to one firm, others buy from lowest price
  2. Bounded rationality: small departure from Nash equilibrium in firms' pricing game can give large price dispersion
     - Quantal response equilibrium, ϵ-equilibrium, mistaken beliefs about price distribution

Fixed search

- Stigler (1961)
- Assumptions:
  1. Distribution of prices on $[\underline{p}, \overline{p}]$, non-degenerate CDF $F(p)$, known by consumers
  2. Each consumer wants to buy K units
  3. Search process: optimally choose fixed number of price quotes, n; buy from firm with lowest price

Fixed search - model implications

- Number of price quotes:

$$\left(\mathrm{E}[p^{(1:n^*-1)}] - \mathrm{E}[p^{(1:n^*)}]\right) K \geq c \geq \left(\mathrm{E}[p^{(1:n^*)}] - \mathrm{E}[p^{(1:n^*+1)}]\right) K$$

  n^* increasing in K
- Firm expected demand:

$$Q(p) = \mu n^* K (1 - F(p))^{n^*-1}$$

- Transaction costs decrease with price dispersion
  - If G is a mean preserving spread of F, then $\mathrm{E}_G[p^{(1:n)}] < \mathrm{E}_F[p^{(1:n)}]$ for $n > 1$
- Expected total costs are lower with greater price dispersion
  - If G is a mean preserving spread of F, then $\mathrm{E}_G[p^{(1:n_G^*)}]K - cn_G^* < \mathrm{E}_F[p^{(1:n_F^*)}]K - cn_F^*$ for $n > 1$

Fixed search - critique I

- Rothschild (1973) critique:
  1. distribution of prices is not endogenous
  2. fixed search may not be optimal for consumers
- For (2) need to be more specific about search environment
  - Fixed search optimal if e.g. waiting time to obtain each price quote
- Diamond (1971) in sequential or fixed search model with homogenous firms and consumers, there is an equilibrium where all firms charge the monopoly price
- Can obtain non-degenerate equilibrium distribution of prices by introducing firm heterogeneity or consumer search cost heterogeneity

Fixed search - endogenous price dispersion I

- Burdett and Judd (1983) : equilibrium price dispersion with ex-ante identical consumers and firms
- Assumptions:
  1. Consumers: unit demand with reservation price v
  2. Fixed sample search
  3. Firms: constant marginal cost m, optimal monopoly price p^*
  4. Consumer utility given price p^* and $n = 1$ is positive
- Equilibrium: price distribution, $F(p)$, and search distribution, $\mathbb{P}(n = i)$ for $i = 1, 2, \ldots$

Fixed search - endogenous price dispersion I

- Implications:
  - If $F(p)$ non-degenerate, then $\mathrm{P}(n=1), \mathrm{P}(n=2) > 0$ and $\mathrm{P}(n>2)=0$, let $\theta = \mathrm{P}(n=1)$, $1-\theta = \mathrm{P}(n=2)$
  - Firm profits:

$$\pi(p) = \begin{cases} (v-m)\theta & \text{if } p = v \\ (p-m)\mathrm{P}(\text{consumer purchases}) & \text{if } p < v \end{cases}$$

$$= \begin{cases} (v-m)\theta & \text{if } p = v \\ (p-m)\left[\theta + (1-\theta)(1-F(p))\right] & \text{if } p < v \end{cases}$$

  - Firms indifferent among prices implies:

$$F(p) = 1 - \frac{v-p}{p-m}\frac{\theta}{1-\theta}$$

Fixed search - endogenous price dispersion II

- Consumers indifferent between $n = 1$ and $n = 2$ pins down θ (generally two equilibria with $\theta \in (0, 1)$ (there's also an equilibrium where firms charge monopoly price and $n = 1$ for all consumers))

Sequential search I

- Sequential search: consumer pays cost c to obtain price $p \sim F$; can either buy at price p (or any previous price) or search again
- Optimal strategy = reservation price $p^* = \min\{\overline{p}, z^*\}$ where

$$c = \int_{\underline{p}}^{z^*} (z^* - p) f(p) dp = \int_{\underline{p}}^{z^*} F(p) dp$$

- With homogeneous firms and consumers unique equilibrium is for firms to charge the monopoly price
- Equilibrium price dispersion with:
  - Heterogeneous firm marginal cost and elastic demand (i.e. not unit demand); or
  - Heterogeneous search costs (and assumptions about distribution of search costs)

