Price dispersion and search

Paul Schrimpf

UBC
Economics 565

November 25, 2015
References

- Structural empirical papers:
1 Introduction

2 Models of price dispersion
   Fixed search
   Sequential search
   Information clearinghouse

3 Hong and Shum (2006)

4 Additional empirical work

5 Koulayev (2013)
Section 1

Introduction
• Homogenous product and competitive market:
  • Theory ⇒ one price
  • Reality ⇒ price dispersion

• Explanations:
  • Unobserved product heterogeneity
    • Likely part of explanation, but it is largely tautological
  • Imperfect information and search costs
    • Stigler (1961)
Price dispersion and search

Paul Schrimpf

Introduction

Models of price dispersion

Fixed search
Sequential search
Information clearinghouse

Hong and Shum (2006)

Additional empirical work

Koulayev (2013)

References

---

Table 1a: Measures of Price Dispersion Reported in the Literature in Offline Markets\(^1\)

<table>
<thead>
<tr>
<th>Study</th>
<th>Data Period</th>
<th>Product Market</th>
<th>Intervals of Estimated Price Dispersion Measures</th>
<th>Dispersion Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Compact Discs</td>
<td>17.6% - 20.0%</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Software</td>
<td>7.1%</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Borenstein and Rose (1994)</td>
<td>1986</td>
<td>U.S. Airline</td>
<td>0.018 - 0.416</td>
<td>Gini coefficient</td>
</tr>
<tr>
<td>Carlson and Pescatello (1980)</td>
<td>1976</td>
<td>Consumer Sundries</td>
<td>3.3% - 41.4%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Baking Powder, Sugar, Salt – 2001</td>
<td>0.0% - 13.4%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>Friborg, Ganslandt and Sandstrom (2001)</td>
<td>1999</td>
<td>$4.00 - $12.00</td>
<td>Standard Deviation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Books</td>
<td>21.94 - 76.20</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compact Discs</td>
<td>20.00 - 40.00</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compact Discs</td>
<td>12.91 - 23.88</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Books (Sweden)</td>
<td>19.00 - 56.00</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compact Discs (Sweden)</td>
<td>21.00 - 46.00</td>
<td>Range</td>
</tr>
<tr>
<td>Lach (2002)</td>
<td>1993 - 1996</td>
<td>Refrigerator (Israel)</td>
<td>4.9%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chicken, Flour, Coffee (Israel)</td>
<td>11.4% - 19.7%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>Marvel (1976)</td>
<td>1964 - 1971</td>
<td>Regular Gasoline</td>
<td>$0.048</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regular Gasoline</td>
<td>$0.015</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Premium Gasoline</td>
<td>$0.048</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Premium Gasoline</td>
<td>$0.017</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Pratt, Wise and Zeckhauser (1978)</td>
<td>1975</td>
<td>Various Products and Services</td>
<td>4.4% - 71.4%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Various Products and Services</td>
<td>11.0% - 567.0%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Various Products and Services</td>
<td>7.2% - 200.0%</td>
<td>Value of Information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fabrics</td>
<td>19.8% - 78.1%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coffee</td>
<td>14.3% - 25.1%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ready-Mixed Concrete</td>
<td>13.2% - 37.2%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Newsprint</td>
<td>4.5% - 8.2%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gasoline</td>
<td>6.2% - 11.8%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tinplate, Steel Cans</td>
<td>25.0% - 31.0%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pan Bread</td>
<td>28.0% - 49.6%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Corrugated Shipping Containers</td>
<td>21.8% - 39.6%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Sundries – 2000</td>
<td>1.6% - 42.0%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Sundries – 2000</td>
<td>5.7% - 28.4%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>Stigler (1961)</td>
<td>1953</td>
<td>Anthracite Coal</td>
<td>$3.46</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td>1959</td>
<td>Anthracite Coal</td>
<td>$1.15</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identical Automobiles</td>
<td>$165.00</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identical Automobiles</td>
<td>$42.00</td>
<td>Standard Deviation</td>
</tr>
</tbody>
</table>

