Clicks, Discontinuities, and Firm Demand Online

Michael R. Baye, Indiana University

J. Rupert J. Gatti, University of Cambridge

Paul Kattuman, University of Cambridge

John Morgan, University of California, Berkeley

November 2006

Abstract

The market values of online platforms, such as Yahoo, stem from their ability to monetize the clicks they generate for firms advertising on their sites. We exploit a unique dataset on clicks from one of Yahoo's price comparison sites to estimate the determinants of clicks received by online retailers. We find that a firm enjoys a 60% jump in its clicks when it offers the lowest price at the site. This discontinuity is consistent with a variety of models that have been used to rationalize the price dispersion observed in online markets. We also show that one may use estimates of the determinants of a firm's clicks to obtain bounds on its underlying demand parameters, including own- and cross-price elasticities. Our results have potentially significant ramifications for online retailers, platforms, and policymakers: Failure to account for discontinuities distorts parameter estimates by 50 to 100 percent.

JEL Classification Numbers: D4, D8, M3, L14

1   Introduction

Clicks are the currency of the realm in online markets. Google derives nearly all of its revenues from clicks, and now boasts a market capitalization of over \$140 billion (as of 11/1/06). Google is not alone in benefiting from this business model; the largest online portal, Yahoo!, derived 87% of its \$5.3 billion in total revenues from pageviews in 2005. Clicks are not only the lifeblood of search providers and portals, they also represent a significant (and often essential) marketing cost to firms competing in the online channel. For example, the largest price listing service in the world (Yahoo's European-based Kelkoo site) charges retailers fees ranging from 40 cents to \$1.90 per click and attracts over 10 million consumers per month in the UK alone. Price comparison sites in the US, including CNet and Pricewatch, charge online retailers similar clickthrough fees and have conversion rates of about 5% in 2002.[1] At a price of \$1 per click, this means that an online retailer pays the platform an average of \$20 to generate a single sale.

At price comparison sites, it is widely recognized that a particular seller's price has a lot to do with its clickthroughs and that having the "low product price is key."[2] But how "key" is it to have the low price? And how do other factors, such as screen location, firm characteristics, the number of rivals and their prices affect its clicks? The answers to these questions are important to a variety of stakeholders including platforms, investors, and the retail firms themselves.

The extant literature on clearinghouse models offers guidance on how "key" it is to have the low price.[3] These models postulate that a mass of consumers, sometimes referred to as "shoppers," consult the complete list of prices displayed on the platform and buy exclusively from the firm offering the lowest price. Other consumers, who may be viewed as "loyal" or simply uninformed, will purchase from a firm even if it does not offer the lowest price. Clearinghouse models have two important implications for online markets. The first—that prices for homogeneous products are dispersed in equilibrium—has been amply

---

[1] "Comparison Search Engines Tested," http://www.marketingexperiments.com on November 14, 2006.

[2] See, for instance, http://www.marketingexperiments.com.

[3] See, for instance, Salop and Stiglitz (1977), Shilony (1977), Rosenthal (1980), Varian (1980), Narasimhan (1988), Spulber (1995), and Ju, Lin, and Zhu (2006). Baye, Morgan, and Scholten (forthcoming) provide a survey of clearinghouse and closely related search-theoretic models which include Stahl (1989), Dana (1994), Brown and Goolsbee (2002), and Janssen and Moraga-Gonzalez (2004).

documented empirically.[4] The second implication of these models is that a firm experiences a discontinuous "jump" in its demand (which corresponds to clicks in these simple models) when it offers the lowest price. Indeed, this discontinuity is a theoretical driver for price dispersion, yet, to the best of our knowledge, whether such a jump even exists—let alone its economic significance—has never been examined empirically. In terms of the relationship between clicks and prices, the size of the jump (if any) determines just how "key" it is to have the lowest price in an online market.

Using a unique dataset from Kelkoo, we estimate the size of the jump and find it is economically and statistically significant. A firm offering the best price gains 60% more clicks than if it had not charged the lowest price. In addition, we estimate that a firm's elasticity of clicks with respect to price (having accounted for the jump) is more elastic in online markets where competition is keener. We find that a monopoly seller faces an elasticity of about −2, while in the most competitive markets we analyze (15 sellers), the elasticity for a representative firm is about −6. We are also able to identify the effect of other determinants of clicks—such as screen location. Our results imply that, other things equal, a firm receives about 17% fewer clicks for every competitor listed above it on the screen.

Failing to account for the discontinuity in clicks at the lowest price generates significant distortions. For example, a firm that competes against the average number of rivals in our sample and ignores the discontinuity will erroneously conclude that a 10% increase in its price will reduce its clicks by 44%. In contrast, a firm that correctly accounts for the discontinuity realizes that a 10% price increase only reduces its clicks by 30% if it is not currently offering the lowest price in the market (or if, after the price increase, it remains the lowest price offering in the market). However, if the 10% price increase moves the firm out of the lowest price position in the market, its clicks will decrease by 90%. Thus, firms that do not take the discontinuity into account will systematically misestimate the impact of a price change on their clickthrough rates. Likewise, a platform (or financial analyst) interested in the impact of a change in the platform's fee structure that raises the prices of all firms by same percentage will overstate the impact on the platform's clicks by 45%.

---

[4] Brynjolfsson and Smith (2000b) were among the first to document this phenomenon; see Table 1b in Baye, Morgan, and Scholten (forthcoming) for a list of nearly 20 different studies that document significant levels of price dispersion in online markets in the US and abroad.

Our analysis takes into account some of the unique features of the online marketplace, including rapid changes in the number of competing firms, the identity of the low-price firm, and firms' screen locations. At price comparison sites, the number of firms listing prices for a given product, and the identity of the firm charging the lowest price, change almost daily.[5] Likewise, "virtual" real estate in the online world changes rapidly. In purchasing "adwords" (advertising space at the side of search queries on Google's site), for instance, 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. Variation in the identity of the low-priced firm enables us to disentangle the jump from other determinants of clicks. Variation in the number of competitors permits us to identify the marginal impact of the number of rivals on a firm's clicks elasticity. Finally, variation in screen locations allows us to identify the value of the virtual real estate separately from other firm characteristics.

To the extent one is interested in estimating a firm's demand rather than its number of clicks, the ideal dataset would include a list of all firms' prices at the price comparison site, their number of clicks and final sales. Unfortunately, the principal-agent problem between the owners of price-comparison sites and e-retailers means that such data generally do not exist.[6] We show that, under plausible assumptions, one may use the estimates of clicks elasticities discussed above to place bounds on the actual demand elasticities for products sold in online markets. Intuitively, if the conversion rate is independent of a firm's price, then the clicks elasticity exactly corresponds to the demand elasticity. Directional assumptions about the relationship between conversion rates and underlying variables of interest, such as a firm's price, allow one to obtain useful bounds on actual demand elasticities using only clicks data. Viewed in this context, our results also suggest that there are significant discontinuities in firms' demand functions in online markets, and that failing to account for

---

[5] See Baye, Morgan and Scholten (2004b) and Ellison and Ellison (2004). Baye and Morgan (2001) show theoretically that variation in the degree of rivalry of a given online market is essential for firms to avoid pure Bertrand competition and for the information "gatekeeper"—the entity running the price comparison site—to profitably operate.

[6] We know of only one demand study (Ellison and Ellison, 2004) that is based on actual sales data. 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' websites.

this may distort estimates of demand elasticities. Since demand elasticities play a central role in welfare analysis, we believe that our finding of significant discontinuities in online markets is potentially important for Internet-related tax policies.[7] In addition to tax policy, demand elasticities also play an important role in antitrust policy and trade policy.

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, which is based on "distribution free" pseudo-maximum likelihood count data methods. Section 4 provides estimates of the determinants of a firm's clicks under the assumption that a firm's number of clicks is a continuous function of its price. These latter estimates are nested as a special case of the discontinuous specification, which is detailed in Section 5. We show that the results discussed above are robust across a variety of specifications that attempt to account for endogeneity, unobserved heterogeneity, and continuous cross-price effects. Section 6 offers conditions under which clicks estimates may be used to obtain bounds on underlying demand parameters, such as a firm's own- and cross-price elasticities of demand. In Section 7, we briefly discuss the potential implications of our results for discrete choice models of individual consumer behavior. Finally, Section 8 discusses the potential ramifications of our results for online retailers, platforms, and policymakers.

2 Data

The proprietary data used in this paper were provided by the UK price comparison site, Kelkoo.com, which is owned by Yahoo! 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. Over 1,800 individual retailers—including 18 of the largest 20 online retailers in the UK—list prices on Kelkoo. According to Yahoo!, Kelkoo is the largest price listing service in the world, operating in seven other European countries besides the UK. It is recognized

---

[7] For instance, it is well-known that the excess burden of a tax is roughly proportional to the demand elasticity for the product. Thus, if demand was estimated to be twice as elastic as the true elasticity, then the evaluation of the excess burden of the tax would be overstated by 100%.

as one of the six most accessed websites in all of Europe.[8]

Consumers interested in purchasing a broad range of products access the Kelkoo site to obtain information about the product and/or to obtain a list of retailers selling the product, together with the prices charged and other relevant information such as shipping charges. Consumers interested in making a purchase must do so from the website of the specific retailer and may easily transfer from the Kelkoo site to a retailer's site by clicking on one of the links provided.

Kelkoo's revenue is generated by charging retailers a fee for each referral made—that is, each time a consumer transfers from the Kelkoo site to a retailer's site. The fees charged vary across products and retailers, but typically range from £0.20 to £1.00 per click. Kelkoo does not charge consumers any fees for using its site.