Information clearinghouse

- Finite number $n > 1$ of homogeneous firms
  - Constant marginal cost c
  - Clearinghouse charges $\phi \geq 0$ to firms to list their prices
- Consumers with unit demand and reservation price v
  - $S > 0$ "shoppers" consult clearinghouse, buy at lowest price if $< v$, else visits one other firm buys if price $< v$, else does not buy
  - $L > 0$ "loyal" consumers visit firm i, buy if $p_i < v$
- Equilibrium with price dispersion if $L > 0$ or $\phi > 0$
  - Non-clearinghouse prices all $= v$
  - Distribution of clearinghouse prices $\leq v$

Section 3

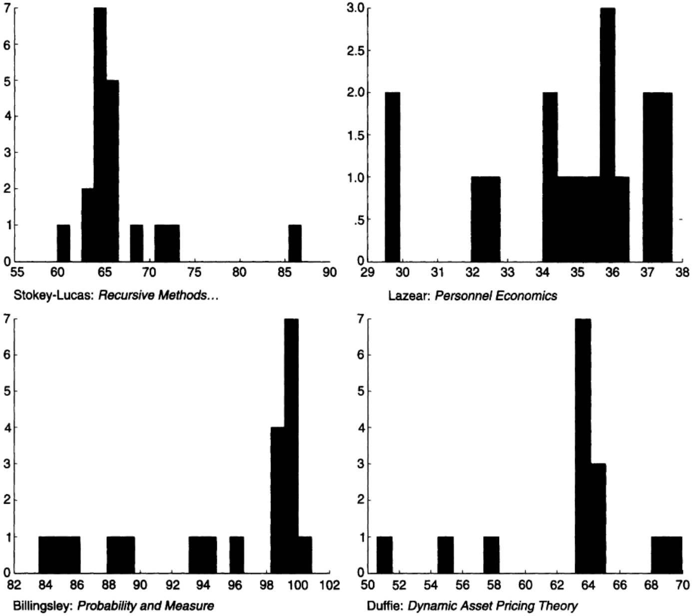
Hong and Shum (2006)

Hong and Shum (2006): "Using price distributions to estimate search costs"

- Goal: estimate consumer search costs
- Environment: online booksellers
  - Homogeneous product
  - Homogeneous firm costs
- Data: distribution of prices
- Method: use distribution of prices + assumption about form of search to estimate distribution of consumer costs

FIGURE 1

RAW HISTOGRAMS OF ONLINE PRICES



x-axis: prices (in dollars)

Fixed search I

- Firm marginal cost r, continuum of firms with equilibrium price distribution F_p
- Consumer search cost $c_i \sim F_c$
- Number of searches

$$\left(\mathrm{E}[p^{(1:n(c_i)-1)}] - \mathrm{E}[p^{(1:n(c_i))}]\right) K \geq c_i \geq \left(\mathrm{E}[p^{(1:n(c_i))}] - \mathrm{E}[p^{(1:n(c_i)+1)}]\right.$$

- Define $\Delta_n = \mathrm{E}[p^{(1:n-1)}] - \mathrm{E}[p^{(1:n)}]$; F_P observed, so Δ_n identified
- Let $\tilde{q}_n = F_c(\Delta_{n-1}) - F_c(\Delta_n) =$ portion of consumers who obtain n prices
  - $\tilde{q}_n$ not observed
  - Assume F_c such that $\tilde{q}_n = 0$ for all $n > K$ (could be relaxed, but complicates econometrics)

Fixed search II

- Firms indifferent among prices $p \in [\underline{p}, \overline{p}]$, so

$$(\overline{p} - r)\tilde{q}_1 = (p - r)\left[\sum_{k=1}^{K} \tilde{q}_k k \left(1 - F_p(p)\right)^{k-1}\right]$$