\(^1\)Table 1a includes studies comparing offline and online price dispersion.
## Table 1b: Measures of Price Dispersion Reported in the Literature in Online Markets Only

<table>
<thead>
<tr>
<th>Study</th>
<th>Data Period</th>
<th>Product Market</th>
<th>Estimated Price Dispersion Measures</th>
<th>Dispersion Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Books (Italy)</td>
<td>€20.00 - €22.88</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compact Discs (Italy)</td>
<td>€2.29 - €2.79</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compact Discs (Italy)</td>
<td>€11.82 - €14.75</td>
<td>Range</td>
</tr>
<tr>
<td>Arbatskaya and Baye (Forthcoming)</td>
<td>1998</td>
<td>Mortgage Interest Rates</td>
<td>&gt; 0.25</td>
<td>Range</td>
</tr>
<tr>
<td>Arnold and Saliba (2002)</td>
<td>2001</td>
<td>Textbooks</td>
<td>10.7% - 52.6%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Textbooks</td>
<td>3.5% - 10.0%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Textbooks</td>
<td>0.2% - 12.5%</td>
<td>Price gap</td>
</tr>
<tr>
<td>Baye, Morgan and Scholten (2004a)</td>
<td>2000 - 2001</td>
<td>Consumer Electronics</td>
<td>9.1% - 9.7%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics</td>
<td>3.79% - 5.38%</td>
<td>Gap</td>
</tr>
<tr>
<td>Baye, Morgan and Scholten (2004b)</td>
<td>1999 - 2001</td>
<td>Consumer Electronics</td>
<td>57.4%</td>
<td>Range</td>
</tr>
<tr>
<td>Baylis and Perloff (2002)</td>
<td>1999</td>
<td>Camera</td>
<td>$342.00</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scanner</td>
<td>$106.00</td>
<td>Range</td>
</tr>
<tr>
<td>Brynjolfsson and Smith (2000)</td>
<td>1998-1999</td>
<td>Books</td>
<td>33.0%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compact Discs</td>
<td>25.0%</td>
<td>Range</td>
</tr>
<tr>
<td>Chevalier and Goolsbee (2003)</td>
<td>2001</td>
<td>Books</td>
<td>8.1% - 12.3%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Books</td>
<td>$7.62</td>
<td>Range</td>
</tr>
<tr>
<td>Clay, Krishnan, Wolff and Femandes (2003)</td>
<td>1999</td>
<td>Books</td>
<td>10.0% - 18.0%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>Ellison and Ellison (2004)</td>
<td>2002</td>
<td>Memory Modules</td>
<td>5.9% - 29.0%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (France)</td>
<td>3.0% - 15.3%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Italy)</td>
<td>4.3% - 14.2%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Netherlands)</td>
<td>5.6% - 20.4%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Spain)</td>
<td>2.2% - 13.3%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Sweden)</td>
<td>6.6% - 14.0%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (UK)</td>
<td>3.5% - 16.2%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Denmark)</td>
<td>6.3% - 20.2%</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (France)</td>
<td>7.8% - 47.4%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Italy)</td>
<td>9.3% - 27.8%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Netherlands)</td>
<td>8.9% - 54.6%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Spain)</td>
<td>3.8% - 32.4%</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consumer Electronics (Sweden)</td>
<td>16.4% - 50.4%</td>
<td>Range</td>
</tr>
</tbody>
</table>
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, $\epsilon$-equilibrium, mistaken beliefs about price distribution
Fixed search

