Brand and Price Advertising in Online Markets*

Michael R. Baye

Indiana University

John Morgan

University of California, Berkeley

August 2008

Abstract

We model an environment where e-retailers sell similar products and endogenously engage in both brand advertising (to create loyal customers) and price advertising (to attract "shoppers"). In contrast to models where loyalty is exogenous, endogenizing the creation of loyal customers by allowing firms to engage in brand advertising leads to a continuum of symmetric equilibria; however, there is a unique equilibrium in secure strategies, and the set of equilibria converges to this unique equilibrium as the number of potential e-retailers grows arbitrarily large. Price dispersion is a key feature of all of these equilibria, including the limit equilibrium. Branding tightens the range of prices and reduces the value of the price information provided by a comparison site, and this reduces profits for platforms (such as an Internet price comparison site) where firms advertise prices. Data from a leading price comparison site are shown to be consistent with several predictions of the model. JEL Nos: D4, D8, M3, L13.
Keywords: Price dispersion

*We are grateful to the editor and two anonymous reviewers for very insightful comments and suggestions. We also thank Rick Harbaugh, Ganesh Iyer, Peter Pan, Ram Rao, Karl Schlag, Michael Schwarz, Miguel Villas-Boas, as well as seminar participants at the IIOC Meetings, the SICS Conference 2004, Berkeley, European University Institute, and Indiana for comments on earlier versions of this paper. We owe a special thanks to Patrick Scholten for valuable input into the data analysis. Baye is grateful to the Universities of Bonn and Oxford for their hospitality. Morgan gratefully acknowledges support from the National Science Foundation through grant NES-0095639, as well as the able research assistance of Jennifer Brown. Address correspondence to: Baye (mbaye@indiana.edu): 1309 East Tenth Street, Kelley School of Business, Indiana University, Bloomington, IN 47401. Morgan (morgan@haas.berkeley.edu): Haas School of Business, Berkeley, CA 94720-1900.

1 Introduction

The size, scope, and persistence of online price dispersion for seemingly identical products has been amply documented.[1] Some have suggested that, while the products sold at price comparison sites may be identical and search costs low, e-retailers go to great pains to be perceived as different. For instance, Brynjolfsson and Smith (2000a) argue that price dispersion in markets for books and CDs is mainly due to perceived differences among retailers related to branding, awareness, and trust—factors influenced by the brand-building activities of online retailers. These activities include the prominent use of logos, clever advertising campaigns, the development of "customized" applications including one-click ordering, custom recommendations, and the development of an online "community" or "culture" loyal to a particular firm. Even on Internet price comparison sites, where consumers are price sensitive, some firms promote their "brand" by featuring their logo along with their advertised price. All of these activities are costly.

How do costly differentiation efforts—what we refer to as brand advertising—interact with firms' pricing and listing decisions—what we refer to as informational advertising—to affect competition and price dispersion in online markets? The existing literature on equilibrium price dispersion does not provide a ready answer; even the most recent literature (see, for instance Iyer, Soberman, and Villas-Boas, 2005) treats the fraction of consumers who are "loyal" to some firm as exogenous. One can imagine that endogenizing brand-building might matter a great deal. If brand advertising ultimately converted all consumers into "loyals," firms would find it optimal to charge the "monopoly" price and price dispersion would vanish. Expressed differently, it is not at all clear that dispersed price equilibria of the sort characterized in the extant literature (see footnote 4) survive when customer loyalty is endogenously determined by firms' marketing mix.

In Section 2 we extend the Baye and Morgan (2001) model to now include endogenous brand advertising—in short, to endogenize several key components of a firm's marketing mix including branding, pricing, and informational advertising. In the model, a fixed number of firms sell similar products. Each firm invests in brand advertising in an attempt to convert some or all consumers

[1] See Elberse, *et al.* (2002); Baye, Morgan, and Scholten (2006); and Ellison and Ellison (2005) for useful surveys of price dispersion in the marketing and economics literatures.

into "loyals." These branding decisions result in an endogenous partition of consumers into "loyals", who are loyal to a specific firm, and "shoppers", who view the products to be identical. Given their stock of loyals, firms independently make pricing decisions as well as decisions about informational advertising.

In Section 3, we characterize the set of symmetric Nash equilibria of the model. In contrast to models where the number of loyal consumers is exogenous, endogenous branding leads to multiple equilibria. In all equilibria branding efforts by firms create a significant number of loyal consumers. However, endogenous branding does not eliminate equilibrium price dispersion, although increased branding is associated with lower levels of price dispersion. We then refine the set of equilibria in two ways. First, we show that, if one restricts attention to secure strategies, there is a unique symmetric equilibrium.[2] Second, we show that if one considers the limit game as the number of competing firms grows arbitrarily large, one again obtains a unique equilibrium in the limit. This equilibrium corresponds to the secure strategies refinement.

Surprisingly, even in the limit equilibrium where the number of potential competitors is "large" (as is the case in global online markets), prices remain dispersed above marginal cost. This finding is in contrast to the models of Varian (1980), Rosenthal (1980), Narasimhan (1988), which all predict that price dispersion vanishes as the number of potential competitors grows large. Our findings for large online markets are also broadly consistent with daily data we have been collecting for several years and post weekly at our website, *Nash-Equilibrium.com.* Price dispersion, as measured by the range in prices, has remained quite stable over the past eight years, at 35 to 40 percent. The stability and magnitude of this dispersion is remarkable from a theoretical perspective, since (1) the products are relatively expensive consumer electronics products for which the average price is about \$500, (2) over the period, the Internet rapidly eliminated geographic boundaries leading to exponential growth in the number of consumers and businesses with direct Internet access, and (3) according to the Census Bureau, there were nearly 10,000 consumer electronics retail establishments in the United States who compete in the consumer electronics market.[3] Our model provides the

[2] Recall that secure branding strategies maximize the minimum possible payoff that can be imposed on a player during the second-stage pricing game.

[3] This figure is based on NAICS classification code 443112, which is comprised of establishments known as consumer electronics stores primarily engaged in retailing new consumer-type electronic products. Source: U.S. Census Bureau,

first equilibrium rationale for how so many firms could compete in such a price sensitive arena and yet have prices remain dispersed above marginal cost.

In Section 4, we derive several key testable comparative static implications of the model, and examine them using data from Shopper.com in Section 5. We find that more intense branding by firms is associated with lower levels of price dispersion and higher prices to loyals and shoppers. These results are robust to a variety of controls.

2 The Environment

Consider an online market where a unit measure of consumers shop for a specific product (e.g., HP LaserJet 1100xi). There are $N \geq 2$ sellers in this market, each having a constant marginal cost of m. Each consumer is interested in purchasing at most one unit of the product, from which she derives value v.[4] The impact of loyalty on consumer behavior is similar to that in the models of Rosenthal (1980), Narasimhan (1988), Raju and Lal (1990), and Iyer *et al.* (2005). Specifically, we assume there are two consumer segments: loyals and shoppers.[5] Shoppers costlessly visit the price comparison site to obtain a list of the prices charged by all firms choosing to list prices there.[6] Since shoppers view sellers as perfect substitutes, they each purchase at the lowest price available at the price comparison site—provided it does not exceed v. If no prices are listed, these shoppers visit the website of a randomly selected firm and purchase if the price does not exceed v.[7] Loyals visit their preferred site directly and purchase if the price does not exceed v.