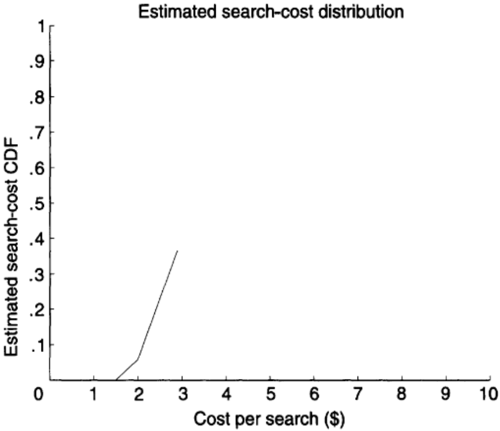
- Observed prices p_j, $j = 1, \ldots, n_f$

$$(\overline{p} - r)\tilde{q}_1 = (p - r)\left[\sum_{k=1}^{K} \tilde{q}_k k \left(1 - \hat{F}_p(p_j)\right)^{k-1}\right]$$

identifies $\tilde{q}_1, \ldots, \tilde{q}_K$ and r

- Knowing $\tilde{q}_1, \ldots, \tilde{q}_K$ can solve for $F_c(\Delta_1), \ldots, F_c(\Delta_K)$
- Estimate using empirical likelihood ($\approx$ efficiently weighted GMM)

- Estimates for Billingsley using 20 prices, $K = 3$, and 5 moments
- $\tilde{q}_1 = 0.633$, $\tilde{q}_2 = 0.309$, $\tilde{q}_3 = 0.058$



Nonsequential search model ($K = 3$, $M = 5$)

Sequential search I

- Consumer search cost $c_i \sim F_c$
- Reservation price, $p_i^* = \overline{p}(c_i) = \min\{z(c_i), \overline{p}\}$ where

$$c_i = \int_{\underline{p}}^{z(c_i)} (z(c_i) - p) f(p) dp = \int_{\underline{p}}^{z(c_i)} F(p) dp$$

  Let $G(p) = \text{CDF of } p_i^*$
- Firm indifference:

$$(\overline{p} - r)(1 - G(\overline{p})) = (p - r)(1 - G(p))$$

- Data: n_f prices, but $n_f - 1$ indifference conditions, so need some restriction
  - Parametric assumption about F_c
  - (in fixed search model, assumption about K played a similar role)
  - Or fix r and estimate F_c nonparametrically
- Estimate by MLE

TABLE 1 **Summary Statistics on Prices for Different Products**

Product	n	List	Mean	Standard Deviation	Median	$\underline{p}$	$\overline{p}$
Stokey-Lucas	19	60.50	66.60	5.64	64.98	59.75	86.80
Lazear	17	31.95	34.73	2.48	35.27	29.51	37.70
Billingsley	20	99.95	95.48	5.87	98.90	83.58	100.87
Duffie	15	65.00	62.71	4.91	63.48	50.58	69.95

Note: Including shipping and handling costs. Price data for all products downloaded from Pricescan.com and MySimon.com: February 5, 2002. Summary price including S&H costs may not exceed the corresponding summary price without S&H costs, since we could not determine the shipping and handling charges from some of the websites.

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.

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)).

Results

- For text books: Stokey-Lucas, Lazear, Billingsley, Duffie
- Fixed search model:
    - Median search cost $\approx$ \$2.50 (quantiles above median not identified)
    - 25%tile \$0.68 - \$2.50
    - Selling cost $r \approx 65\%$ of median price
- Sequential search model:
    - Median search cost \$9.22-\$29.40
    - Search cost such that $z(c_i) = \overline{p}$, $\$4.56 - \$19.19$
    - Selling cost $r \approx 40\%$ of median price
- Check whether parametric assumption driving sequential results: fix r and estimate nonparametrically

TABLE 4 **Nonparametric Estimates of Sequential-Search Model, Holding *r* Fixed**

Product	r	Search-Cost Estimates
***r* fixed at sequential-model estimates**[b]		
Stokey-Lucas	22.90	$F_c^{-1}(0.42) = 20.20$
Lazear	11.31	$F_c^{-1}(0.31) = 2.97$
Billingsley	65.37	$F_c^{-1}(0.49) = 5.39$
Duffie	28.24	$F_c^{-1}(0.46) = 7.24$
***r* fixed at nonsequential-model estimates**		
Stokey-Lucas	49.52	$F_c^{-1}(0.5) = 4.52$[d]
Lazear	27.76	$F_c^{-1}(0.5) = 0.26$
Billingsley	69.73	$F_c^{-1}(0.5) = 3.21$
Duffie	35.48	$F_c^{-1}(0.5) = 3.51$

[a] For fixed r, quantiles of search-cost distribution are obtained nonparametrically using equation (14).