The products in our dataset consist of 18 of the most popular models of PDAs sold by 19 different retailers. These include models by Palm, HP, Sony, and Toshiba and span a wide range in prices. The lowest priced item is the Palm Handspring Treo, which has a median price of about £130, while the highest priced item is the Sony Clie nz90, with a median price of about £537.

Figure 1 shows a typical return from a price search on Kelkoo, which lists twelve retailers selling the HP iPAQ H5550 PDA. The information displayed includes a brief description of the product, the names of retailers selling the product, and price information detailed into item price, shipping charges ("P&P" in Kelkoo's terminology) and the total price inclusive of sales tax (VAT). A consumer interested in purchasing the item may click on the "More" button, or the retailer's name or logo, to be transferred directly to the retailer's website. Figure 1 illustrates the heterogeneity in the types of retailers using the site and the wide range of prices charged for an identical product. Some retailers, such as Comet and PC World, are "bricks-and-clicks" retailers who have physical stores in addition to an online presence. Others, such as Amazon and Dell, are well-known pure e-retailers, while firms such as Big Gray Cat are less well-known specialty e-retailers.

The Kelkoo dataset has certain advantages relative to clicks data from a US price comparison site. First, because of differences in the sales tax rates on interstate versus intrastate

---

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

transactions as well as differences in tax rates across states, it is difficult to identify the firm that offers the lowest price in US data. For instance, a California-based firm offering the lowest (pre-tax) price may offer the lowest final price to a customer domiciled in Indiana (where sales tax would typically be absent from the transaction) but not to a consumer from California (where sales tax would be charged). In contrast, the tax component of a firm's price is identical for all UK consumers shopping at the Kelkoo site regardless of whether they are domiciled in Scotland, England, or Wales.

Second, unlike many other price comparison sites, the order in which retailers' prices are displayed —that is, screen locations— are neither auctioned nor sold directly to retailers, and are independent of the price charged. Consequently, as far as both consumers and retailers are concerned, the order of price quotations for any specific screen is random. Finally, the complete list of prices is always displayed on a single page in our data. Thus, a change in a firm's screen location is never associated with a consumer having to click to "page two" in order to view the listing.

Kelkoo maintains an information log for each "referral" generated at its site.[9] 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. Kelkoo provided us with information extracted from their log files for the 18 PDA models for the period from 18 September 2003 to 6 January 2004, a period which generated over 40% of Kelkoo's annual traffic.[10]

For the 18 PDAs in our sample, this traffic amounted to 39,568 clicks generated by 20,509 separate cookies. The majority (60.1%) of cookies generated only one click, while a small number of cookies (0.56%) generated more than ten clicks.[11] Over the period of our study, there were 6,151 individual product, retailer and day specific price listings across the 18 PDA models. Our analysis is based on these observations, along with the number of last clicks generated for each PDA during the day of each listing.

---

[9] Throughout the paper we use the terms "referral", "lead", and "click" interchangeably.

[10] Kelkoo is bound to protect the anonymity of retailers and consumers in disclosing information about the referrals they obtain. So 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.

[11] Following Brynjolfsson, Dick, and Smith (2002), in instances where a consumer's cookie generated multiple clicks, we use the consumer's last click as an indicator of her final choice. We also performed the analysis reported below using all clicks as well as only first-clicks data, and the results are qualitatively similar.

The price used in our analysis is total purchase price—the actual cost to a consumer (including shipping and taxes) of purchasing a specific PDA model. To ensure an "apples to apples" comparison, we cleaned the data such that our analysis is based on listings of products that are identical in every respect (including condition). We note that Kelkoo verifies and updates the information it displays daily.

Table 1 provides descriptive statistics for the data. The PDAs in our sample are somewhat pricey, and shipping accounts for only a small fraction of the total purchase price. The average total price of a PDA in our sample is £309.04 ($497.55 at exchange rates as of 9/03), of which an average of £4.16 is accounted for by shipping charges. The number of sellers for a given product on any given day ranges from 1 to 15, with a mean of 4. The median number of clicks per day for a firm selling a specific PDA in our sample is 2. Consistent with the pattern observed in many traditional retail environments, referrals occur disproportionately in the fourth quarter of the year. However, as shown in Figure 2, online shopping disproportionately occurs on weekdays rather than weekends—opposite the pattern observed in many traditional retail environments.

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 clicks enjoyed by firms offering less favorable prices. Likewise, consumers tend to frequent firms that are listed above others on the screen. While the screen location and price rank are uncorrelated in our data, it is possible that the results displayed in Figure 3 are the result of joint correlation with a third variable.[12] We deal with this issue formally in Section 4 of the paper.

3   Estimation Methodology

We now describe our methodology for estimating the factors that determine a firm's number of clicks at the Kelkoo site.

---

[12] Note that the relationship in Figure 3 does not control for the fact that screens with only n listings never have price ranks or locations greater than n. Controlling for the number of listings, the weighted average correlation between price rank and screen location is -0.075 in our data.

3.1 How Firms Obtain Clicks

The number of clicks a firm receives for a given product on a given date depends on many factors, the most important of which are highlighted in Figure 4. As the figure shows, a firm's clicks depend on the price it charges, the number and prices of rival firms offering the same product, the identity of the firm and its rivals, the location of the firm's listing on the "page," and the date. Formally, let X denote this and other information that consumers obtain 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 quantity of clicks 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. The information in X—and this information alone—determines the number of clicks a firm obtains from the Kelkoo site. To estimate the parameters associated with the determinants of a firm's clicks, we use a pseudo-maximum likelihood approach that does not require us to make specific assumptions about the underlying distribution generating a firm's number of clicks; instead, we initially assume the underlying stochastic process has finite mean, given by

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

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 1, over 50% of the data consist of days in which a firm selling a particular PDA received two or fewer clicks. For this reason, analysis of these data requires regression techniques suitable for count data.

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 estimation (MLE) 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 (see 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 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 is misspecified. Since the Poisson distribution 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), we use the Poisson-based PML approach to obtain consistent estimates of β.[13] Following Hall and Ziedonis (2001), we use robust standard errors to obtain consistent estimates of the variance-covariance matrix. We also provide MLE estimates 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 Hausman, Hall, and Griliches 1984). Our results are robust to these alternative specifications.

4   Continuous Models

In this section we provide estimates of the determinants of a firm's clicks under the assumption that its expected number of clicks is a continuous function of its price.

---

[13] 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. Similarly, if one used a discrete choice framework to estimate behavior at the individual consumer level rather than the firm level, one would likewise obtain more efficient estimates if the underlying distributional assumption on the random utility component of the choice generating process were correct.

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 PML 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 β_j may be interpreted as the own price elasticity of clicks for a representative seller of a model j PDA. The vector $X_{1,ijt}$ consists of the following controls:

Position on Screen. As we showed earlier in Figure 3, when a firm's price is listed above its rivals, it tends to receive more clicks. Clicks tend to decrease as the position on the screen gets lower. Hence, we include a linear position on screen variable to capture this effect.[14]

Weekend. As displayed in Figure 2, firms systematically receive fewer clicks on weekends. Hence, we include a weekend dummy variable to control for this effect on clicks.

Month. We include month dummies to control for seasonal effects on clicks.

Table 2 reports the results of the individual product regressions. Notice that 13 of the estimated own price elasticities in Table 2 are statistically significantly different from zero at the 1% 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 clicks elasticities — not market clicks elasticities. One of the key theoretical determinants of a firm's price 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. For instance, it is well-known that 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): $E_F = nE_M$.

---

[14] We investigated nonlinear specifications of screen location as well and obtained similar results to those reported here. We chose the linear specification for parsimony.

One might expect clicks behavior to have a similar property.[15] If this is the case and different numbers of firms sold different types of PDAs, firms' clicks elasticities would vary widely across PDA models even if the market elasticity of clicks were the same for each model of PDA.

Thus, it seems useful to investigate the relationship between the clicks elasticity estimates reported in Table 2 and the average number of firms selling each PDA. This relationship is plotted in Figure 5. The estimates are divided into those that are not statistically significant at conventional levels (shown as open circles) and those that are significant at the 1% level (shown as filled-in diamonds). As Figure 5 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 to control for the number of sellers were one to pool across all products.

4.2 Pooled Estimates

We now report estimates obtained by pooling across firms (i), dates (t), and different models of PDAs using the PML procedure. 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:

Product. 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 product dummies for each of the 18 PDA models.

Product-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

---

[15] For instance, if a firm's demand were proportional to clicks, a firm's clicks elasticity would correspond to its demand elasticity and thus be a function of the number of sellers in the market.

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 product dummies with the month dummies mentioned above. This, in principle, allows PDA "life cycles" to differ during the time horizon of our study.

Bricks and Clicks Retailer. Some of the firms in our dataset have an established physical presence in addition to their online presence. Clearly, the reputation as well as the ease of returns and accumulated brand equity of these retailers may differ from pure online sellers. Thus, we include a dummy variable for whether a particular firm is a bricks and clicks retailer.

With these controls in place, we report PML estimates (Table 3, Model 1) based on the mean specification in equation (4). The bottom of Table 3 also reports the results of a likelihood ratio test for overdispersion of the negative binomial (2) type (cf. 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 variance-covariance estimates, we employ the standard error correction techniques of Rogers (1993), Huber (1967) and White (1980,1982).[16] The corresponding z-statistics are reported in Table 3.[17]

The results show a price elasticity of clicks of -4.61, which is fairly close to the average over the individual product elasticities reported in Table 2. More favorable screen positions lead to increased clicks: All else equal, a firm that moves up one screen position enjoys an 18.6% increase in clicks. These results confirm what we saw earlier in Figure 3: There is

---

[16] 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.

[17] 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.

a strong tendency for consumers to click on firms listed at the top of the display screen, all else equal. This may also explain why search engines that auction screen positions, such as Google, receive significant premia for positions located near the top of the screen. 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 in this continuous specification.