• Stigler (1961)
• Assumptions:
  1. Distribution of prices on \([p, \bar{p}]\), 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:
  \[
  \left( E[p^{(1:n^*-1)}] - E[p^{(1:n^*)}] \right) K \geq c \geq \left( E[p^{(1:n^*)}] - E[p^{(1:n^*+1)}] \right) K
  \]
  \( n^* \) 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
    \[
    E_G[p^{(1:n)}] < E_F[p^{(1:n)}] \text{ for } n > 1
    \]
  - Expected total costs are lower with greater price dispersion
    - If \( G \) is a mean preserving spread of \( F \), then
      \[
      E_G[p^{(1:n^*_G)}]K - cn^*_G < E_F[p^{(1:n^*_F)}]K - cn^*_F \text{ for } n > 1
      \]
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
Price dispersion and search
Paul Schrimpf

Introduction
Models of price dispersion
Fixed search
Sequential search
Information clearinghouse
Hong and Shum (2006)
Additional empirical work
Koulayev (2013)
References

Fixed search - endogenous price dispersion

- 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, \( P(n = i) \) for \( i = 1, 2, \ldots \)
Fixed search - endogenous price dispersion

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

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

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

- 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 $\theta$ (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 $= reservation\ price\ p^* = \min\{\bar{p}, z^*\}$
  where

  \[
  c = \int_{\bar{p}}^{z^*} (z^* - p)f(p)dp = \int_{\bar{p}}^{z^*} F(p)dp
  \]

- 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
Price dispersion and search
Paul Schrimpf

Introduction
Models of price dispersion
Fixed search
Sequential search
Information clearinghouse

Hong and Shum (2006)
Additional empirical work
Koulayev (2013)
References

FIGURE 1
RAW HISTOGRAMS OF ONLINE PRICES

Stokey-Lucas: Recursive Methods...

Lazear: Personnel Economics

Billingsley: Probability and Measure

Duffie: Dynamic Asset Pricing Theory

x-axis: prices (in dollars)
Fixed search

• Firm marginal cost $r$, continuum of firms with equilibrium price distribution $F_p$
• Consumer search cost $c_i \sim F_c$
• Number of searches
  \[
  \left( E[p^{(1:n(c_i)-1)}] - E[p^{(1:n(c_i))}] \right) K \geq c_i \geq \left( E[p^{(1:n(c_i))}] - E[p^{(1:n(c_i)+1)}] \right)
  \]
• Define $\Delta_n = E[p^{(1:n-1)}] - E[p^{(1:n)}]$; $F_P$ observed, so $\Delta_n$ identified
• Let $\tilde{q}_n = F_c(\Delta_{n-1}) - F_c(\Delta_n) = \text{portion of consumers who obtain } n \text{ prices}$
  • $\tilde{q}_n$ 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
  \[
  (\overline{p} - r)\tilde{q}_1 = (p - r) \left[ \sum_{k=1}^{K} \tilde{q}_k k (1 - F_p(p))^{k-1} \right]
  \]
  
- Observed prices $p_j, j = 1, \ldots, n_f$
  \[
  (\overline{p} - r)\tilde{q}_1 = (p - r) \left[ \sum_{k=1}^{K} \tilde{q}_k k (1 - \hat{F}_p(p_j))^{k-1} \right]
  \]
  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
  - $\tilde{q}_1 = 0.633$, $\tilde{q}_2 = 0.309$, $\tilde{q}_3 = 0.058$
Sequential search I

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

  Let $G(p) = \text{CDF of } p_i^*$
- Firm indifference:
  