[b] As reported in Table 2.

[c] As reported in Table 3.

[d] Median obtained by linear interpolation.

Section 4

Additional empirical work

Additional empirical work I

- Moraga-González and Wildenbeest (2008):
  - Oligopoly version of Hong and Shum (2006) fixed search model
  - MLE instead of nonparametric EL
- Chen, Hong, and Shum (2007):
  - Model selection test to choose between fixed and sequential search
  - Test is inconclusive
- Moraga-González, Sándor, and Wildenbeest (2012)
  - Hong and Shum (2006)/Moraga-González and Wildenbeest (2008) fixed search model with multiple markets
  - Data: multiple markets with common search cost distribution, but different reservation prices, firm costs, and/or number of firms
  - Semi-nonparametric estimator
  - Application: memory chips

Additional empirical work II

- De los Santos, Hortaçsu, and Wildenbeest (2012)
    - Data on web browsing and purchases to test sequential vs fixed search
    - Key difference: behavior in sequential model depends on prices observed so far; in fixed model it does not
    - Context: online book stores
    - Results: favor fixed search model; also evidence of unobserved product heterogeneity (store loyalty)
- Hortaçsu and Syverson (2004)
    - Context: mutual funds
    - Model with search frictions and product heterogeneity
    - Results:
        - Investors value observable nonportfolio product attributes
        - Small search costs can rationalize price dispersion
- Wildenbeest (2011)
    - Vertical product differentiation and search frictions

Additional empirical work III

  - Fixed search model
  - ML estimation
  - Context: grocery items
  - Results: supermarket heterogeneity more important than search frictions
- Honka (2014): search & switching costs in auto insurance
  - Fixed search model
  - Consumer knows price of current insurer, and prices of k others
  - Pays switching cost if change insurer
  - Finds search costs more important than switching costs for customer retention & consumer welfare
- Search with learning: De los Santos, Hortacsu, and Wildenbeest (2012), Koulayev (2013)

Section 5

Koulayev (2013)

Koulayev (2013)

- Search model with unknown distribution
- Model based on Rothschild (1974)
- Applied to S&P 500 mutual funds
- Highlights differences with search model with known price distribution

Model I

- N products with utilities $S_N = \{u_1, \ldots, u_N\}$, where $u_1 > u_2 > \cdots > u_N$
- Consumer believes possible utilities $S_G = \{\tilde{u}_1, \ldots, \tilde{u}_G\}$ with $S_N \subseteq S_G$
- Search technology: each search independent and $\tilde{u}_g$ drawn with probability p_g
- Consumer does not know p_g, has Dirichlet prior with parameters $\alpha_1, \ldots, \alpha_G$,

$$f(\tilde{p}_1, \ldots, \tilde{p}_g) = \frac{\Gamma(\sum \alpha_g)}{\prod \gamma(\alpha_g)} \prod \tilde{p}_g^{\alpha_g - 1}$$

which implies

$$\mathrm{E}[\tilde{p}_j] = \frac{\alpha_j}{\sum \alpha_g}$$

Model II

- Bayesian updating: after seeing $\tilde{u}_g$, n_g times,

$$
\begin{aligned}
f(p|n) \propto & f(n|p)f(p) \\
& \propto \frac{(\sum n_g)!}{\prod n_g!} \prod p_g^{n_g} \frac{\Gamma(\sum \alpha_g)}{\prod \gamma(\alpha_g)} \prod \tilde{p}_g^{\alpha_g - 1} \\
& \propto \frac{\Gamma(\sum \alpha_g + n_g)}{\prod \gamma(\alpha_g + n_g)} \prod \tilde{p}_g^{\alpha_g + n_g - 1}
\end{aligned}
$$

so

$$
\mathrm{E}[\tilde{p}_j | n_1, \ldots, n_g] = \frac{\alpha_j + n_j}{\sum \alpha_g + n_g}
$$

- Sequential search and at end buy best good found
- Search cost c, best good found so far u_{r^*}