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

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

where n_{jt} denotes the number of sellers of type j PDA on date t. Notice that, in this specification, the clicks elasticity for a representative firm is given by

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

Thus, the coefficient of total price (β_0) represents the clicks elasticity for a monopoly seller, $\beta_0 + \beta_1$ represents the elasticity in duopoly PDA markets, and more generally, β_1 represents the impact on a firm's clicks elasticity 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 clicks elasticity 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. We include a linear term for the number of sellers.[18]

The resulting estimates are displayed in the Model 2 column of Table 3. As the table shows, the number of sellers has a significant effect—both in terms of levels as well as on

---

[18] We also ran dummy specifications for the number of sellers and obtained similar results.

clicks elasticities. Controlling for the number of firms listing prices, we find that the price elasticity of clicks for a monopoly seller is -3.761, which implies that a 10% increase in price leads to a 38% decline in clicks. Adding a second firm to the market raises the elasticity to around -4.049. When ten firms list prices, the estimated elasticity becomes -6.64, which implies that a 10% increase in price results in a 66% decline in a firm's number of clicks.

What is the impact of a change in the number of sellers on a firm's overall number of clicks? 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

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

or about -0.04 ($p = .0155$). It is useful to contrast the magnitude of this "rivalry" effect with that of a change in a firm's screen position. As Table 3 shows, a reduction of one screen position decreases the firm's clicks 17.5%. Thus, our estimates suggest that the impact of screen position is more than four times larger than the impact of an additional competitor appearing on the price comparison site.

4.3   Potential Misspecification

One may have a number of concerns regarding the estimates based on the continuous clicks specification in equation (5). As we have emphasized, the PML approach is robust against alternative distributional assumptions but not to the misspecification of the underlying mean of the stochastic process.

First and foremost, price comparison sites are often used by consumers looking to obtain a given product at the best price. For instance, Brynjolfsson and Smith (2000) have provided evidence that 49% of consumers using price comparison sites in the U.S. make purchase decisions based purely on price. The results of Ghose, *et al.* (2006) 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% of the clicks are at the lowest price. These observations, coupled with the recent literature discussed in the introduction that rationalizes the observed levels of price

dispersion in online markets suggests that a firm lowering its price from the second-lowest to the lowest price enjoys a discontinuous jump in its number of clicks.

To see the potential ramifications of a discontinuity at the lowest price (as implied by clearinghouse models) on clicks estimation, suppose there is a unit mass of consumers, half of which are "shoppers" who click on the firm offering the lowest price and the other half are "loyals" who have a preference for a particular seller. Consumers within each group have identical and constant clicks elasticities, θ; thus, a firm's number of clicks is given by $Q = p^{-\theta}$. A firm that charges the lowest price in the market enjoys clicks from both groups, while a firm charging a price above the minimum price in the market only receives clicks from its loyal customers. Figure 6 illustrates the ramifications on clicks estimation. The slope of the two steep lines through the data are the same, and represent the true clicks elasticity, $-\theta$, for prices above or below the minimum price. At the minimum price, there is a discontinuous jump in clicks 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 that occurs when the firm charges the lowest price. Notice that, by ignoring the jump in clicks at the lowest price, one obtains an elasticity estimate that overstates how responsive consumers are to a change in price.

In addition to the potential problem caused by using a continuous 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 clicks) 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.

Second, while we have controlled for one 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 account for unobserved firm characteristics in estimating a firm's clicks.

We address these and other issues in the next section.

5 Discontinuous Models

Clearinghouse models (e.g., the seminal models of Varian (1980) and Rosenthal (1980)) postulate that firms compete for two types of consumers: price-sensitive shoppers (who always purchase at the lowest price) and loyals (who do not). In this section we use these models to derive functional forms that allow for a possible jump in a firm's clicks when it charges the lowest price. We then describe the estimating equation and report results. Finally, we examine issues associated with endogeneity, unobserved firm characteristics, and potential continuous cross-price effects.

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 total price of firm i in this market. A firm in this market competes for two types of consumers: Loyals, who click on their preferred firm, and shoppers, who click on the firm charging the lowest price. 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 as:

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

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

$$Q_{ijt} = \begin{cases} Q_{ijt}^S + Q_{ijt}^L & \text{if } i \in B_{jt} \\ Q_{ijt}^L & \text{if } i \notin B_{jt} \end{cases}$$

Thus, firm i faces a jump in clicks for product j when it is among those firms offering the best price for product j on date t.

To account for the discontinuity in clicks 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 number of firms offering the best price for product j at time t.

Suppose that firm i 's elasticity when it sells product j is θ_{jt}, so that we may write

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

where $\alpha_{ijt}^L(X)$ and $\alpha_{ijt}^S(X)$ represent the non-price determinants of clicks (such as screen location) on loyals and shoppers, respectively. To ease the notational burden, we suppress the X argument where it is clear. Hence we may rewrite Q_{ijt} as:

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

where

$$\lambda_{ijt} = \frac{\alpha_{ijt}^S}{\alpha_{ijt}^L}.$$

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

$$\ln Q_{ijt} = \lambda_{ijt} \frac{\mathbf{I}_{jt}}{\#B_{jt}} + \ln \alpha_{ijt}^L - \theta_{jt} \ln p_{ijt} \qquad (6)$$

Estimation requires imposing additional structure on the parameters in equation (6). We assume

$$\theta_{jt} = \beta_0 + (n_{jt} - 1)\beta_1 \qquad (7)$$

As in the previous section, this parsimonious specification allows a firm's clicks elasticity for product j to depend on the number of sellers at time t. 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

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

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

$$\lambda_{ijt} = \lambda = \frac{a^S}{a^L}. \qquad (8)$$

Under these assumptions, the mean specification is:

$$E[Q_{ijt}|X] = \exp\left[(\beta_0 + (n_{jt} - 1)\beta_1)\ln p_{ijt} + \beta_2 n_{jt} + \lambda \frac{\mathbf{I}_{jt}}{\#B_{jt}} + \gamma X_{1,ijt}\right], \qquad (9)$$

where $X_{1,jt}$ is the matrix of controls discussed earlier (position on screen, bricks and clicks retailer, weekend, month, and product dummies as well as product-month interaction dummies). As above, we may interpret the expression $\beta_0 + (n_{jt} - 1)\beta_1$ as the price elasticity of clicks for a firm that faces $n_{jt} - 1$ rivals. Similarly, we may interpret λ as the jump from shoppers; that is, the size of the discontinuous jump in clicks that a firm enjoys when it offers the "best" price. Notice that the continuous model is nested in the specification of equation (9) when $\lambda = 0$; thus, we may readily test the null hypothesis implied by the continuous clicks specification.

5.2   Parameter Estimates

Model 1 in Table 4 reports PML estimates of the parameters in equation (9). Recall that under the nested continuous clicks specification, the coefficient associated with the jump from shoppers (λ) is equal to zero. The alternative hypothesis, predicted from the clearinghouse models, 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 model in favor of the (one-sided) alternative discontinuous specification at the 1% significance level. In short, we find considerable evidence for a jump in clicks when a firm offers the lowest price.

Figure 6 suggested that, in the presence of such a jump, estimates based on the continuous specification imply a greater sensitivity of a firm's clicks to its price than in the discontinuous specification. The results shown in Model 1 of Table 4 compared to those in Model 2 of Table 3 are consistent with this observation. Accounting for the discontinuity in clicks, the estimated elasticity for a monopoly seller becomes less elastic—going from -3.761 (in Model 2 of Table 3) to -2.459 (in Model 1 of Table 4). The difference in the elasticity estimates is greater for markets with more than one seller: The effect of an additional rival on the clicks elasticity is reduced by around 12.5% (from -0.288 in Model 2 of Table 3 to -0.252 in Model 1 of Table 4). We also note that, in contrast to the continuous specification, the effect of a change in the number of firms on a firm's overall number of clicks $\left(\partial \ln E\left[Q_{ijt}|X \right] / \partial n_{jt} |_{\bar{p}_{ijt}} \right)$ is not statistically different from zero ($p = .4674$).

It is of some interest to note the economic relevance of our estimate of the jump from shoppers ($\lambda = 0.603$). Other things equal, a firm that sets the lowest price in the market

enjoys a 60.3% increase in clicks. In contrast, notice that the "position on screen" coefficient implies that a firm would have to move up 3 screen positions to generate the same increase in clicks that results from setting the lowest price in the market. Also, note that setting the lowest price in the market entails a jump in clicks that is about twice as large as the 32.1% shift associated with being a bricks and clicks retailer.

One may use our estimates of λ to obtain a very crude estimate of the fraction of consumers using the Kelkoo site who are shoppers. 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

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

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

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

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

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

which implies (given the estimate of $\lambda = .603$ reported in Model 1 of Table 4 and the mean number of listings (4.05) in our data) that about 13% of consumers at Kelkoo are shoppers. While the symmetry assumptions used to calculate this crude estimate are at odds with the data (among other things, the estimates suggest that bricks-and-clicks sellers receive 32.1% more clicks than pure online sellers), it nonetheless illustrates that even in an online market where 87% of the consumers are "loyal" to a particular firm, discontinuities arising from shoppers can significantly impact estimates of a firm's clicks elasticity.

5.3   Potential Misspecification

While the PML approach used to obtain the estimates reported above does not make specific distributional assumptions about the underlying clicks generating process (apart from assuming that the conditional mean specification is correct), it is nonetheless useful to compare the shape of the empirical distribution with the distribution based on the number of clicks predicted by the model. As Figure 7 reveals, the distribution of predicted clicks (based on Model 1 in Table 4) resembles that observed in the data. While this is somewhat reassuring, it does not rule out the possibility that Model 1 in Table 4 is misspecified due to endogeneity, unobserved heterogeneity across firms, or the failure to incorporate additional cross price effects.