  $$(\bar{p} - r) (1 - G(\bar{p})) = (p - r) (1 - G(p))$$

- 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

<table>
<thead>
<tr>
<th>Product</th>
<th>$n$</th>
<th>List</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Median</th>
<th>$\bar{p}$</th>
<th>$\bar{p}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stokey-Lucas</td>
<td>19</td>
<td>60.50</td>
<td>66.60</td>
<td>5.64</td>
<td>64.98</td>
<td>59.75</td>
<td>86.80</td>
</tr>
<tr>
<td>Lazear</td>
<td>17</td>
<td>31.95</td>
<td>34.73</td>
<td>2.48</td>
<td>35.27</td>
<td>29.51</td>
<td>37.70</td>
</tr>
<tr>
<td>Billingsley</td>
<td>20</td>
<td>99.95</td>
<td>95.48</td>
<td>5.87</td>
<td>98.90</td>
<td>83.58</td>
<td>100.87</td>
</tr>
<tr>
<td>Duffie</td>
<td>15</td>
<td>65.00</td>
<td>62.71</td>
<td>4.91</td>
<td>63.48</td>
<td>50.58</td>
<td>69.95</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Product</th>
<th>$K^a$</th>
<th>$M^b$</th>
<th>$\tilde{q}_1^c$</th>
<th>$\tilde{q}_2$</th>
<th>$\tilde{q}_3$</th>
<th>Selling</th>
<th>MEL Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stokey-Lucas</td>
<td>3</td>
<td>5</td>
<td>.480 (.170)</td>
<td>.288 (.433)</td>
<td></td>
<td>49.52 (12.45)</td>
<td>102.62</td>
</tr>
<tr>
<td>Lazear</td>
<td>4</td>
<td>5</td>
<td>.364 (.926)</td>
<td>.351 (.660)</td>
<td>.135 (.692)</td>
<td>27.76 (8.50)</td>
<td>84.70</td>
</tr>
<tr>
<td>Billingsley</td>
<td>3</td>
<td>5</td>
<td>.633 (.944)</td>
<td>.309 (.310)</td>
<td></td>
<td>69.73 (68.12)</td>
<td>199.70</td>
</tr>
<tr>
<td>Duffie</td>
<td>3</td>
<td>5</td>
<td>.627 (1.248)</td>
<td>.314 (.195)</td>
<td></td>
<td>35.48 (96.30)</td>
<td>109.13</td>
</tr>
</tbody>
</table>

**Parameter estimates and standard errors: nonsequential-search model**

<table>
<thead>
<tr>
<th></th>
<th>$\Delta_1$</th>
<th>$F_c(\Delta_1)$</th>
<th>$\Delta_2$</th>
<th>$F_c(\Delta_2)$</th>
<th>$\Delta_3$</th>
<th>$F_c(\Delta_3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stokey-Lucas</td>
<td>2.32</td>
<td>.520</td>
<td>.68</td>
<td>.232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lazear</td>
<td>1.31</td>
<td>.636</td>
<td>.83</td>
<td>.285</td>
<td>.57</td>
<td>.150</td>
</tr>
<tr>
<td>Billingsley</td>
<td>2.90</td>
<td>.367</td>
<td>2.00</td>
<td>.058</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duffie</td>
<td>2.41</td>
<td>.373</td>
<td>1.42</td>
<td>.059</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

* $a$ 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 $\Delta_k$ computed as $E(P_{1,k}) - E(P_{1,k+1})$ (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

<table>
<thead>
<tr>
<th>Product</th>
<th>$\delta_1$</th>
<th>$\delta_2$</th>
<th>Median$^a$</th>
<th>Selling Cost</th>
<th>$\alpha^b$</th>
<th>$F_{c}^{-1}(1 - \alpha; \Theta)$</th>
<th>Log-L Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stokey-Lucas</td>
<td>.46 (.02)</td>
<td>1.55 (.03)</td>
<td>29.40 (1.45)</td>
<td>22.90 (1.31)</td>
<td>.58</td>
<td>19.19</td>
<td>31.13</td>
</tr>
<tr>
<td>Lazear</td>
<td>.40 (.01)</td>
<td>1.15 (.01)</td>
<td>16.37 (1.00)</td>
<td>11.31 (.79)</td>
<td>.69</td>
<td>4.56</td>
<td>34.35</td>
</tr>
<tr>
<td>Billingsley</td>
<td>.25 (.01)</td>
<td>2.01 (.04)</td>
<td>9.22 (.94)</td>
<td>65.37 (.83)</td>
<td>.51</td>
<td>8.43</td>
<td>23.73</td>
</tr>
<tr>
<td>Duffie</td>
<td>.21 (.02)</td>
<td>4.57 (.29)</td>
<td>10.57 (2.01)</td>
<td>28.24 (1.63)</td>
<td>.54</td>
<td>7.00</td>
<td>18.93</td>
</tr>
</tbody>
</table>

