A Comparison of Online and Offline Consumer Brand Loyalty

Peter J. Danaher • Isaac W. Wilson • Robert A. Davis

Department of Marketing, University of Auckland, Private Bag 92019, Auckland, New Zealand
DataMine Ltd., Wellington, New Zealand
Department of Marketing, University of Auckland, Private Bag 92019, Auckland, New Zealand
p.danaher@auckland.ac.nz • Isaac@datamine.co.nz • r.davis@auckland.ac.nz

In this study we compare consumer brand loyalty in online and traditional shopping environments for over 100 brands in 19 grocery product categories. The online purchase data come from a large traditional grocery retailer that also operates an online store for its products. The offline data corresponds to the exact same brands and categories bought in traditional stores by a panel of homes operated by ACNielsen for purchases made in the same city and over the same time period. We compare the observed loyalty with a baseline model, a new segmented Dirichlet model, which has latent classes for brand choice and provides a very accurate model for purchase behavior. The results show that observed brand loyalty for high market share brands bought online is significantly greater than expected, with the reverse result for small share brands. In contrast, in the traditional shopping environment, the difference between observed and predicted brand loyalty is not related to brand share.
(*Brand Choice; Probability Models; Internet Shopping*)

1. Introduction

The rapid growth of electronic commerce provides a challenge for marketers because "...as consumers adopt new technologies, their behaviors change" (Zinkhan and Watson 1998, p. 6). It is widely recognized by academics and practitioners that transacting through a virtual medium is different from traditional shopping environments (Alba et al. 1997). Some key differences include the means of obtaining product information, the greater perceived risk, and the ability for consumers to repurchase the same product through the use of a savable personal shopping list. It is particularly important for managers to be able to determine how these differences influence consumer brand choice because this will ultimately impact brand loyalty and, in turn, profitability (Aaker 1991, Keller 1998, Kapferer 1998). The focus of this study is the role of brands in an online environment. Our findings show that high share

(and therefore better-known) brands have greater-than-expected loyalty when bought online compared with an offline environment, and conversely for small share brands. This has important implications for established brands that can leverage their existing awareness when offered online. It also highlights the difficulties that fledgling online brands can have when trying to compete with better-known brands.

2. Relevant Literature

2.1. Consumer Behavior in a Virtual Environment
Andrews and Currim (2000), Bakos and Brynjolfsson (2000), Brynjolfsson and Smith (2000), Burke et al. (1992), Degeratu et al. (2000), Häubl and Trifts (2000), Lynch and Ariely (2000), Pauwels and Dans (2001), Shankar et al. (2002, 2003), and Ward and Lee (2000) have all conducted empirical research exploring consumer behavior in an online environment.

These studies center on understanding the manner in which the process of information search differs between traditional and online environments. Information search is split into the subgroups of price sensitivity and brand choice, and research thus far shows there is mixed evidence for the role of information search in an online environment. For example, price sensitivity is sometimes higher in online versus offline settings, depending on the computer screen graphics (Burke et al. 1992).

Of all these studies, only Degeratu et al. (2000) and Andrews and Currim (2000) specifically compare online and offline purchase behavior in the grocery sector, as we do. Degeratu et al. (2000) show that for some categories the brand name is more important online than in a traditional shopping environment, but this might depend on the available attribute information. Andrews and Currim (2000) find that the brand loyalty coefficient in a multinomial logit model is lower for online versus offline grocery shopping, but online shoppers select from a smaller consideration set of brands, thereby remaining loyal to a smaller number of brands. Both these studies are limited in that they do not fully account for the demographic differences between online and offline shoppers, which may well be causing the difference in brand loyalty between the two environments. Their studies use just three or fewer product categories. In contrast, we compare brand loyalty for over 100 brands in 19 categories. Our online data come from a large nationwide traditional grocery retailer that offers the same in-store products online (much like the successful Tesco model used in Europe). The offline data is from a Nielsen Home-Scan panel that exactly matches the online data in terms of location, time period, categories, and brands. To handle the problem of the different demographic composition of online and offline shoppers, we compare an observed brand loyalty measure against a model-expected value rather than comparing online and offline directly.

2.2. The Role of the Brand in an Online Environment

A useful way of explaining the role of the brand in a virtual environment is to use the classification of search and experience attributes used by consumers in the decision-making process (Nelson 1974). Search attributes can be determined by inspection prior to the purchase of the brand, whereas experience attributes can only be determined after the purchase has occurred (Nelson 1974). The brand plays a crucial role in helping the consumer to infer the consumption benefits pertaining to a specific product. This has significant implications in the virtual shopping context as "...retail formats differ greatly in their capability to provide information about attributes linked to consumption benefits" (Alba et al. 1997, p. 43). This means that if a consumer purchasing a product in a traditional environment is able to evaluate the quality of the product prior to purchase, the product can be categorized as a search good. However, if the same product is sold in an online environment, the physical cues that are available in the traditional environment are not present and the product could be reclassified as an experience good (Alba et al. 1997, Moore and Andradi 1996).

This transition from a search to an experience good means that an important cue for inferring quality online is the product brand (Moore and Andradi 1996). As such, the brand name within a virtual environment, through the brand value process, converts experience attributes to search attributes that are communicated visually (Alba et al. 1997). This reliance on the brand is due to the increased information required to transform a good from an experience into a search classification. Furthermore, the increased perceived risk of transacting in the online medium heightens the effect of the product brand name (Ernst and Young 1999).

