Estimating Firm-Level Demand at a Price Comparison Site: Accounting for Shoppers and the Number of Competitors

Baye, Michael
GATTI, RUPERT J
Kattuman, Paul
et al.

Publication Date
2004-12-01

Working Papers
Competition Policy Center
University of California, Berkeley

Working Paper No. CPC05-050

Estimating Firm-Level Demand at a Price Comparison Site: Accounting for Shoppers and the Number of Competitors

Michael R. Baye
Kelley School of Business, Indiana University

J. Rupert J. Gatti
Trinity College, University of Cambridge

Paul Kattuman
The Judge Institute of Management, University of Cambridge

John Morgan
Haas School of Business and Department of Economics, University of California, Berkeley

December 2004

JEL Classification: D400, C250, C810, M300, L100

Keywords: Internet, Price Dispersion, Advertising

Abstract:
Clearinghouse models of online pricing---such as Varian (1980), Rosenthal (1980), Narasimhan (1988), and Baye-Morgan (2001)---view a price comparison site as an "information clearinghouse" where shoppers and loyals obtain price and product information to make online purchases. These models predict that the responsiveness of a firm's demand to a change in its price depends on the number of sellers and whether the price change results in the firm charging the lowest price in the market. Using a unique firm-level dataset from Kelkoo.com (Yahoo!'s European price comparison site), we examine these predictions by providing estimates of the demand for PDAs. Our results indicate that the number of competing sellers and both the firm's location on the screen and relative ranking in the list of prices are important determinants of an online retailer's demand. We find that an online monopolist faces an elasticity of demand of about -2, while sellers competing against 10 other sellers face an elasticity of about -6. We also find empirical evidence of a discontinuous jump in a firm's demand as its price declines from the second-lowest to the lowest price. Our estimates suggest that about 13% of the consumers at Kelkoo are "shoppers" who purchase from the seller offering the lowest price.

Estimating Firm-Level Demand at a Price Comparison Site: Accounting for Shoppers and the Number of Competitors*

Michael R. Baye J. Rupert J. Gatti Paul Kattuman

John Morgan

December 2004

Abstract

Clearinghouse models of online pricing—such as Varian (1980), Rosenthal (1980), Narasimhan (1988), and Baye-Morgan (2001)—view a price comparison site as an "information clearinghouse" where shoppers and loyals obtain price and product information to make online purchases. These models predict that the responsiveness of a firm's demand to a change in its price depends on the number of sellers and whether the price change results in the firm charging the lowest price in the market. Using a unique firm-level dataset from Kelkoo.com (Yahoo!'s European price comparison site), we examine these predictions by providing estimates of the demand for PDAs. Our results indicate that the number of competing sellers and both the firm's location on the screen and relative ranking in the list of prices are important determinants of an online retailer's demand. We find that an online monopolist faces an elasticity of demand of about -2, while sellers competing against 10 other sellers face an elasticity of about -6. We also find empirical evidence of a discontinuous jump in a firm's demand as its price declines from the second-lowest to the lowest price. Our estimates suggest that about 13% of the consumers at Kelkoo are "shoppers" who purchase from the seller offering the lowest price.

JEL Classification Numbers: D400, C250, C810, M300, L100

*The authors wish to thank Glen Drury and Rene Reimer of Kelkoo and Fru Hazlett of Yahoo! for providing access to this data. We thank Alex Chicesco, Steven Lewis, and Jason Snyder for their invaluable research assistance. We also thank Bronwyn Hall for her helpful advice. Finally, we are grateful for financial suppoort from the ESRC E-Society program and the National Science Foundation.

1 Introduction

A firm's demand in the online marketplace—particularly at price comparison sites such as Shopper.com, Nextag, and Kelkoo.com—fundamentally differs from demand at most physical marketplaces. One of the key differences—first noted by Baye and Morgan (2001) but certainly anticipated by the early works of Varian (1980), Rosenthal (1980), Shilony (1977), and Narasimhan (1988)—stems from the fact that consumers typically obtain a complete list of the prices charged by different sellers before making their purchase decision. As a consequence, 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. Moreover, unlike traditional retail markets where firms like Wal Mart and Target compete and there is little turnover in the identity of the firm charging the lowest price (it is almost always Wal Mart), the identify of the low-price seller frequently changes in online markets (Baye, Morgan, and Scholten, 2004b; Ellison and Ellison, 2004). Thus, it would seem to be important to account for the impact of shoppers when estimating online demand.

The online marketplace also differs from its physical cousin in the volatility in the number of competing sellers. In conventional retail markets, the number of firms competing for customers in (say) Walnut Creek, California change infrequently owing to the barriers to entry and exit associated with setting up a physical retail location. In the online world, change comes faster. This is particularly true in the marketplace defined by a price comparison site. Here, the number of firms listing prices for a given product changes almost daily (see Baye, Morgan, and Scholten 2004a). Indeed, as pointed out by Baye and Morgan (2001), this variation in the degree of rivalry of a given online market is essential for firms to avoid pure Bertrand competition in these markets and for the information "gatekeeper"—the entity running the price comparison site—to profitably operate its site. Thus, it would seem to be important to account for the degree of rivalry—the number of competing sellers—in estimating demand online. Among other things, theory indicates that the number of competitors is a key determinant of firm-level elasticities of demand: The greater the number of rivals, the more elastic is a firm's demand.

A third way online markets differ from conventional markets is in the changing location of firms. In conventional retail markets, the physical real estate a retailer occupies changes infrequently, and identifying the component of the "value-added" by its physical location is difficult to disentangle from other elements of the retailer's characteristics. In contrast, "virtual" real estate in the online world changes rapidly. For instance, in purchasing "ad-words" (advertising space at the side of search queries on Google's site), retailers realize the advantage conferred to being the first listing on the page—and bid aggressively to obtain such a position. At any moment, a retailer can find itself displaced from this "prime" real estate to a less favorable screen location. Similar locational advantages presumably accrue to firms with the topmost listings at price comparison sites, such as Kelkoo.com.

The rapidly changing nature of the online marketplace—the numbers of competing firms, identity of the low-price firm, and a firm's screen locations—presents both a challenge and an opportunity in estimating demand in online markets. The challenge is that if one fails to properly account for discontinuities in a firm's demand when it charges the lowest price, and the impact on a firm's demand of its screen location and the number of rivals, one might obtain biased estimates of demand. The opportunities stem from the dynamic nature of the data. Variation in the identity of the low priced firm enables one to disentangle the demand jump stemming from price sensitive shoppers from other determinants of demand. In conventional retail, such as in the competition between Target and Wal Mart, this identification is most difficult. A second opportunity arises from the frequently changing number of competitors, which enables identification of the marginal impact of the number of rivals on a firm's demand elasticity (and hence its markup). Obtaining these types of estimates from the physical marketplace is more difficult. Finally, the variation in screen locations enables identification of the value of the virtual real estate separately from firm characteristics.

Before summarizing how we seek to overcome these challenges and take advantage of the opportunities described above, it is important to point out a final challenge typically faced by researchers seeking to estimate demand in online markets. Clearly, the ideal dataset for demand estimation would include actual quantity data; however, owing to the fragmented nature of the e-retail marketplace and the proprietary nature of data, this is often a formidable challenge. Indeed, we know of only one paper, Ellison and Ellison (2004), that has been able to obtain quantity data, and even here the study is limited by the fact that the data is only for one of the many sellers competing at a particular price comparison site. More often, all that is available is clicks (or leads) data at a price comparison site—data on the number of customers that clicked on a particular seller's price displayed at the price comparison site. Customers clicking through, in this fashion, are then redirected to the seller's site to purchase the product.

We address these challenges and opportunities. Our paper demonstrates that theoretical "clearinghouse" models of online competition can be used, in conjunction with identifying restrictions we set forth in Proposition 1, to estimate online demand using clicks data. In particular, we show that one can use existing pseudo-maximum likelihood as well as standard maximum likelihood techniques specifically designed for count data to obtain consistent estimates of underlying demand parameters (including demand elasticities). We then apply these techniques to unique clickthrough data for 18 personal digital assistants (PDAs) obtained from Yahoo!s European price comparison site, Kelkoo.com. Consistent with what one might conjecture based on the challenges we identified above, we find evidence that it is indeed important to account for demand discontinuities, the number of rival sellers, and other determinants of demand such as screen location when estimating a firm's demand in online markets.

More specifically, our econometric results reveal that it is important to account for the "jump" in a firm's demand when it offers the lowest price. We find that a firm offering the best price enjoys a 60 percent increase in demand compared to what it would have enjoyed had it not charged the lowest price. Perhaps more importantly, failing to account for the jump in demand leads to elasticity estimates that are about twice those obtained allowing for demand discontinuities. Our results also reveal that a firm's elasticity of demand (apart from the jump discussed above) is more elastic in online markets where competition is keener. A monopoly seller faces an elasticity of demand of about -2.5, while in the most competitive markets we analyzed (15 sellers), the elasticity of demand for a representative firm's product is about -6.0. We are also able to identify the effect of other determinants of demand—such as screen location—on firm demand. Our results imply that firms lose about 15% of their business for every competitor listed above them on the screen.

Our results are related to a variety of papers in the literature. As noted earlier, Ellison

& Ellison (2004) use sales data for computer memory chips obtained from a single store, which listed on Pricewatch.com. The emphasis of their paper is on obfuscation and cross price elasticities between low and higher qualities of the same product offered by the firm; however they also indirectly obtain price elasticity estimates ranging from -25 to -40. In contrast, our analysis uses data from the complete set of firms listing prices across a broader selection of products, and where obfuscation is not prevalent. Ghose, Smith and Telang (2004) impute the sales of used books from the website of Amazon.com, which lists price offers for used books from many alternative and independent retailers. Using a multinomial logit model they estimate a price elasticity of -4.7, and note that the lowest priced offers receive a discontinuously higher proportion of sales. Chevalier and Goolsbee (2003) rather ingeniously impute price elasticities for new books at two bookstores (Amazon and Barnes & Noble) using prices and relative sales rankings - obtained directly from the retailers' web sites rather than through a price comparison site (where price sensitivity may be expected to be greater). They estimate price elasticities of -0.6 for Amazon and -4 for Barnes and Noble, but these estimates do appear to be sensitive to the particular estimating technique adopted. Using a similar methodology, Ghose et al. (2004) estimate the price elasticity for new books at Amazon to be -1.2.

Our analysis complements these studies by offering a methodology to directly estimate elasticities and to disentangle the impact on demand elasticities of: (1) discontinuous jumps at the lowest price; (2) variation in the number of competing firms; and (3) variation in screen location of price quotes.[1] Our paper rationalizes the disparate elasticity estimates (which range from -0.6 to -40) obtained in various online markets.

The remainder of the paper proceeds as follows: The next Section describes our data and provides an overview of the shopping environment at Kelkoo. In Section 3 we present the theory underlying our estimation methodology. Section 4 provides demand estimates based on individual as well as pooled products under the assumption that demand is continuous. These latter estimates are nested as a special case of the discontinuous demand specification, which is detailed in Section 5. Finally, Section 6 concludes.

[1] Pricewatch lists retailers in order of price, with the cheapest at the top of the screen. Thus, it is not possible to separately identify price ranking and screen location effects. Ellison and Ellison do not monitor the total number of firms listing prices at any particular time.

2 Data

The data used in this paper was obtained from the price listing service, Kelkoo, which is owned by Yahoo! According to Yahoo!, Kelkoo is the largest price listing service in world, operating in eight European countries and is recognized as one of the six most accessed web-sites in Europe. Within the UK, Kelkoo is the third largest retail website and attracts over 10 million individual users per month—more than twice that of its closest rival. Kelkoo would seem to be representative of the overall UK e-retail marketplace. Over 1,800 individual retailers—including 18 of the 20 largest online retailers in the UK—list prices on Kelkoo.[2]

Consumers interested in purchasing a broad range of products can access the Kelkoo site to obtain information about the product, a list of retailers selling the product, and the total prices charged for the product, including taxes and shipping charges. Consumers interested in making a purchase must 'click' on the firm's link at the Kelkoo site. They are then sent to the retailer's site, where the final purchase is made. Figure 1 illustrates a typical screenshot for the HP iPAQ H5550 PDA. Notice that consumers are provided with a brief description of the product, a list of retailers selling the product, along with the total price including VAT and shipping charge ('P&P' in Kelkoo's terminology). Our analysis is based on the total price—the actual final amount a consumer purchasing the product would be charged, inclusive of taxes and shipping charges. We cleaned the data obtained from Kelkoo to ensure that our analysis is based on listings of products that are identical in every respect (including condition).[3]

Kelkoo's revenue is generated by charging retailers a fee for each referral made; that is, each time a consumer clicks on the link at the Kelkoo site to a retailer's site. The fee charged by Kelkoo varies across products and retailers but is typically in the range of £0.20 to £1.00 per click, and is independent of whether a sale is subsequently made. Consumers are not charged for using Kelkoo's site.