In contrast to Baye and Morgan (2001), who assume that all consumers view firms as identical, and Raju and Lal (1990), who assume that all consumers have a preference for a product sold by a particular firm, we allow for the possibility that some consumers have a preference for particular

1997 Economic Census, January 5, 2001, p. 217.

[4] The model readily extends to the case where there are positive fixed costs and downward sloping demand.

[5] An alternative approach (see Raju and Lal, 1990) assumes that all consumers are loyal to some firm, but consumers vary in the price differential at which they will switch to the competing brand. More recently, Johnson, Bellman and Lohse (2003) have suggested that loyalty can also arise from network effects.

[6] Baye and Morgan (2001) show that a monopoly "gatekeeper" that owns a price comparison site has an incentive to set consumer subscription fees sufficiently low in an attempt to induce all consumers to utilize the site. Hence, we assume that all shoppers have access to the comparison site at no cost. This assumption is consistent with empirical evidence; virtually all price comparison sites—including Shopper.com, Nextag, and Dealtime—permit consumers to browse their sites at no charge.

[7] See Baye and Morgan (2001) for a detailed description of the mild assumptions required for this to comprise an optimal sequential search strategy.

sellers while other do not. There is considerable evidence that this is indeed the case. For instance, many consumers prefer to purchase books from Amazon rather than Barnes and Noble—even at higher prices.[8] To capture these effects, let β_i denote the proportion of consumers who are loyal to firm i. Thus, the total measure of consumers loyal to some firm is $B = \sum_{i=1}^{N} \beta_i$. The remaining $1 - B$ shoppers view the sellers as identical. Brand advertising creates loyal customers through the advertising response function $\beta(a_i, A_{-i})$,where $A_{-i} = \sum_{j \neq i} a_j$ denotes aggregate brand advertising by all firms other than i. To obtain closed-form solutions for firms' equilibrium branding levels, we assume

$$\beta(a_i, A_{-i}) = \delta \frac{a_i}{A_{-i} + a_i} + a_i \sigma \tag{1}$$

and where $\sigma > 0$ and $1 > \delta > 0$ are parameters.[9] The "δ" term in equation (1) captures potential "brand stealing" effects of brand advertising—brand advertising that steals loyal customers from other sellers. The "σ" term captures "brand expansion" effects—brand advertising that converts some shoppers into loyals. The response function in equation (1) is a generalization of Moorthy's (1993) oligopoly version of the Little (1970) ADBUDG model; the Moorthy response function allows only for brand stealing effects, and obtains as the special case of equation (1) when $\sigma = 0$.

There are three components to a firm's strategy: Firm i must decide its price (denoted p_i), its informational advertising strategy, which is modeled as a binary decision to spend $\phi > 0$ to list its price on the price comparison site (or not) in each period, and its current brand advertising level which influences its stock of loyal consumers through equation (1).

Firms' incentives to engage in branding activities depend on the sensitivity of β_i to branding efforts (that is, the magnitude of δ, σ, and the aggregate branding efforts of rival firms) and brand advertising costs. For simplicity, we assume that the marginal cost of a unit of brand advertising is $\tau > 0$, so that the total cost to firm i of a_i units of brand advertising is τa_i. Finally, we assume that $a_i \in \left[0, \frac{1-\delta}{N\sigma}\right]$, which guarantees that aggregate branding efforts do not lead to more loyals than is feasible given the unit mass of consumers and the specification in equation (1).

[8] For instance, Chevalier and Goolsbee (2003) provide evidence that the demand for books at Barnes and Noble is about 8 times more elastic than that at Amazon.

[9] When $A_{-i} > 0$, positive branding effort is required for firm i to sustain *any* loyal consumers. For the case where $a_j = 0$ for all j, we assume that $\beta(a_i, A_{-i}) = \delta/N$; that is, we take the limit of $\beta(a_i, A_{-i})$ as the branding expenditures of all firms go to zero at the same rate to obtain the expression for $\beta(0, 0)$.

3 Equilibrium

To facilitate comparisons with standard models with a fixed number of loyals, we model branding and pricing decisions as a two-stage game. In the first stage, firms simultaneously choose brand advertising levels, a_i, in an attempt to create a stock of loyal consumers. In the second stage, after having observed first stage decisions, firms simultaneously make pricing and listing decisions. This sequencing of moves is consistent with evidence that in many online markets, firms adjust prices frequently and quickly, and there is considerable turnover in the identity of the firm offering the lowest price; see Ellison and Ellison (2005). In contrast, branding decisions typically require substantial up-front investments, which take time to mature into a sizeable base of loyal customers.[10] This structure also captures the "strategic uncertainty" present in firms' branding and pricing decisions. In particular, the value to a firm committing up-front resources on branding activity critically depends on its view of the competitiveness of the market for shoppers in the second-stage game. We first show in Proposition 1 that the strategic uncertainty present in this setting leads to a continuum of symmetric Nash equilibria. We then offer two methods for refining the set of equilibria to obtain a unique prediction. Specifically, we show that (1) there exists a unique symmetric equilibrium in which players employ *secure branding strategies*, and (2) all symmetric equilibria converge to the unique equilibrium in secure branding strategies as the number of competing firms grows arbitrarily large.

3.1 The Set of Symmetric Equilibria

As will be apparent in our characterization of equilibria, it is useful to define

$$a_L \equiv \frac{1}{(\tau - \sigma(v-m))}\frac{1}{N^2}\left(\sqrt{(N-1)\,\delta\,(v-m)} - \sqrt{\frac{N\phi}{N-1}}\right)^2$$

and

$$a_H \equiv \frac{1}{(\tau - \sigma(v-m))}\frac{1}{N^2}\left(\sqrt{(N-1)\,\delta\,(v-m)} + \sqrt{\frac{N\phi}{N-1}}\right)^2$$

We focus on equilibria in which firms actively utilize *both* informational and brand advertising. Of course, there are products where firms choose not to use both channels actively. In the context of

[10] The model may readily be extended to allow for dynamics. Details are available from the authors upon request.

the present model, these situations are simple to analyze. For products in which it is relatively easy (or inexpensive) to create loyal consumers—or it is too costly to advertise prices— firms will find it in their interest to use brand advertising exclusively and there will be no competition in prices. In this case all firms will charge the monopoly price, and the only question is the classical one regarding the optimal level of brand advertising. There are also products where firms eschew the brand advertising channel entirely owing to an unfavorable cost-benefit calculus of creating loyalty through brand advertising. In terms of our model, these markets are essentially a special case of Baye and Morgan (2001) and may be readily analyzed with that framework. Product markets in which both brand and price advertising channels are used require that the informational advertising channel be sufficiently attractive that firms find it in their interest to periodically advertise prices at the price comparison site, and that brand advertising be sufficiently expensive that firms do not find it in their interest to use this channel exclusively. For this reason, we assume:

Condition 1 $\phi \in \Omega \equiv \left\{\phi : \phi < \frac{N-1}{N}(v-m)(1-\delta-N\sigma a_H)\right\}$ *and* $\tau > \frac{(v-m)\sigma}{1-\delta}$.