Therefore, brands that are capable of creating additional search components will be advantaged within the virtual environment. Conceptually, it has been posited that this task will favor large share brands due to the fact that they provide the salient attributes of familiarity, a signal of presence, commitment, and substance (Moore and Andradi 1996). Larger brands are therefore capable of providing sufficient information for consumers to predict satisfaction without experiencing the merchandise (Alba et al. 1997). For this reason, larger brands might have an advantage over smaller, less well-known brands in an

online environment (Moore and Andradi 1996). This reasoning is consistent with the study by Ernst and Young (1999), in which 82% of respondents indicated that a product's brand name is important in their decision to buy online.

2.3. Research Method

Our research approach proceeds as follows. We begin by establishing baseline brand loyalty levels, which we hope to achieve by fitting the Dirichlet[1] model of Goodhardt et al. (1984) to both online and offline grocery purchase data for matched brands. Our measure of brand loyalty is Share of Category Requirements (SCR), defined to be each brand's market share among triers of the brand[2] (Fader and Schmittlein 1993). SCR indicates how much the customers of each brand satisfy their product needs by purchasing a particular brand rather than buying competing alternatives (Uncles et al. 1994). We realize that defining brand loyalty in this behavioral sense is not without its problems. For instance, Allenby and Rossi (1991) note that someone may have the appearance of being loyal to a brand through repurchase of the same brand when it is being price promoted. It is not the purpose of this paper to reconcile various definitions of brand loyalty, which is a complex concept. We chose SCR to measure brand loyalty due to its simplicity, widespread industry use, and because it can be directly calculated from the parameters of the Dirichlet (Bhattacharya 1997, Fader and Schmittlein 1993). Consequently, SCR is a good measure of brand loyalty for our application.

While the Dirichlet is a very good model of grocery brand choice and incorporates some consumer heterogeneity, several authors (Bhattacharya 1997, Fader and Schmittlein 1993) have noted that it cannot adequately model "excess brand loyalty" to high share brands, nor the high loyalty to niche

brands. Therefore, we generalize the Dirichlet by incorporating latent segments in the brand choice component. We find this generalized Dirichlet captures the excess brand loyalty previously unaccounted for by the single-segment Dirichlet, to the point where a brand's market share is no longer related to the difference between actual and model-estimated SCR in an *offline* setting. However, when we examine *online* purchases for the same brands, the market share effect still persists, indicating that the continued excess brand loyalty for online purchases is likely to be due to the different shopping environment.

3. Method for Comparing Online and Offline Brand Loyalty

3.1. Establishing a Baseline for Brand Loyalty

There are three ways we could establish whether online shoppers are more loyal to big brands than offline shoppers. The first way is to compare a loyalty measure, such as SCR, across datasets comprising online and offline grocery purchases. The second is to compare online and offline purchases for the same people over time. The third way is to compare online brand loyalty with predictions from a baseline model. All these methods have advantages and disadvantages, which we now discuss.

While it seems reasonable to compare brand loyalty on the basis of online versus offline shopping data, there are potentially huge problems due to selectivity bias. This arises because the sort of person who shops online is different from someone who buys groceries from a traditional store (Degeratu et al. 2000, Emmanouilides and Hammond 2000). For instance, Degeratu et al. (2000) report that homes using Peapod's online shopping service are younger, better educated, more affluent, and more likely to have children than the average U.S. household. They attempt to correct for these differences by using education level as a factor when selecting offline grocery-shopping data from IRI panel households. Although this goes some way towards aligning online and offline shoppers, differences will remain, making it difficult to conclude whether brand loyalty differences are due to the shopping mode or the type of

---

[1] The Dirichlet is more correctly known as the Dirichlet multinomial-negative binomial distribution as it is the compound of the Dirichlet multinomial distribution for brand choice and the negative binomial distribution for number of category purchases. However, the name Dirichlet has become common in the marketing literature, so we also use this terminology.

[2] We base this market share on purchase incidents rather than purchase quantity so as not to introduce biases known to affect the Dirichlet (Bhattacharya et al. 1996, Bhattacharya 1997).

person using the alternative modes. For the data in our study, a further problem is that we have only limited demographic information on each customer, and not all customers supply this information. Therefore, we cannot apply a pseudo matched sample approach, as used by Degeratu et al. (2000).

The second comparison method could track a panel of people that shop both online and offline to see how their brand loyalty differs when they shop in these different environments. Such a panel is not available in the market we study, but even if it were, there are still potential problems with this approach, as we now illustrate. Suppose a person is a regular buyer of a particular brand of butter when they shop online. However, one day they run out of butter and rather than wait for an Internet delivery they go to their local grocery store only to find their usual brand is unavailable, so they buy another brand. This has the appearance of lower brand loyalty on the offline environment, when it is more a stockout/distribution issue. Such issues are much less prevalent with Internet grocery shopping.

