1 Consumer Information and Search: Topics and Main Findings of Key Papers

Sorensen (2000)

- Establishes the empirical importance of price dispersion due to costly consumer search by examining retail prices for prescription drugs.

- Analyses two geographically distinct markets and finds considerable price variation within both (i.e. among pharmacies within the same market) even after controlling for pharmacy fixed effects.

- Pharmacy heterogeneity accounts for at most one-third of the observed price dispersion.

- **Observed price distributions are consistent with the predictions of models based on consumer search**. The empirical approach hinges on the observation that incentives to price-shop are strongest for prescriptions that must be purchased frequently, such as medications used to treat chronic conditions.

- Prices for repeatedly purchased prescriptions (i.e. items warranting search) exhibit significant reductions in both dispersion and price-cost margins.

- Other findings: Pharmacies cannot be easily sorted into "low-price" and "high-price" categories (suggestive of search, not pharmacy characteristics); frequency of purchase reduces price dispersion, after factoring out pharmacy fixed effects; also finds lower margins on high-frequency purchases (approx. 37% lower); rules out alternative hypotheses that pharmacy heterogeneity (e.g. better location) and cost heterogeneity (multi-source versus single-source drugs) drive results rather than search intensity.

De Los Santos, Hortaçsu & Wildenbeest (2012)

- Utilize novel data on web browsing and purchasing behaviour for a large panel of consumers to test classical models of consumer search – Stigler's (1961) fixed sample size search model and the McCall (1970) – Mortensen (1970) model of sequential search.

- Key difference between the theories: under **fixed sample size search** consumers pick the number of retailers they'll visit in advance, whereas under **sequential search** the number of retailers considered depends on the outcome of the search. Hence under fixed sample size search the outcome of the previous draw should not influence the decision to keep searching; it should with sequential search.

- Substantial descriptive evidence against sequential search model – **consumers tend to buy from retailers visited previously**, despite not having visited all stores they're aware of; little evidence that high early price draws correlated with further search. Find similar results looking at search *within* retailers (e.g. multiple sellers listed on Amazon).

- Structural fixed sample size search model that allows for product differentiation and asymmetric sampling (i.e. some stores more likely to be drawn than others). In baseline model, prices are the only source of uncertainty for consumers, and they are discovered upon selection of the bundle. Alternative model wherein consumers remain uncertain about retailer-specific utility until visiting (due e.g. to differences in shipping times).

- Larger own-price elasticities than in standard DC differentiated product models that incorrectly assume full information – **traditional DC models mistake unawareness for low price elasticity**.

Sovinsky-Goeree (2008)

- **Advertising** affects options appearing in consumer choice sets.

- Finds markup of 19% over production costs, with top firms advertising more than average and earning higher than average markups.

These notes are based on Howard Smith's lectures from the 2018-2019 academic year.
Of course, this is our interpretation of the material Howard presented. Good content is his; mistakes are ours.

- Markups mainly due to information asymmetry – full information model predicts markups one-quarter of those under asymmetric information.

- **Traditional models over-predict price elasticity of demand and under-predict industry markups**.

2   De Los Santos, Hortaçsu & Wildenbeest (2012) – Fixed Sample Size Search

Model

- Assume consumers have indirect utility

$$u_{ij} = \delta_{ij} + \alpha_i p_j \tag{2.1}$$

where δ_{ij} is the consumer's gross utility from each store, p_j is store j's price for the book, and α_i is a consumer-specific price coefficient. The gross utility is given by

$$\delta_{ij} = \mu_j + X_i\beta_j + \epsilon_{ij} \tag{2.2}$$

where we allow this utility to depend on a store fixed effect μ_j, consumer characteristics X_i and an idiosyncratic utility draw ϵ_{ij}. Consumers know their gross utility δ_{ij}, but are uncertain about prices, which they learn when they visit the store.

- Implications of fixed sample search strategy: Consumers decide which subset of the J online stores to visit and then make a purchase decision among the visited stores. Consumers sample stores at a cost $c_i = c + X_i\beta$ – i.e. the cost of sampling can depend on consumer characteristics X_i. The net expected benefit to consumer i for visiting all online stores in a subset S, denoted m_{iS}, is

$$m_{iS} = \mathbb{E}_p\left[\max_{j \in S}\{u_{ij}\}\right] - k \cdot c_i \tag{2.3}$$

where k is the number of stores in subset S. To smooth the choice set probabilities the authors add a mean-zero stochastic noise term ς_{iS} to m_{iS} (they interpret this as reflecting errors in an individual's assessment of the net expected gain of visiting all stores in subset S or an idiosyncratic total cost of sampling subset S), assumed to be i.i.d. EVT-1 with scaling parameter σ_ς. The probability that consumer i finds it optimal to sample the set of stores S from the store set space $\mathscr{S}$ is then

$$P_{iS} = \frac{\exp[m_{iS}/\sigma_\varsigma]}{\sum_{S' \in \mathscr{S}} \exp[m_{iS'}/\sigma_\varsigma]} \tag{2.4}$$