Model III

- Continue searching if

$$\mathrm{E}[\max\{\tilde{u}, u_{r^*}\}|n] - u_{r^*} > c$$

$$\sum_{\tilde{u}_g > u_{r^*}} (\tilde{u}_g - u_{r^*})\mathrm{E}[\tilde{p}_g|n] > c$$

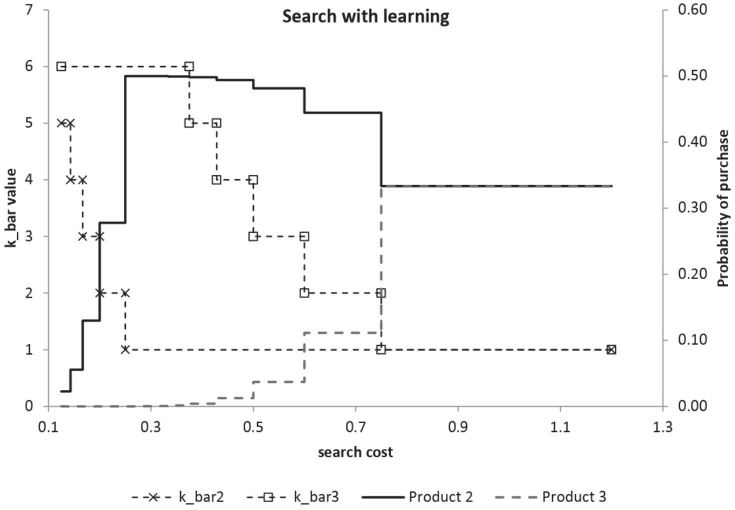
Market shares I

- Observe: market shares, product characteristics
- Consumers have different search costs $c_i \sim F(c)$
- Challenge: many search histories can lead to the same choice; need to integrate over all search histories to compute market shares
    - Define $\bar{k}_r$ = longest a consumer with best draw u_r will continue searching

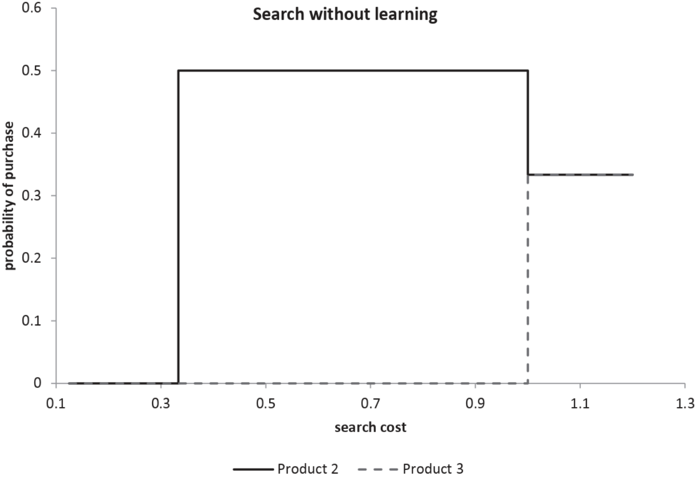
$$\bar{k}_r = \max\left\{1, \left\lfloor \frac{1}{c} \sum_{\tilde{u}_g > u_{r^*}} (\tilde{u}_g - u_{r^*})\alpha_g - \sum_g \alpha_g \right\rfloor\right\}$$

    - Show that market shares can be written as a function of just the $\bar{k}_1, ..., \bar{k}_N$
    - $\bar{k}_r$ is integer valued and decreasing in c

Example


Search with learning
k_bar value
Probability of purchase
search cost
0
1
2
3
4
5
6
7
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.1
0.3
0.5
0.7
0.9
1.1
1.3
k_bar2
k_bar3
Product 2
Product 3

Example


Search without learning
probability of purchase
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
0.3
0.5
0.7
0.9
1.1
1.3
search cost
Product 2
Product 3

- Paper has simulations comparing price elasticity in search with learning versus search without learning models
- Simulations also show that ignoring learning can lead to bias