The third way of establishing possible stronger brand loyalty for online shoppers is to compare observed loyalty with predicted loyalty from a baseline model. The Dirichlet is just such a model, as it has been shown to be very accurate at predicting brand SCR, especially for packaged goods (Bhattacharya 1997, Fader and Schmittlein 1993, Goodhardt et al. 1984, Uncles et al. 1994). Furthermore, the Dirichlet model is designed to provide the theoretical market position of a brand in relation to other brands (East 1997) and is regarded as one of the best empirical generalizations in marketing (Uncles et al. 1995). Consequently, a reasonable procedure is to predict a brand loyalty measure like SCR using the Dirichlet and compare these predictions with the actual values observed for both offline and online purchases. However, Bhattacharya (1997) and Fader and Schmittlein (1993) showed that even though the Dirichlet is a very good model, it has one weakness in not being able to adequately model SCR for some high share brands. Bhattacharya (1997) also shows that the difference between actual and estimated SCR is a function of whether or not the brand is a niche product, plus some marketing-mix factors. Therefore, it appears that

the "raw" Dirichlet may not be robust enough as a baseline model, because it underestimates brand loyalty for high share brands. Hence, if we observe higher-than-expected brand loyalty for large share brands bought online, it will be unclear whether this is due to the online shopping environment or a known limitation of the Dirichlet model. To alleviate this problem, we propose using a segmented Dirichlet model as suggested (but not implemented) by Fader and Schmittlein (1993). We now discuss such a segmented Dirichlet model.

3.2. Segmented Dirichlet Model

Goodhardt et al. (1984) detail the derivation of the Dirichlet model. Among its useful features is the ability to capture the phenomenon of double jeopardy, whereby large share brands have more buyers and those buyers purchase the brand more frequently. The reverse occurs for small share brands. Fader and Schmittlein (1993) demonstrate that the Dirichlet model predicts that a brand's repeat purchase rate is proportional to its market share. That is, the Dirichlet is sufficiently general to accommodate the double jeopardy effect. However, Fader and Schmittlein (1993) go on to show that even though the Dirichlet facilitates double jeopardy, it does not go quite far enough. They produce several examples of high share brands with "market share premiums" which have repeat purchase rates and SCRs in excess of that predicted by the Dirichlet.

Fader and Schmittlein (1993) analyzed the assumptions underlying the Dirichlet to see which ones (if any) might be violated, thereby explaining the Dirichlet's inability to fully model consumer loyalty to high share brands. Their analysis uncovered two factors that can give rise to this shortcoming of the Dirichlet.

The first factor is due to uneven distribution of brands across a market, with larger brands being available in more stores compared with small brands. The Dirichlet model implicitly assumes that all brands in the market are available at every purchase occasion. This is unlikely to be true for traditional grocery shopping (Farris et al. 1989). However, for online grocery shopping it is usually the case that all brands are available in the online "store"—at least they are

listed, although they may be out of stock at the time.[3] Hence, unequal distribution by brand market share should not be an issue for the online grocer used in this study.

The second factor identified by Fader and Schmittlein (1993) concerns a particular heterogeneity that is not explicitly accounted for by the Dirichlet, namely, the existence of a segment of consumers highly loyal to high share brands. This situation seems very plausible, and given that we have already eliminated the uneven distribution factor, Fader's and Schmittlein's (1993) analysis indicates that such a segmentation effect is the only remaining reason for observing excess brand loyalty. A relatively simple way to correct for this deficiency of the Dirichlet is to create latent segments via a finite mixture model (Kamakura and Russell 1989). In the case of the Dirichlet model, we propose a finite mixture of up to L segments, given by

$$f(X_1, X_2, \ldots, X_g)$$
$$= \sum_{l=1}^{L} \lambda_l \sum_{n=0}^{\infty} \Pr(X_1, X_2, \ldots, X_g \mid n, \boldsymbol{\alpha}_l) \Pr(N = n \mid r, a)$$
$$= \sum_{l=1}^{L} \lambda_l \mathrm{Dir}(\boldsymbol{\alpha}_l, S_l, r, a) \tag{1}$$

where $(X_1, X_2, \ldots, X_g)$ represents the vector of number of purchases of brand j, $j = 1, \ldots, g$. $\Pr(X_1, X_2, \ldots, X_g \mid n, \boldsymbol{\alpha}_l)$ is modeled with a Dirichlet-multinomial distribution having parameters $\boldsymbol{\alpha}_l = (\alpha_{l1}, \alpha_{l2}, \ldots, \alpha_{lg})$ and $S_l = \sum_j \alpha_{lj}$ for known n, while $\Pr(N = n \mid r, a)$ is modeled by a negative binomial distribution with parameters r, the mean number of purchases, and a, the shape parameter. Note that $\sum_{l=1}^{L} \lambda_l = 1$ and $\lambda_l \geq 0$, $\forall l$, while $\mathrm{Dir}(\boldsymbol{\alpha}_l, S_l, r, a)$ denotes the Dirichlet model of Goodhardt et al. (1984). Notice also that r and a do not vary by segment, as the Dirichlet assumes that total category purchases

and brand choice are independent. The heterogeneity we might observe concerns just brand choice, so the segment structure is created for just this component of purchase behavior. The mixture model in (1) is no longer a Dirichlet model, as noted by Fader and Schmittlein (1993, p. 488). The number of segments is selected by varying L upwards from one until the Bayesian Information Criterion (BIC) is maximized (Bucklin and Gupta 1992). We show later that one or two segments are often enough to achieve this. When there is more than one segment, it always happens that one of the segments is strongly loyal to one or more of the large share brands.