Note: Including shipping and handling charges. Standard errors in parentheses. $\delta_1$ and $\delta_2$ 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 $\bar{p}$, implied by estimate of $r$ (see equation (11)).
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 $z(c_i) = \bar{p}$, $4.56 - 19.19$
- 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

<table>
<thead>
<tr>
<th>Product</th>
<th>$r$</th>
<th>Search-Cost Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$ fixed at sequential-model estimates$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stokey-Lucas</td>
<td>22.90</td>
<td>$F_c^{-1}(0.42) = 20.20$</td>
</tr>
<tr>
<td>Lazear</td>
<td>11.31</td>
<td>$F_c^{-1}(0.31) = 2.97$</td>
</tr>
<tr>
<td>Billingsley</td>
<td>65.37</td>
<td>$F_c^{-1}(0.49) = 5.39$</td>
</tr>
<tr>
<td>Duffie</td>
<td>28.24</td>
<td>$F_c^{-1}(0.46) = 7.24$</td>
</tr>
</tbody>
</table>

$r$ fixed at nonsequential-model estimates

<table>
<thead>
<tr>
<th>Product</th>
<th>$r$</th>
<th>Search-Cost Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stokey-Lucas</td>
<td>49.52</td>
<td>$F_c^{-1}(0.5) = 4.52^d$</td>
</tr>
<tr>
<td>Lazear</td>
<td>27.76</td>
<td>$F_c^{-1}(0.5) = 0.26$</td>
</tr>
<tr>
<td>Billingsley</td>
<td>69.73</td>
<td>$F_c^{-1}(0.5) = 3.21$</td>
</tr>
<tr>
<td>Duffie</td>
<td>35.48</td>
<td>$F_c^{-1}(0.5) = 3.51$</td>
</tr>
</tbody>
</table>

$^a$ 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)

  - 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 $\alpha_1, \ldots, \alpha_G$,

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

which implies

$$E[\tilde{p}_j] = \frac{\alpha_j}{\sum \alpha_g}$$
Bayesian updating: after seeing $\tilde{u}_g, n_g$ times,

$$f(p|n) \propto f(n|p)f(p)$$

$$\propto \frac{(\sum n_g)!}{\prod n_g!} \prod p_g^{n_g} \frac{\Gamma(\sum \alpha_g)}{\prod \gamma(\alpha_g)} \prod \tilde{p}_g^{\alpha_g-1}$$

$$\propto \frac{\Gamma(\sum \alpha_g + n_g)}{\prod \gamma(\alpha_g + n_g)} \prod \tilde{p}_g^{\alpha_g + n_g - 1}$$

So

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

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

$$\mathbb{E}[\max\{\tilde{u}, u_r^*\}|n] - u_r^* > c$$

$$\sum_{\tilde{u}_g > u_r^*} (\tilde{u}_g - u_r^*) \mathbb{E}[\tilde{p}_g|n] > c$$
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 $\bar{k}_r = \text{longest a consumer with best draw } u_r \text{ will continue searching}$

$$
\bar{k}_r = \max \left\{ 1, \left[ \frac{1}{c} \sum_{\tilde{u}_g > u_r^*} (\tilde{u}_g - u_r^*) \alpha_g - \sum_g \alpha_g \right] \right\}
$$

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

Paul Schrimpf

Introduction

Models of price dispersion

Fixed search
Sequential search
Information clearinghouse

Hong and Shum (2006)

Additional empirical work

Koulayev (2013)

References

Example
Price dispersion and search
Paul Schrimpf

Introduction
Models of price dispersion
- Fixed search
- Sequential search
- Information clearinghouse

Hong and Shum (2006)
Additional empirical work
Koulayev (2013)

