Search Costs, Demand Structure and Long Tail in Electronic Markets: Theory and Evidence

Anindya Ghose
Stern School of Business
New York University

Bin Gu
McCombs School of Business
University of Texas at Austin

1. Introduction

An important advantage that internet channels posit over physical markets is a reduction in search costs for prices. The reduction in search cost has rekindled interests in examining the Law of One Price (Brynjolfsson and Smith 2000). Empirical evidence shows that firms have been able to maintain their pricing power and there is significant price dispersion online (Brynjolfsson and Smith 2000, Baye et al. 2005). These price dispersions are often attributed to the presence of search costs in the online world. In a similar vein, theoretical research has also shown that search costs create imperfect information about sellers' prices among consumers, which leads to equilibrium price dispersion in otherwise homogeneous product markets (Reinganum 1979).

However, despite its importance, little empirical evidence exists so far on how search costs actually affect demand structures faced by retailers. To empirically investigate whether search costs are present in online retail markets, our research uses prior theoretical insights that relate the nature of consumer demand structure to the presence of search costs. The presence of search costs also implies that that it takes time for information about price changes to dissipate among consumers in the market. This leads to lower short-term price elasticity and higher long-term price elasticity, especially for information on price decreases. A number of empirical studies have tried to quantify search costs both in online and offline settings. These include Sorensen (2000), Bajari and Hortacsu (2003), Brynjolfsson, Dick and Smith (2004), Hong and Shum (2006).The empirical work on search costs has shown that online search costs, although smaller than search costs in physical markets, are still significant.

The objective of our study is therefore to consider how such search costs affect market competition, how they influence consumer demand structure faced by online retailers and how fast price change information dissipates in the online market. An additional contribution of this study is that we do so with only aggregate level data by exploring a unique insight of Stiglitz (1989) that search costs create kinkedness in aggregate demand when firms change prices. The significance of the kink reflects the magnitude of online search costs and the kinked demand function provides information on how search costs affect competition in the online market. One advantage of using aggregate level data is that we can estimate the impact of search cost for an entire market. This approach compliments prior studies that mainly focus on estimating search costs for individual consumers. Further, our approach allows us to estimate the change in consumer surplus due to the presence of search costs.

2. Theory

A well-known feature of search costs in a competitive marketplace is that it creates a kinked demand function (Stiglitz 1989). He suggested that when the underlying consumer demand function is kinked, the demand elasticity for price decreases is different from the demand elasticity for price increases. The significance of the kinkiness is determined by the magnitude of the search cost faced by consumers. When search cost is high, consumers are only aware of the price of the retailer they visit, but are unaware of the prices for the same product in the retailers they do not visit. So when a retailer increases its price, its own immediate customers (who know about the increase) are induced to search for lower prices amongst rival retailers and the retailer loses customers accordingly. But when a retailer decreases its price, then unless it expends resources on advertising, its action induces no new customers to launch a search. Hence, while it will able to keep its existing customers, it does not gain a proportionate number of new customers. This leads to lower price elasticity for price decreases. Under such circumstances, the gains to lowering prices may be markedly lower than the losses from raising prices. On the other hand, when search cost is low, a reduction in product price by a retailer has the potential to attract customers from its competitors, but a price increase only affects the firm's current customers. This leads to higher price elasticity for price decreases.[1] Thus, by studying the kinkedness of consumer demand function, we can infer the level of search costs in online markets.

[1] In such a situation a firm enjoys a discontinuous jump in demand when it succeeds in charging the lowest price because it instantly attracts the price-sensitive "shopper" segment of the market.

3. Data

We estimate our models using a panel data set compiled from publicly available information about product prices and sales rankings, gathered using automated Java scripts. These scripts access and parse HTML and XML pages downloaded from Amazon.com and Barnes and Noble.com between September 2005 and April 2005. The panel includes over 3210 books for this study across all major book categories. These products include a mix of best sellers, new releases, random selected titles and less popular books selected from the different genres such as fiction, non-fiction, business, text books, computer books and so on. Each observation from Amazon and Barnes and Noble contains the product's list price, its retail price, its sales rank (which serves as a proxy for units of demand, as described later), the date the product was released into the market, the average customer rating for the product, and the number of reviews based on which the average rating was computed. We also collected secondary market data including the number of used copies available for sale and the minimum price of the used good. Thus, our data also takes into account some of the competitive effects on prices. In our data we find that on an average, Amazon changes a book's price every 151 days and BN change prices every 49 days.