3.3. Testing for Market Share Effects Online and Offline

A key objective of this study is to look for possible excess brand loyalty for high share brands in a virtual-shopping environment. The segmented Dirichlet model will be used to benchmark online and offline brand loyalty. We follow the method of Bhattacharya (1997) and Fader and Schmittlein (1993), using a linear regression model to analyze the deviations between actual and estimated SCR to see if they are related to brand market share. Such a model for SCR is written as

$$\mathrm{SCR}_i^{(a)} - \mathrm{SCR}_i^{(e)} = \alpha + \beta \, \mathrm{MS}_i + \varepsilon_i, \tag{2}$$

where $\mathrm{SCR}_i^{(a)}$ and $\mathrm{SCR}_i^{(e)}$ are the actual and estimated SCR for brand i, respectively; MS_i is the market share; and ε_i is a random error with zero mean. Fader and Schmittlein (1993) fitted a separate regression model for each category and tested whether or not $\beta > 0$ as evidence of excess brand loyalty. In their case, 8 of 28 categories had positive slope terms at the 10% level of significance. This is more categories than would be expected if no excess brand loyalty existed. As Fader and Schmittlein (1993) found, there are difficulties conducting separate regressions for each category due to the small number of brands within each category. Instead, we propose a single regression across all online and offline brands, as we now explain.

To compare brand loyalty for the same brands in both an online and offline shopping environment, we extend the Fader and Schmittlein (1993) model in Equation (2) to incorporate two slope coefficients

---

[3] In the case of the online service offered by the grocery retailer used in this study, when a brand is out of stock the shopper is offered a substitute. In most cases this works out fine (e.g., Heinz baked beans substituted for Oak baked beans). In cases where the product is not easily substituted, the shopper is then either called by phone to decide on a substitute or the product is left out and an apology made.

for market share, one when brands are bought in the online environment and one for the offline environment. That is, we examine the *interaction* between market share and purchase environment. In addition, we include a dummy variable indicating whether or not the brand was bought online. This is intended to capture any structural differences between the online and offline environments that are not related to market share.

Bhattacharya (1997) showed that in addition to market share, factors such as price, depth of price cut, promotion frequency, and whether a brand is niche or change of pace (Kahn et al. 1988) are related to the difference between observed and estimated SCR. A niche brand has low penetration but a high purchase frequency, while a change-of-pace brand is one with low purchase frequency relative to its penetration. While niche and change-of-pace brands are not commonplace, both types of brands can potentially produce deviations from the Dirichlet model (Fader and Schmittlein 1993), but in opposite directions, with niche brands having higher observed loyalty and change-of-pace brands having lower loyalty than predicted by the Dirichlet. We use the same method as Bhattacharya (1997) and Kahn et al. (1988) to operationalize our definition of niche and change-of-pace brands by calculating a niche index as follows:

$$\text{niche\_index}_j = \frac{w_j(1 - b_j)}{\sum_{j'} \text{MS}_{j'} \, w_{j'}(1 - b_{j'})},$$

where w_j is the purchase frequency and b_j is the penetration of brand j. Values of the niche index greater than one indicate the brand has niche qualities; namely, it has a high purchase frequency relative to its penetration and vice versa for change-of-pace brands. Following Bhattacharya (1997) and Kahn et al. (1988), we define niche brands to be those with a niche index greater than 1.1 and change-of-pace brands to be those with a niche index less than 0.9.

We use the same marketing-mix factors as used by Bhattacharya et al. (1996) and Bhattacharya (1997): average unpromoted price, denoted as Price; and average price-cut depth divided by Price, denoted as Price Cut.[4] To obtain comparable measures across

categories, we standardize each of these marketing-mix factors by subtracting the category mean and dividing by the category standard deviation (precisely as done by Bhattacharya et al. 1996 and Bhattacharya 1997).[5] Hence, our final regression model for analyzing brand loyalty is

$$\begin{aligned} \text{SCR}_i^{(a)} &- \text{SCR}_i^{(e)} \\ &= \alpha + \beta_1 \text{ON}_i + \beta_2 (1 - \text{ON}_i) \times \text{MS}_i + \beta_3 \text{ON}_i \times \text{MS}_i \\ &\quad + \beta_4 \text{niche}_i + \beta_5 \text{chpace}_i + \beta_6 \text{Price}_i \\ &\quad + \beta_7 \text{Price\_cut}_i + \varepsilon_i, \end{aligned} \tag{3}$$

where ON_i is one if the brand was purchased online and zero otherwise.

If there is excess loyalty among high share brands, then $\beta_2 > 0$ and $\beta_3 > 0$, as observed by Bhattacharya (1997) and Fader and Schmittlein (1993) for the offline environment, at least. Indeed, our primary interest centers on these two parameters. For instance, if they are the same we can conclude there is no difference in the way market share affects brand loyalty between the online and offline environments. On the other hand, if $\beta_3 > 0$, but $\beta_2 = 0$, then we have evidence that market share affects brand loyalty in the online but not offline environment.