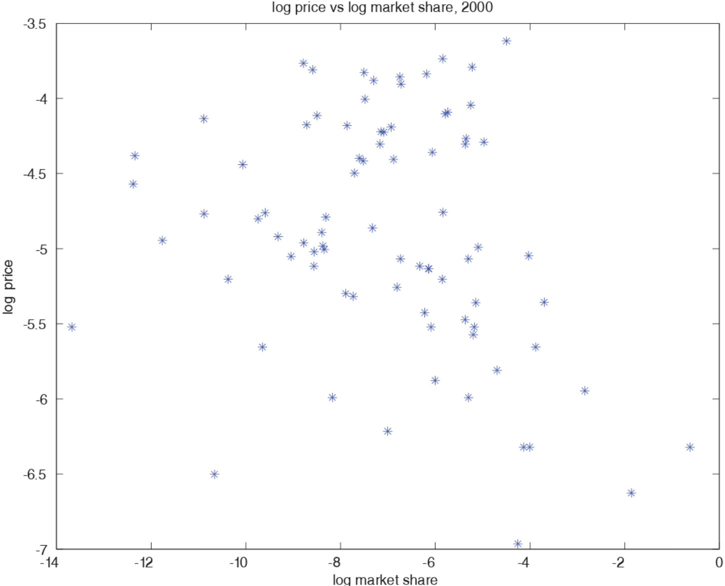
Application: S&P 500 mutual funds

- $u_g = -$price (fixed fee per \$10,000 invested)
- $\log c_i \sim N(\mu_0 + \mu_1 t, \delta_0 + \delta_1 t)$
- Search probabilities depend on fund age:

$$\rho_{jt} = \frac{A_{jt}^{\gamma}}{\sum A_{kt}^{\gamma}}$$

- Rational prior: $\alpha_{jt} = \rho_{jt} N_0$
  - Consumers' prior not identified from market share data alone

Application: S&P 500 mutual funds

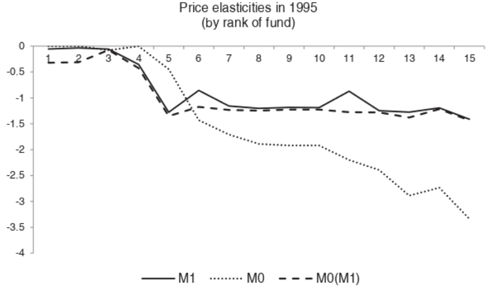


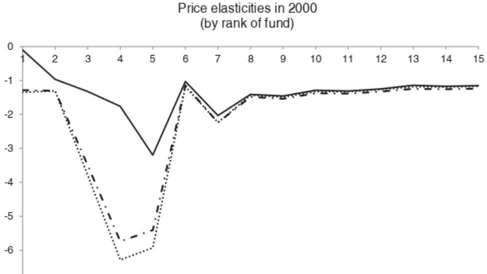
Application: S&P 500 mutual funds

Table 1. Estimates of the distribution of search costs from models of search with (M1) and without (M0) learning. Search costs are expressed in basis points, that is, dollars per 10,000 investment. Bootstrapped standard errors in the columns are labeled "SE"

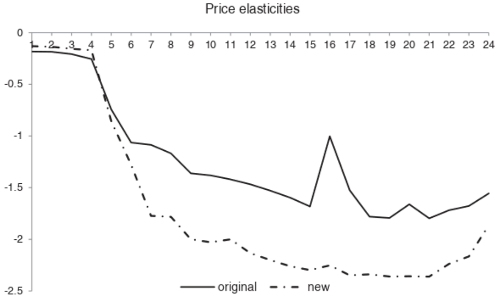
Parameter estimates	M0	SE	M1	SE
Elog(c)	2.74	0.55	2.25	0.38
Trend of Elog(c)	−0.34	0.28	−0.25	0.15
SDlog(c)	0.95	0.22	1.86	0.32
Trend of SDlog(c)	0.37	0.25	0.22	0.22
Age effect	2.31	0.70	2.35	0.89
R2	0.99		0.98	
Median search cost by year	**M0**	**M1**		
1995	15.44	9.50		
1996	10.96	7.43		
1997	7.79	5.81		
1998	5.53	4.54		
1999	3.93	3.55		
2000	2.79	2.77		
Mean search cost by year	**M0**	**M1**		
1995	24.20	53.85		
1996	26.10	64.98		
1997	32.27	82.31		
1998	45.70	109.42		
1999	74.15	152.69		
2000	137.87	223.65		
Interquartile range of search costs	**M0–25**	**M0–75**	**M1–25**	**M1–75**