Our data comes from referrals made through "menu prompted" results screens. A consumer arrives at these screens by navigating the products menu, moving from broader to

[2] Data taken from Hitwise Statistics and company information provided by Kelkoo.

[3] Arguably, one could simply add product controls for the small number of "variants" (e.g. refurbished, extra memory, etc.) of a given PDA in the sample. However, the number of these items was to permit this approach.

narrower product classes until arriving at the particular product in which she is interested.[4] For example, from Computers and Software in the opening page menu, a consumer would click and move to PDAs, where the menu has groupings by brand, price range, memory, weight and operating system. From this list, she would choose a specific PDA and be presented with a list of sellers and their prices. There are a number of reasons we opted for these data. First, Kelkoo generates roughly 45% of all its leads through consumers comparing prices in this fashion. Second, consumers using menu prompted searches are presented with price listings where the set and order of retailers displayed is unilaterally determined by Kelkoo and is identical for all consumers at any given time. Finally, the information displayed is verified and updated daily. Unlike many other price comparison sites, the order of retailers listed on these pages is neither auctioned or sold directly to retailers and is independent of the price quoted and the speed with which the retailers respond to a price request. Consequently, as far as both consumers and retailers are concerned, the order of price quotations on any specific screen is exogenous.

It is clear from Figure 1 that there is considerable variation in both prices charged and the characteristics of retailers on the site. Retailers listing on Kelkoo include retailers with a large brick and mortar presence, such as Comet and PC World, well-known online retailers such as Amazon and Dell, as well as less known online retailers such as Big Grey Cat.

Kelkoo maintains a log of information on each 'referral' made.[5] The log registers the retailer name, product name, price information, time of referral, location of the retailer on the screen and a cookie specific reference. For this study, Kelkoo provided us with daily data extracted from their log for 18 specific PDAs for the period from 18 September 2003 to 6 January 2004, a period which generated over 40% of Kelkoo's annual traffic.[6]

This traffic amounted to 39,568 leads generated via 20,509 separate "cookies." The majority (60.1%) of cookies generated only single leads, while a small number of cookies (0.56%)

[4] One might worry about a potential selection issue using this data since the search technique adopted by the consumer is endogenous. For this reason, we examined the robustness of our results by using data based on referrals generated using alternative methods. The results are qualitatively similar to those reported in the paper.

[5] Throughout the paper we use the terms 'referral,' 'lead,' and 'click' interchangeably.

[6] Kelkoo is bound to protect the anonymity of retailers in disclosing information about the referrals they obtain. In providing the information from their log files, the retailers were identified in the dataset by codes, and by some key characteristics, such as whether they had a brick and mortar presence.

generated a great many (>10) leads each. To avoid the potential problem of over-weighting consumers who generated multiple leads, we restrict our analysis only to the last clicks data for each cookie. [7] Over the period of the study there were 6,151 individual product, retailer and day specific price listings across the 18 PDAs monitored. For each of these daily listings, we determined the number of (last-click) referrals per day, for each PDA sold by each seller. Finally, we note that for all products in our sample, the complete list of price offers was always displayed on a single page. Thus, no consumer was required to click more than once to view the entire list of firms selling a specific PDA on a specific day.

2.1 Summary Statistics

Our analysis is based on the 18 models of PDAs listed in Table 1. These include different models of Palm, HP, Sony, and Toshiba PDAs, and span a wide range in prices. The lowest priced item is the Palm Handspring Treo, which has a median price of about £130. The most expensive product in our dataset is the Sony Clie nz90, which has a median price of about £537. An initial inspection of the prices in our dataset suggested that prices in the UK were considerably higher than in the US. To examine this hypothesis, we obtained data from the US price comparison site Shopper.com for the same time period and the same set of PDAs. A comparison of the US and UK prices listed in Table 1 confirm that UK prices are indeed higher than those in the US. Part of the difference in prices between the two countries stems from the fact that the UK prices include 17.5% sales tax (VAT), unlike US prices which are quoted exclusive of taxes. But even after deducting the VAT, the median retail price for PDAs sold online in the UK is about 20% higher than in the US. Part of the explanation may lie in differing market structures—the average number of retailers listing a PDA at Kelkoo is less than 4, while the corresponding number of sellers at Shopper.com is about 24.

Table 2 summarizes our data, which consists of daily prices and referrals for 18 PDAs offered by 19 retailers. Referrals are fairly evenly distributed across the period of our study. Firms with a brick and mortar presence obtained 29% of the total number of referrals.

[7] We also performed the analysis reported below using data on all clicks as well as only first-clicks data and obtained qualitatively similar results.

Weekdays generated the most traffic, with an average of 350 referrals each day, while the weekends generated about three-quarters of this number (cf. Figure 2).

There are very substantial differences in the number of referrals enjoyed by different retailers, ranging on a per-product basis from zero to 36 per day. The median number of referrals a firm received on a given product was 2. In a substantial number of instances, a firm received 0, 1, or 2 referrals for a given product on a given date. The average price was about £305 and the average shipping charge was about £4. The number of firms listing prices on a given product-date varies considerably, ranging from 1 to 15 firms with a standard deviation of 2.93.

Figure 3 suggests that price and screen location play a potentially important role in determining the business enjoyed by particular online retailers. Consumers appear to be very sensitive to price, as is evidenced by the dramatic decline in leads enjoyed by firms offering less favorable prices. Likewise, consumers tend to frequent firms that are listed above others on the screen. While screen location is not determined by price, it is possible that the results displayed in Figure 3 are the result of spurious correlation between screen location and price. We deal with this issue formally in Section 4 of the paper.

3 Estimation Methodology

In this section, we describe our methodology for estimating the impact of various explanatory variables on firm demand at the Kelkoo site. Given that we only observe clicks and not final purchases, we offer identification restrictions on demand that, if satisfied, allow us to estimate firm demand elasticities even though we do not have final sales data. In the second subsection, we describe the pseudo maximum likelihood procedures which permit us to obtain consistent estimates of various elasticities of demand.

3.1 Identification Restrictions

Recall that, to purchase a product, a consumer visiting the Kelkoo site must first process the information contained on the site and decide whether, and on which firm, to click. Following this, a consumer clicking through to the merchant's site obtains additional information about the desirability of purchasing the product and ultimately decides whether to buy or not. Thus, demand can be decomposed into two parts: the click generating process and the process of converting leads to sales. The process of generating leads depends on a number of factors, the most important of which are highlighted in Figure 4. As the figure shows, leads depend on the price a firm charges, market structure—the number of other firms offering the same product and the prices these firms are charging, the characteristics of the firm, timing, and the location of the firm's listing on the "page." Formally, let X denote this and other information a consumer obtains directly from the Kelkoo site. Note that X may include dummy variables to control for product-specific characteristics (some products are more popular and receive more clicks, on average, than others), firm characteristics (some firms may have a brick-and-mortar presence while others do not), and time effects (firms may receive fewer clicks on weekends or products may exhibit life-cycle effects that cause clicks to vary systematically over time). Let the number of leads that firm i receives, Q_i , be drawn from some distribution $F_i(\cdot|X)$. Thus,

$$E[Q_i|X] = \int q dF_i(q|X) \tag{1}$$

where we use a Lebesque integral to account for the fact that Q_i is discrete. Based on the information in X—and this information alone—a representative consumer can decide to close his or her window or to click through to a particular merchant.

A consumer who clicks through to a firms' site then receives additional information (denoted Z_i) that influences her decision to purchase. This information might include the firm's attempt at obfuscation along the lines described by Ellison and Ellison (2004), the visual attractiveness and usability of the firm's site, whether the firm is offering any guarantees on the product over and above those provided by the manufacturer, the exact restocking and return policies of the firm, and so on. Of course, a consumer's perceptions of these factors may be colored by the previous information, X, obtained on the Kelkoo site. Thus, the probability that a click is converted into a sale is

$$\Pr(sale_i|Z_i, X) = G_i(Z_i, X)$$

and, combining these expressions, the expected demand for a given product sold by firm i is

$$E[D_i|X, Z_i] = G_i(Z_i, X) \times E[Q_i|X]$$

Suppose that we are interested in the impact on the demand for a given product sold by firm i from some explanatory variable, x_i. It is useful to rewrite $X = (x_i, X_1)$ where X_1 represents all components of X other than x_i. The following proposition offers conditions in which one may identify the impact of x_i on firm i's demand when only leads data are available.

Proposition 1 *Suppose that $G_i(Z, (x_i, X_1)) = G_i(Z, (x_i', X_1))$ for all x_i, x_i'. Then one can use clicks data to identify the elasticities of demand with respect to x_i (when demand is differentiable with respect to x_i) as well as the percentage change in demand resulting from a given change in x_i (when demand is not differentiable with respect to x_i).*

Proof. We first prove the result for the differentiable case. Recall that log expected demand is given by

$$\ln E[D_i|X, Z_i] = \ln G_i(Z_i, X) + \ln E[Q_i|X]$$

Differentiating with respect to x_i yields

$$\frac{\partial \ln E[D_i|X, Z_i]}{\partial \ln x_i} = \frac{\partial \ln G_i(Z_i, X)}{\partial \ln x_i} + \frac{\partial \ln E[Q_i|X]}{\partial \ln x_i}$$

and since $G_i(Z, (x_i, X_1)) = G_i(Z, (x_i', X_1))$ for all x_i, x_i'., then $\frac{\partial \ln G_i(Z_i, X)}{\partial \ln x_i} = 0$. Hence

$$\frac{\partial \ln E[D_i|X, Z_i]}{\partial \ln x_i} = \frac{\partial \ln E[Q_i|X]}{\partial \ln x_i}.$$

Next, we prove the result for the non-differentiable case.

$$\begin{aligned} \%\Delta E[D_i|(x_i, X_1), Z_i] &= \frac{G_i(Z_i, (x_i, X_1)) E[Q_i|x_i, X_1] - G_i(Z_i, (x_i', X_1)) E[Q_i|x_i', X_1]}{G_i(Z_i, (x_i, X_1)) E[Q_i|x_i, X_1]} \\ &= \frac{E[Q_i|x_i, X_1] - E[Q_i|x_i', X_1]}{E[Q_i|x_i, X_1]} \end{aligned}$$

where we have again used the fact that $G_i(Z, (x_i, X_1)) = G_i(Z, (x_i', X_1))$ for all x_i, x_i'. This completes the proof.

Two special cases of Proposition 1 are worth noting. First, when x_i is firm i's price we can, in principle, estimate own price demand elasticities purely through leads data. Second, when x_i is a discrete variable (such as a dummy variable), the proposition implies that we can use estimates based on clicks data to infer the percentage impact of a discrete change in x_i on demand.

It is important to note, however, that even if the identifying restriction stated in Proposition 1 is not satisfied, we are still in a position to examine the impact of various aspects of the leads generating process described in Figure 4 on the leads a firm receives from Kelkoo. Even in this latter case, one may be able to use *a priori* information along with clickthrough elasticities to obtain bounds on demand elasticities. For instance, if one has reason to believe that conversion rates are increasing in x_i, it follows that elasticities based on clickthrough data provide a lower bound on the relevant demand elasticity (since in this case, $\frac{\partial \ln G_i(Z_i,X)}{\partial \ln x_i} + \frac{\partial \ln E[Q_i|X]}{\partial \ln x_i} > \frac{\partial \ln E[Q_i|X]}{\partial \ln x_i}$).