Endogeneity

One concern that might be raised with the preceding analysis is that two of the key variables of interest—price and the number of listing firms—might be endogenous in the sense that the regressors may be correlated with omitted variables, thus creating the possibility of inconsistent parameter estimates. This problem might arise through unobserved factors that influence clicks, such as variations in a product's popularity. In addition, changes in popularity would also be correlated with the number of sellers (an increase in popularity would presumably induce more sellers to enter) as well as the prices charged for the product (firms would raise prices for "hot" products and reduce them for "cold" products).

To be concrete, suppose that the correct clicks generating process is given by:

$$Q_{ijt} = \exp\left[(\beta_0 + (n_{jt} - 1)\beta_1)\ln p_{ijt} + \beta_2 n_{jt} + \lambda \frac{\mathbf{I}_{jt}}{\#B_{jt}} + \gamma X_{1,ijt} + pop_{jt}\right] + \varepsilon_{ijt} \qquad (10)$$

where pop_{jt} denotes an unobserved latent variable and ε_{ijt} is a zero mean error term. Then, the correct conditional mean specification is:

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

and hence equation (9) is misspecified. However, the parameter estimates based on equation (9) remain consistent provided:

$$E\left[Q_{ijt}|X_{ijt}, pop_{jt}\right] = E\left[Q_{ijt}|X_{ijt}\right] \tag{11}$$

That is, to the extent that fluctuations in pop_{jt} only occur *inter-month*, any influence of pop_{jt} on $E\left[Q_{ijt}|X_{ijt}, pop_{jt}\right]$ will be absorbed in the product-time interaction dummies contained in $X_{1,ijt}$, and equation (11) will hold—the mean specification given in equation (9) will be correct. If there is *intra-month* variation in pop_{jt}, however, this variation will not be absorbed in the product-time interaction dummies. Furthermore, for the reasons described above, pop_{jt} is likely to be correlated with p_{ijt} and n_{jt}. In this case, equation (11) will not hold, and the parameter estimates reported above will be inconsistent.

To deal with this potential endogeneity problem, we collected additional data from the US to serve as a proxy for unobserved factors that influence clicks in the UK (i.e. the variable pop_{jt}). In particular, for each of the PDAs in our UK sample and for each date, we obtained data on that PDA's product popularity ranking (for that same day) from Shopper.com—a US price comparison site.[19] The product popularity ranking are integers (1 represents the most popular product, 2 the second most popular product, and so on), and are constructed by Shopper.com based on consumer pageviews at its site. It seems plausible that the product popularity ranking of an identical PDA model in the US on a given date is a valid proxy for its popularity in the UK for that same date, as it is likely correlated with pop_{jt} but uncorrelated with the other UK regressors included in equation (10).

Formally, suppose that

$$pop_{jt} = \eta D_{jt}^{USPOP} + \nu_{jt}$$

where D_{jt}^{USPOP} is a matrix of dummy variables for the popularity of product j at time t from the US data, η is a vector of parameters associated with each product rank, and ν_{jt} is an

---

[19] The US data is discussed in more detail in Baye, Morgan and Scholten (2004a).

22

error term. Substituting this relation into equation (10), one obtains

$$Q_{ijt} = \exp\left[(\beta_0 + (n_{jt} - 1)\beta_1)\ln p_{ijt} + \beta_2 n_{jt} + \lambda\frac{\mathbf{I}_{jt}}{\#B_{jt}} + \gamma X_{1,ijt} + \eta D_{jt}^{USPOP}\right]\exp\left[\nu_{jt}\right] + \varepsilon_{ijt}$$

Assuming ν_{jt} is independent of the other regressors, this implies a conditional mean specification of

$$E\left[Q_{ijt}|X_{ijt}, pop_{jt}\right] = \exp\left[(\beta_0 + (n_{jt} - 1)\beta_1)\ln p_{ijt} + \beta_2 n_{jt} + \lambda\frac{\mathbf{I}_{jt}}{\#B_{jt}} + \gamma X_{1,ijt} + \eta D_{jt}^{USPOP}\right]$$
$$(12)$$

where we have assumed, without loss of generality, that $E\left[\exp\left[\nu_{jt}\right]\right] = 1$.

Under the stated assumptions, PML estimation of equation (12) gives consistent parameter estimates. We report the results of this specification as Model 2 in Table 4. As the table shows, controlling for potential endogeneity does little to the magnitude or the significance of the coefficient estimates.[20]

Unobserved Firm Heterogeneity

Another potential shortcoming of the PML approach used in Models 1 and 2 of Table 4 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 clicks are also driven by unobserved factors. For this reason, we also report in Table 4 results that allow for the effects of unobserved firm heterogeneity.

Model 3 in Table 4 reports maximum likelihood estimates of the discontinuous model based on the random effects specification for unobserved firm heterogeneity pioneered by Hausman, Hall and Griliches (1984), while Model 4 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 4 and discussed below are similar if one uses the likelihood function for a negative binomial (2) specification.

---

[20] There is some evidence that the US popularity dummies address the endogeneity issues discussed above. In all specifications that include US popularity dummies, we reject the null hypothesis that the coefficients on D_{jt}^{USPOP} are jointly equal to zero ($p < 0.001$).

Notice that, in both the random effects (Model 3) and fixed effects (Model 4) specifications, the coefficients of interest are roughly comparable to those obtained ignoring potential unobserved heterogeneity (Models 1 and 2). As the coefficient on the jump from shoppers reveals, we still reject the null hypothesis of the continuous model in favor of the discontinuous specification. Further, the economic value of the coefficient associated with the jump is largely unchanged by allowing for potential unobserved heterogeneity. Likewise, the coefficient associated with the clicks elasticity for a monopoly firm remains at about -2.5, similar to the estimate obtained in Model 1.

In contrast, both the effect of a change in the number of rivals on a firm's clicks elasticity as well as the effect of change in screen position on a firm's clicks are dampened in Models 3 and 4 compared to Model 1. This suggests that part of the effect that was previously attributed to changes in screen position or rivalry is more properly accounted for by unobserved firm heterogeneities. Nonetheless, the coefficients on these variables remain economically and statistically significant.

Continuous Cross-Price Effects

A final potential concern with the clicks estimates presented above is that, although the discontinuous specification is consistent with theoretical clearinghouse models, it does not allow for cross-price effects (except for the discontinuous jump in clicks a firm enjoys when the minimum rival's price rises above its price). One might conjecture that our relatively inelastic estimates of clicks elasticities, as well as the significant discontinuity in clicks (jump from shoppers) stem from misspecification – that is, a failure to allow for continuous cross-price effects.

To investigate this alternative hypothesis, we generalized the specifications to allow for continuous cross-price effects by including the logarithm of the geometric average of rivals' prices on a given product-date as a regressor.[21] In this specification, the estimated coefficient for the logarithm of the geometric mean of rival's prices represents a firm's cross price elasticity of clicks with respect to the (geometric) average of rivals prices.

The results are displayed in Table 5, and reveal that the findings discussed above are

---

[21] Since cross-price effects are only relevant for product-dates in which two or more firms compete, for product dates in which only a single firm listed a price, we coded the logarithm of the geometric average of rivals' prices as zero and included a dummy that equaled one for these events and zero otherwise.

robust to the inclusion of continuous cross-price effects. For the continuous specification (Model 1), the cross-price effects are not statistically different from zero at conventional levels (even after controlling for potential endogeneity, as both specifications in Table 5 do). For the discontinuous specification (Model 2), note that although the coefficient for the cross price elasticity of clicks is positive, it is relatively small (0.38) and not significantly different from zero at conventional levels.[22] More importantly, our finding that a firm enjoys a significant 60% jump in clicks when it succeeds in offering the lowest price is robust to the inclusion of continuous cross-price effects.

6 Implications for Firm Demand

In this section, we offer conditions under which the above parameter estimates may be used to recover or place bounds on the elasticities of demand facing individual firms operating in the Kelkoo marketplace.

Toward this end, recall the process by which a click leads to a sale shown in Figure 4. After having observed information X at the Kelkoo site, consumers clicking through to a firm's site receive additional information (denoted Z_i) that influences their 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, consumers' perceptions of these factors may be colored by the previous information, X, obtained on the Kelkoo site. To account for the possibility that consumers may observe Z_i's for all firms before making purchase decisions, let $Z = (Z_1, Z_2...Z_n)$ denote the vector of all such information. In this case, the probability that a click on firm i is converted into a sale, given (Z, X), is

$$\Pr\left(sale_i | Z, X\right) = G_i\left(Z, X\right)$$

---

[22] Note, however, that one can reject the joint hypothesis that the coefficients of absence of rivals dummy and log geometric mean of rival's prices are zero at the 10 percent level in Model 1 and the 1 percent level in Model 2.