- In the second stage the uncertainty about prices for the selected stores is resolved and consumer i purchases from the store j that provides the highest utility in her sample S_i: $\arg\max_{j \in S_i} u_{ij}$. This happens with probability

$$P_{ij|S_i} = \Pr(u_{ij} > u_{ik} \; \forall \; k \neq j \in S_i). \tag{2.5}$$

- Finally, the probability of observing a consumer i selecting a choice set S and buying product j is

$$P_{ijS_i} = P_{iS_i} P_{ij|S_i} \tag{2.6}$$

Estimation

- To obtain a closed form expression for $\mathbb{E}_p[\max_{j \in S}\{u_{ij}\}]$ the authors assume prices are (you guessed it) distributed type I extreme value, with known store-specific location parameter γ_j and common scale parameter σ, implying by the log-sum result that

$$\mathbb{E}_p\left[\max_{j \in S}\{u_{ij}\}\right] = \alpha_i\sigma \log\left(\sum_{j \in S}\exp\left[\frac{\delta_{ij} + \alpha_i\gamma_j}{\alpha_i\sigma}\right]\right) \tag{2.7}$$

- This is then plugged in (2.3); (2.4) follows. Estimation proceeds via Maximum Simulated Likelihood:

$$\hat{\theta} \in \arg\max_\theta \ln\mathscr{L}(\theta) \equiv \sum_i \log \check{P}_{ijS}(\theta) \equiv \sum_i \log \check{P}_{iS}(\theta)\check{P}_{ij|S}(\theta) \tag{2.8}$$

3 Sovinsky-Goeree (2008) – Informative Advertising

Model

- Extends BLP by allowing nonrandom differences in information based on consumer observables.

- An individual chooses from J products, indexed $j = 1, \ldots, J$, where a product is a PC model defined as a (brand $\times$ CPU type $\times$ CPU speed $\times$ form) combination. Product j characteristics are price p, non-price observed attributes x (CPU speed, Pentium CPU, firm, laptop form factor, etc.), and attributes unobserved to the researcher but known to consumers and producers ξ. The indirect utility consumer i obtains from j at time t is

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt} \tag{3.1}$$

  where $\delta_{jt} = x'_j \beta + \xi_{jt}$ captures the base utility every consumer derives from j and mean preferences for x_j are captured by β. The composite random shock $\mu_{ijt} + \epsilon_{ijt}$ captures heterogeneity in consumers' tastes for product attributes; ϵ_{ijt} is a mean zero i.i.d. EVT-1 stochastic term.

- The term μ_{ijt} includes interactions between observed consumer attributes D_{it}, unobserved (to the econometrician) consumer tastes v_i, and x_j. Specifically, defining income by y_{it},

$$\mu_{ijt} = \alpha \ln(y_{it} - p_{jt}) + x'_j(\Omega D_{it} + \Sigma v_i) \quad v_i \sim N(0, I_k) \tag{3.2}$$

  where the Ω matrix measures how tastes vary with x_j.

- The consumers' outside option is non-purchase, purchase of a used PC, or purchase of a new PC from a firm not in the 15 included firms. Normalizing p_{0t} to zero, the indirect utility from the outside option is

$$u_{i0t} = \alpha \ln(y_{it}) + \xi_{0t} + \epsilon_{i0t} \quad \text{with } \xi_{0t} \text{ normalised to zero} \tag{3.3}$$

- Information Technology: Since there is much innovation in this industry, full information is a strong assumption. Assuming consumers are aware of the outside option with probability 1, the (conditional) probability that consumer i purchases j is

$$s_{ijt} = \sum_{S \in C_j} \prod_{l \in S} \phi_{ilt} \prod_{k \notin S} (1 - \phi_{ikt}) \frac{\exp\{\delta_{jt} + \mu_{ijt}\}}{y^a_{it} + \sum_{r \in S} \exp\{\delta_{rt} + \mu_{irt}\}} \tag{3.4}$$

  where C_j is the set of all choice sets that include product j. The ϕ_{ijt} term is the probability i is informed about j. The y^a_{it} term is from the presence of the outside good. The outside sum is over all the choice sets that include product j. Reading this equation from left to right, we're summing over all the choice sets containing j, multiplying by the probability consumer i is aware of the items in choice set S and unaware of all other items (this gives the probability that S is consumer i's choice set) and the final term is the conditional probability of choosing j given S is the consumer's choice set.