Bhattacharya (1997) found that niche brands tend to have larger differences between actual and estimated SCR. Although he did not include change-of-pace brands in his study, it would be reasonable to assume that change-of-pace brands exhibit the reverse pattern, with smaller (possibly negative) differences between actual and estimated SCR. For this reason we expect niche (change-of-pace) brands to be positively (negatively) related to the difference between actual and estimated SCR, in which case $\beta_4 > 0$ (as observed by Bhattacharya 1997) and $\beta_5 < 0$. Lastly, for the marketing-mix factors, Bhattacharya et al. (1996) found significant negative effects for price and price cut. Bhattacharya (1997) found both factors to be significant and negative. That is, β_6 and β_7 are likely to be negative.

However, we found this variable caused some multicollinearity problems and so we had to omit it.

[5] We also standardize the difference between actual and estimated SCR and the market shares within each category, as was done by Bhattacharya et al. (1996) and Bhattacharya (1997).

---

[4] Bhattacharya et al. (1996) and Bhattacharya (1997) also used the percent of a brand's sales that are made during a promotion.

Suppose that high share brands have greater positive differences between actual and estimated SCR, with small share brands having negative values of $SCR_i^{(a)} - SCR_i^{(e)}$. Then Equation (3) is unsuitable for testing this possible switch in sign for $SCR_i^{(a)} - SCR_i^{(e)}$ from large to small brands. All it can test is that $SCR_i^{(a)} - SCR_i^{(e)}$ increases (or decreases) with market share. Hence, we test the *sign* of the difference between the actual and estimated SCR with a logistic regression model analogous to Equation (3) and defined as

$$\log\left(\frac{\Pr(D_i = 1)}{\Pr(D_i = 0)}\right)$$
$$= \gamma + \delta_1 ON_i + \delta_2(1 - ON_i) \times MS_i + \delta_3 ON_i \times MS_i$$
$$+ \delta_4 niche_i + \delta_5 chpace_i + \delta_6 Price_i + \delta_7 Price\_cut_i, \quad (4)$$

where $D_i = 1$ if $SCR_i^{(a)} - SCR_i^{(e)} \geq 0$ and 0 otherwise. If high market share brands have positive values of $SCR_i^{(a)} - SCR_i^{(e)}$ both online and offline, and vice versa for small market share brands, then $\delta_2 > 0$ and $\delta_3 > 0$, which can easily be tested via Equation (4).

4. Data

4.1. Online Data

Selection of Households. As mentioned earlier, the online data come from a large grocery retailer in New Zealand that has both a nationwide network of stores and an online service which mimics these stores. That is, the categories and brands available online are exactly the same as what a customer could get when shopping in one of their traditional stores. Furthermore, prices and price discounts in-store are mirrored exactly in the online environment. Although the online service is now offered in the metropolitan areas of Auckland, Wellington, and Christchurch, we use data just from Auckland city, which was the first to obtain the service in early 1997. The data used in this study span the 12 months of January through December 1998, during which the service was well established and had a reasonable customer base. Using a twelve-month period is consistent with other studies that have used the Dirichlet (Uncles

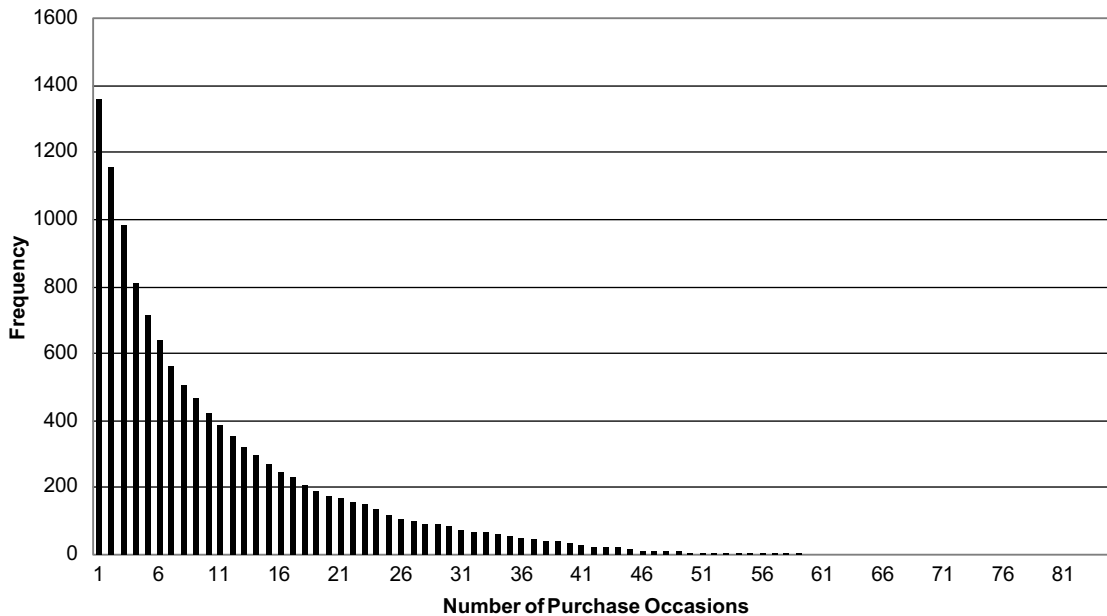and Ehrenberg 1990a, b; Fader and Schmittlein 1993; Uncles et al. 1994; Bhattacharya 1997).