It is straightforward to show that the set of parameter values satisfying Condition 1 is non-empty—even in the limit as N goes to infinity. Note, however, that if either part of Condition 1 fails, i.e. when $\phi > \frac{N-1}{N}(v-m)(1-\delta-N\sigma a_H)$ or $\tau < \frac{(v-m)\sigma}{1-\delta}$, the equilibrium is degenerate: Firms find it in their interest to exclusively use the brand advertising channel and charge the monopoly price (v) in equilibrium.

We now provide a complete characterization of the set of symmetric equilibria arising when Condition 1 holds. In the sequel, let α_i denote the probability that a firm engages in informational advertising (i.e., lists its price on the comparison site), and let $F_i(p)$ to represent the distribution of firm i's listed price when this channel is active.

Proposition 1 *There exists a continuum of symmetric equilibria when brand and informational advertising is endogenous. In any symmetric equilibrium:*

Each firm chooses branding level $a \in [a_L, a_H]$*, which generates*

$$\beta_i = \beta = \frac{\delta}{N} + \sigma a$$

loyal consumers per firm. The total measure of loyal customers in the market is $B \equiv N\beta \in (0,1)$. *Each firm lists its price on the price comparison site with probability*

$$\alpha_i = \alpha \equiv 1 - \left(\left(\frac{\phi}{(v-m)(1-N\beta)}\right)\left(\frac{N}{N-1}\right)\right)^{\frac{1}{N-1}} \tag{2}$$

and, conditional on listing, selects a price from the cumulative distribution function

$$F_i(p) = F(p) \equiv \frac{1}{\alpha}\left(1 - \left(\frac{(v-p)\beta + \phi\frac{N}{N-1}}{(1-N\beta)(p-m)}\right)^{\frac{1}{N-1}}\right) \tag{3}$$

over the support $[p_0, v]$ *where*

$$p_0 = m + \frac{(v-m)\beta + \frac{\phi}{(N-1)}N}{(1-(N-1)\beta)}.$$

Firms that do not advertise a price at the price comparison site charge a price of $p_i = v$ *on their own websites. Each firm earns equilibrium profits of*

$$E\pi_i = E\pi = (v-m)\beta + \frac{\phi}{N-1} - \tau a. \tag{4}$$

Proposition 1, which is proved in the online supplement, shows that multiple equilibria arise in the presence of endogenous branding. Nonetheless, all of the equilibria have the property that branding efforts by firms convert some but not all consumers into loyals; in equilibrium, there remain $1 - B > 0$ shoppers who purchase from the firm charging the lowest price listed at the comparison site. This prediction appears consistent with empirical findings that some, but not all, online consumers buy at the lowest listed price.

The equilibria identified above share features present in the models of Iyer *et al.*, Raju and Lal, Varian, Rosenthal, Narasimhan, and Baye-Morgan—as well as some important differences. Similar to all of these models, equilibria in the present model require any firm advertising a price on the price comparison site to use a pricing strategy that prevents rivals from being able to systematically predict the price offered to consumers who enjoy the information posted at the site (hence the distributional strategy, $F(p)$). Like Baye-Morgan, our model permits firms to endogenously determine whether to utilize the price comparison site (the other models constrain all firms to list prices at the site with probability one, and Baye-Morgan essentially show that this is not an equilibrium when it is costly for firms to advertise prices at the site). As a consequence, in any equilibrium firms must randomize the timing of price advertisements to preclude rivals from systematically determining the number of listings at the price comparison site (hence, the informational advertising propensity, $\alpha \in (0,1)$).

In contrast to Iyer *et al.*, Raju and Lal, Narasimhan, and Rosenthal, the present model relaxes the assumption that firms are costlessly endowed with an exogenous number of brand-loyal consumers. In the present model, a firm that spends nothing to promote its "brand" or "service" in the face of positive expenditures by rivals enjoys no loyal consumers. In contrast to Varian and Baye-Morgan, the present model does not impose the assumption that all consumers view the products sold by different firms to be identical; indeed, in equilibrium, each firm enjoys a strictly positive measure of loyal consumers—thanks to the positive level of branding activity that arises in equilibrium. A key managerial implication is that the price comparison site attracts fewer consumers than in the Baye-Morgan model. Expressed differently, the branding efforts of firms reduce the traffic enjoyed by the "information gatekeeper" operating the price comparison site.

Another difference between existing models and the present one is that, in the former, there is a unique symmetric equilibrium while, in the latter, endogenous branding leads to a continuum of symmetric equilibria. As discussed below, however, only one of these equilibria survives two reasonable refinements.

3.2 The Unique Equilibrium in "Secure" Strategies

The presence of a continuum of equilibria gives rise to a coordination problem: How do firms determine which "branding equilibrium" to play? The set of symmetric equilibria can be payoff-ordered from highest $(a = a_L)$ to lowest $(a = a_H)$, and the equilibria differ in terms of the payoff risk to which firms are exposed. In this respect, these equilibria resemble those of the coordination games studied both theoretically and experimentally by Van Huyck, Battalio, and Beil (1990).[11] They find experimental evidence that subjects tend to adopt secure strategies when faced with coordination games of this type; thus, it seems natural to compare the symmetric equilibria identified in

[11] Specifically, Van Huyck *et al.* study behavior in minimum action games where the "lowest" choice made by any of the players determines the payoff to all players. Our second-stage game shares exactly this feature. Since shoppers only purchase from the firm offering the lowest price, the "minimum action" of firms in terms of price setting determines the payoffs to all firms in this stage of the game.

Proposition 1 in terms of their security properties.

Notice that, when rivals choose branding levels $a_j = a$ in the first stage, the lowest payoff that can be imposed on firm i is

$$E\pi_i^{\text{secure}}(a_i, A_{-i}) = (v-m) \times \beta(a_i, A_{-i}) - \tau a_i.$$

That is, firm i can do no worse than to eschew informational advertising ($\alpha_i = 0$) and charge the monopoly price to its loyal customers ($p_i = v$) regardless of its perceptions about the competitiveness of the market for shoppers. Substituting for $\beta(a_i, A_{-i})$ yields

$$E\pi_i^{\text{secure}} = \left(\delta \frac{a_i}{(N-1)a + a_i} + \sigma a_i\right)(v-m) - \tau a_i.$$

The brand advertising level that maximizes i's secure payoff satisfies the first-order condition

$$\left(\delta \frac{(N-1)a}{(a_i + (N-1)a)^2} + \sigma\right)(v-m) - \tau = 0. \tag{5}$$

It is routine to show that these first-order conditions imply:

Proposition 2 *There exists a unique symmetric equilibrium in secure branding strategies, denoted* $a^* \in (a_L, a_H)$. *Specifically, (1) firms choose brand advertising levels*

$$a_i = a^* \equiv \delta \frac{(N-1)(v-m)}{N^2(\tau - (v-m)\sigma)}$$

to obtain

$$\beta_i = \beta^* = \frac{\delta}{N}\left(\frac{N\tau - (v-m)\sigma}{N(\tau - (v-m)\sigma)}\right)$$

loyal consumers per firm; and (2) firms follow the second stage pricing and informational advertising strategies described in Proposition 1.