- This is a computationally demanding problem (2^{2111} choice probabilities given that correspond to each choice set for each individual and product); as a solution, the **author simulates the choice set** facing i, thereby making only one purchase probability computation per individual necessary – the one corresponding to i's simulated choice set. The data used to form the choice sets are those used to construct the ϕ_{ijt} "information technology" term ϕ_{ijt}.

- The "information technology" term ϕ_{ijt} describes the effectiveness of advertising at informing consumers about products. Suppressing time notation, it is given by

$$\phi_{ij}(\theta_\phi) = \frac{\exp(\gamma_j + \lambda_{ij})}{1 + \exp(\gamma_j + \lambda_{ij})} \tag{3.5}$$

  where

$$\gamma_j = a'_j(\varphi + \rho a_j + i_m \Psi_f) + \vartheta x^{age}_j \tag{3.6}$$

  where the vectors φ and ρ, measure the effectiveness of advertising media at informing consumers. The author also includes fixed effects for every product quarter (the Φ), but does not estimate a fixed effect

for each medium, so i_m is a column vector of 1's. Consumers may be more likely to know a product the longer it has been on the market: this is captured by ϑ, where x_j^{age} is the PC age measured in quarters. The author does not have data on individual ad exposure, but allows/controls for variation in the household ad exposure via λ_{ij}, where

$$\lambda_{ij} = a_j'(\Upsilon D_i^s \zeta + \kappa_i) + \widetilde{D}_i'\widetilde{\lambda} \quad \ln \kappa_i \sim N(0, I_m) \tag{3.7}$$

The Υ matrix captures how advertising media's effectiveness varies by observed consumer characteristics, D_i^s is a set of consumer demographic characteristics; $\Upsilon_m D_i^s$ is the exposure of individual i to medium m, and $a'j\Upsilon D_i^s$ is the exposure of i to ads for product j. The parameter ζ measures the effect of this ad exposure on the information set, and the vector κ_i denotes unobserved (to the econometrician) consumer heterogeneity with regard to ad medium effectiveness (assumed independent of all other variables). In the absence of advertising, consumers may still be differentially informed ($\phi(a = 0) > 0$). The $\widetilde{D}$ (a subset of D) proxy for the opportunity costs of acquiring information, and the magnitude of ψ_{ij} when no advertising occurs depends on $\widetilde{D}_i\widetilde{\lambda} + \vartheta x_j^{age}$.

- Note: In the slides, the model is substantially simplified compared with what's been detailed here. There's also the supply side of the model, which I won't discuss in any technical detail since it wasn't prioritized in lectures: firms are allowed to advertise in groups, FOCs of firm profit are used to estimate marginal costs (used to calculate markups), and the equilibrium concept is non-cooperative Bertrand-Nash competition (as discussed in lecture 2).

Identification and estimation

- Identification essentially follows BLP; need a bunch of instruments that are orthogonal to product-specific unobserved utility (ξ_{jt}), the marginal product cost residual from the pricing first order condition, and the marginal advertising cost residual obtained from an advertising first order condition. Uses five sets of moments [not that important for the exam, but the additional sets are from the purchase decisions, which matches model probabilities that consumers buy from specific firms (not products) to empirical counterparts, and the media exposure decisions, which match the model's predictions from exposure to media m (conditional on observed characteristics) to the observed exposure].

- Estimation also follows BLP: GMM and simulation are used.

Findings

- Full information models of the PC industry would indicate that the industry is quite competitive with modest median markups (of e.g. 5%).

- Under asymmetric information with informative advertising, the industry looks very different: cross-price elasticities indicate products are not as substitutable as full information suggests; estimated median markups of 19% over production costs in 1998, where top firms engage in higher than average advertising and earn higher than average markups.

- The results suggest that

  - Allowing for heterogeneity in consumers' choice sets yields more realistic estimates of substitution patterns between goods;

  - Assuming full information may result in incorrect conclusions regarding the intensity of industry competition, with firms benefiting from limited consumer information (of which there is evidence);

  - Exposure to advertising significantly affects the composition of consumers' choice sets, but the effects can differ substantially across individuals, media and firms.

  - Also some interesting evidence showing that advertising one product can negatively affect sales of other products by the same firm, though not as much as it negatively affects the sales of other firms.

  - Finally, the author finds evidence of economies of scope in group advertising, with some firms finding group advertising profitable.