The data required to fit the Dirichlet normally come from a panel of households (Uncles and Ehrenberg 1990a, Fader and Schmittlein 1993, Bhattacharya 1997). Unlike Peapod (Degeratu et al. 2000), this particular online grocery service has no subscription cost, instead charging a flat delivery fee of $7.50.[6] Therefore, there is no entry, ongoing, or exit cost to users of the service, and consequently the customer base is dynamic. We can only be certain when someone uses the service for the first time and never know when or if they stop using the service. For instance, if a household shops online just twice a year on average and we first observe them in October 1998, we will probably not observe a second purchase occasion in that same calendar year. Hence, to us they will appear as a single-purchase household in 1998.

Figure 1 shows a histogram of the number of purchase occasions for all homes that used the service at least once in 1998. It can be seen that a large proportion of homes used the service just once or twice. This could be due to a lot of homes using the service very infrequently, a lot of homes joining the service late in the year, or many homes trying the service once or twice, then discontinuing. We believe the latter is the most likely reason for the pattern shown in Figure 1, given the high labor intensity of the first shopping occasion, where customers usually make a lot of purchase decisions, often from a long list of brands, many of which provide additional nutrition information.

To get as near a "panel" as we can from the online shopper customer database, we selected just those homes that had at least six shopping occasions in the first six months of the year. This eliminated infrequent shoppers and recent entrants to the service, leaving 601 homes that are used in all subsequent analyses. These 601 households spend an average of $90 on each shopping occasion and average 14 days between purchase occasions. The median interpurchase time is nine days, indicating the distribution is right skewed, as might be expected. For the offline panelists, the average spend per visit is $33 and customers visit a store about seven times per month. Degeratu et al.

---

[6] All currency is expressed in U.S. dollars.

Figure 1   Number of Online Shopping Occasions per Household

(2000) report that in the United States the average offline grocery order is very similar, being valued at $23.50 with customers visiting a store about eight times per month. However, online Peapod shoppers spend considerably more, at $125 per order, but shop less frequently, about twice per month. The online shoppers in our sample are similar to Peapod shoppers in that they spend more, but do so much less often than traditional grocery shoppers.

Selection of Product Categories. We now describe the method used to select a representative market basket of frequently purchased packaged goods. Our original dataset contained purchase information on 221 product categories. However, many of these had only infrequent purchases and could not be used. Our criteria for category selection required that the final market basket be both representative of frequently purchased packaged goods and that products be purchased in high enough volume to provide sufficient data points for reliable estimation of the Dirichlet parameters.

To select categories that provided a representative market basket of regularly purchased packaged goods, we adopted the selection process developed by Fader and Lodish (1990) and Bell and Lattin (1998).

They used cluster analysis to select only categories with high penetration and short purchase cycles. We conducted a similar cluster analysis and obtained 165 such categories. The next refinement of the categories centered on estimation issues. To obtain reliable parameter estimates for the Dirichlet model, Fader and Schmittlein (1993) stipulated that a category must, on average, be purchased a minimum of three times per year and have at least 1,000 repeat purchases. Further constraints were also applied, namely, (i) brand share must exceed 1%, (ii) a minimum of three eligible brands must be present in each category, (iii) a minimum of 80% overall category volume must be represented by the eligible brands, and (iv) product sizes within categories had to be the same. These constraints reduced the number of eligible categories to 19, which are listed in Table 1. Our final categories exhibit a good representation of typical frequently purchased packaged goods categories. The total number of brands across these categories is 129.

4.2. Offline Data
The offline data were obtained from ACNielsen's HomeScan panel, also for the 12 months of 1998 and from the same city as the online data, namely,

468

Table 1   Categories Used for Analysis

Online Categories	Online Penetration, %	Offline Penetration, %	No. of Brands
Soda drink	83	92	12
Butter	83	88	6
Canned cat food	48	43	5
Canned fruit	76	85	5
Canned tuna	64	49	6
Cheese	86	59	6
Eggs	91	57	3
Cookies	63	57	4
Frozen vegetables	78	80	6
Fruit juice	86	69	9
Ice cream	67	73	6
Instant coffee	59	80	8
Laundry powder	75	75	10
Potato chips	73	70	5
Sliced brown bread	76	87	11
Milk	86	82	6
Bathroom tissue	96	69	7
Toothpaste	89	94	7
Yogurt	82	53	7
Total	—	—	129

Auckland, New Zealand. The panel comprised 443 households selected to represent the household profile of the city. All these households contributed data for at least 25 weeks in 1998. HomeScan households purchase their goods in supermarkets in the usual way, then scan the barcode of each product when they return home. Incentives are given to ensure the panelists participate reliably. As Auckland has about 400,000 homes and only 601 homes were in our online "panel," the chance of a home being in both the online and offline panels is extremely low.

We obtained from Nielsen the household-level purchase history for the exact same categories and brands as selected for the online products. The HomeScan data logged purchases for all of the grocery retailers in Auckland, not just the retailer that provides the online service. The retailer offering online and traditional grocery shopping had about a 30% share of total grocery sales in Auckland city. Our offline data did not have all the brands used for the online study, as some brands had so few offline sales to panelists. The final number of matching brands from the offline data was 119. All of the same categories were represented, however.

Hence, our offline data are an exact match on region, time period, and categories, and a very near match on brands.