Note that, while Proposition 2 is derived for a particular advertising response function and assumes a constant marginal cost of brand advertising, similar results obtain for alternative advertising response functions, $\beta(a_i, A_{-i})$, and nonlinear brand advertising costs, $C(a_i)$—provided $E\pi_i^{\text{secure}}(a_i, A_i)$ is strictly concave in a_i. In short, one can extend the model to allow for a variety of alternative functional forms but at the possible cost of the absence of closed-form solutions for equilibrium levels of brand advertising.

For future reference, we let $B^* = N\beta^*$ and use α^*, F^*, p_0^* and $E\pi^*$ to denote the relevant second-stage components of the equilibrium identified in Proposition 2. Together, these components comprise what we shall hereafter refer to as an a^* equilibrium.

3.3 The Unique Limit Equilibrium

As noted in the introduction, there are over 10,000 firms in the U.S. that could list prices for consumer electronics products at a price comparison site like Shopper.com. It is therefore of interest to examine characteristics of online markets where a "large" number of firms compete. We next show, in Proposition 3, that studying limiting behavior refines the set of equilibria—there is a unique symmetric equilibrium level of brand advertising converging to the a^* equilibrium as $N \to \infty$. That is, the coordination problem is less severe in "large" online markets: all symmetric equilibria are arbitrarily close to the equilibrium identified in Proposition 2. This proposition is proved in the online supplement as well.

Proposition 3 *In any symmetric equilibrium, first-stage branding levels converge to a^* as the number of competing firms (N) grows arbitrarily large. Formally, let $\langle a_N, \alpha_N, F_N \rangle$ be an arbitrary sequence of symmetric equilibria. Then*

(1) $\lim_{N\to\infty} a_N = \lim_{N\to\infty} a^*$, *and*

(2) $\lim_{N\to\infty} Na_N = \lim_{N\to\infty} Na^* = \frac{\delta(v-m)}{\tau-\sigma(v-m)}$.

Importantly, the unique limit equilibrium is nontrivial in the sense that it displays both price dispersion and finite numbers of firms (in expectation) using the informational advertising channel. To see this, note that the number of potential competitors, N, generally exceeds the actual number of firms listing prices at any instant in time. In particular, given that each firm lists a price with probability α^*, the actual number of listings is a binomial random variable with mean,

$$\overline{n} = N\alpha^* < N.$$

It is straightforward to verify that

$$\lim_{N\to\infty} \beta^* = \lim_{N\to\infty} E\pi^* = 0.$$

This implies that, in markets where N is large, each firm enjoys a negligible share of the loyal consumers and essentially earns zero economic profits. Thus, the environment we study in this section shares two features of competitive markets: (1) each firm is small relative to the total market, and (2) firms earn zero equilibrium profits.

However, even though firms earn zero economic profits in the limit, the resulting equilibrium does not entail marginal cost pricing. In fact, prices remain dispersed and exceed marginal cost with probability one when the number of competitors becomes arbitrarily large. The reason is that each firm engages in less branding and attracts fewer loyals as N increases, Proposition 3 implies that *aggregate* branding converges to

$$A^L = \lim_{N\to\infty} Na^* = \delta \frac{(v-m)}{\tau - (v-m)\,\sigma} > 0.$$

This, in turn, implies that the aggregate number of loyals is given by

$$B^L = \lim_{N\to\infty} N\beta^* = \frac{\delta\tau}{\tau - (v-m)\,\sigma} < 1.$$

Thus, a positive measure of shoppers remains in the market even as the number of competing firms engaging in branding grows arbitrarily large. Furthermore, in the limit, the expected number of price listings at the comparison site is

$$\overline{n}^L = \lim_{N\to\infty} \overline{n} = \ln\left(\frac{(v-m)\,((1-\delta)\,\tau - (v-m)\,\sigma)}{\phi\,(\tau - (v-m)\,\sigma)}\right),$$

which is positive and finite since

$$\begin{aligned} \phi &< \lim_{N\to\infty}\left(\frac{N-1}{N}\,(v-m)\,(1-\delta-N\sigma a_H)\right) \\ &= (v-m)\left(\frac{\tau\,(1-\delta)-\sigma\,(v-m)}{\tau-\sigma\,(v-m)}\right) \end{aligned}$$

by Condition 1. In other words, even in online environments where a large number of firms could potentially advertise prices at the comparison site, the average number of listings on the site at any point in time is quite modest in size.

Finally, note that prices remain dispersed and above marginal cost even as the number of firms grows arbitrarily large. The limiting distribution of advertised prices is given by

$$F^L\,(p) = \lim_{N\to\infty} F^*\,(p) = \frac{\ln\left(\frac{\phi(\tau-(v-m)\sigma)}{(p-m)((1-\delta)\tau-(v-m)\sigma)}\right)}{\ln\left(\frac{\phi(\tau-(v-m)\sigma)}{(v-m)((1-\delta)\tau-(v-m)\sigma)}\right)}$$

on $\left[p_0^L, v\right]$, where

$$p_0^L = m + \frac{\phi\left(\tau - (v-m)\,\sigma\right)}{(1-\delta)\,\tau - (v-m)\,\sigma}.$$

To summarize:

Proposition 4 *In online markets where an arbitrarily large number of firms endogenously engage in both brand and informational advertising, there is a unique symmetric equilibrium. Furthermore:*

(1) *The average number of prices listed at the price comparison site is finite and is given by $\bar{n}^L$.*

(2) *The aggregate demand for brand advertising is finite and given by A^L.*

(3) *A non-negligible fraction of shoppers, $1 - B^L > 0$, remains in the market.*

(4) *Prices listed at the comparison site are dispersed according to F^L on a non-degenerate interval above marginal cost, $\left[p_0^L, v\right]$.*

Our calibrations (discussed in detail in the online supplement) reveal that convergence to the unique symmetric equilibrium is smooth and very rapid; the distribution of advertised prices in an a^*equilibrium when $N = 20$ is virtually indistinguishable from the limit equilibrium. In short, when the number of sellers that could potentially list prices at a comparison site is large, equilibrium branding levels are necessarily close to a^*. For these reasons, we summarize the key comparative static properties of the unique a^* equilibrium in the table below (supporting details are provide in the online supplement):

Variable	δ	σ	τ	N	ϕ	v	m
$E\pi^*$	$+$	0	0	$-$	$+$	$+$	$-$
a^*	$+$	$+$	$-$	$-$	0	$+$	$-$
β^*	$+$	$+$	$-$	$-$	0	$+$	$-$
B^*	$+$	$+$	$-$	$+$	0	$+$	$-$
p_0^*	$+$	$+$	$-$	?	$+$	$+$	?
α^*	$-$	$-$	$+$	?	$-$	?	?

where "?" denotes comparative static results that are ambiguous in sign.

4 Testable Hypotheses

We are now in a position to derive four testable implications of our model for the patterns of prices observed in online markets. Two of these predictions are general properties of the model while the other two are derived based on calibrated parameter values. Details of the calibration are offered in the online supplement.

We begin by showing that the equilibrium distribution prices listed on the price comparison site may be stochastically ordered by the intensity of branding activity.

Proposition 5 *In any symmetric equilibrium, the distribution of advertised prices in markets where firms create more loyal consumers first-order stochastically dominates that in markets where firms create fewer loyal consumers.*

Proposition 5, which is proved in the online supplement, implies that both the average price and, for a given number of price listings, the expected minimum price listed at a price comparison site are increasing in the branding efforts of firms.