Using equation (2), we may write the expected demand for a given product sold by firm i, conditional on (X, Z), as

$$
\begin{aligned}
E\left[D_i | X, Z\right] &= G_i\left(Z, X\right) \times E\left[Q_i | X\right] \\
&= G_i\left(Z, X\right) \times \exp\left[X\beta\right]
\end{aligned}
$$

The multiplicative separability of the probability of conversion, G_i, and the clicks generating process follows naturally from the Kelkoo search and buying environment. Of central interest is the effect of information obtained from the Kelkoo site (various components of X) on final demand. Suppose one wished to measure the effect on firm i's demand of a change in x_i (some component of i's information posted at Kelkoo). It is useful to rewrite $X = (x_i, X_1)$ where X_1 represents all components of X other than x_i. If x_i influences firm i's clicks but does not impact its conversion rate for that product, one may recover the firm's elasticity of demand with respect to x_i from its elasticity of clicks. Formally,

Proposition 1 *Suppose that $G_i\left(Z, (x_i, X_1)\right) = G_i\left(Z, (x_i', X_1)\right)$ for all x_i, x_i'. Then*

$$
\frac{E\left[D_i | x_i, X_1, Z\right] - E\left[D_i | x_i', X_1, Z\right]}{E\left[D_i | x_i, X_1, Z\right]} = \frac{E\left[Q_i | x_i, X_1\right] - E\left[Q_i | x_i', X_1\right]}{E\left[Q_i | x_i, X_1\right]},
$$

and furthermore, if demand is differentiable,

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

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

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

Differentiating with respect to x_i yields

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

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

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

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

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

where we have again used the fact that $G_i\left(Z,\left(x_i, X_1\right)\right) = G_i\left(Z,\left(x_i', X_1\right)\right)$ for all x_i, x_i'. ∎

Notice that the conditions of the Proposition allows firm i's conversion rate for product j to depend on additional information obtained from *all* firms' individual websites (Z), as well as all the other information posted at the Kelkoo site (X_1)—including information other than x_i posted by firm i. Provided that firm i's conversion rate is insensitive to x_i—that is, conditional on (x_i, X_1) being sufficiently "favorable" to induce the consumer to click through to firm i's site first place, the level of x_i does not influence the likelihood that the clickthrough is converted into a sale—the Proposition shows how to recover the associated demand parameter from clicks estimates.

The proposition has several potentially useful implications. First, when x_i is firm i's price listed on the Kelkoo site, and the conversion rate is independent of firm i's price, then one may interpret a firm's elasticity of "clicks" with respect to price as its own price elasticity of demand. Second, when x_i is a discrete variable (such as a dummy variable), one may use estimates based on clicks data to infer the percentage impact of a discrete change in x_i on a firm's demand.

More importantly, one may use clicks estimates to obtain bounds for underlying demand parameters. For instance, if firm i's conversion rate is increasing in x_i, then the estimated clickthrough elasticity is a lower bound for the associated demand elasticity (since in this case, $\frac{\partial \ln G_i(Z,X)}{\partial \ln x_i} + \frac{\partial \ln E[Q_i|X]}{\partial \ln x_i} > \frac{\partial \ln E[Q_i|X]}{\partial \ln x_i}$). Consequently, the *clicks* elasticity of -2.091 reported in Model 2 of Table 5 implies that a 10% increase in a monopoly seller's price reduces its *quantity demanded* by at least 20.91% if conversion rates are non-increasing in price. Similarly, our estimates imply that bricks and clicks retailers enjoy at least 32.6% higher sales than their purely online competitors if conversion rates for these firms are at least as high as their purely online rivals.

7 Implications for Discrete Choice Models of Individual Behavior

Our analysis has focused on the determinants of the total numbers of clicks that *firms* receive in an online marketplace, treating the firm as the unit of observation. One possible worry concerns the relationship between the specification of clicks at the firm level, given in equation (12), and underlying individual behavior. However, we showed that this specification arises when one aggregates individual behavior as it is specified in standard clearinghouse models: Exogenous differences in preferences induce some consumers to only click the low-price firm and others to click their preferred firm. These models predict that there exists a jump in a firm's clicks (stemming from the choice behavior of shoppers) when it charges the lowest price in the market. As documented above and in the Appendix, there is considerable evidence that this jump is economically and statistically significant. Moreover, failing to account for the jump distorts the estimates of the parameters associated with the determinants of a firm's total number of clicks.

To the extent that one is interested in a more detailed analysis of individual consumer choice rather than the total number of clicks that a firm (or platform) receives, discrete choice models of individual behavior (such as the multinomial logit family of models) are appropriate alternatives. Standard versions of these models postulate that each consumer has the same propensity to click on a particular firm, given that firm's characteristics and the characteristics of the other firms in the choice set. Variation in choices across consumers is assumed to arise from a random utility component, which is typically assumed to be *iid* across consumers. Our results suggest that it might be appropriate to account for jumps when estimating click propensities at the individual consumer level. In particular, a "shopper's" propensity to click a given firm critically depends on whether it offers the lowest price, while a "loyal's" propensity does not. Given appropriate assumptions on the structure of loyal and shopper behavior, one could, in principle, use a discrete choice framework to obtain estimates of the determinants of individual consumer's propensities to click on a particular firm. Our results based on analysis at the firm-level suggest that accounting for such asymmetries in individual choice is a potentially fruitful avenue for future research.

8 Discussion

Standard models of online price competition predict a discontinuous jump in a firm's clicks when it succeeds in charging the lowest price in the market. These models also predict that the responsiveness of a firm's clicks to a change in its price depends on the number of rivals. Our results indicate that both the number of competing sellers and the firm's rank in the list of prices are important determinants of an online retailer's clicks. We find that an online monopolist faces a clicks elasticity of about -2, while sellers competing against 10 other sellers face an elasticity of about -6. We also find empirical evidence of a 60% jump in a firm's clicks as its price declines from the second-lowest to the lowest price. We conclude by illustrating the potential relevance of our findings for stakeholders in online markets, including firms that advertise prices at comparison sites, the sites themselves, and policymakers.

Consider the low-priced firm in a market with one other rival that is contemplating a price increase. Based on the estimates in Model 1 of Table 5, if the firm ignored the discontinuity in clicks it would conclude that a 10% price increase would reduce clicks by 37%. A firm that properly accounted for the jump would recognize that the impact of a price increase critically depends on whether it remained the low priced firm. If the firm can increase its price by 10% and retain low-priced status, Model 2 of Table 5 indicates that its clicks would decline by only 24%. However, if the 10% price increase moves the firm out of the low-price position, the firm's clicks would decline by 85%. Regardless of whether conversion rates are independent of the firm's price, this translates into a discrepancy in the firm's estimated clicks (and clickthrough fees) of 13% to 48%.

Next, consider the situation from the viewpoint of Yahoo's price comparison site. Suppose Yahoo correctly forecasts that a 3% increase in its clickthrough fees would lead to a 1% increase in the prices charged by each firm in a duopoly market. If the platform used Model 1 of Table 5 to estimate the clicks elasticity, it would erroneously conclude that the 3% increase in clickthrough fees would reduce its clicks by 3.7%—an unprofitable move. On the other hand, if it used Model 2 in Table 5, the comparison site would (correctly) conclude that the 3% increase in clickthrough fees would reduce clicks by only 2.4%—a profitable move.

Finally, consider a policymaker contemplating the imposition of a national *ad valorem* tax on Internet transactions that would increase each firm's price by a factor of $1 + \tau$. The policy-maker's first-order approximation of the excess burden of the tax paid by a consumer k who purchased q_k units at the initial price of p_k is given by the well-known formula

$$EBT = \frac{1}{2}\theta\tau^2 p_k q_k,$$

where p_k and q_k are the initial price and quantity data for this particular consumer and θ is the own-price elasticity of demand. Our results suggest that, since the elasticity estimates under the continuous specifications are significantly more elastic than in the discontinuous specifications, the failure to account for discontinuities leads to potential biases in the welfare analysis of such a tax. For example, if the policymaker's estimate of the elasticity of demand were based on the continuous specification in Model 1 of Table 5 and a duopoly market, the ratio of the excess burden under the two specifications is

$$\frac{EBT_{\text{Continuous}}}{EBT_{\text{Discontinuous}}} = \frac{\theta_{\text{Continuous}}}{\theta_{\text{Discontinuous}}} = \frac{-3.7}{-2.4} = 1.54$$

In other words, failure to account for discontinuities in estimation and policy analysis could lead to excess burden calculations that are 54% higher than that actually experienced by *any* individual consumer in the market (shopper or loyal). While a proper analysis of taxes and Internet transactions obviously requires more detailed analysis, this simple calculation illustrates that failing to account for discontinuities in online data may lead to significant biases in policy analysis.

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, "Information, Search, and Price Dispersion," *Handbook of Economics and Information Systems* (forthcoming), T. Hendershott, ed., North Holland: Elsevier.

[3] 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.

[4] 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.

[5] Brown, J.R. and A. Goolsbee, "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy* (1994) 110(4), pp. 481-507.

[6] Brynjolfsson, Erik, Astrid Dick, and Michael Smith, "Consumer Heterogeneity and Search Costs on the Internet," Working Paper, Carnegie Mellon University, 2002.

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

[8] Brynjolfsson, E. and M.D. Smith, "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science* (2000b) 46(4), 563-585.

[9] 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.

[10] 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.

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

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

[13] Dana, J.D., "Learning in an Equilibrium Search Model," *International Economic Review* (1994) 35, pp. 745-771.

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

[15] Ghose, A., M. Smith, and R. Telang, "Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact," *Information Systems Research* (2006) 17(1), pp. 3-19.

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

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

[18] Hausman, J., 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.

[19] 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.

[20] 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.

[21] Janssen, M. and J.L. Moraga-González, "Strategic Pricing, Consumer Search , and the Number of Firms," *Review of Economic Studies* (2004) 71 (4), pp. 1089-1118.

[22] Ju, J., S. Linn, and Z. Zhu, "Price Dispersion in a Model with Middlemen and Oligopolistic Market Makers: A Theory and an Application to the North American Natural Gas Market," *mimeo*, 2006.

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

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

[25] 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.

[26] Salop, S.C. and J.E. Stiglitz, "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion," *Review of Economic Studies* (1977) 44, pp. 493-510.

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

[28] Spulber, D., "Bertrand Competition when Rivals' Costs are Unknown," *Journal of Industrial Economics* (1995) 43 (1), pp. 1-11.