5.   Results

5.1.   Fitting the Dirichlet Model

As explained above, we use a segmented Dirichlet model rather than the unsegmented Dirichlet model[7] of Goodhardt et al. (1984). The key reason for this is to explicitly allow for customer heterogeneity arising from consumers who are strongly loyal to one or more of the large market share brands within a category. Fader and Schmittlein (1993) showed that such heterogeneity is very likely to be the cause of excess behavioral loyalty for offline purchases, which is not well modeled by an unsegmented Dirichlet.

Instead of the ad hoc parameter estimation method employed by Goodhardt et al. (1984), we use maximum likelihood to obtain estimates of the parameters for the finite mixture of Dirichlet models. The estimated value of SCR was obtained using the formula supplied by Goodhardt et al. (1984, p. 629), with a modification of the brand-level beta binomial distribution to account for our finite mixture modification given in Equation (1).

Table 2 gives the results for the ice cream category in the online environment, which has six brands. It can be seen that the BIC for the two-segment model is higher than that of the one-segment Dirichlet, indicating that the second segment is meaningful and improves model fit[8] (Bucklin and Gupta 1992). Judging by the estimated market shares in the two-segment model (given by α_{ij}/S_i), the first segment is largely comprised of households purchasing the largest brand, Tip Top. This is exactly the effect that Fader and Schmittlein (1993) conjectured may be apparent in many categories and should give rise

---

[7] Another possible model is one which allocates a segment to each brand to allow for sole buying to each brand (Dillon and Gupta 1996). We do not pursue this model here.

[8] The three segment solution had a BIC of $-1,094.0$, lower than for the two-segment model. Hence, we retain just two segments for ice cream. Indeed, none of the online categories could support more than two segments.

Table 2 Parameter Estimates for the One- and Two-Segment Dirichlet Models—Online Ice Cream Category

	One-Segment Dirichlet			Two-Segment Dirichlet				
Log-likelihood	−1,073.9			−1,042.6				
BIC	−1,091.3			−1,080.2				
				Segment One Proportion (42%)		Segment Two Proportion (58%)		Two-Segment
	Market Share, %	α_j	α_j/S, %	α_{1j}	α_{1j}/S_1, %	α_{2j}	α_{2j}/S_2, %	Estimated Market Share
Brand								
Tip Top	63.1	0.843	60.8	0.565	86.2	0.933	43.2	61.4
First Choice*	14.0	0.209	15.1	0.018	2.7	0.518	24.0	15.0
Talleys	10.0	0.133	9.6	0.041	6.2	0.265	12.3	9.7
Generic	5.9	0.088	6.3	0.001	0.1	0.217	10.1	5.8
Blue Ribbon	5.2	0.092	6.6	0.011	1.6	0.207	9.6	6.2
Litelicks	1.8	0.021	1.5	0.021	3.1	0.018	0.8	1.8
S parameter	—	1.386	—	0.655	—	2.158	—	—
$1/(1+S)$	—	0.42	—	0.60	—	0.32	—	—

*First Choice is the store brand for this grocery retailer.

to brand loyalty over and above that already predicted by an unsegmented Dirichlet. In addition, it is known that the inverse of the Dirichlet S parameter is related to consumer heterogeneity. Specifically, a value of $1/(1+S)$ near zero indicates homogeneity, while a value near one indicates heterogeneity. Using this criterion, Table 2 shows that Segment 1 is more heterogeneous than Segment 2. This is precisely the situation where Fader and Schmittlein (1993) show that the unsegmented Dirichlet is most vulnerable to understating brand loyalty to big brands. Hence, it makes sense to allow for strong loyalty to big brands by using a finite mixture of Dirichlets, as we do. Indeed, we found that a two-segment Dirichlet revealed a segment strongly loyal to big brands in 12 of the 19 categories, as demonstrated by higher BIC values for the two compared with the one-segment model. We also used the segmented Dirichlet for the offline data. Here, 13 of the 19 categories had two or more segments.

Table 3 compares the actual and estimated SCR for the online ice cream category. Recall that Fader and Schmittlein (1993) predicted that the unsegmented Dirichlet would be likely to understate loyalty to high share brands and that some form of segmentation is required. We see evidence of this in the SCR measure in Table 3, where the one-segment model underestimates the actual SCR for the highest-share brand

(Tip Top) by five percentage points. However, the two-segment model predicts a higher level of SCR for Tip Top, apparently attempting to correct for the SCR underestimation known to occur for high share brands with an unsegmented Dirichlet. Therefore, using a finite mixture of Dirichlets helps correct for the limitations of the unsegmented Dirichlet highlighted by Fader and Schmittlein (1993).

Apart from the generic brand, the two-segment Dirichlet gives improved SCR estimates compared with the one-segment model. One brand that deserves further mention is Litelicks, which has the smallest share but the second-highest purchase frequency in the category, indicating it is a niche brand. This is borne out by the niche index (rightmost column of Table 3) being 1.16 for Litelicks. The one-segment model clearly underestimates the SCR for Litelicks, while the two-segment model estimates are much closer to the observed values. Returning to Table 2, we see that Litelicks is weighted relatively much more heavily to Segment 1 rather than Segment 2. Therefore, Segment 1 is comprised of those not just loyal to the highest-share brand, but those loyal to any brand, no matter what the share.