Next, we turn to the impact of branding on the level of online price dispersion. Recall that an a^* equilibrium entails a nondegenerate distribution of prices. One of the more widely used measures of dispersion for online markets is the range, which we operationalize as the support of the price distribution. This may be written (using Propositions 1 and 2) as

$$R^* = v - p_0^* = \frac{(v-m)(1-\beta^* N) - \frac{\phi}{(N-1)}N}{(1-(N-1)\beta^*)}.$$

This permits us to establish:

Proposition 6 *In an a^* equilibrium, price dispersion, measured by the range, is greater in online markets where (1) it is less costly to advertise prices at the comparison site; or (2) it is more costly or more difficult to create loyal customers through brand advertising. More generally, in any symmetric equilibrium, price dispersion, measured by the range, is greater in markets where firms create fewer loyal consumers.*

Part (1) of this proposition follows from the fact that, other things equal, a reduction in the cost of informational advertising increases the potential profits from listing prices at the price comparison site relative to the profits from not listing. Since in equilibrium firms are indifferent between listing prices and not, firms compete away these potential profits by pricing more aggressively at the price comparison site. This reduces the lower support of the price distribution, thus increasing the range in prices.

Part (2) stems from the impact of reduced branding incentives on the total number of loyal consumers in the online marketplace. Increases in τ (or decreases in δ and/or σ) induce each firm to spend less on branding. In equilibrium, this reduces the total number of loyal consumers in the market, thereby heightening competition for the resulting larger number of shoppers. This heightened competition reduces the lower support of the price distribution and again the price range increases. In short, higher levels of price dispersion (measured by the range) are associated with *more* competitive pricing online.

How do these observations translate into testable hypotheses? We begin by considering price dispersion. It is worth noting that even in markets where there are no branding activities (when $\delta = 0$), the model predicts that prices are nonetheless dispersed: The range of observed prices is predicted to be non-degenerate even for products in which there are no loyal consumers. Recall that Proposition 6 implies that the range in prices, defined as the difference between the upper and lower supports of the equilibrium price distribution, is decreasing in firms' branding activities. While one cannot directly observe the upper and lower supports of the distribution, one can observe the *sample range*, which is defined as the difference between the highest and lowest prices listed on the comparison site. In the online supplement, we show that for calibrated parameter values of the model, the sample range is also decreasing in firms' branding activities (see Figure 1 in the online supplement). Thus,

Prediction 1 *For calibrated parameter values of the model, when brand advertising intensity is higher, price dispersion is lower.*

Next, recall that Proposition 5 implies that advertised prices are stochastically ordered. Hence, the average price listed at the price comparison site, as well as the average minimum price, is an increasing function of firms' branding intensities. Thus,

Prediction 2 *When brand advertising is higher, average listed prices are also higher.*

Prediction 3 *When brand advertising is higher, the average minimum listed price is also higher.*

The economic motivation for focusing on these two predictions stems from the fact that the average listed price and the average minimum price are related to the prices paid by loyal consumers and shoppers. Other things equal, higher average listed prices imply higher transactions prices for loyal consumers, and higher average minimum prices imply higher prices paid by shoppers who purchase products online. Note that the difference in these two average prices reflects the average savings of a consumer who purchases at the "best" listed price rather than the average listed price. Thus, $Ep - Ep_{\min}$ provides one measure of the value of the price information provided by a price comparison site. The calibrations in the online supplement also imply that this measure of the value of information is decreasing in firms' branding activities (see Figure 1 in the online supplement). Thus,

Prediction 4 *For calibrated parameter values of the model, when brand advertising intensity is higher, the value of price information is lower .*

5 Evidence from Shopper.com

To examine these predictions, we assembled a dataset for 90 of the best-selling products sold at Shopper.com during the period from 21 August 2000 to 22 March 2001. While this dataset is rich in price and product information, it suffers from the limitation that cannot directly observe branding intensity by firms. Instead, we observe a crude proxy—whether the firm has chosen to display a logo on the Shopper.com site or not—that is likely to be correlated, albeit imperfectly, with other branding activities. Moreover, while firms are symmetric in the theory model, this is unlikely to be true of firms in the data. Before we describe the econometric approach we use to attempt to deal with these limitations, we describe the data and the Shopper.com environment in more detail.

During this period, Shopper.com was the top price comparison site for consumer electronics products (including specific brands of printers, PDAs, digital cameras, software, and the like). A consumer wishing to purchase a specific product (identified by a unique part number) could query the site to obtain a page view that includes a list of sellers along with their advertised price. "Shoppers" can easily sort prices from lowest to highest and, with a few mouse clicks, order the product from the firm offering the lowest price. "Loyals," on the other hand, can easily sort sellers alphabetically or scan the page for their preferred firm's logo and click through to purchase the item from that firm.

We used a program written in PERL to download all the information returned in a page view for each of the products each day, which amounted to almost 300,000 observations over the period. During the period of our study, firms uploaded their prices into Shopper.com's database, which then fetched the uploaded data at specified times twice each day. Thus, daily pricing decisions reflect simultaneous moves. Moreover, there is a minimum twelve hour lag for any firm to "answer" a pricing move by its rival owing to the upload/refresh cycle. To advertise a product price, a merchant was required to pay a fixed fee of $1,000 to set up an account at Shopper.com, plus an additional fee of $100 per month. This fee structure provides merchants incentives to post accurate prices; a firm advertising a bogus price in an attempt to lure customers to its own website would generate many qualified leads, but would likely alienate potential customers and incur additional costs.[12] We also verified the accuracy of prices via an audit; more than 96 percent of the prices audited at Shopper.com were accurate within \$1.

In addition to Predictions 1-4, the equilibrium characterization offered in Proposition 1 suggests a number of other stylized facts about equilibrium pricing and listing decisions on the comparison site. These implications, which are shared with many "clearinghouse models" (see Baye, Morgan, and Scholten, 2004a), have been shown to be consistent with pricing patterns observed at Shopper.com as well as other price comparison websites. These include: (a) ubiquitous and persistent price dispersion using a variety of price dispersion measures; (b) turnover in the identity of the firm offering the lowest price; (c) discontinuities in a firm's demand above and below the lowest price offered by a rival; (d) turnover in the identities of the firms listing on the site ($\alpha < 1$). Baye, Morgan, and Scholten (2006) offer a survey of these and other empirical findings related to clearinghouse models. In light of the existing evidence, we focus on Predictions 1-4, which are unique to the introduction of branding decisions.

Table 1 provides basic summary statistics for these data averaged over all products and dates; henceforth, product-dates.[13] On average, 29 firms listed prices for each product and, on average, 8.29 percent of these firms advertised using a logo along with their price listing.[14] While the average price of a product was \$458.86, there is considerable variation in the prices different firms charge for a given product. The average lowest price is \$387.58, while the average highest price charged is \$555.11. The average level of price dispersion is substantial, with an average range of \$167.53. Finally, the average range is fairly stable and quite sizeable during the period of our study.

5.1 Estimation Strategy and Results

The theory presented above suggests that, for each product i and date t, the range (R_{it}) and average prices (Ep_{it} and $Ep_{\min,it}$) are nonlinear functions of product characteristics (such as the marginal cost of the product, m_{it}), consumer demand characteristics (such as v_{it}), the level of branding

[12] The \$100 monthly fee entitled sellers to up to 200 free clickthroughs from consumers per month. Sellers who exceed this threshold incur a cost on the order of 50 cents per clickthrough.