[29] Stahl, D.O., "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review* (1989) 79, pp. 700-712.

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

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

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

Figure 1: Kelkoo Screenshot

Figure 2: Average Number of Clicks by Day of Week
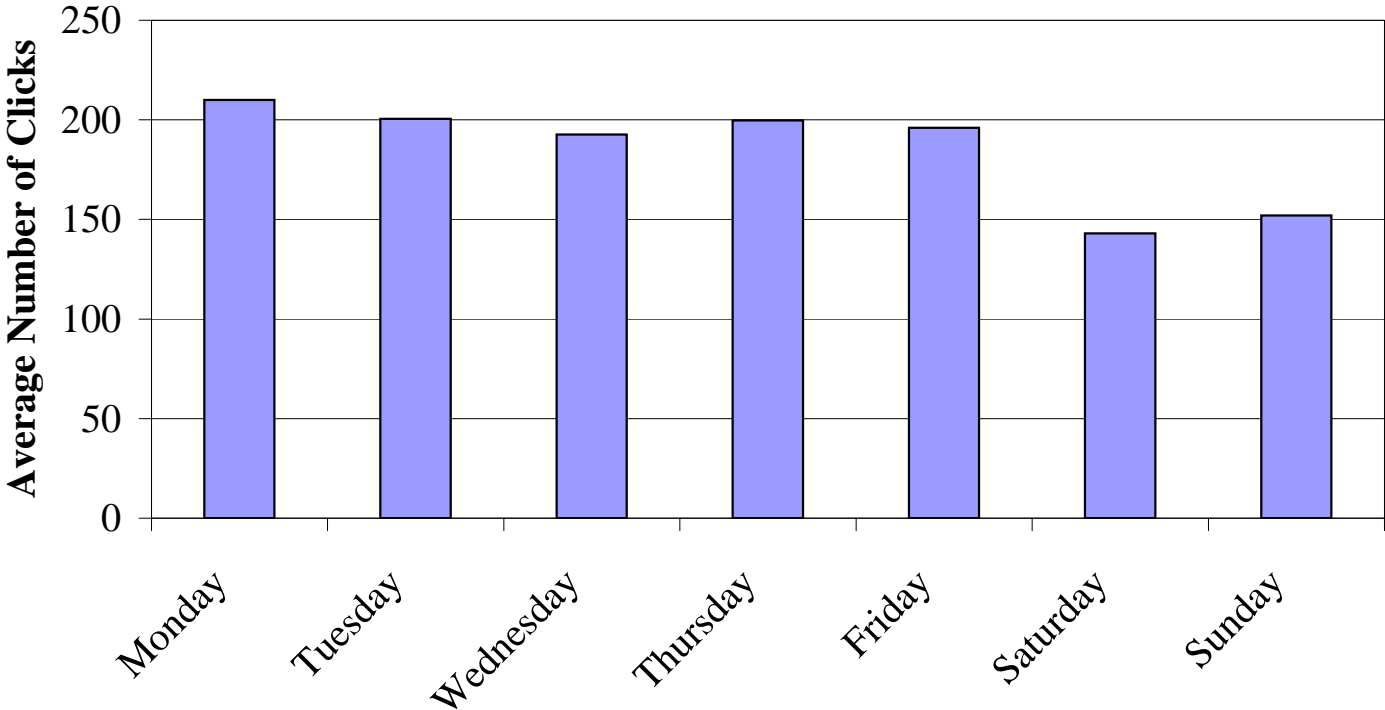
Figure 3: Histogram of Clicks by Price Rank and Screen Location
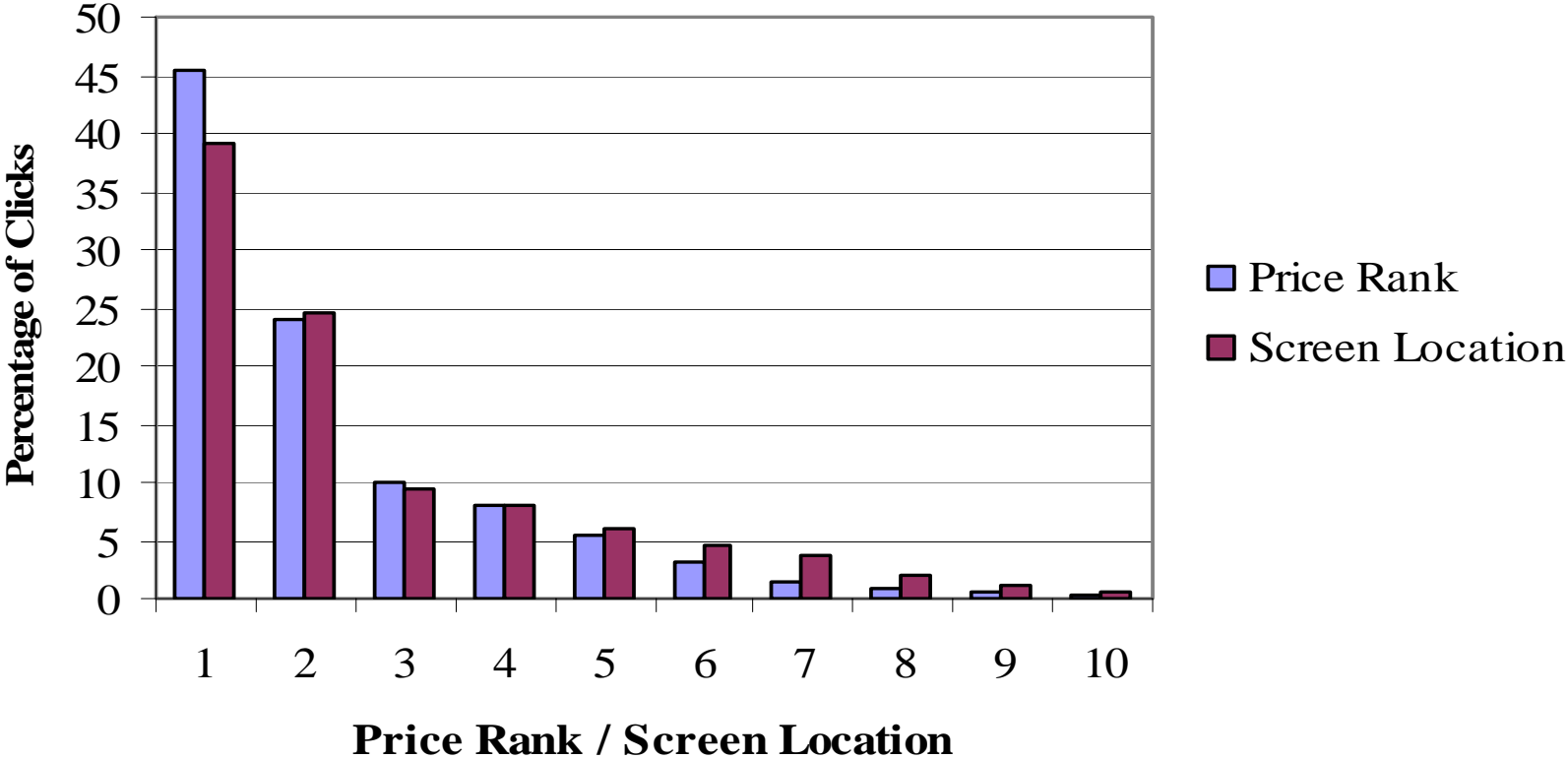
Figure 4: Factors Influencing Numbers of Clicks
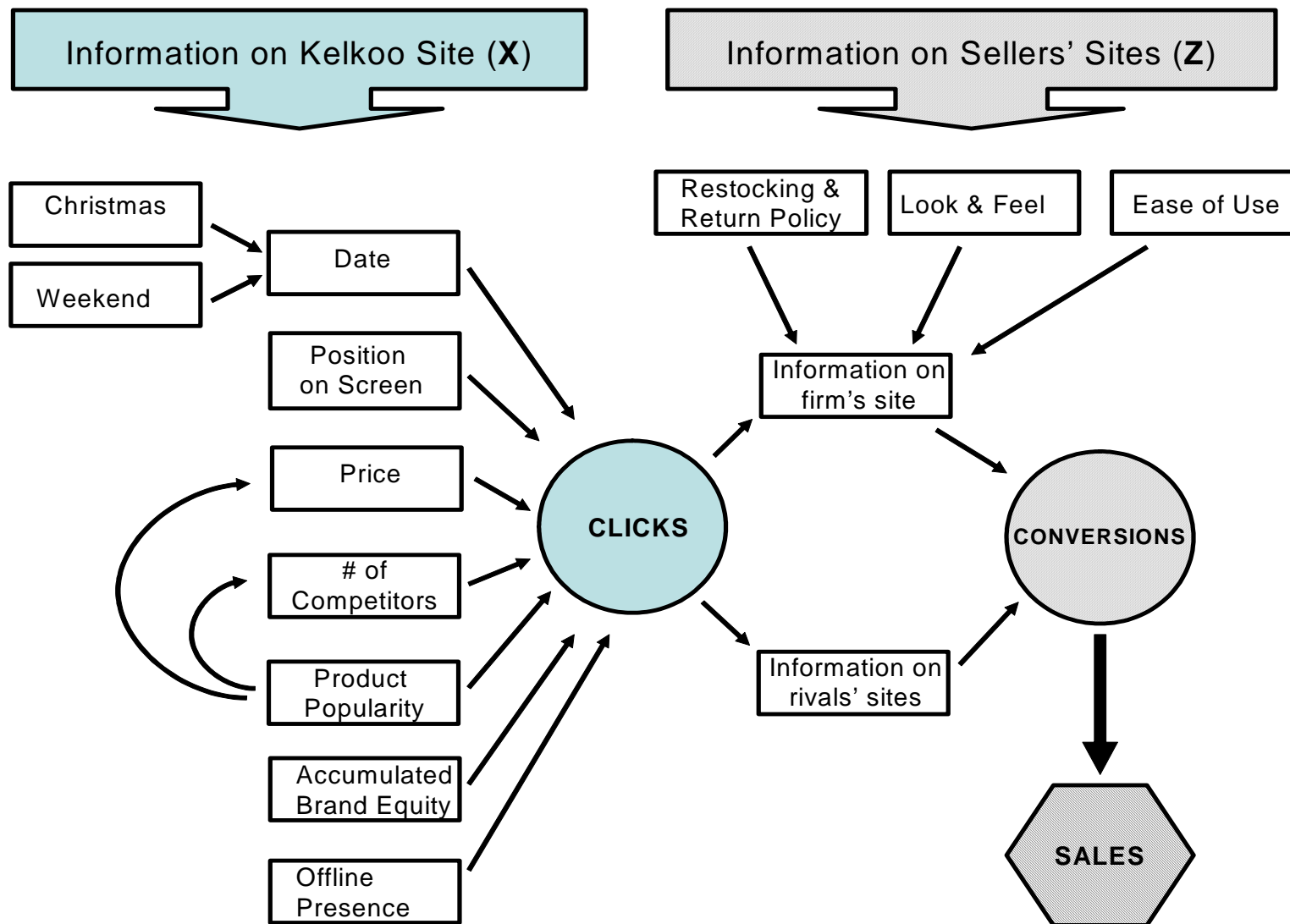
Figure 5: Estimated Clicks Elasticity and Number of Sellers
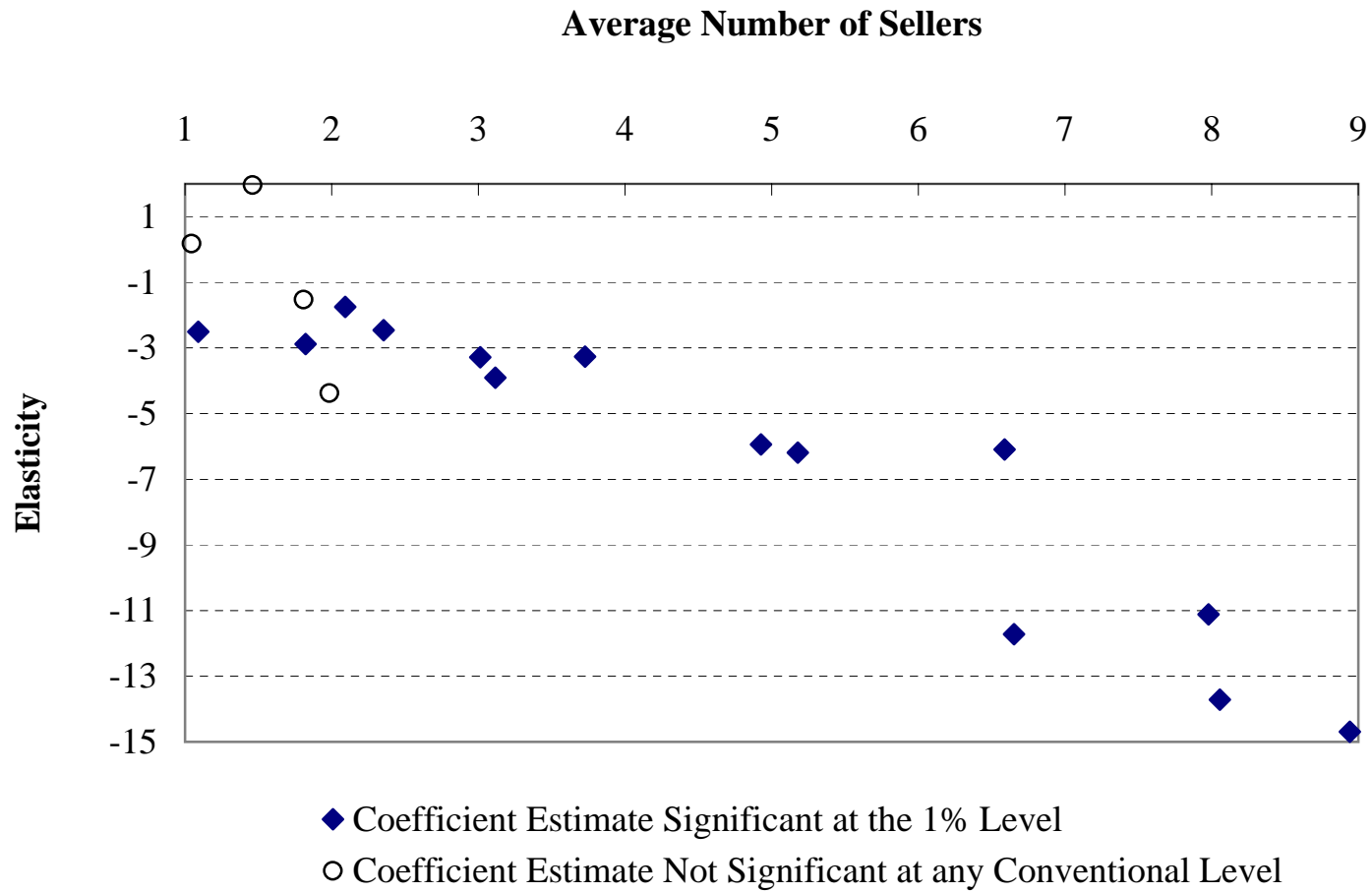
Figure 6: Misspecification from using Continuous Clicks Model in Discontinuous Market Setting