In the sequel, we assume that the condition of Proposition 1 are satisfied, so that we can identify relevant demand parameters purely through clicks data.

3.2 Data Generating Process for Leads

In light of the identification restrictions in Proposition 1, the next step in estimating demand parameters is the specification of the underlying stochastic process generating leads. For the reasons discussed below, we use a pseudo-maximum likelihood approach that does not require us to make specific assumptions about the underlying distribution generating Q_i; instead, we merely assume the underlying stochastic process has finite mean, given by

$$E\left[Q_i|X\right] = \exp\left[X\beta\right] \tag{2}$$

In order to estimate the vector of unknown parameters, β, one must account for the fact our clicks data consist of integer numbers of clicks. In fact, as shown in Table 2, over 50 percent of the data consist of observations where firms selling a given product received two or fewer clicks on a given day. For this reason, analysis of these data require regression techniques suitable for count data. Thanks to recent advances in the econometrics of count data, a variety of estimation techniques are available. One approach is to make a specific distributional assumption regarding the underlying stochastic process (Poisson or negative binomial, for instance), and use standard maximum likelihood methods to obtain estimates of the underlying parameters, β. Conditional on the underlying distributional assumption being correct, one obtains consistent estimates and standard errors and may perform standard hypothesis tests on β. Unfortunately, even if the mean specification in equation (2) is correct, it is known (cf. Gourieroux, *et al.* (1984a,b); Cameron and Trivedi, 1986) that the resulting maximum likelihood estimates of β and/or the standard errors will be inconsistent if the true stochastic process is different from that used to obtain maximum likelihood estimates.

For this reason, we adopt the pseudo-maximum likelihood (PML) approach due to Gourieroux, *et al.* (1984a,b) that has received renewed interest due to Cameron and Trivedi (1998) and Hall and Ziedonis (2001). Roughly, Gourieroux, *et al.* (1984a) show that so long as the mean specification in equation (2) is correct, any maximum likelihood estimator for β obtained by maximizing the likelihood function based on the linear exponential class will be consistent for β even if the underlying distribution used to obtain the MLE is misspecified. Since the Poisson MLE is in the linear exponential class but the negative binomial and other common specifications used for count data are not (when the parameters of the assumed distribution are unknown), Hall and Ziedonis (2001) use the Poisson-based PML approach to obtain consistent estimates of β.

While the Gourieroux, *et al.* results imply that the ML estimates of β based on a Poisson distribution are consistent even when the underlying data generating process for the Q_i 's are not Poisson, the resulting ML estimates of the variance-covariance matrix are not consistent if the underlying distribution is not, in fact, Poisson. For this reason, Gourieroux, *et al.* propose what they call pseudo-maximum likelihood estimation: The estimator of β is based on the first-order conditions for maximizing the likelihood function based on a Poisson distribution, but the distribution of the estimator is not based on the Poisson distributional assumption. Our approach is similar to that taken by Hall and Ziedonis (2001): we obtain consistent estimates of the variance matrix without specifying a specific functional form for the variance (as would be required if one used the negative binomial) by using robust standard errors.

In short, by using pseudo-maximum likelihood estimates based on a Poisson distributional assumption, we obtain a consistent estimate of β even if the underlying distribution is not Poisson. By using robust standard errors, we obtain consistent estimates of the variance of this estimate. In contrast, maximum likelihood methods based on a specific distributional assumption (such as the negative binomial) would lead to more efficient estimates if the specification of the data generating process is correct, but inconsistent estimates if the distribution is not correct. Thus, as in Hall and Ziedonis (2001), our preference for the Poisson pseudo-maximum likelihood approach stems from its robustness to potential misspecification. As discussed below, we also provide MLE estimates in the Appendix based on specific distributional assumptions, including the negative binomial (see Cameron and Trivedi, 1998), as well as specifications that allow for unobserved firm heterogeneity (using both random and firm specific effects, as in Griliches, Hausman, and Hall, 1986). It is reassuring that our results are robust to these alternative specifications.

4 Continuous Demand Models

In this section we provide estimates of a representative firm's demand under the assumption that each firm's demand is a continuous function of its price. As noted in the introduction, this is the conventional way of estimating demand, and has led to firm elasticity estimates ranging from –0.6 (in the relatively concentrated online market for books) to about -40 (in a highly competitive online market for computer memory). The main message of this section is to show that differences in seller concentration across different PDAs is useful in explaining variations in firm elasticities.

4.1 Estimates by Product

As a starting point, we pool across firms (i) and dates (t), but estimate separate elasticities for each of the 18 different models of PDAs in our data using the pseudo-maximum likelihood procedure described above. Specifically, we assume

$$E\left[Q_{ijt}|X_{ijt}\right] = \exp\left[\beta_j \ln p_{ijt} + \gamma_j X_{1,ijt}\right], \tag{3}$$

where Q_{ijt} is the number clicks firm i received on product j at time t, p_{ijt} is the total price (including VAT and shipping) firm i charged for product j at time t, and $X_{1,ijt}$ is a vector of controls. Notice that, under our maintained hypothesis that the identifying restrictions in Proposition 1 hold, β_j is the own price elasticity of demand for a representative seller of a model j PDA. Table 3 reports the results, which include the following controls:

Screen Location. As we discussed above, when a price listing is located nearer to the top of the screen, it tends to receive more clicks. Clicks decrease monotonically as the position on the screen gets lower. Hence, we have included a linear screen location control to absorb this effect.

Weekend. As displayed in Figure 2, there are systematically fewer clicks on weekends than on weekdays. Hence, we have included a weekend dummy variable to control for this effect on demand.

Month. Our dataset covers, in part, the fourth quarter of the year, which traditionally the strongest part of the retail season. As a result, we include month dummies to control for seasonal effects on demand.

Notice that 13 of the estimated own price elasticities in Table 3 are statistically significantly different from zero at the 1 percent level, with values ranging from -1.75 (for the Toshiba E770) to -14.691 (for the HP Compaq iPAQ 1940). These estimates vary widely across PDAs. In interpreting these results, and to better understand the widely different estimates obtained for different models of PDAs, it is important to recognize that these estimates are firm elasticities — not market elasticities. One of the key determinants of the elasticity of demand is the availability of substitutes – the more sellers offering the same product, the more elastic is the demand facing a firm selling that product. It is well-known, for instance, in a symmetric n-firm capacity-constrained price-setting environment, the elasticity of demand facing an individual firm (E_F) is n times the market elasticity (E_M): $E_F = nE_M$. If this is the case and different numbers of firms sold different types of PDAs, the firm elasticities of demand would vary widely across PDA models even if the market elasticity of demand were the same for each model of PDA.

Thus, it seems useful to investigate the relationship between our elasticity estimates and the average number of firms listing a price, across PDA models. This relationship is plotted in Figure 5. The estimates are divided into those that do not obtain statistical significance at conventional levels, which are shown as open circles and those that attain significance at the 1% level, which are shown as filled-in diamonds. As the figure shows, there is a strong negative relationship between the elasticity estimates for each of the products and the average number of firms offering price quotes for the product. This suggests the need for controlling for “market structure” among PDAs, were one to pool across all products. Thus, even if the market elasticity of demand were identical for each of the 18 models of PDAs, firms selling different PDAs would face differing elasticities of demand given the inverse relationship between the elasticities of demand the number of firms selling each product.

4.2 Pooled Estimates

We now report estimates obtained by pooling across firms (i) and dates (t), and different models of PDAs using the pseudo-maximum likelihood procedure described earlier. Here we consider two models: a baseline model that does not allow elasticities to vary with the number of sellers, and a more general model that takes into account our preliminary findings in the individual product specifications. The baseline model assumes

$$E\left[Q_{ijt}|X_{ijt}\right] = \exp\left[\beta \ln p_{ijt} + \gamma X_{1,ijt}\right]. \tag{4}$$

The controls for this specification include all of those in equation (3) as well as following:

PDA Model. As the previous specification revealed, there are differences in clicks for each of the different PDA models. For instance, PDAs differ from one another in terms of their popularity, their operating system, various performance characteristics, add-on software, and so on. Thus we include dummy variables for each of the 18 PDA models.

PDA Model-Month Interactions. In addition, the popularity of a PDA varies depending on new entrants in the PDA product space. As technology and performance improve with the introduction of new models, the popularity of an existing PDA can decline—sometimes dramatically. To control for these effects, we include dummies interacting each of the 18 PDAs with the month dummies mentioned above. This, in principle, allows for differing PDA “life cycle” effects.

Bricks and Clicks Retailer. Some of the firms in our dataset have an established physical presence in addition to their online presence. These are commonly referred to as “bricks and clicks” retailers. Clearly, the reputation as well as the ease of returns and accumulated brand equity of these retailers are likely to be different from those with only an online presence. Thus, we include a dummy variable for whether a particular firm is a bricks and clicks retailer or not.

With these controls in place, we report pseudo-maximum likelihood estimates (based on a Poisson likelihood function) using equation (4) in the column labeled "Model 1" in Table 4. The bottom of Table 4 also reports the results of a Lagrange multiplier test for over dispersion of the negative binomial (2) type (see Cameron and Trivedi, 1990). This is a test of the null hypothesis that the mean and variance of the click generating process are equal, as would be the case were the data generating process truly coming from a Poisson distribution. As the table shows, we overwhelmingly reject this hypothesis, indicating that the underlying distribution is not Poisson. As discussed above, the parameter estimates are nonetheless consistent (provided the mean specification in equation (4) is correct), but the overdispersion test indicates that Poisson-based maximum likelihood estimates of their standard errors are *not* consistent. To obtain consistent estimates of the standard errors, we employ the techniques of Rogers (1993), Huber (1967) and White (1980,1982).[8] The corresponding z-statistics are reported in Table 4.[9]

The results show a price elasticity of -4.61, which is fairly close to the average over the individual product elasticities reported in Table 3. More favorable screen positions are also shown to lead to increased clicks. Using the maintained hypothesis that conversion rates are independent of screen position at the comparison site, one can interpret the effect on demand of screen position as follows: All else equal, a firm which is shifted by one rank in its screen location experiences an 18.6% decrease in demand. These results confirm what we saw earlier in Figure 3: There is a strong tendency for consumers to click on firms in higher positions, all else equal. This is consistent with the observation that, for search engines such as Google, Overture, and Nextag, who auction screen position, there is a significant premium associated with being located in the highest position.

Interestingly, while the coefficient associated with being a bricks and clicks retailer has the expected positive sign (0.262), it is not significant at conventional levels. One potential reason for this is that, in light of the relationship we observed in Figure 5, the baseline model

[8] Specifically, we use the grouping technique of Rogers (1993) to relax the independence of observations for a given firm i across products and time. This allows potential autocorrelation and heteroskedasticity in the errors.

[9] Some researchers have taken the view that the rejection of the null hypothesis of no overdispersion warrants the use of a negative binomial specification. For this reason, we report ML estimates based on the negative binomial (2) specification in Table A1. As that table shows, the parameter estimates are very similar.

is potentially misspecified because it assumes a representative firm's elasticity of demand is independent of the number of firms. As we shall see in Section 5, when we account for this, together with the potential discontinuity in demand for the firm offering the lowest price, the coefficient on bricks and clicks retailers remains positive and becomes significant.

To account for a potential relationship between a firm's elasticity of demand and the number of competing seller in the pooled model, we modify equation (4) to allow individual firm elasticities to depend on the number of listing firms as follows:

$$E\left[Q_{ijt}|X_{ijt}\right] = \exp\left[\left(\beta_0 + \left(n_{jt} - 1\right)\beta_1\right)\ln p_{ijt} + \beta_2 n_{jt} + \gamma X_{1,ijt}\right], \tag{5}$$

where n_{jt} denotes the number of firms listing prices. Notice that, in this specification, the elasticity of demand facing a representative firm is given by

$$\beta_0 + \left(n_{jt} - 1\right)\beta_1.$$

Thus, β_0 represents the elasticity of demand facing a monopoly seller, $\beta_0 + \beta_1$ represents the elasticity of demand in duopoly PDA markets, and more generally, β_1 represents the impact on a firm's elasticity of demand of facing an additional competitor. In addition to our earlier controls, we include the following:

The Number of Sellers. Besides the theoretical rationale for permitting a representative firm's elasticity of demand to depend on the number of sellers, one might expect the number of clicks received by a particular firm to directly depend on the number of sellers. For a given consumer base, adding additional sellers would tend to reduce the expected number of clicks enjoyed by any particular firm. In addition, one might speculate that consumers are more likely to click and purchase PDAs that are sold by more firms, as additional firms might stimulate online sales by making the market appear more credible in the eyes of consumers. As we will see below, our framework permits one to disentangle these two competing effects.