[13] The number of product-dates listed in Table 1 is less than what simple math would suggest (90 products × 205 days = 18,400 product dates) due to product life-cycle effects. That is, products naturally drop out of the sample over time due to the introduction of new models or product upgrades.

[14] Since our identification strategy relies on variation in branding intensity, cross-sectional variation in firm branding decisions is essential. Below we provide an econometric rationale for cross-sectional variation in the "symmetric" model.

(or alternatively, β_{it}), and the number of firms in the market for product i in period t (N_{it}). For example, using the distribution of advertised prices in an a^* equilibrium and integrating by parts yields the following structural expression for the expected advertised price of product i in period t as a function of the relevant explanatory variables:

$$Ep_{it} = v_{it} - \int_{m_{it}+\frac{(v_{it}-m_{it})\beta_{it}+\frac{\phi_{it}}{(N_{it}-1)}N_{it}}{(1-(N_{it}-1)\beta_{it})}}^{v_{it}} \left[\frac{1-\left(\frac{(v_{it}-p)\beta_{it}+\phi_{it}\frac{N_{it}}{N_{it}-1}}{(1-N_{it}\beta_{it})(p-m_{it})}\right)^{\frac{1}{N_{it}-1}}}{1-\left(\left(\frac{\phi_{it}}{(v_{it}-m_{it})(1-N_{it}\beta_{it})}\right)\left(\frac{N_{it}}{N_{it}-1}\right)\right)^{\frac{1}{N_{it}-1}}} \right] dp \tag{6}$$

In light of the gross nonlinearities involved—and the fact that we only have proxies for some potentially important explanatory variables—our estimation strategy is to attempt to isolate the impact of branding on the variables of interest (Predictions 1-4) by controlling for other variables that theory suggests might influence the observed levels of price dispersion, average prices, and value of information. In what follows, we estimate a logarithmic first-order Taylor's series approximation of the nonlinear functional forms for the expected price, minimum price, and range of prices for product i at time t. Specifically, in light of the cross-sectional time series nature of our data, we use product dummies to control for the fact that consumers are likely to have very different reservation prices (v_{it}) for different products and firms most likely incur different marginal costs (m_{it}) in selling different products. To further control for potential heterogeneities in demand across products, we also include dummy variables for product popularity. Among other things, this controls for possibility that consumers have higher reservation prices for popular products, as well as the possibility that firms are more eager to sell such products. In order to control for the possibility that the general costs of e-retailing, the number of consumers with Internet access, or overall consumer demand for consumer electronics products (and hence reservation prices) temporally varied during the period of our study, we also include date dummies to control for potential systematic temporal differences in reservation prices and/or firms' cost. One of the advantages of the size of our dataset is that it permits us to include 205 date dummies for each day in our sample, 100 dummy variables to control for product popularity (the most popular product, the second most popular product, and so on), as well as 90 product dummies.

More specifically, assume

$$y_{jt} = \gamma_{0jt} + \gamma_1 X_{jt} + \gamma_2 a_{jt} + \eta_{jt} \tag{7}$$

where y_{jt} is the measure of interest (e.g., the logarithm of price range), X_{jt} is a set of the controls discussed above, and a_{jt} is the equilibrium level of branding advertising for product j at date t.

Even though, in equilibrium, each firm is hypothesized to engage in the same level of branding for a product, heterogeneities observed by firms but not the econometrician are likely to induce firms to use differing types of marketing activities to achieve a given level of branding. For instance, one firm might opt to pay programmers to develop web 2.0 features that improve shoppers' experiences on its site, while another firm might opt to advertise its brand on television or other old media outlets. By their very nature, marketing strategies involve horizontal differentiation to achieve branding effectiveness. While we cannot directly observe the full range of marketing activities firms use to create loyals,[15] we do observe one component—the decision to post a logo on the Shopper.com site.[16] Each firm's decision to display a logo or not on the Shopper.com site is viewed as an optimizing decision over its entire marketing mix as a function of various environmental and cultural variables observed by the firms but not by the econometrician.

The total brand advertising of firm i for product j at time t is composed of a part we observe (L_{ijt}) and an unobservable component $\left(\xi_{ijt}\right)$:

$$a_{jt} = \rho_{1jt} + \rho_2 L_{ijt} + \xi_{ijt} \tag{8}$$

where $\rho_2 > 0$ and ξ_{ijt} is an artifact of the unobserved heterogeneity. Here, $L_{ijt} = 1$ if firm i displayed its logo at Shopper.com for product j at time t and zero otherwise. The assumption that $\rho_2 > 0$ reflects the idea that logoing contributes positively to overall branding effort, while the presence ξ_{ijt} reflects the fact that it is not the only contributor. Notice that, while ξ_{ijt} is clearly correlated with L_{ijt}, from the view of the econometrician, these ξ_{ijt}'s are independently and identically distributed across firms with zero mean (without loss of generality, as any non-zero product-level component is absorbed in ρ_{1jt}).

Since equation (8) holds for all n_{jt} firms selling product j at date t,we can average over all i to obtain

$$\begin{aligned} a_{jt} &= \rho_{1jt} + \rho_2 \frac{\sum_{i=1}^{n_j} L_{ijt}}{n_{jt}} + \frac{\sum_{i=1}^{n_j} \xi_{ijt}}{n_{jt}} \\ &= \rho_{1jt} + \rho_2 FRAC_{jt} + \frac{\sum_{i=1}^{n_j} \xi_{ijt}}{n_{jt}} \end{aligned} \tag{9}$$

where $FRAC_{jt}$ is the fraction of firms posting logos for product j at date t.

[15] The majority of firms in our sample are privately held, and thus, the total amount of money firms spent on brand advertising is unobservable.

[16] Logo branding is consistent with the classical marketing definition in Keller (2002): "A brand is a name, term, sign, symbol, or combination of them that is designed to identify the goods or services of one seller or group of sellers and to differentiate them from those of competitors." (Keller (2002, page 152).

Substituting equation (9) into equation (7) yields

$$\begin{aligned} y_{jt} &= \gamma_{0jt} + \gamma_1 X_{jt} + \gamma_2 \left(\rho_{1jt} + \rho_2 FRAC_{jt} + \frac{\sum_{i=1}^{n_j} \xi_{ijt}}{n_{jt}} \right) + \eta_{jt} \\ &= \beta_{0jt} + \beta_1 X_{jt} + \beta_2 FRAC_{jt} + \frac{\sum_i \varepsilon_{ijt}}{n_{jt}} + \eta_{jt} \end{aligned} \tag{10}$$

where $\beta_{0jt} = \gamma_{0jt} + \gamma_2 \rho_{1jt}$, $\beta_1 = \gamma_1, \beta_2 = \gamma_2 \rho_2$, and $\varepsilon_{ijt} = \gamma_2 \xi_{ijt}$. This is the equation we estimate in the tables below.

Before turning to the results, we discuss several potential concerns stemming from the fact that (1) we only observe one component of a firm's mix of brand advertising; (2) errors are likely to exhibit autocorrelation and heteroskedasticity (3) the number of *potential* competitors in the market is unobservable and potentially endogenous; and (4) the dataset is restricted to firms listing prices at a specific price comparison site over a seven month period.