4 Empirical Methodology and Analysis

Until recently it was difficult to calculate the price elasticity for products sold on the Internet because, while the price of individual items could be readily observed, the quantity sold was generally unobservable. However, an emerging stream of work (Chevalier and Goolsbee 2003, Brynjolfsson, Hu and Smith 2003) has addressed this problem by providing a way to map the observable Amazon.com sales rank to the corresponding number of books sold by using the following Pareto relationship: $Quantity = \delta \cdot Rank^{\theta}$.

Stiglitz (1989) suggested that search costs lead to a kinked demand function. When search costs are high, price decrease information disseminates slowly among potential customers, but price increases are immediately observed by current customers. As a result, a firm in a high search costs environment faces lower price elasticity for price decreases than for price increases. This often happens for relatively smaller and lesser known firms whose price information is not well followed by the market. On the other hand, when search costs are low, price decrease information disseminates quickly and attracts not only a firm's own regular customers but its competitors' customer as well. This results in higher price elasticity for price decreases than for price increases. This phenomenon is often observed for market leaders whose price information is well followed by customers and competitors.

We empirically investigate this theory by allowing demand elasticity for price decreases to vary from that for price increases. To do so, we construct a dummy variable *Price Decrease* which takes the value of 1 if the most recent action on product *i* is a price decrease. The presence of search costs also suggests that it takes time for information on price decreases to spread. To quantify this information diffusion process, we consider how demand elasticity evolves after a price change. We allow demand elasticity to vary from week to week for up to 4 weeks after the initial price decrease. We also include book specific fixed effects to control for difference in book popularity. In keeping with prior work (Chevalier and Mayzlin 2006), we use a log-log specification for the demand function and estimate the relative demand as a function of prices charged by retailers. The final empirical model therefore takes the following form:

$$\log(SalesRank_{AMAZON_{it}}) - \log(SalesRank_{BN_{it}}) = \alpha_i + \beta_1 \log\left(P_{A_{it}}\right) + \sum_{j=1}^{4} \beta_{2j}\left(\log\left(P_{A_{it}}\right) - \log\left(\underline{P}_{A_{it}}\right)\right) \times PriceDecrease_{A_{it}} \times Week_{A_{itj}} +$$
$$\beta_3 \log\left(P_{B_{it}}\right) + \sum_{j=1}^{4} \beta_{4j}\left(\log\left(P_{B_{it}}\right) - \log\left(\underline{P}_{B_{it}}\right)\right) \times PriceDecrease_{B_{it}} \times Week_{B_{itj}} + \gamma_i \log\left(T_{it}\right) + \Omega' X_{it} + \varepsilon_{it} \quad (1)$$

where *i* and *t* index product and time. The dependant variable is the log of sales rank on Amazon minus the log of sales rank at BN. The independent variables are log Amazon price (log(P_A)), log BN price (log(P_B)), level of price decreases (in log terms) for both Amazon and BN, the number of days since the product was released (T_{it}) and a vector of other control variables (X). X include the log of the lowest used product price for a given product, the consumer rating for the product, the log of the number of reviews, and the log of the number of used products offered for sale. Here β_1 and β_3 represent demand elasticity of Amazon and BN for price increases. β_2 denotes the difference between Amazon demand elasticity for price decreases and that for price increases. A negative value of β_2 thus implies high search cost for consumers at Amazon wherein demand elasticity for price decreases is smaller than that for price increases. If β_2 is positive, it indicates a low search cost for Amazon consumers where demand elasticity for price decreases is larger than that for price increases. The interpretation of β_4 is exactly the opposite of β_2 due to our definition of relative demand. Negative β_4 indicates low search cost for BN customers while positive β_4 indicates high search cost for BN customers