References

Example

Search without learning

- probability of purchase
- 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 = -\text{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:
  $$\rho_{jt} = \frac{A_{jt}^\gamma}{\sum_k A_{kt}^\gamma}$$
- Rational prior: $\alpha_{jt} = \rho_{jt} N_0$
  - Consumers’ prior not identified from market share data alone
Price dispersion and search

Paul Schrimpf

Introduction

Models of price dispersion

Fixed search
Sequential search
Information clearinghouse

Hong and Shum (2006)

Additional empirical work

Koulayev (2013)

References

Application: S&P 500 mutual funds
Introduction

Models of price dispersion

Fixed search
Sequential search
Information clearinghouse

Hong and Shum (2006)

Additional empirical work

Koulayev (2013)

References

Price dispersion and search

Paul Schrimpf

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.

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>M0</th>
<th>SE</th>
<th>M1</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elog(c)</td>
<td>2.74</td>
<td>0.55</td>
<td>2.25</td>
<td>0.38</td>
</tr>
<tr>
<td>Trend of Elog(c)</td>
<td>-0.34</td>
<td>0.28</td>
<td>-0.25</td>
<td>0.15</td>
</tr>
<tr>
<td>SDlog(c)</td>
<td>0.95</td>
<td>0.22</td>
<td>1.86</td>
<td>0.32</td>
</tr>
<tr>
<td>Trend of SDlog(c)</td>
<td>0.37</td>
<td>0.25</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Age effect</td>
<td>2.31</td>
<td>0.70</td>
<td>2.35</td>
<td>0.89</td>
</tr>
<tr>
<td>R²</td>
<td>0.99</td>
<td></td>
<td>0.98</td>
<td></td>
</tr>
</tbody>
</table>

Median search cost by year

<table>
<thead>
<tr>
<th>Year</th>
<th>M0</th>
<th>M1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>15.44</td>
<td>9.50</td>
</tr>
<tr>
<td>1996</td>
<td>10.96</td>
<td>7.43</td>
</tr>
<tr>
<td>1997</td>
<td>7.79</td>
<td>5.81</td>
</tr>
<tr>
<td>1998</td>
<td>5.53</td>
<td>4.54</td>
</tr>
<tr>
<td>1999</td>
<td>3.93</td>
<td>3.55</td>
</tr>
<tr>
<td>2000</td>
<td>2.79</td>
<td>2.77</td>
</tr>
</tbody>
</table>

Mean search cost by year

<table>
<thead>
<tr>
<th>Year</th>
<th>M0</th>
<th>M1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>24.20</td>
<td>53.85</td>
</tr>
<tr>
<td>1996</td>
<td>26.10</td>
<td>64.98</td>
</tr>
<tr>
<td>1997</td>
<td>32.27</td>
<td>82.31</td>
</tr>
<tr>
<td>1998</td>
<td>45.70</td>
<td>109.42</td>
</tr>
<tr>
<td>1999</td>
<td>74.15</td>
<td>152.69</td>
</tr>
<tr>
<td>2000</td>
<td>137.87</td>
<td>223.65</td>
</tr>
</tbody>
</table>

Interquartile range of search costs

<table>
<thead>
<tr>
<th></th>
<th>M0–25</th>
<th>M0–75</th>
<th>M1–25</th>
<th>M1–75</th>
</tr>
</thead>
</table>

Bootstrapped standard errors in the columns are labeled “SE”
Application: S&P 500 mutual funds

Price elasticities in 1995
(by rank of fund)

Price elasticities in 2000
(by rank of fund)

References
Price dispersion and search

Paul Schrimpf

Introduction

Models of price dispersion
  Fixed search
  Sequential search
  Information clearinghouse

Hong and Shum (2006)

Additional empirical work

Koulayev (2013)

References

Application: S&P 500 mutual funds

Price elasticities

Predicted market shares


Introduction

Models of price dispersion

Fixed search
Sequential search
Information clearinghouse

Additional empirical work

Koulayev (2013)

References