First, since $\beta_2 = \gamma_2 \rho_2$, the coefficient on $FRAC$ describes the marginal effect on the dependent variable of branding via the logoing channel—ρ_2 is not separately identified. However, since β_2 has the same sign as ρ_2, we can readily test the directional hypotheses derived above. Also, measurement error in a_{jt} stemming from the unobserved heterogeneity in the composition of branding decisions across firms implies that $\sum_i \varepsilon_{ijt}/n_{jt}$ may be correlated with $FRAC_{jt}$. This would produce attenuation bias and hence lead to coefficient estimates of β_2 that are biased towards zero. While attentuation bias disappears when the number of firms selling a given product on a given date is large (since $\lim_{n_{jt} \to \infty} \sum_i \varepsilon_{ijt}/n_{jt} = 0$), both attenuation and composition (when $\gamma_2 < 1$) lead to estimates of the marginal effect of brand advertising on the dependent variable that understate the magnitude of the true effect. Thus, the magnitude of the coefficient estimates for β_2 should be viewed as a lower bound of the economic significance of brand advertising on the various measures of pricing and price dispersion we consider.

Second, note that the error structure in equation (10) is likely to exhibit heteroskedasticity and potential autocorrelation owing to differences in the number of firms (and their patterns over time). For this reason, we adjust standard errors to account for these effects using the Newey-West error correction with a lag length of one period.

Third, while the number of potential firms is unobservable, it is statistically related to the observed number of listings on a given date. For this reason, we use the number of listings for product i on date t as a proxy for N_{it}. It is important to stress, however, that while the theoretical model presented above is an oligopoly model in which the number of sellers is taken to be exogenous, we are sympathetic to the possibility that firms' decisions to enter the online market for a particular product might be endogenous. Unfortunately, we do not have available instruments to correct for this potential endogeneity. However, the potential problem is mitigated to some extent by the fact that we include product rank dummies (which control to some extent for the possibility that more popular products attract more firms) and by the fact that every firm at Shopper.com must make its period t pricing decisions before it knows how many other firms have decided to compete on that date.

Finally, several factors led us to use data from the Shopper.com price comparison site during the seven month (205 day) period discussed above. First, there is considerable cross-sectional and time series variation in the use of logos by firms over this period, and this is necessary for identification in the presence of product fixed effects. Since then, both the online strategies of firms and the structure of the Shopper.com site have evolved in ways that make it more difficult to test the above predictions. Today, there is less cross-sectional variation in the use of logos (many more firms advertise their logos at Shopper.com). While this "structural change" in logoing would provide additional time-series variation that would be helpful in identifying the effects of branding, the composition of products sold on Shopper.com changed after April 2001, as searches began returning mixtures of new and refurbished products. This makes it difficult to determine whether any observed changes in price dispersion (or average prices) stem from increased product heterogeneity (comparing new versus used product prices) or increased brand advertising by firms. To eliminate this problem, we focus on the seven months in the present study, where Shopper.com treated new and refurbished versions of otherwise identical products as different products. In short, by restricting the sample to this set of products and time period, we ensured that all of the 90 products in our sample are new products (see the online supplement for a complete description of the products). However, a concern one might have about the Shopper.com dataset is that it consists entirely of firms that have self-selected to participate on the site. This concern, however, ignores the fact that the predictions of the model are valid only for firms and products where both the informational and brand advertising channels are active (i.e., where Conditions 1 and 2 hold). Hence, an appropriate sample is one that *selects* for firms actively using the informational advertising channel. In contrast, a sample drawn at random from the population of all online firms would not be appropriate for examining the hypotheses of the model developed above, as it would include firms that do not actively use the informational advertising channel.

With these caveats, we turn to the data analysis. In Tables 2-5 we report semi-log regression results that summarize the estimated impact of branding on, respectively, the sample range, average price, average minimum price, and the value of information.[17] For the reasons discussed above, all specifications include product dummies to control for unobserved components of branding and other factors that might give rise to systematic differences in the levels of prices across different products. We also include a variety of other controls to account for the impact of market structure, product life cycles, and other factors. In each table, Model 1 represents a baseline regression in which the dependent variable associated with product i at time t is regressed on branding activity, the number of firms listing prices on that date, and product dummies. Models 2 through 4 add controls for nonlinear number of firm effects, product popularity dummies, and date dummies, respectively. Popularity dummies are based on Shopper.com's Product Rank (which ranges from 1 to 100 for the products in our sample).

Table 2 examines whether price dispersion varies systematically with firms' branding efforts. Here, the dependent variable is the (log) sample range. In all specifications, the results indicate that, at the 1 percent significance level, price dispersion *negatively* covaries with branding. These results indicate that an increase in the fraction of logos from 8.29% to 9.29% decreases the price range by \$2.05 in Model 1 and \$3.20 in Model 4. These findings are consistent with Prediction 1.

Table 3 summarizes results for the (log) average price regressions. With the exception of Model 4, the estimates suggest that average prices positively covary with branding. The semi-log regression coefficients imply that an increase in the fraction of logos from the mean (8.29%) to 9.29% increases the average price by 42 cents in Model 1 and increases it by 41 cents in Model 3. The most general specification, Model 4, is at odds with Prediction 2. While the coefficient associated with branding is negative in that specification, it is not statistically significant.

Table 4 summarizes results for the (log) minimum price regressions. Minimum prices positively covary with branding and are significant at the one percent level in Models 1 through 3. These results indicate that an increase in the fraction of logos from 8.29% to 9.29% increases average minimum prices by 99 cents in Model 1 and \$1.02 in Model 3. These results are consistent with Prediction 3; however, the coefficient associated in the most general specification, Model 4, remains positive but loses statistical significance.

Why does the coefficient associated with branding in the most general specification lose significance and, in the case of the average price regressions, change sign? One possibility is that logo ads constitute only a small component of a firm's portfolio of branding activities, and the inclusion of the date dummies absorbs the remaining variation in the data. The key here is that the use of logos

[17] The results are robust to regressions based on levels rather than logs.

decreases over time in our sample. At the same time, price levels decline over the course of the sample, presumably due to the relatively short life cycles of consumer electronics products. Absent date dummies, the branding coefficient captures this time variation in prices thus giving rise to the positive coefficients in Models 1 through 3. Model 4 illustrates the importance of controlling for product life-cycle effects. Adding this control absorbs the time series variation in overall prices, reducing the precision of the estimated branding coefficient.

Notice that this issue does not arise in Model 4 of Table 2. In particular, this specification is based on the difference in the highest and lowest prices at each product date. To the extent that the life cycle effects for a given product are similar for both the highest and lowest prices, differencing the data eliminates individual product life cycle effects. Thus, the specification in Model 4 of Table 2 allows for life cycle differences across products, but does not in Tables 3 and 4.

Table 5 summarizes the results for the (log) value of information regressions. Since the value of information is the difference between the average and minimum price for each product date, this specification (like that in Table 2) allows for heterogenous product life cycle effects. The coefficient on branding indicates that the value of information negatively covaries with branding in all four specifications. The coefficient estimates are significant at the 1% level—even in Model 4. These results indicate that an increase in the fraction of logos from 8.29% to 9.29% decreases that value of price information at Shopper.com by \$1.33 in Model 1 and \$1.60 in Model 4. In short, all specifications in Table 4 lead to results that are consistent with Prediction 4: branding efforts appear to reduce the value of the price comparison site.