The resulting estimates are displayed in the Model 2 column of Table 4. As the table shows, the number of sellers has a significant effect on clicks—both in terms of levels as well as on price elasticities. Controlling for the number of firms listing prices, we find that the price elasticity of a monopoly seller is -3.761, which implies a gross margin of 26.6%. Adding a second firm to the market raises the price elasticity to around -4.049 and cuts the gross margin to 24.7%. When ten firms list prices, the estimated elasticity becomes -6.641 or about 15.1% gross margins. These results suggest that the UK online market for PDAs is extremely competitive. By way of comparison, the average gross margin for US electronic shopping and mail order retailers (NAICS 4541) was 38.5%.[10] The sign of the coefficient capturing the impact of the number of firms listing prices on elasticity is also consistent with the simple capacity constrained price setting model described above.

What is the impact of a change in the number of competing firms on the overall numbers of clicks for a specific PDA? As we mentioned above, there is a direct effect as well as an indirect effect from increased competitiveness. Taking the derivative of equation (5) and evaluating it at the mean of our data yields

$$\begin{aligned} \frac{\partial \ln E\left[Q_{ijt}|X\right]}{\partial n_{jt}}|_{\bar{p}_{ijt}} &= \hat{\beta}_1 \ln \bar{p}_{ijt} + \hat{\beta}_2 \\ &= -.288\,(5.67) + 1.593, \end{aligned}$$

or about -0.04 ($p = .0155$). It is useful to contrast this effect with the effect on numbers of clicks of a change in screen position. As Table 4 shows, moving down one screen position decreases the number of clicks by 17.5%. Thus, our estimates suggest that the impact on clicks of screen position is more than four times larger than the impact on clicks of an additional competitor appearing on the price comparison site.

While the above results are of some interest, there is reason to believe that the underlying continuous demand model on which the estimated demand elasticities as well as the marginal effects of screen location and the number of listings is misspecified. As we have emphasized, the pseudo-maximum likelihood approach is robust against distributional assumptions but not misspecification of the underlying mean of the stochastic process. We conclude this section with a discussion of several potential problems that we address below.

4.3 Potential Misspecification

One may have a number of concerns regarding the estimates based on the continuous demand specification in equation (5). First and foremost, price comparison sites are often used by

[10] Source: Table 6. Estimated gross margin as a percent of sales by kind of business. US Census Bureau, revised June 1, 2001.

consumers looking to obtain a given product at the best price. For instance, Brynjolfsson and Smith (2000) have provided evidence that 49 percent of consumers using price comparison sites in the U.S. make purchase decisions based purely on price. The results of Ghose, *et al.* seem to indicate a jump in a firm's demand when it sets the lowest price. Moreover, recall that in our data (see Figure 3), 45 percent of the clicks are at the lowest price. These observations, coupled with the recent literature that rationalizes the observed levels of price dispersion in online markets (see Baye and Morgan, 2001; Baye, Morgan, and Scholten, 2004a) suggests that a firm lowering its price from the second-lowest to the lowest price enjoys a discontinuous jump in demand.

To see the potential ramifications of this on demand estimation, suppose there is a unit mass of consumers, half of which are "shoppers" who purchase at the lowest price and the other half are "loyals" who have a preference for a particular seller. Suppose consumers within each group have identical demand functions given by $D = p^{-\theta}$. A firm that charges the lowest price in the market enjoys demand from both groups, while a firm charging a price above the minimum price in the market sells only to its loyal customers. Figure 6 illustrates the ramifications on demand estimation. The slope of the two steep lines through the data are the same, and represent the true elasticity of demand, $-\theta$, for prices above or below the minimum price. At the minimum price, there is a discontinuous jump in demand owing to the fact that the firm attracts all of the shoppers at this price.

The dashed line through the data represents the elasticity estimate that results from failing to take into account the discontinuous jump in demand that occurs when the firm charges the lowest price. Notice that, by ignoring the jump in demand at the lowest price, one obtains an estimate of the true elasticity that overstates how responsive consumers are to a change in price.

In addition to the potential problem caused by using a continuous demand specification in the presence of "shoppers," two additional econometric issues are potentially relevant. First, while there are sound theoretical reasons for elasticities (and per-firm demand) to depend on the number of firms listing prices, the estimates may be biased due to potential endogeneity. In particular, popular products are likely to (for a given number of firms) result in a firm receiving more clicks, and this may encourage additional firms to enter the market.

We attempted to control for this by including product dummies and interactions between product and month dummies. However, endogeneity of the sort described above could, in principle, still be a problem.

Second, while we have controlled for what seemed an important firm characteristic—whether a firm is a bricks and clicks retailer—a variety of unobserved firm characteristics, such as the degree of accumulated brand equity or differences in consumers' perceptions of firm quality, could also potentially bias our results. Thus, it may be important to try to control for unobserved firm characteristics in estimating demand.

We address these and other issues in the next section.

5 Discontinuous Demand and Unobserved Heterogeneity

We now turn to estimating demand in the presence of a mix of price-sensitive shoppers and "loyals". We first sketch the theory underlying the demand estimation. We then describe the estimating equation and report results. Finally, we examine issues associated with endogeneity and unobserved firm characteristics.

5.1 Theory and Estimation Strategy

Suppose that n_{jt} firms numbered $i = 1, 2, ..., n_{jt}$ sell product j at a price comparison site on date t. Let p_{ijt} denote the price of firm i in the market for product j. A firm in this market sells to two types of consumers: Shoppers, who always purchase from the firm charging the lowest price, and loyals, who purchase from their preferred firm. Because of the extreme price sensitivity of shoppers, it is useful to define the set of firms offering the "best" (lowest) price for product j at time t . Define the set

$$B_{jt} = \{i : p_{ijt} \leq p_{kjt} \text{ for all } k \neq i\},$$

which is the set of firms offering the "best" price on this product date.

Let Q^S_{ijt} and Q^L_{ijt} denote the product j leads firm i obtains from shoppers and loyals, respectively, when charging the price p_{ijt}. Recall that firm i obtains product j leads from shoppers only if it is in the set B_{jt}; that is, if it offers one of the best prices. Thus, the clicks firm i obtains when it charges a price p_{ijt}, given the prices charged by other firms is

$$Q_{ijt} = \begin{cases} Q^S_{ijt} + Q^L_{ijt} & \text{if} \quad i \in B_j \\ Q^L_{ijt} & \text{if} \quad i \notin B_j \end{cases}$$

Thus, firm i faces a "jump in demand" for product j when it is among those firms offering the "best" price for product j.

We utilize the following functional approach that facilitates structural estimation of demand in a clearinghouse model. To account for the discontinuity in demand when the firm offers one of the best prices in the market, let $\mathbf{I}_{jt}$ be an indicator function that equals unity when $i \in B_{jt}$ and zero otherwise, and let $\#B_{jt}$ denote the cardinality of B_{jt}; that is, the number of firms offering the best price for product j. Suppose that firm i 's elasticity of demand when it sells product j is θ_{jt}, so that we may write

$$Q_{ijt} = \alpha^L_{ijt}(X)\, p_{ijt}^{-\theta_{jt}} + \mathbf{I}_j \frac{1}{\#B_{jt}} \alpha^S_{ijt}(X)\, p_{ijt}^{-\theta_{jt}}$$

where $\alpha^L_{ijt}(X)$ and $\alpha^S_{ijt}(X)$ represent the non-price determinants of leads (such as screen location) on loyals and shoppers, respectively. To ease the notational burden, we suppress the X argument where it is clear.

$$\begin{aligned} Q_{ijt} &= \left(\alpha^L_{ijt} + \frac{\mathbf{I}_{jt}}{\#B_{jt}} \alpha^S_{ijt}\right) p_{ijt}^{-\theta_{jt}} \\ &= \left(1 + \frac{\mathbf{I}_j}{\#B_j} \lambda_{ijt}\right) \alpha^L_{ijt} p_{ijt}^{-\theta_{jt}} \end{aligned}$$

where

$$\lambda_{ijt} = \frac{\alpha^S_{ijt}}{\alpha^L_{ijt}}.$$

Taking logs (and noting that $\ln[1 + \lambda_{ijt} \frac{\mathbf{I}_{jt}}{\#B_{jt}}] \approx \lambda_{ijt} \frac{\mathbf{I}_{jt}}{\#B_{jt}}$) yields

$$\ln Q_{ijt} = \lambda_{ijt} \frac{\mathbf{I}_{jt}}{\#B_{jt}} + \ln \alpha^L_{ijt} - \theta_{jt} \ln p_{ijt} \tag{6}$$

In addition to the identifying restriction in Proposition 1, estimation requires imposing additional structure on the parameters in equation (6). We assume a firm's elasticity of demand for product j in period t is given by

$$\theta_{jt} = \left(\beta_0 + (n_{jt} - 1)\,\beta_1\right) \tag{7}$$

As in the previous section, this parsimonious specification takes into account the theoretical relation between a firm's elasticity and the number of competing firms. In addition, we allow different firms to have different numbers of loyals and shoppers, and also permit the number of each to vary over time and across products. However, we assume

$$\begin{aligned} \alpha_{ijt}^{S}\left(X_{ijt}\right) &= a^{S}\alpha_{ijt}\left(X_{ijt}\right) \\ \alpha_{ijt}^{L}\left(X_{ijt}\right) &= a^{L}\alpha_{ijt}\left(X_{ijt}\right) \end{aligned}$$

so that the ratio of these two expressions is constant. In particular, this assumption implies

$$\begin{aligned} \lambda_{ijt} &\equiv \frac{\alpha_{ijt}^{S}\left(X_{ijt}\right)}{\alpha_{ijt}^{L}\left(X_{ijt}\right)} \\ &= \frac{a^{S}}{a^{L}} \\ &\equiv \lambda, \end{aligned}$$

which yields

$$E\left[Q_{ijt}|X\right] = \exp\left[\left(\beta_0 + \left(n_{jt} - 1\right)\beta_1\right)\ln p_{ijt} + \beta_2 n_{jt} + \lambda\frac{\mathbf{I}_{jt}}{\#B_{jt}} + \gamma X_{1,ijt}\right], \tag{8}$$

where $X_{1,jt}$ is the matrix of controls discussed earlier (position on screen, bricks and clicks retailer, weekend, product dummies, month dummies, and product-month interaction dummies). We continue to take as our maintained hypothesis that the identifying restrictions in Proposition 1 hold so that $(\beta_0 + (n_{jt} - 1)\beta_1)$ represents the elasticity of demand of a firm that faces $n_{jt} - 1$ competitors. If, in addition, conversion rates are independent of t, $\mathbf{I}_{jt}$, and the cardinality of the number of firms offering the lowest price $\#B_{jt}$, one may interpret λ as the size of the jump in demand that a firm enjoys when it offers the "best" price. Notice that the continuous demand model is nested in the specification of equation (8) when $\lambda = 0$. Thus, the null hypothesis that the continuous "restriction" on the discontinuous demand model is true lends itself readily to testing.

5.2 Discontinuous Demand Estimates

Model 1 in Table 5 reports Pseudo-maximum Likelihood (PML) estimates of the parameters in equation (8). Recall that under the nested model of continuous demand, the coefficient associated with the demand shift from shoppers is predicted to equal zero. The alternative hypothesis, predicted from the clearinghouse literature, is that this coefficient should be positive. The coefficient estimate for this effect is 0.603. Moreover, we can reject the null hypothesis of the continuous demand model in favor of the (one-sided) alternative of discontinuous demand at the 1% significance level. In short, we find considerable evidence for a demand shift when a firm offers the minimum price.