Ln(Price)

Best Price

Firm i's
Estimated Clicks
Behavior

Firm i's Actual
Clicks
Behavior

Ln(Clicks)

Jump from
'Shoppers'
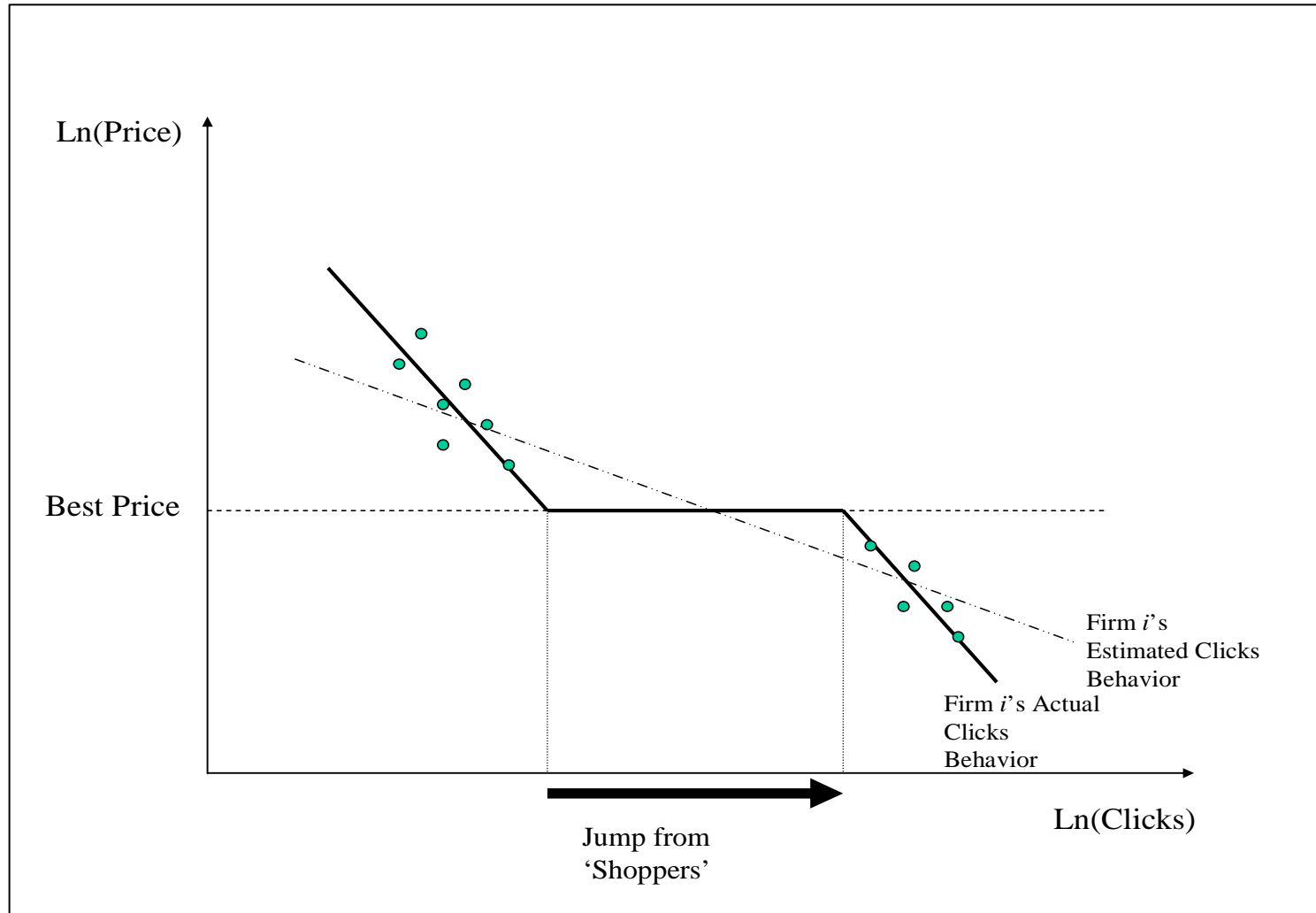
Figure 7: Actual versus Predicted Clicks Frequency
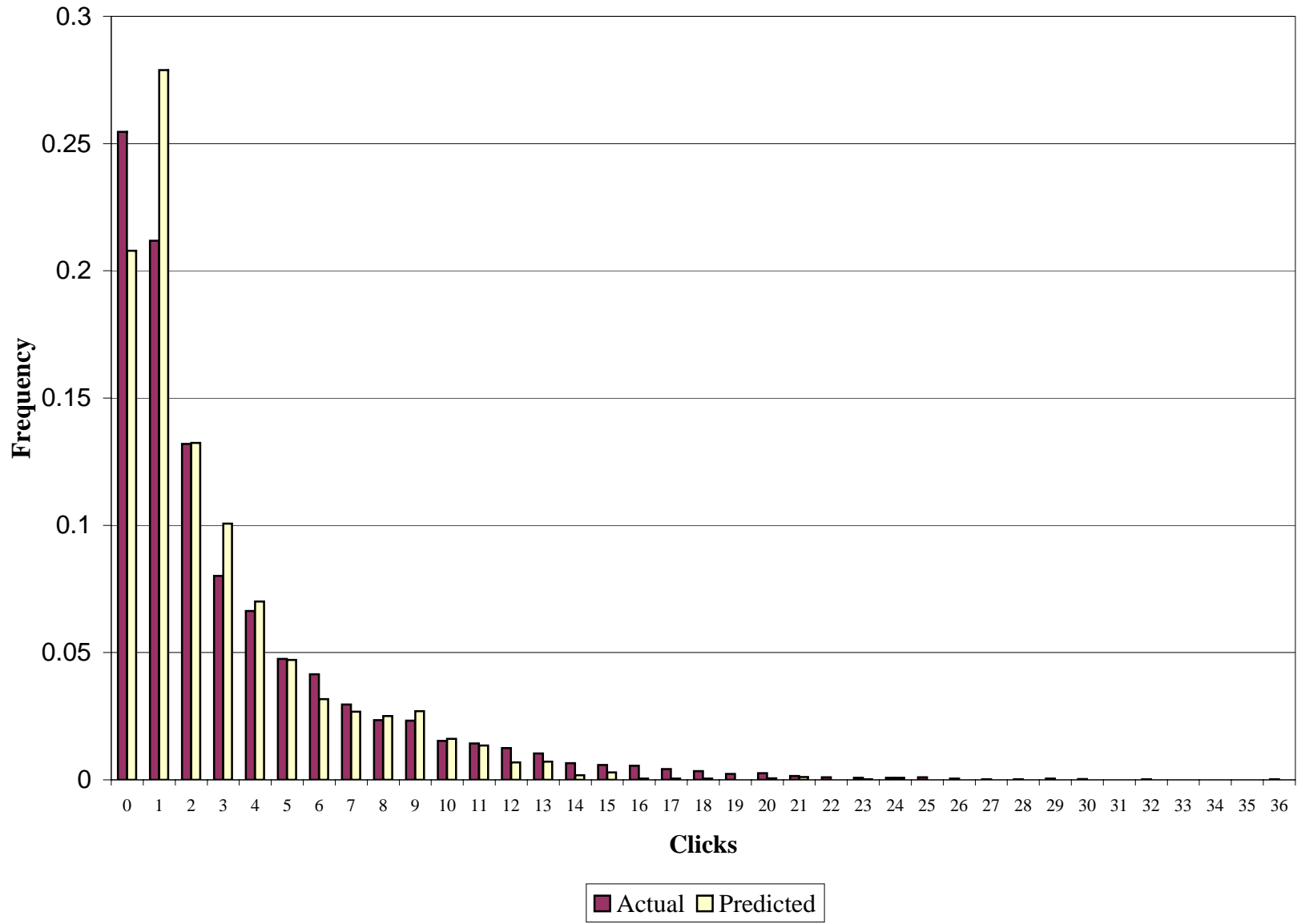
Table 1: Descriptive Statistics