Independent Variable	All Books	All Books	Popular Books	Unpopular Books
Log(P_{Amazon})	1.71*** (0.11)	2.17*** (0.14)	0.88*** (0.25)	2.64*** (0.18)
Log(P_{BN})	-1.76*** (0.10)	-1.84*** (0.11)	-2.56*** (0.22)	-1.31*** (0.13)
ΔLog($_{Amazon}$)*$PriceDecrease_{Amazon}$	0.37*** (0.11)			
ΔLog(P_{BN})*$PriceDecrease_{BN}$	0.74*** (0.10)			
ΔLog(P_{Amazon})*$PriceDecrease_{Amazon}$*$OneWeek_{Amazon}$		0.08 (0.15)	0.62*** (0.27)	-0.23 (0.18)
ΔLog(P_{Amazon})*$PriceDecrease_{Amazon}$*$TwoWeeks_{Amazon}$		0.28* (0.15)	1.29*** (0.29)	-0.15 (0.18)
ΔLog(P_{Amazon})*PriceDecrease*$ThreeWeeks_{Amazon}$		0.38** (0.15)	1.66 (0.29)	-0.1 (0.18)
ΔLog(P_{Amazon})*$PriceDecrease_{Amazon}$*$FourWeeks_{Amazon}$		0.40*** (0.12)	2.24*** (0.23)	-0.17 (0.15)
ΔLog(P_{BN})*$PriceDecrease_{BN}$*$OneWeek_{BN}$		0.99*** (0.12)	-0.46** (0.22)	1.52*** (0.14)
ΔLog(P_{BN})*$PriceDecrease_{BN}$*$TwoWeeks_{BN}$		0.88*** (0.15)	0.22 (0.26)	1.22*** (0.19)
ΔLog(P_{BN})*$PriceDecrease_{BN}$*$ThreeWeeks_{BN}$		0.89*** (0.17)	0.3 (0.28)	1.16*** (0.21)
ΔLog(P_{BN})*$PriceDecrease_{BN}$*$FourWeeks_{BN}$		0.95*** (0.14)	0.3 (0.25)	1.13*** (0.17)
No. of Observations	89045	62994	20877	41404
R-square	69%	68%	79%	62%

Table 1: The dependent variable is ln(sales $rank_{Amazon}$) - ln(sales $rank_{BN}$). Standard errors are listed in parenthesis; ***, **, * denote significance at 0.01, 0.05 and 0.10, respectively. All models use product-level fixed effects.

The results are listed in Columns 1 and 2 in Table 1. Our results suggest that the demand function is indeed kinked and that search costs are low on Amazon but higher on BN. We observe that price elasticity for books increases by 8%, 28%, 38%, and 40% in the first, second third and fourth week respectively after a price reduction on Amazon. This result verifies Stiglitz (1989) theoretical prediction that a reduction in price produces higher price elasticity if search costs are low. Our result also shows that the price elasticity on Amazon follows a gradually increasing curve. That is, price elasticity increases over the first few weeks. The initial gradual increase in price elasticity is likely due to the presence of Amazon bargain shoppers who have been waiting for price decreases. Their presence produces the initial demand increase right after price decrease. With the passage of time, price information spreads out and attracts more consumers from Amazon's competitors. This increase in demand contributes to the gradual increase in price elasticity. Interestingly the price elasticity on BN remains more or less constant over the time period of four weeks. This result indicates that BN faces high consumer search cost and its price reduction information is rarely spread out even after 4 weeks.

4.2 Search Costs and Product Popularity

Since the inception of online retailing in the late 1990s, product assortments on the web have increasingly become broader and deeper. This enables them to offer convenient access to a larger selection of products than brick-and-mortar retailers. For example, small stores stock approximately 20,000 unique titles, and large independent booksellers stock approximately 40,000 unique titles (Brynjolfsson, Hu and Smith 2003). Prior work has argued that the Internet reduces search costs for rare and relatively unpopular books. However, the search costs on such unpopular books are likely to be higher than those on popular books. This is because popular books are more likely to be advertised and prominently featured by bookstores in both online and offline stores. These promotional actions reduce search costs for price information of popular books. Moreover price changes on popular books by one retailer are more likely to be matched by its competitor, leading to a faster dissemination of price change information and lower search costs for consumers. As an example, while it takes Amazon on average 74 days to response to BN price changes, if we separate books by their popularity, then it takes Amazon an average of 55 days to respond to changes in prices for the top 20,000 books, 58 days to respond to changes in prices for the top 40,000 books and 65 days to respond to changes in prices for the top 100,000 books. On the other hand, for books with sales rank higher than 100,000, Amazon responds to a BN price change after 90 days.

In order to test the hypothesis that search costs are higher on unpopular and rare books compared to popular books, we conduct a similar analysis as before but on a select sample. Specifically, we split the sample by salesrank into smaller samples. In keeping with the findings of prior work, we use 20,000 as the median salesrank cut-off for denoting books that are relatively popular and are likely to be stocked by offline retailers such as discount stores and specialty book stores. The results are reported in Columns 3 and 4 in Table 1. Comparing the coefficients on popular books and unpopular books against the results for all books in Column 2, we note that Amazon price elasticity increases significantly for popular books after a price decrease. For unpopular books, Amazon price elasticity however decreases after a price decrease. The coefficients on BN price decreases show a similar pattern. Overall, the results suggest that popular books face much less consumer search costs than unpopular books. Results using a median salesrank of 40,000 and 100,000 are very similar. As a robustness check we also used as the cut-off books which had a median salesrank of 20,000, 40,000 or 100,000 at Barnes and Noble. Our results are consistent across different specifications.