Figure 6 suggested that, were such a demand shift present, we should estimate demand as being more elastic under the continuous demand model than when one accounts for the potential demand discontinuity. Accounting for the discontinuity in demand, the estimated elasticity goes from -3.761 (in Model 2 of Table 4) to -2.459 (in Model 1 of Table 5). Translating this into gross margins, accounting for discontinuous demand changes raises the estimated gross margin of a monopoly online seller from around 27% to around 40%. This estimate seems quite reasonable in view of the 38.5% gross margin reported in the Census data described above. Turning to the effect on elasticity of increasing numbers of firms one again sees the same directional bias in the estimates of the continuous demand model compared to the discontinuous demand model. Specifically, the incremental effect of an additional firm on elasticity is reduced by around 12.5 percent (from -0.288 to -0.252) when the demand shift from "shoppers" is accounted for. Taken together, the coefficient estimates on log total price and log total price $\times$ number of listings are consistent with the effect on demand illustrated in Figure 6: The continuous demand model tends to yield more elastic demand estimates than the discontinuous demand model when a demand shift is present. We also note that, in contrast to the continuous demand specification, the effect of a change in the number of firms on a firm's overall demand $\left(\partial \ln E\left[Q_{ijt}|X\right]/\partial n_{jt}|_{\bar{p}_{ijt}}\right)$ is not statistically different from zero ($p = .4674$).

It is of some interest to note the economic relevance of our estimate of λ (0.603). Other things equal, a firm that sets lowest price in the market enjoys a 60.3 percent increase in demand, compared to the case where its price is not the lowest price. At the individual firm level, λ may be interpreted as the ratio of the number of shoppers to the number of consumers loyal to a particular firm. Thus, for every 100 consumers loyal to a particular firm, the representative firm gains an additional 60 shoppers when it sets the lowest price.

In contrast, notice that the position on screen coefficient is only -0.175. This implies that a firm would have to move up in its screen position more than 3 positions to generate the same demand increase that results by setting the lowest price in the market. Finally it is worth noting that, accounting for discontinuous demand, the effect of being a bricks and clicks retailer is statistically significant at the 5 percent level. The demand for a bricks and clicks retailer is about 32.1 percent higher than that of a firm that only has an online presence.

One may use the estimate of λ to obtain a crude gauge of the fraction of consumers using the Kelkoo site who are shoppers. Note that the total number of clicks for product j on a given date is

$$\sum_{i=1}^{n_{jt}} Q_{ijt} = \sum_{i=1}^{n_{jt}} \left(\alpha_{ijt}^{L} + \frac{\mathbf{I}_{jt}}{\# B_{jt}} \alpha_{ijt}^{S} \right) p_{ijt}^{-\theta_{jt}}$$

while the corresponding number of clicks stemming from shoppers is

$$\sum_{i=1}^{n_{jt}} Q_{ijt}^{S} = \frac{\mathbf{1}}{\# B_{jt}} \sum_{i \in B_{jt}} \alpha_{ijt}^{S} p_{ijt}^{-\theta_{jt}}$$

Hence, shoppers as a fraction of all consumers is given by

$$\begin{aligned}
\frac{S}{S+L} &= \frac{\sum_{i=1}^{n_{jt}} Q_{ijt}^{S}}{\sum_{i=1}^{n_{jt}} Q_{ijt}} \\
&= \frac{\mathbf{1}}{\# B_{jt}} \frac{\sum_{i \in B_{jt}} \alpha_{ijt}^{S} p_{ijt}^{-\theta_{jt}}}{\sum_{i=1}^{n_{jt}} \left(\alpha_{ijt}^{L} + \frac{\mathbf{I}_{jt}}{\# B_{jt}} \alpha_{ijt}^{S} \right) p_{ijt}^{-\theta_{jt}}} \\
&= \frac{\mathbf{1}}{\# B_{jt}} \frac{\sum_{i \in B_{jt}} a^{S} \alpha_{ijt}\left(X_{ijt}\right) p_{ijt}^{-\theta_{jt}}}{\sum_{i=1}^{n_{jt}} a^{L} \alpha_{ijt}\left(X_{ijt}\right) p_{ijt}^{-\theta_{jt}} + \frac{\mathbf{1}}{\# B_{jt}} \sum_{i \in B_{jt}} a^{S} \alpha_{ijt}\left(X_{ijt}\right) p_{ijt}^{-\theta_{jt}}}
\end{aligned}$$

Imposing symmetry across firms (so that all of the above terms are independent of i), one obtains

$$\begin{aligned}
\frac{S}{S+L} &= \frac{\mathbf{1}}{\# B_{jt}} \frac{\sum_{i \in B_{jt}} a^{S} \alpha_{jt}\left(X_{jt}\right) p_{jt}^{-\theta_{jt}}}{\sum_{i=1}^{n_{jt}} a^{L} \alpha_{jt}\left(X_{jt}\right) p_{jt}^{-\theta_{jt}} + \frac{\mathbf{1}}{\# B_{jt}} \sum_{i \in B_{jt}} a^{S} \alpha_{jt}\left(X_{jt}\right) p_{jt}^{-\theta_{jt}}} \\
&= \frac{a^{S} \alpha_{jt}\left(X_{jt}\right) p_{jt}^{-\theta_{jt}}}{n_{jt} a^{L} \alpha_{jt}\left(X_{jt}\right) p_{jt}^{-\theta_{jt}} + a^{S} \alpha_{jt}\left(X_{jt}\right) p_{jt}^{-\theta_{jt}}} \\
&= \frac{\lambda}{n_{jt} + \lambda}
\end{aligned}$$

which implies (given the estimate of $\lambda = .603$ reported in Model 1 of Table 5 and the mean number of listings (4.05) in our data) that about 13 percent of consumers at Kelkoo are Shoppers. While the symmetry assumptions used to obtain this crude estimate are at odds with the data (among other things, the estimates suggest that bricks-and-clicks sellers receive 32.1 percent more clicks than pure online sellers), it nonetheless illustrates that even in online markets where only 13 percent of the consumers are "shoppers," the discontinuity arising from these consumers can significantly impact elasticity estimates. Indeed, a firm that reduces its price just a penny below that of its rivals enjoys a 60 percent increase in demand and an extremely "elastic" overall response.

The results reported for Model 1 in Table 5, like those presented earlier for the continuous demand specifications, suffer from a number of potential problems. We conclude by addressing some of these concerns.

Unobserved Firm Heterogeneity. One potential shortcoming of the PML approach used in Model 1 of Table 5 is that the specification presumes there is no unobserved heterogeneity across firms. While we have attempted to control for differences across firms that stem from their having different online and offline presences, as well as different screen locations, it is still possible that a particular firm's demand is also driven by unobserved factors. For this reason, we also report in Table 5 results that allow for the effects of unobserved firm heterogeneity.

Model 2 in Table 5 reports maximum likelihood estimates of the discontinuous demand model based on the random effects specification for unobserved firm heterogeneity pioneered by Hausman, Hall and Griliches (1984), while Model 3 reports conditional maximum likelihood estimates based on a fixed effects specification for unobserved firm heterogeneity. Note that these results require the specification of the actual likelihood function, which we have take to be Poisson. However, Table A2 in the appendix shows that the results reported in Table 5 and discussed below are similar if one uses the likelihood function for a negative binomial (2) specification.

Notice that, in both the random effects (Model 2) and fixed effects (Model 3) specifications, the coefficients of interest are roughly comparable to those obtained ignoring potential unobserved heterogeneity (Model 1). The continuous demand model is once again nested (and rejected) in these specifications. Further, the economic value of the coefficient associated with the demand shift is largely unchanged by allowing for potential unobserved heterogeneity. Likewise, the coefficient associated with the elasticity of demand for a monopoly firm remains at about -2.5, similar to the estimate obtained in Model 1.

In contrast, the coefficient associated with the marginal effect on price elasticity of changing the number of rivals is reduced in Models 2 and 3 compared to Model 1. One possibility is that some firms tend to sell in markets where there are a large number of rivals while other firms tend to sell products where there are only a small number of rivals. Allowing for unobserved firm heterogeneity soaks up this variation which was previously ignored in Model 1. Nonetheless, the results continue to suggest that it is important to account for the degree of rivalry in online markets: firms that face more rivals continue to face more elastic demand.

Endogeneity. Another potential concern, which we alluded to earlier, is that the number of firms listing prices may be endogenous —firms may be more eager to list prices for products where consumer demand is high, and thus the number of firm effects documented earlier might stem from spurious correlation. To examine this possibility, we used the test procedure suggested by Wooldridge (1997); cf. Terza, (1998). The idea is to obtain instruments that are correlated with the number of listings in our data but uncorrelated with the number of clicks enjoyed by a particular firm. One then uses standard techniques to regress the number of listings on the instruments and the remaining control variables, compute the resulting residuals from this regression, and then include these residuals as an *additional* explanatory variable in the maximum likelihood estimation of the underlying count model. The test is a simple LM test, which is conducted after estimating the model under the null assumption that the coefficients on these residual terms is zero.

As is the case with any endogeneity test, the power of this test depends on the availability of good instruments. For this reason, we collected additional data from the U.S. which we believe is likely to satisfy these conditions. In particular, for each of the PDAs in our UK sample and for each date, we obtained data on that PDAs product popularity rating from Shopper.com, a US price comparison site.[11] This product popularity ranking is based on

[11] The interested reader should consult Baye, Morgan, and Scholten (2004a) for additional details con-

the lagged number of U.S. clickthroughs on the U.S. price comparison site. It seems likely that the product popularity of an identical model PDA in the US is correlated with its popularity in the UK. However, it is not likely to be correlated with the actual number of clicks that particular U.K. sellers listing at the Kelkoo site received on any given date. Among other things, it would be unusual for a US consumer using Shopper.com to cross the "virtual" border and shop at UK price comparison sites such as Kelkoo. Indeed, as Table 1 showed, there would be little incentive for a US consumer to do so during the period of our study. Since the effect of number of firms listing prices enters both directly as well as through an interaction term with $ln(p_{ijt})$ in the specification given in equation (8), we used US product rank and product rank squared as instruments. Residuals were obtained based on simultaneous estimation to obtain residuals for both number of firms and the interaction term.

Table 5 reports the results of the endogeneity tests for Models 1 through Model 3. In all cases, there appears to be little evidence against the null hypothesis that the number of firms is not endogenous. In short, it does not appear that the estimated effect of the number of rival firms on an online seller's elasticity of demand is driven by an endogeneity problem or unobserved heterogeneity.

6 Conclusions

We have developed identification restrictions which enable us to estimate demand elasticities in online markets using only clicks data. We then showed that one can use pseudo-maximum likelihood methods on these count data to obtain consistent estimates of an online seller's demand. Applying there methods to a unique data set for 18 models of PDAs sold at Yahoo!'s European price comparison service, Kelkoo, we report estimates of the elasticity of demand facing online retailers in the UK, as well as estimates of several other key determinants of online demand.

In addition to providing a theoretical basis for demand estimation using clicks data, there are three main messages that emerge from our analysis. First, in estimating an online

cerning data from this site.

retailer's elasticity of demand, it appears to be important to control for variation in the number of rival firms in the relevant online market. Our results suggest that a monopoly seller faces an elasticity of demand of about -2.5, while a firm's demand becomes increasing elastic as the number of rival sellers increases. In the most competitive markets we analyzed (15 sellers), the elasticity of demand for a representative firm's product is about -6.0. This finding may partially explain the wide array of elasticity estimates obtained in online markets for books and computer memory. One potential explanation for these diverse estimates is differences in the number of firms—i.e., the rivalry of the industry: the online market for books is considerably more concentrated than the online retail market for computer memory.

Second, our results provide some support for clearinghouse models that have been widely used to model online competition. These models predict a "jump" in a firm's demand when it offers the lowest price. Indeed, our results indicate that such a jump is present in the data and that the economic impact of the jump is significant. In the UK online market for PDAs, a firm offering the lowest price enjoys a 60 percent increase in demand compared to what it would have enjoyed had it not charged the lowest price. Moreover, we show that failing to account for the jump in demand leads to biased elasticity estimates: Elasticity estimates in the continuous demand specification are almost twice those obtained allowing for demand discontinuities. Interestingly, we showed that estimates of the size of the "jump" may be used to obtain a crude estimate of the fraction of price sensitive "shoppers" in the market. For the case of the PDA market at Kelkoo, approximately 13 percent of online consumers are shoppers who purchase at the lowest price, while it appears that 87 percent of the consumers are loyal to particular sellers. Our results thus suggest that even a relatively small number of "shoppers" can result in sizable jumps in firm-level demand and a significant bias (if ignored) in elasticity estimates.