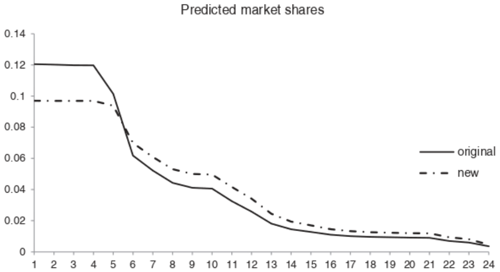
Application: S&P 500 mutual funds





Application: S&P 500 mutual funds





Baye, Michael R., John Morgan, and Patrick Scholten. 2006. "Information, Search, and Price Dispersion." Working Papers 2006-11, Indiana University, Kelley School of Business, Department of Business Economics and Public Policy. URL http://EconPapers.repec.org/RePEc:iuk:wpaper:2006-11.

Burdett, Kenneth and Kenneth L. Judd. 1983. "Equilibrium Price Dispersion." *Econometrica* 51 (4):pp. 955–969. URL http://www.jstor.org/stable/1912045.

Chen, Xiaohong, Han Hong, and Matthew Shum. 2007. "Nonparametric likelihood ratio model selection tests between parametric likelihood and moment condition models." *Journal of Econometrics* 141 (1):109 – 140. URL http://www.sciencedirect.com/science/article/pii/S0304407607000103. <ce:title>Semiparametric methods in econometrics</ce:title>.

De los Santos, Babur, Ali Hortacsu, and Matthijs R Wildenbeest. 2012. "Search with Learning." Tech. rep. URL http://www.bus.indiana.edu/riharbau/RePEc/iuk/wpaper/bepp2012-03-DelosSantos-Hortacsu-Wildenbeest.pdf.

De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest. 2012. "Testing models of consumer search using data on web browsing and purchasing behavior." *The American Economic Review* 102 (6):pp. 2955–80. URL http://www.aeaweb.org/articles.php?f=s&doi=10.1257/aer.102.6.2955.

Diamond, Peter A. 1971. "A model of price adjustment." *Journal of Economic Theory* 3 (2):156 – 168. URL http://www.sciencedirect.com/science/article/pii/0022053171900135.

Hong, Han and Matthew Shum. 2006. "Using Price Distributions to Estimate Search Costs." *The RAND Journal of Economics* 37 (2):pp. 257–275. URL http://www.jstor.org/stable/25046241.

Honka, Elisabeth. 2014. "Quantifying search and switching costs in the US auto insurance industry." *The RAND Journal of Economics* 45 (4):847–884. URL http://dx.doi.org/10.1111/1756-2171.12073.

Hortaçsu, Ali and Chad Syverson. 2004. "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds." *The Quarterly Journal of Economics* 119 (2):403–456. URL http://qje.oxfordjournals.org/content/119/2/403.abstract.

Koulayev, Sergei. 2013. “Search With Dirichlet Priors: Estimation and Implications for Consumer Demand.” *Journal of Business & Economic Statistics* 31 (2):226–239. URL http://dx.doi.org/10.1080/07350015.2013.764696.

Moraga-González, José Luis and Matthijs R. Wildenbeest. 2008. “Maximum likelihood estimation of search costs.” *European Economic Review* 52 (5):820 – 848. URL http://www.sciencedirect.com/science/article/pii/S001429210700102X.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest. 2012. “SEMI-NONPARAMETRIC ESTIMATION OF CONSUMER SEARCH COSTS.” *Journal of Applied Econometrics* :n/a–n/aURL http://dx.doi.org/10.1002/jae.2290.

Rothschild, Michael. 1973. "Models of Market Organization with Imperfect Information: A Survey." *Journal of Political Economy* 81 (6):pp. 1283–1308. URL http://www.jstor.org/stable/1830741.

———. 1974. "Searching for the Lowest Price When the Distribution of Prices Is Unknown." *Journal of Political Economy* 82 (4):689–711. URL http://www.jstor.org/stable/1837141.

Stigler, George J. 1961. "The Economics of Information." *Journal of Political Economy* 69 (3):pp. 213–225. URL http://www.jstor.org/stable/1829263.

Wildenbeest, Matthijs R. 2011. "An empirical model of search with vertically differentiated products." *The RAND Journal of Economics* 42 (4):729–757. URL http://dx.doi.org/10.1111/j.1756-2171.2011.00152.x.