Variable	Mean	Standard Deviation	First Quartile	Median	Third Quartile	Maximum	Minimum
Clicks	3.33	4.27	0	2	5	36	0
Price	304.88	106.84	229.99	279.98	396.63	601.95	104.57
Shipping	4.16	4.50	0	3.95	5.82	17.63	0
Total Price	309.04	107.01	234.42	283.94	396.63	607.77	108.10
Number of Sellers	4.05	2.93	2	3	6	15	1
Location on Screen	3.40	2.43	1	3	5	15	1
Bricks and Clicks Retailer	0.29						
Weekend	0.28						
September	0.11						
October	0.29						
November	0.29						
December	0.27						
January	0.05						
Total Number of Products	18						
Total Number of Firms	19						
Total Number of Dates	111						
Total Number of Observations	6151						

Table 2: Product Specific Estimates

Likelihood Specification for Clicks: Poisson PML

Product	Log Total Price	Position on Screen	Weekend	Month Dummies	# of Obs.	Average Number of Sellers
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%.

<div align="center">Table 3: Continuous Clicks Specifications</div>

	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 Sellers − 1)		-0.288
		(4.14)**
Number of Sellers		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
Overdispersion Test Test Statistic	2656.46	2488.77
P-Value	0	0

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

<div align="center">Table 4: Discontinuous Clicks Specifications</div>

	Model 1	Model 2	Model 3	Model 4
Likelihood Specification for Clicks	Poisson PML	Poisson PML	Poisson CML	Poisson CML
Log Total Price	-2.459	-2.386	-2.446	-2.449
	(9.11)**	(9.64)**	(23.78)**	(23.75)**
Log Total Price x (Number of Sellers − 1)	-0.252	-0.289	-0.175	-0.173
	(4.60)**	(5.44)**	(9.94)**	(9.83)**
Jump from Shoppers	0.603	0.591	0.599	0.6
	(7.11)**	(6.61)**	(26.60)**	(26.55)**
Number of Sellers	1.415	1.614	0.98	0.97
	(4.52)**	(5.31)**	(9.93)**	(9.82)**
Position on Screen	-0.175	-0.174	-0.149	-0.149
	(4.37)**	(4.46)**	(21.61)**	(21.32)**
Bricks and Clicks Retailer	0.321	0.317	0.367	
	(2.41)*	(2.43)*	(-1.87)	
Weekend	-0.268	-0.272	-0.263	-0.263
	(13.79)**	(14.24)**	(15.44)**	(15.43)**
Product Dummies	17	17	17	17
Month Dummies	4	4	4	4
Product x Month Dummies	55	55	55	55
US Product Rank Dummies	No	52	52	52
Robust Standard Errors Clustered by Firm	Yes	Yes	No	No
Controls for Unobserved Firm Heterogeneity	No	No	19	19
			Random Effects	Fixed Effects
Observations	6151	6151	6151	6151
Overdispersion Test — Test Statistic	1942.27	1822.4		
Overdispersion Test — P-Value	0	0		

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

Table 5: Clicks Specifications with Cross Price Effects

	Model 1	Model 2
Likelihood Specification for Clicks	Poisson PML	Poisson PML
Log Total Price	-3.311	-2.091
	(5.97)**	(5.99)**
Log Total Price x (Number of Sellers − 1)	-0.358	-0.309
	(4.74)**	(5.12)**
Jump from Shoppers		0.609
		(6.51)**
Log Geometric Mean of Rival's Prices	0.809	0.38
	(1.89)	(1.21)
Absence of Rivals Dummy	4.537	1.839
	(1.92)	(1.04)
Number of Sellers	1.973	1.718
	(4.60)**	(4.99)**
Position on Screen	-0.174	-0.175
	(4.55)**	(4.49)**
Bricks and Clicks Retailer	0.237	0.326
	(1.73)	(2.48)*
Weekend	-0.271	-0.272
	(12.20)**	(13.76)**
Product Dummies	17	17
Month Dummies	4	4
Product x Month Dummies	55	55
US Product Rank Dummies	52	52
Robust Standard Errors Clustered by Firm	Yes	Yes
Observations	6151	6151
Joint test of continuous cross-price effects		
Test Statistic	4.76	10.96
P-Value	0.093	0.004

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

Table A1: Continuous Clicks - 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 Sellers − 1)		-0.343
		(5.54)**
Number of Sellers		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

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

Table A2: Discontinuous Clicks - Alternative Specifications

	Model 1	Model 2	Model 3	Model 4
Likelihood Specification for Clicks	Negative Binomial ML	Negative Binomial ML	Negative Binomial CML	Negative Binomial CML
Log Total Price	-2.343	-2.304	-2.334	-2.333
	(8.18)**	(8.89)**	(16.72)**	(16.64)**
Log Total Price x (Number of Sellers − 1)	-0.314	-0.342	-0.168	-0.165
	(5.38)**	(5.95)**	(7.22)**	(7.09)**
Jump from Shoppers	0.619	0.608	0.578	0.579
	(8.24)**	(7.78)**	(18.92)**	(18.84)**
Number of Sellers	1.77	1.912	0.932	0.918
	(5.30)**	(5.82)**	(7.15)**	(7.02)**
Position on Screen	-0.166	-0.165	-0.145	-0.144
	(4.31)**	(4.35)**	(15.93)**	(15.57)**
Bricks and Clicks Retailer	0.324	0.319	0.272	
	(2.66)**	(2.64)**	(2.81)**	
Weekend	-0.29	-0.297	-0.247	-0.247
	(14.86)**	(15.13)**	(10.96)**	(10.95)**
Product Dummies	17	17	17	17
Month Dummies	4	4	4	4
Product x Month Dummies	55	55	55	55
US Product Rank Dummies	No	52	52	52
Robust Standard Errors Clustered by Firm	Yes	Yes	No	No
Controls for Unobserved Firm Heterogeneity	No	No	19	19
			Random Effects	Fixed Effects
Observations	6151	6151	6151	6151

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