Finally, we have identified several other potentially important determinants of online demand. It appears that online consumers in the UK favor firms that have both an online as well as a brick and mortar presence: These firms enjoy about 30 percent more sales than their pure online competitors. Our results also suggest that UK consumers are clicking more often at work than at home: online sales are about 27 percent lower on weekends than during the week. We speculate that this is an artifact of the low (relative to the US) broadband penetration at residences, and probably overstates the weekend effect present in US data and probably in the UK data going forward. Finally, as is the case in the physical marketplace, location matters in the online world. We find that firms listed at the top of the screen of price quotes systematically receive more clicks than firms listed further down the screen. Indeed, we estimate that being shifted down one position on the screen costs a firm about 15% of its clicks, even after controlling for price and other characteristics.

All of the results discussed above are robust to alternative assumptions regarding the specification of likelihood functions, controls for potential unobserved firm heterogeneity, as well as endogeneity. Nonetheless, we stress that our interpretation of these results as representing demand rather than clicks estimation critically depends on the assumption that the identification restriction in Proposition 1 holds. Roughly, this requires that the rate at which clicks are converted into final sales are, on average, independent of the underlying price and other determinants observable at the Kelkoo site generating a click in the first place.

While we believe these identification assumptions are reasonable, we conclude by noting that our results are of economic interest even if the identifying restrictions do not hold. The business models of many of the most successful online firms such as Google and Yahoo! are built on revenues derived from *clicks* (not from the conversion of clicks into sales). Thus, understanding the determinants of click behavior in the online marketplace is arguably as important as understanding the demand facing individual firms selling in the online marketplace.

References

[1] Baye, Michael R. and John Morgan, "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets," *American Economic Review* (2001) 91(3), pp. 454-474.

[2] Baye, Michael R., John Morgan, and Patrick Scholten, "Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site," *Journal of Industrial Economics* (2004a) 52, pp. 463-496.

[3] Baye, Michael R., John Morgan, and Patrick Scholten, "Temporal Price Dispersion: Evidence from an Online Consumer Electronics Market," *Journal of Interactive Marketing* (2004b) 18, pp. 101-115.

[4] Brynjolfsson, Erik and Michael D. Smith, "The Great Equalizer? Consumer Choice Behavior at Internet Shopbots." MIT Working Paper, 2000.

[5] Chevalier, J. and A. Goolsbee, "Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com," *Quantitative Marketing and Economics*, (2003) 1(2), pp. 203 - 222.

[6] Cameron, A. and P. Trivedi, "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators," *Journal of Applied Econometrics* (1986) 1, pp. 29-53.

[7] Cameron, A. and P. Trivedi, "Regression-based tests for overdispersion in the Poisson model," *Journal of Econometrics* (1990) 46(3), pp. 347-364.

[8] Cameron, A. and P. Trivedi, *Regression Analysis of Count Data,* Econometric Society Monograph No. 30 (1998).

[9] Ellison, Glenn and Sara Fisher Ellison, "Search, Obfuscation, and Price Elasticities on the Internet." *mimeo,* 2004.

[10] Ghose, A., M. Smith, and R. Telang, "Internet Exchanges for Used Books: An Empirical Analysis of Welfare Implications," *mimeo*, 2004.

[11] Gourieroux, Christian, A. Monfort, and A. Trognon, "Pseudo Maximum Likelihood Methods: Theory," *Econometrica* (1984a) 52(3), pp. 681-700.

[12] Gourieroux, Christian, A. Monfort, and A. Trognon, "Pseudo Maximum Likelihood Methods: Applications to Poisson Models," *Econometrica* (1984b) 52(3), pp. 701-20.

[13] J. Hausman, B. Hall, and Z. Griliches, "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica* (1984) 52, pp. 909-938.

[14] Hall, B. and R. Ziedonis, "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-1995," *RAND Journal of Economics* (2001) 32(1), pp. 101-128.

[15] Huber, P., "The Behavior of Maximum Likelihood Estimates under Non-standard Conditions," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (1967) 1, pp. 221-233.

[16] Narasimhan, Chakravarthi, "Competitive Promotional Strategies." *Journal of Business*, (1988) 61, pp. 427-449.

[17] Rogers, W. "Regression Standard Errors in Clustered Samples," *Stata Technical Bulletin* (1993) 13, pp. 19-23.

[18] Rosenthal, Robert W., "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price," *Econometrica* (1980) 48(6), pp. 1575-1580.

[19] Shilony, Yuval, "Mixed Pricing in Oligopoly," *Journal of Economic Theory* (1977) 14, pp. 373-388.

[20] Terza, J., "Estimating Count Data Models with Endogenous Switching: Sample Selection and Endogenous Switching Effects," *Journal of Econometrics* (1998) 84, pp. 129-137.

[21] Varian, Hal, "A Model of Sales," *American Economic Review* (1980) 70, pp. 651-659.

[22] White, H. "A Heteroskedasticty-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* (1980) 48, pp. 817-830.

[23] White, H., “Maximum Likelihood Estimation for Misspecified Models,” *Econometrica* (1982) 50, pp. 1-25.

[24] Wooldridge, J., “Quasi-likelihood Methods for Count Data,” in *Handbook of Applied Econometrics, Vol. 2* (M. Pesaran and P. Schmidt, eds.) Oxford: Blackwell (1997), pp. 352-406.

Figure 1: Kelkoo Screenshot

Product	Shop	Price	Total Price inc P & P	More
HP IPAQ H5550 COLOUR POCKET PC	PC WORLD Shop info	£429.98 + P & P: £3.95	£433.93	More
HP iPAQ H5550 Pocket PC	[illegible] Shop info	£417.95 + P & P: Free	£417.95	More
HP iPAQ Pocket PC h5550 - RAM: 128 MB - ROM: 48 MB - Windows Mobile 2003 - display 3.8" TFT	dabs.com Shop info	£464.12 + P & P: £5.88	£470.00	More
FA107A HP iPAQ H5550	BigGreyCat Shop info	£486.45 + P & P: £5.20	£491.65	More
IPAQ 5550 INTEL 400MHZ 48MB ROM 128MB 3.8 TFT 802.1X SDIO BIOME	SiMPLY computers Shop info	£504.00 + P & P: £5.29	£509.29	More
HP iPAQ h5550 Pocket PC (English)	expansys Shop info	£382.71 + P & P: £6.98	£389.69	More
iPAQ H5550 400MHz Bluetooth FA107A#ABU	Technoworld Shop info	£491.83 + P & P: £8.17	£500.00	More
iPAQ H5550/EN 400MHz 48MB128MB Bluetooth	easyetrader Shop info	£551.63 + P & P: Free	£551.63	More
HEWLETT PACKARD IPAQ H5550	COMET comet.co.uk Shop info	£429.98 + P & P: Free	£429.98	More
HP iPAQ H5550 Pocket PC (English)	amazon.co.uk Shop info	£429.99 + P & P: Free	£429.99	More
Hewlett Packard iPAQ 5550 Pocket PC	C&A Electronics Shop info	£424.99 + P & P: £6.99	£431.98	More
HP iPAQ H5550 New	totalpda Shop info	£445.33 + P & P: £17.63	£462.95	More