5. Implications for Consumer Welfare

The presence of online search costs has significant implication for consumer welfare since search costs reduce consumers' incentives to look for lower prices. Some consumers may drop out of the market due to their inability in finding low prices and other consumers may buy at a higher price than they could have got in a market without such costs. Using our estimates of demand, we measure consumer welfare loss by considering the gap between demand function with search costs and the corresponding demand function without search costs. Stiglitz (1989) indicates that price elasticity will increase after price decrease in a low search costs environment, if the firm attracts customers from competing firms. However, if the competing firms enjoy high customer loyalty, the increase in price elasticity would be minimal as few customers will migrate from the competing firms. We apply this principle to the competition between Amazon and BN. Prior studies show that in the competition between Amazon and BN, Amazon enjoys strong consumer loyalty while BN customers are relatively more price sensitive. To quantify consumer welfare loss, we also note that equation (1) only estimates relative demand. This is inappropriate for estimating changes in consumer welfare. We therefore revise the equation to estimate the true demand function faced by Amazon and BN. Our approach is similar to Chevalier and Goolsbee (2006) in that we also explore the time-series feature of the dataset. The empirical model for consumer welfare estimations takes the following simultaneous equation form:

$$\begin{aligned}&\log\left(SalesRank_{AMAZON_{it}}\right)-\log\left(\underline{SalesRank}_{AMAZON_{it}}\right)=\alpha_i+\beta_1\left(\log\left(P_{A_{it}}\right)-\log\left(\underline{P}_{A_{it}}\right)\right)+\\&\sum_{j=1}^{4}\beta_{2j}\left(\log\left(P_{A_{it}}\right)-\log\left(\underline{P}_{A_{it}}\right)\right)\times \Pr iceDecrease_{A_{it}}\times Week_{A_{itj}}+\beta_3\left(\log\left(P_{B_{it}}\right)-\log\left(\underline{P}_{B_{it}}\right)\right)+\gamma_i\left(\log\left(T_{it}\right)-\log\left(\underline{T}_{it}\right)\right)+\Omega' X_{it}+\varepsilon_{it}\end{aligned} \tag{2}$$

$$\begin{aligned}&\log\left(SalesRank_{BN_{it}}\right)-\log\left(\underline{SalesRank}_{BN_{it}}\right)=\alpha_i+\beta_1\left(\log\left(P_{B_{it}}\right)-\log\left(\underline{P}_{B_{it}}\right)\right)+\\&\sum_{j=1}^{4}\beta_{2j}\left(\log\left(P_{B_{it}}\right)-\log\left(\underline{P}_{B_{it}}\right)\right)\times \Pr iceDecrease_{A_{it}}\times Week_{A_{itj}}+\beta_3\left(\log\left(P_{A_{it}}\right)-\log\left(\underline{P}_{A_{it}}\right)\right)+\gamma_i\left(\log\left(T_{it}\right)-\log\left(\underline{T}_{it}\right)\right)+\Omega' X_{it}+\varepsilon_{it}\end{aligned} \tag{3}$$

where the dependent variable is the log of demand changes since the most recent price change.

Equations (2) and (3) represent the actual demand functions faced by Amazon and BN. Given the discussion above, we can also derive the hypothetical demand functions faced by Amazon and BN if there were no search cost in the market. The equations for the hypothetical demand functions are:

$$\begin{aligned}&\log(SalesRank_{AMAZON_{it}})-\log(\underline{SalesRank}_{AMAZON_{it}})=\hat{\alpha}_i+\hat{\beta}_1\left(\log\left(P_{A_{it}}\right)-\log\left(\underline{P}_{A_{it}}\right)\right)+\\&\hat{\beta}_{24}\left(\log\left(P_{A_{it}}\right)-\log\left(\underline{P}_{A_{it}}\right)\right)\times \Pr iceDecrease_{A_{it}}+\hat{\beta}_3\left(\log\left(P_{B_{it}}\right)-\log\left(\underline{P}_{B_{it}}\right)\right)+\hat{\gamma}_i\left(\log(T_{it})-\log(\underline{T}_{it})\right)+\hat{\Omega}' X_{it}\end{aligned} \tag{4}$$

$$\begin{aligned}&\log(SalesRank_{BN_{it}})-\log(\underline{SalesRank}_{BN_{it}})=\hat{\alpha}_i+\hat{\beta}_1\left(\log\left(P_{B_{it}}\right)-\log\left(\underline{P}_{B_{it}}\right)\right)+\hat{\beta}_3\left(\log\left(P_{A_{it}}\right)-\log\left(\underline{P}_{A_{it}}\right)\right)+\\&\hat{\gamma}_i\left(\log(T_{it})-\log(\underline{T}_{it})\right)+\hat{\Omega}' X_{it}\end{aligned} \tag{5}$$