6 Concluding Remarks

Branding is obviously a central element of any firm's marketing strategy; however, many models of pricing and informational advertising treat this as exogenous. The present paper represents a modest attempt to remedy this deficiency by extending the Baye-Morgan (2001) model of pricing and informational advertising to endogenize firms' branding decisions. The extended model highlights the interplay between various elements of firm marketing strategies–average prices, best prices, the amount of informational advertising and price dispersion all hinge critically on the intensity of branding in a firm's marketing mix. Moreover, the basic predictions of the model—some of which are general theoretical properties of the model and others based on a calibration—were shown to be consistent with data from a popular price comparison site.

Our results represent a beginning, rather than the last word, on this important area. From an empirical perspective, additional econometric work that uses additional datasets and more refined measures of firm branding efforts are likely to enhance our understanding of the managerial implications of the interplay of these marketing strategies. Indeed, the results based on the imperfect proxy used in our econometric approach to the data problem suggest that something is going on in the data, and future research is warranted to refine and examine the robustness of our findings. From a theoretical perspective, an important next step is to extend the model to asymmetric environments. As Baye, Morgan, and Scholten (2006) have noted, even in the absence of endogenous branding, little is known about asymmetric models within this class. Breakthroughs on this front would not only constitute a major theoretical advance, but permit a tighter fit between the underlying theory and empirics.

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(4), pp. 463-96.

[3] Baye, Michael R., John Morgan, and Patrick Scholten, "Information, Search and Price Dispersion," in *Handbook of Economics and Information Systems* (T. Hendershott, ed.), Elsevier Press, (2006).

[4] Brynjolfsson, Erik and Michael D. Smith,"Frictionless Commerce? A Comparison of Internet and Conventional Retailers." *Management Science* (2000a) 46(4), pp. 563-585.

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

[6] Elberse, Anita, Patrick Barwise, and Kathy Hammond, "The Impact of The Internet on Horizontal and Vertical Competition: Market Efficiency and Value Chain Reconfiguration," *Advances in Applied Microeconomics* (2002) 11, pp. 1-28.

[7] Ellison, Glenn and Sara Fisher Ellison, "Lessons about Markets from the Internet." *Journal of Economic Perspectives,* (2005) 19(2), pp. 139-158.

[8] Iyer, G., Soberman, D, and Villas-Boas, J.M, "The Targeting of Advertising." *Marketing Science* (2005) 24, pp. 461-476.

[9] Johnson, E. J., S. Bellman, and G. L. Lohse "Cognitive Lock-In and the Power Law of Practice." *Marketing Science* (2003) 67(2), pp. 62-75.

[10] Keller, K., "Branding and Brand Equity," in *Handbook of Marketing*, B. Weitz and R. Wensley eds., Sage Publications, London: 2002.

[11] Little, J., "Models and Managers: The Concept of a Decision Calculus," *Management Science* (1970) 16(8), pp. 466-484.

[12] Moorthy, K. S., "Competitive Marketing Strategies," in *Marketing Handbooks and Operations Research in Management Science, Vol. 5,* J. Eliashberg and G. L. Lilien, North Holland, Amsterdam: 1993.

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

[14] Raju, J. S., and R. Lal, "The Effects of Brand Loyalty on Competitive Price Promotional Strategies," *Management Science*, (1990) 36(3), pp. 276-304.

[15] 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.

[16] Van Huyck, J. B., R. C. Battalio, and R. O. Beil, "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review* (1990) 80 (1), pp. 234-248.

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

Table 1: Data Summary

Total Observations			
Number of Products		90	
Number of Dates		205	
Number of Prices		291,039	
Product Summary Statistics	Mean	Std. Dev	Median
Price			
Average Price	\$458.86	\$496.64	\$325.94
Lowest Price	\$387.58	\$412.07	\$282.00
Highest Price	\$555.11	\$586.76	\$404.25
Advertising Levels			
Number of Advertised Prices	29.07	17.23	29.00
Percentage of Listings with Logos	8.29%	6.49%	8.11%
Price Dispersion			
Price Range	\$167.53	\$229.75	\$104.35

Table 2: Log Range Regressions

Dependent variable: Log Range	Model			
	1	2	3	4
Branding	-1.224	-1.290	-1.272	-1.912
	(4.31)**	(4.67)**	(4.72)**	(6.35)**
# of Firms	0.024	0.050	0.049	0.043
	(18.01)**	(13.02)**	(12.24)**	(9.80)**
(# of Firms)²		-0.0004	-0.0004	-0.0003
		(8.94)**	(8.56)**	(6.60)**
Product Dummies	yes	yes	yes	yes
Popularity Dummies			yes	yes
Date Dummies				yes
# of observations	9980	9980	9980	9980

Notes: HAC adjusted t statistics in parentheses.
* significant at 5%; ** significant at 1%

Table 3: Log Average Price Regressions

Dependent variable: Log Average Price	Model			
	1	2	3	4
Branding	0.091	0.093	0.090	-0.069
	(3.37)**	(3.48)**	(3.43)**	(1.55)
# of Firms	-0.001	-0.001	-0.002	-0.005
	(3.03)**	(1.98)*	(2.09)*	(5.63)**
(# of Firms)²		0.00001	0.00001	0.00004
		(1.39)	(1.52)	(4.20)**
Product Dummies	yes	yes	yes	yes
Popularity Dummies			yes	yes
Date Dummies				yes
# of observations	10013	10013	10013	10013

Notes: HAC adjusted t statistics in parentheses.
* significant at 5%; ** significant at 1%

Table 4: Log Minimum Price Regressions

Dependent variable: Log Minimum Price

	Model 1	Model 2	Model 3	Model 4
Branding	0.256	0.263	0.262	0.127
	(5.56)**	(5.79)**	(5.82)**	(1.80)
# of Firms	-0.002	-0.005	-0.005	-0.009
	(6.99)**	(5.18)**	(5.14)**	(8.76)**
(# of Firms)²		0.00004	0.00005	0.00008
		(3.68)**	(3.73)**	(6.91)**
Product Dummies	yes	yes	yes	yes
Popularity Dummies			yes	yes
Date Dummies				yes
# of observations	10013	10013	10013	10013

Notes: HAC adjusted t statistics in parentheses.
* significant at 5%; ** significant at 1%

Table 5: Log Value of Information Regressions

Dependent variable: Log Value of Information

	Model 1	Model 2	Model 3	Model 4
Branding	-1.870	-1.924	-1.902	-2.247
	(6.82)**	(7.13)**	(7.28)**	(7.40)**
# of Firms	0.014	0.035	0.035	0.040
	(11.07)**	(9.08)**	(8.65)**	(9.19)**
(# of Firms)²		-0.0003	-0.0003	-0.0004
		(7.28)**	(7.03)**	(7.99)**
Product Dummies	yes	yes	yes	yes
Popularity Dummies			yes	yes
Date Dummies				yes
# of observations	9980	9980	9980	9980

Notes: HAC adjusted t statistics in parentheses.
* significant at 5%; ** significant at 1%