Figure 2: Average Number of Leads by Day of Week

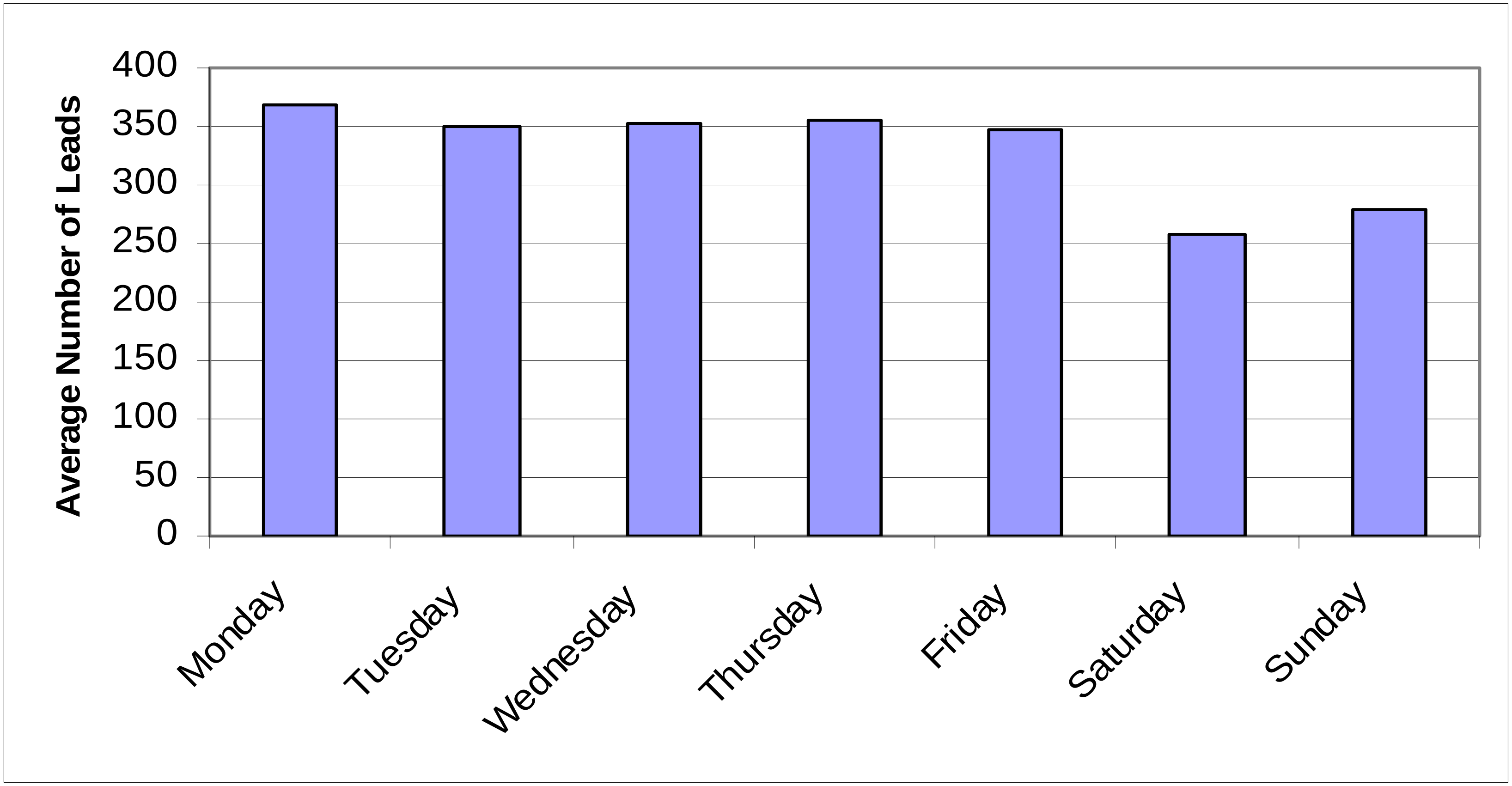

Figure 3: Histogram of Leads by Price Rank and Screen Location

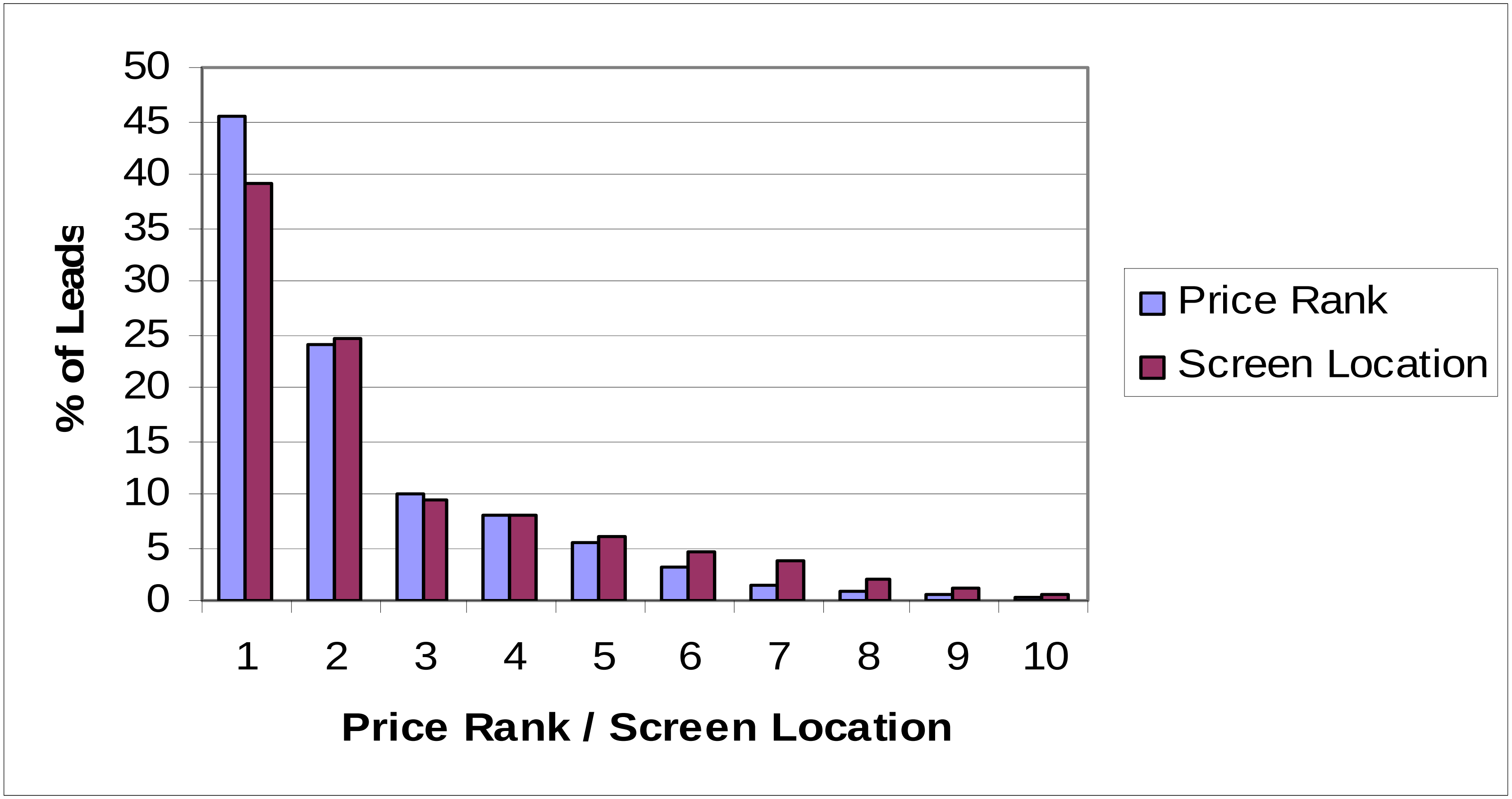

Figure 4: Factors Influencing Numbers of Leads

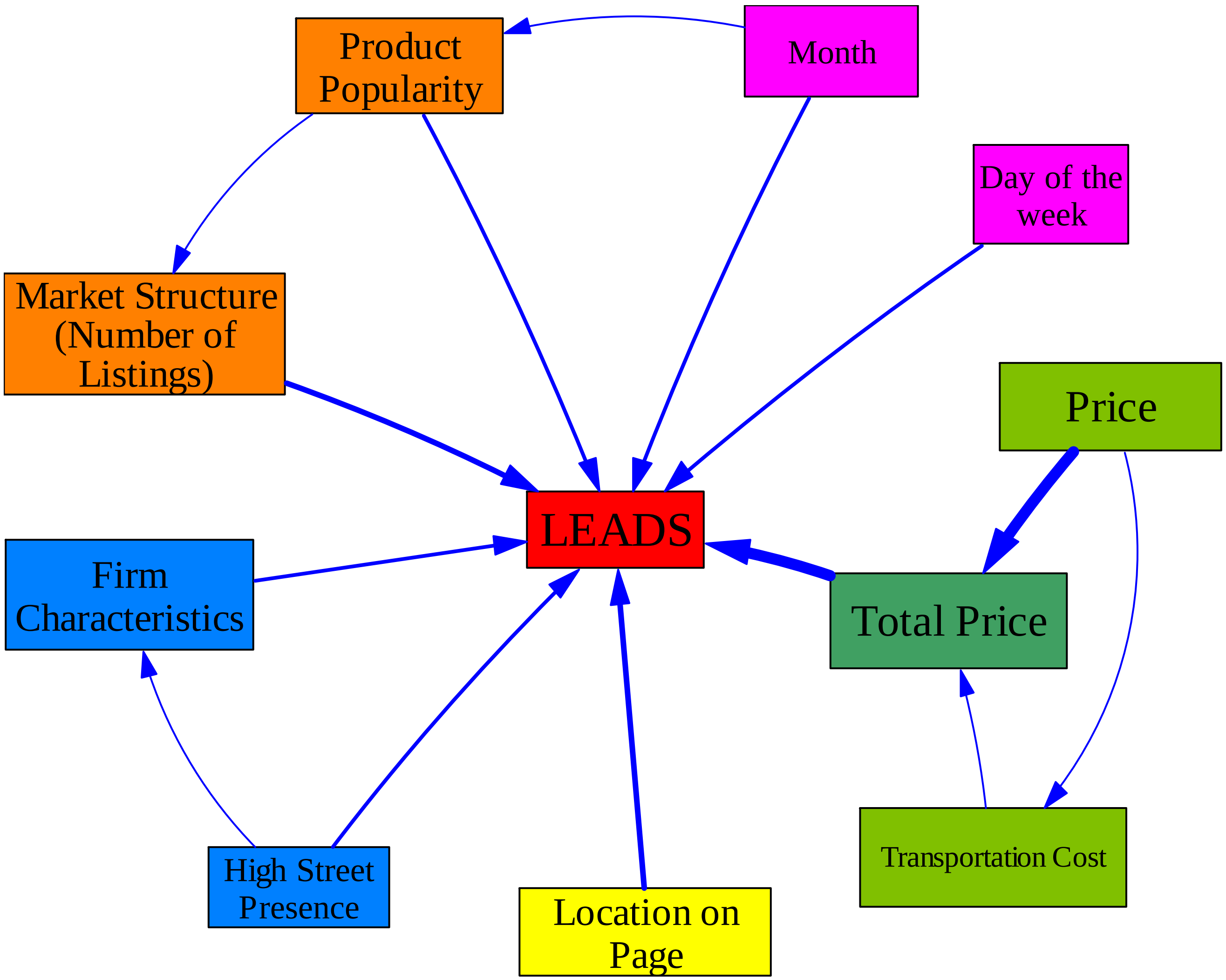

Figure 5: Estimated Demand Elasticity and Number of Firms

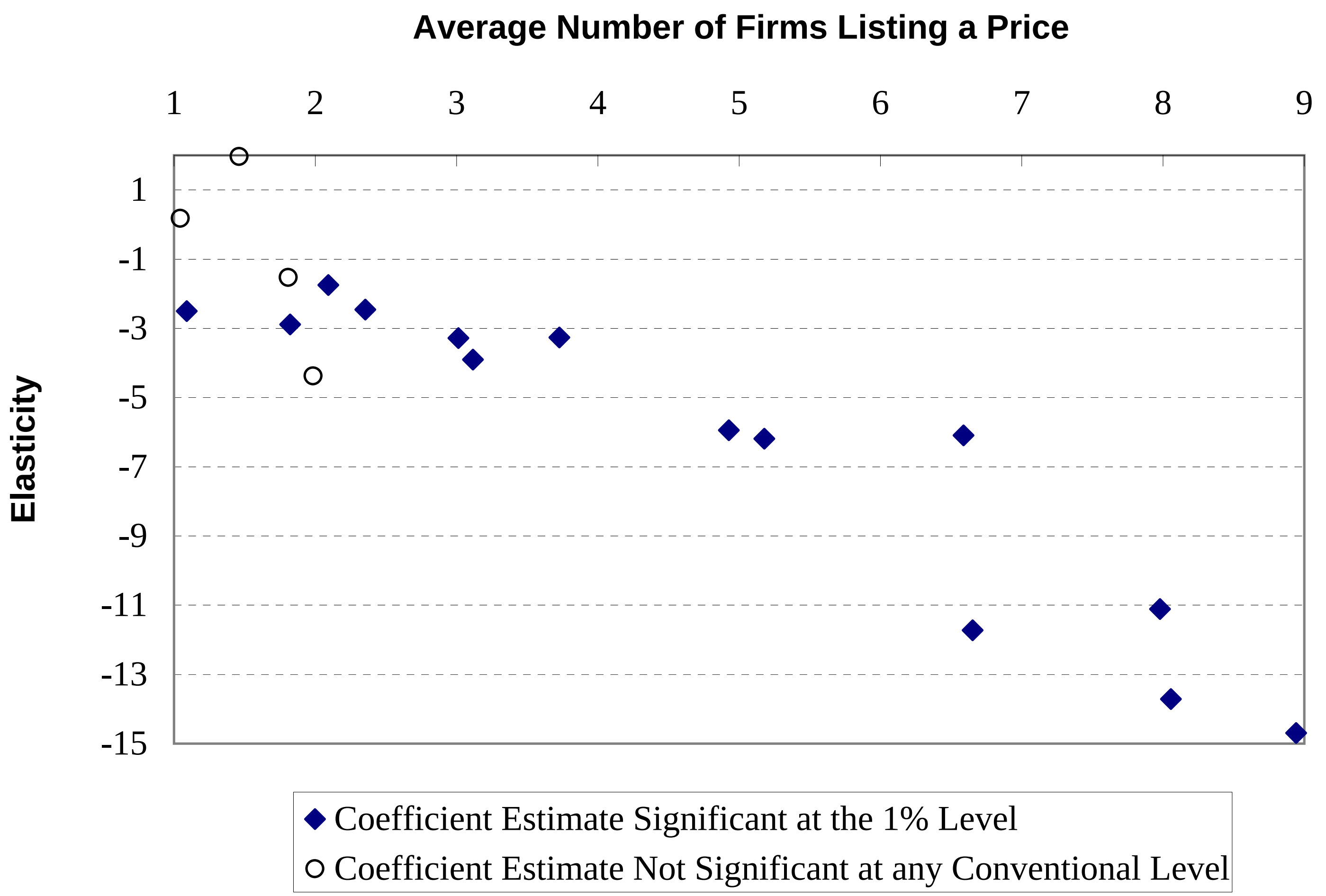

Figure 6: Misspecification from using Continuous Demand Model in Discontinuous Demand Setting