Based on the results from estimation of equations (2)-(5) we can calculate loss of consumer welfare. We first identify all instances where Amazon and BN reduce product prices. We then calculate the actual welfare gain after the price decrease using estimated demand function (2) and (3) and the hypothetical welfare gain using equations (4) and (5). The difference between the two represents loss in consumer welfare due to search costs. Given that the demand data is estimated for daily sales, the loss of consumer welfare calculated is the daily loss for the particular product after a particular price decrease. To provide a standardized measure of loss in consumer welfare, we consider average consumer welfare loss for each $1 reduction in price per book sold. After standardizing the measure, the average daily social welfare loss on Amazon and BN is 2 and 8 cents respectively for each $1 price decrease per book sold. Using these statistics, we can aggregate them to quantify the total consumer loss by multiplying them with total annual sales and the average price decrease level. We know that total industry new book sales are about 2 billion units in 2005, of which Amazon sold about 101 million books. Using our data, we find that the average price decrease level for books is $2.02. The total annual consumer loss on Amazon is therefore 101 million* $2.02*0.02 = $4.08 million.

6. Discussion and Ongoing Research

We provide the first known empirical analysis of the nature of search cost on consumer demand curve in electronic markets. First, we show that as predicted by theory, consumer demand is indeed kinked, and this highlights that consumers face positive search costs even in online markets. Unlike prior work, our results use aggregate price and demand data from large online retailers, thus making the result more applicable to the general online markets. Our analysis also suggests that price elasticity increases after a price decrease for Amazon, but it decreases after a price decrease for BN. This implies that while search costs are ubiquitous in online markets, consumers face lower search costs on Amazon than they face on Barnes and Noble. Second, we consider the dynamics of how search costs change over time and quantify the time it takes for pricing information to percolate in the market. We demonstrate that price elasticity after a price decrease increases over time across a period of 4 weeks, suggesting that search costs do reduce over time. Third, we quantify the magnitude of the loss in consumer surplus from the presence of search costs. Our estimates reveal that the first order impact of search costs resulted in a decrease in consumer surplus by $4.08 million for Amazon customers. Finally, our analysis corroborates anecdotal evidence that popular books have lower search costs compared to less popular books. Since the latter are unlikely to be stocked by physical retailers and are only available in the online world, this finding can have implication for the Long Tail and optimal size of product assortments in the online markets. Search costs encourage a user to migrate toward more coherent content like recommendations and top sellers, thereby affecting the Long Tail of sales. Our ongoing work is aimed at econometrically analyzing how search costs affect the distribution of the Long Tail in online markets. We will be able to present the full paper by WISE 2006, if we are given the opportunity.

References

Baye, M. R., R. Gatti., P. Kattuman and Morgan, J. "Estimating Firm-Level Demand at a Price Comparison Site: Accounting for Shoppers and the Number of Competitors", 2005, Working paper.

Brynjolfsson, E., Y. Hu, and Smith, M. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety", *Management Science* 49(11), 2003, 1580-1596.

Brynjolfsson, Erik, Astrid A. Dick, and Smith, M. "Search and Product Differentiation at an Internet Shopbot", Working Paper, 2004, Carnegie Mellon University.

Chevalier, J. and Mayzlin, D. "The Effect of Word of Mouth on Sales: Online Book Reviews," forthcoming, *Journal of Marketing Research,* 2006.

Hong, H., and M. Shum. 2006. Using price distributions to estimate search costs. Forthcoming, *Rand Journal of Economics.*

Reinganum, J.F., "A Simple Model of Equilibrium Price Dispersion", *Journal of Political Econom*y, 87(4), 1979, 851-858.

Sorensen, A.T. "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs", *Journal of Political Economy*, 108, 2000, 833-850.

Stiglitz, J.E. "Competition and the Number of Firms in a Market: Are Duopolies More Competitive Than Atomistic Markets?", *Journal of Political Economy*, 95(5), 1987, 1041-1061.