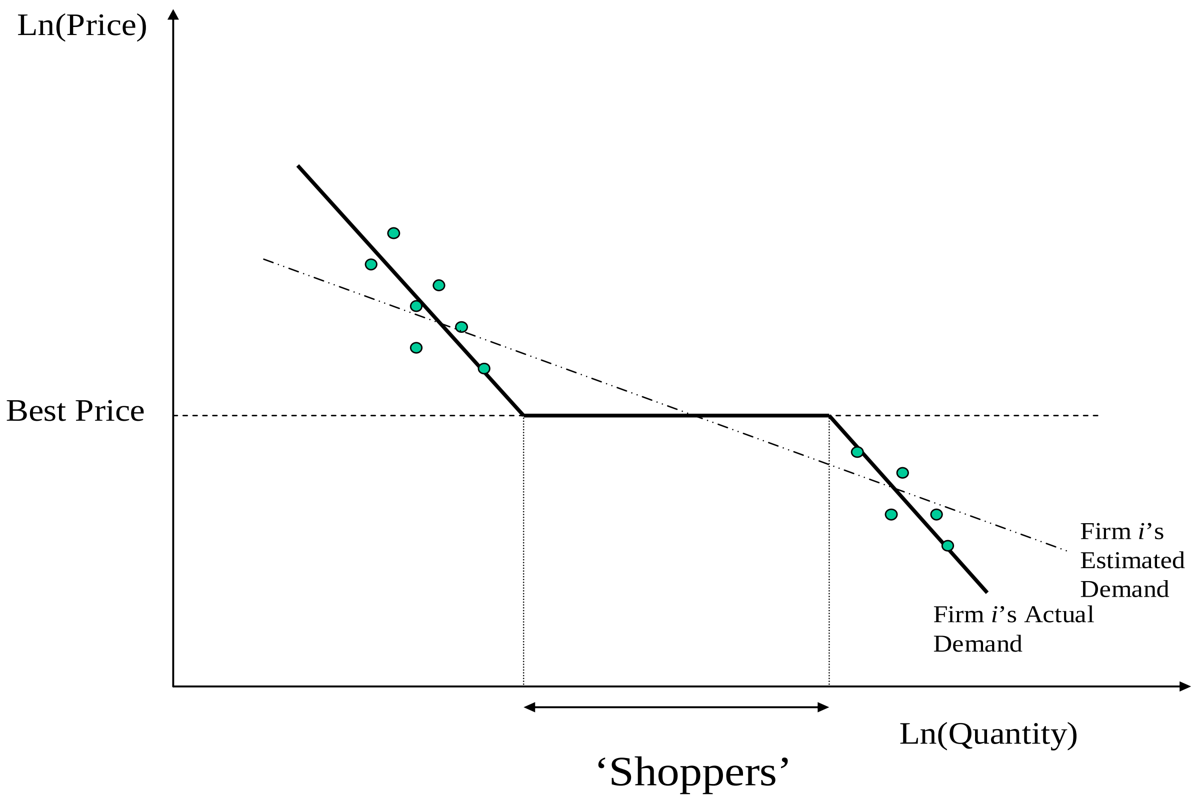

Table 1: Descriptive Statistics-UK and US Price Level Comparisons

	Median Price (£)		Mean # of Listings		Coefficient of Variation	
	UK	US	UK	US	UK	US
Handspring Treo 90	130	95	1.94	22.23	0.25	0.10
HP Compaq IPAQ 1910	216	N/A	2.63	N/A	0.11	N/A
HP Compaq IPAQ 1940	267	174	8.09	27.20	0.08	0.07
HP Compaq IPAQ 2210	319	228	6.57	27.95	0.07	0.08
HP Compaq IPAQ 3950	279	172	1.30	4.95	0.01	0.13
HP Compaq IPAQ 3970	317	N/A	1.35	N/A	0.01	N/A
HP Compaq IPAQ 5550	452	372	7.67	25.99	0.05	0.08
Palm m515	195	191	1.05	6.56	0.01	0.20
Palm Tungsten T2	269	198	6.11	31.79	0.07	0.12
Palm Tungsten W	390	229	2.66	27.72	0.08	0.14
Palm Zire 71	200	167	7.21	31.93	0.07	0.10
Sony Clie NX70V	283	272	2.13	15.58	0.21	0.25
Sony Clie NX73V	381	245	4.51	31.58	0.08	0.09
Sony Clie NZ90	537	435	1.51	27.35	0.03	0.09
Sony Clie SJ22	147	87	3.32	26.51	0.09	0.10
Sony Clie SJ33	171	121	1.02	10.94	0.00	0.10
Sony Clie TG50	269	174	3.86	32.72	0.06	0.10
Toshiba E740 WIFI	435	195	1.95	12.40	0.20	0.27
Overall	305	206	3.87	24.36	0.09	0.12

Note: Median price denotes the average over all dates of the median price on each date. US prices are denominated in £ at the daily USD/£ exchange rate. Coefficients of variation include listings in which there is only a single firm. If we omit single firm listings, unweighted overall coefficient of variation in the UK increases to 11% and in the US to 12.1%.

Table 2: Summary Statistics

Variable Name	Mean	Standard Deviation	Median	First Quartile	Third Quartile	Maximum	Minimum
Clicks	3.33	4.27	2	0	5	36	0
Price	304.88	106.84	279.98	229.99	396.63	601.95	104.57
Shipping	4.16	4.50	3.95	0	5.82	17.63	0
Total Price	309.04	107.01	283.94	234.42	396.63	607.77	108.10
Number of listings of the product	4.05	2.93	3	2	6	15	1
Location on Screen	3.40	2.43	3	1	5	15	1
D(Bricks and Clicks Retailer = 1)	0.29						
D(Weekend =1)	0.28						
D(Month = September)	0.11						
D(Month = October)	0.29						
D(Month = November)	0.29						
D(Month = December)	0.27						
D(Month = January)	0.05						

Total Number of Products = 18
Total Number of Firms = 19
Total Number of Dates = 111
Total Number of Observations = 6151

Table 3: Product Specific Demand Estimates

Product	Log total price	Position on Screen	Weekend	Month Dummies	# of Obs.	Average Number of Firms
	Likelihood Specification for Clicks: Poisson PML					
Toshiba E740 WIFI	-1.75 (8.64)**	0.272 (3.23)**	-0.214 (2.35)*	4	216	2.093
HP Compaq IPAQ 1910	-3.281 (5.68)**	-0.591 (4.73)**	-0.215 (2.29)*	2	171	3.012
HP Compaq IPAQ 1940	-14.691 (20.39)**	-0.165 (13.98)**	-0.255 (4.45)**	4	898	8.942
HP Compaq IPAQ 2210	-11.725 (10.54)**	-0.058 (2.04)*	-0.251 (2.43)*	1	184	6.652
HP Compaq IPAQ 3950	1.961 (1.56)	-0.351 (1.02)	-0.152 (0.62)	3	91	1.462
HP Compaq IPAQ 3970	-1.53 (1.91)	-0.262 (3.10)**	-0.12 (1.14)	4	131	1.809
HP Compaq IPAQ 5550	-13.712 (22.97)**	-0.153 (13.92)**	-0.288 (5.17)**	4	851	8.055
Palm m515	-2.503 (3.88)**	-0.458 (0.99)	-0.444 (2.82)**	2	44	1.091
Sony Clie NX70V	-2.455 (9.41)**	-0.227 (2.99)**	-0.116 (0.91)	3	164	2.354
Sony Clie NX73V	-5.941 (10.82)**	-0.258 (7.18)**	-0.163 (1.73)	4	501	4.928
Sony Clie NZ90	-2.884 (1.51)	-0.144 (0.82)	-0.331 (1.60)	4	151	1.821
Sony Clie SJ22	-3.263 (8.65)**	-0.085 (3.04)**	-0.278 (3.54)**	4	368	3.728
Sony Clie SJ33	0.182 (0.08)		-0.215 (1.51)	2	44	1.045
Sony Clie TG50	-6.188 (6.28)**	-0.049 (1.22)	-0.202 (1.87)	4	428	5.178
Handspring Treo 90	-4.375 (1.67)	-0.723 (0.79)	-0.225 (2.92)**	2	136	1.985
Palm Tungsten T2	-6.096 (11.90)**	-0.153 (6.30)**	-0.265 (3.04)**	4	678	6.587
Palm Tungsten W	-3.902 (4.37)**	-0.328 (4.08)**	-0.406 (2.30)*	4	295	3.115
Palm Zire 71	-11.115 (11.47)**	-0.157 (7.71)**	-0.316 (3.65)**	4	800	7.978

Note: Robust z statistics in parentheses. * Significant at 5%. ** Significant at 1%.

Table 4: Continuous Demand Estimates

		Model 1	Model 2
Likelihood Specification for Clicks		Poisson PML	Poisson PML
Log Total Price		-4.61	-3.761
		(8.91)**	(7.45)**
Log Total Price x Number of Listings			-0.288
			(4.14)**
Number of listings of the product			1.593
			(4.05)**
Position on Screen		-0.186	-0.175
		(4.54)**	(4.47)**
Bricks and Clicks Retailer		0.262	0.236
		(1.58)	(1.67)
Weekend		-0.242	-0.265
		(10.82)**	(11.46)**
Product Dummies		17	17
Month Dummies		4	4
Product x Month Dummies		55	55
Robust Standard Errors Clustered by Firm		Yes	Yes
Observations		6151	6151
z statistics in parentheses			
* significant at 5%; ** significant at 1%			
Overdispersion Test			
	Chi-Square	2656.46	2310.09
	P-Value	0	0

Table 5: Discontinuous Demand Estimates

Likelihood Specification for Clicks		Model 1 Poisson PML	Model 2 Poisson ML	Model 3 Poisson CML
Log Total Price		-2.459	-2.487	-2.49
		(9.11)**	(24.78)**	(24.74)**
Log Total Price x Number of Listings		-0.252	-0.146	-0.145
		(4.60)**	(8.82)**	(8.71)**
Demand Shift from Shoppers		0.603	0.62	0.621
		(7.11)**	(27.84)**	(27.80)**
Number of listings of the product		1.415	0.833	0.824
		(4.52)**	(8.94)**	(8.84)**
Position on Screen		-0.175	-0.153	-0.152
		(4.37)**	(22.08)**	(21.81)**
Bricks and Clicks Retailer		0.321	0.373	
		(2.41)*	(1.87)	
Weekend		-0.268	-0.26	-0.26
		(13.79)**	(15.51)**	(15.50)**
Product Dummies		17	17	17
Month Dummies		4	4	4
Product x Month Dummies		55	55	55
Controls for Unobserved Firm Heterogeneity		No	19 - Random Effects	19 - Fixed Effects
Robust Standard Errors Clustered by Firm		Yes	No	No
Observations		6151	6151	6151
z statistics in parentheses				
* significant at 5%; ** significant at 1%				
Overdispersion Test				
	Chi-Square	1818.99		
	P-Value	0		
Endogeneity Test				
	Chi-Square	2.64	1.44	1.45
	P-Value	0.267	0.488	0.485

Table A1: Continuous Demand Estimates - Alternative Specifications

	Model 1	Model 2
Likelihood Specification for Clicks	Negative Binomial ML	Negative Binomial ML
Log Total Price	-4.81	-3.696
	(10.29)**	(8.66)**
Log Total Price x Number of Listings		-0.343
		(5.54)**
Number of listings of the product		1.897
		(5.37)**
Position on Screen	-0.178	-0.166
	(4.70)**	(4.46)**
Bricks and Clicks Retailer	0.316	0.272
	(2.26)*	(2.23)*
Weekend	-0.263	-0.288
	(11.62)**	(13.42)**
Product Dummies	17	17
Month Dummies	4	4
Product x Month Dummies	55	55
Robust Standard Errors Clustered by Firm	Yes	Yes
Observations	6151	6151

z statistics in parentheses

* significant at 5%; ** significant at 1%

Table A2: Discontinuous Demand Estimates - Alternative Specifications

	Model 1	Model 2	Model 3
Likelihood Specification for Clicks	Negative Binomial ML	Negative Binomial ML	Negative Binomial CML
Log Total Price	-2.343	-2.372	-2.359
	(8.18)**	(17.31)**	(17.12)**
Log Total Price x Number of Listings	-0.314	-0.139	-0.137
	(5.38)**	(6.28)**	(6.18)**
Demand Shift from Shoppers	0.619	0.594	0.597
	(8.24)**	(19.49)**	(19.56)**
Number of listings of the product	1.77	0.782	0.774
	(5.30)**	(6.31)**	(6.21)**
Position on Screen	-0.166	-0.146	-0.144
	(4.31)**	(15.92)**	(15.36)**
Bricks and Clicks Retailer	0.324	0.266	
	(2.66)**	(2.81)**	
Weekend	-0.29	-0.246	-0.245
	(14.86)**	(10.98)**	(10.97)**
Product Dummies	17	17	17
Month Dummies	4	4	4
Product x Month Dummies	55	55	55
Controls for Unobserved Firm Heterogeneity	No	19 - Random Effects	19 - Fixed Effects
Robust Standard Errors Clustered by Firm	Yes	No	No
Observations	6151	6151	6151

z statistics in parentheses
* significant at 5%; ** significant at 1%

Endogeneity Test	Model 1	Model 2	Model 3
Chi-Square	4.67	0.55	0.64
P-Value	0.0967	0.7578	0.7261