Table 3 shows lower average absolute errors between actual and estimated SCR across the six brands in the online ice cream category. The next row in the table shows that the two-segment model results

Table 3   Actual and Estimated SCR for the One- and Two-Segment Dirichlet Models for the Online Ice Cream Category

Ice Cream Category	Market Share	SCR—Actual	SCR—Estimated One-Segment Dirichlet Model	SCR—Estimated Two-Segment Dirichlet Model	Niche Index
Brand					
Tip Top	63.1	72.0	67.0	68.4	1.04
First Choice	14.0	37.8	35.8	37.0	0.89
Talleys	10.0	32.5	32.8	32.0	1.03
Generic	5.9	31.4	31.1	27.9	0.92
Blue Ribbon	5.2	25.2	31.2	28.7	0.77
Litelicks	1.8	36.2	28.6	34.4	1.16
Average absolute difference—ice cream	—	—	3.5	2.3	—
Average absolute difference—all online categories	—	—	7.2	5.7	—
Average absolute difference—all offline categories	—	—	6.5	5.2	—
Average absolute difference—all offline categories for proxy Internet homes	—	—	6.3	5.6	—

in much lower average absolute errors (a 21% reduction) between actual and estimated SCR across all 129 online brands. The penultimate row of Table 3 gives the equivalent information for all the offline categories, where the reduction in absolute error is 20% when going from a single to a multisegment Dirichlet model.

Remember that a key component of our research method when comparing online and offline brand loyalty is that we compare observed SCRs for each dataset against a segmented Dirichlet baseline, rather than directly comparing observed SCRs between the two datasets. We do this as we anticipate that comparing SCRs for online and offline shoppers will be confounded by the knowledge that two respective groups of buyers are demographically different, and this might be causing any apparent brand loyalty differences. The purpose of using the baseline model is to eliminate demographic effects, since the segmented Dirichlet benchmark will adjust to whatever dataset is used to fit the parameters, be it Internet grocery shoppers or a panel of shoppers purchasing in traditional grocery stores.

To provide some evidence that the segmented Dirichlet benchmark is robust against changes in demographic composition, we examined our offline database to find a group of homes that could act as a proxy for homes with Internet access. We had only limited demographic information for the Home-Scan panel, but two available demographic descriptors that are known to be associated with Internet access are age and household income. Using information from Nielsen's e-ratings panel in New Zealand, we set a criterion that approximately matches that of homes with Internet access, namely, the annual household income had to exceed $40,000 and the age of the main grocery buyer had to be less than 65 years. Some 45% (199 homes) of the original 443 HomeScan panel homes met this criterion, being very close to the known Internet penetration of 47% in 1998 for New Zealand.

We now pair up the standardized values of $SCR_i^{(a)} - SCR_i^{(e)}$ for the full offline panel and the subset of proxy Internet-access homes. The correlation between $SCR_i^{(a)} - SCR_i^{(e)}$ values is 0.88 and the paired-sample t-test figure is -1.31, having a p-value of 0.19, indicating there is no significant difference between the standardized $SCR_i^{(a)} - SCR_i^{(e)}$ values for the differing demographic groups when purchasing offline. This shows that even though the demographic groups are different, the segmented Dirichlet can adjust to these changes and produce consistent $SCR_i^{(a)} - SCR_i^{(e)}$ values. This helps validate our method, which relies on the segmented Dirichlet to provide an accurate benchmark to compare against observed brand SCRs.

5.2. Market Share Effects on Brand Loyalty

Recall that the regression equation (3) is constructed to facilitate a test of possible market share effects in the offline, online, or both environments. Table 4 gives the results of the regression analysis, firstly (center column of Table 4) for the conventional unsegmented Dirichlet and secondly (on the rightmost column of Table 4) for the latent class generalization of the Dirichlet given in Equation (1).[9]

For the unsegmented Dirichlet the market share for brands bought *offline* is significantly and positively related to the difference between the actual and estimated SCR, thereby corroborating the findings of Bhattacharya (1997) and Fader and Schmittlein (1993). A positive and significant regression coefficient for brands bought in the *online* environment is also evident. That is, a single-segment Dirichlet cannot adequately account for excess brand loyalty for high share brands in either environment. Note also that the regression coefficient for market share in the online environment is nearly twice that for the offline environment. A test of the equality of these two coefficients has a *p*-value of 0.066, being significant at the 10% level. This is some indication that market share effects are stronger online than offline.

We now turn our attention to the multisegment Dirichlet baseline model on the rightmost column of Table 4, which is the main focus of our study. It can be seen that when the Dirichlet is generalized to have latent segments for brand choice, offline market share is no longer significant, while the *online market share effect is still positive and significant.* This demonstrates two things—first, that excess brand loyalty for high share brands in the offline environment is captured by the generalized Dirichlet model. Hence, we can be confident in saying that the segmented Dirichlet is a good baseline model for grocery buying. Second, and more importantly for this study, even when the Dirichlet is corrected for its known "excess market share" deficiency, the market share effect in the *online*

environment persists. The fact that brand market share is related to the difference between observed and baseline SCR for online, but not offline, purchases is reasonably convincing evidence that higher-share brands have greater loyalty online compared with the traditional environment, with the reverse effect for low share brands. Remember also that factors like category, brand, time period, and geographic location are the same across the two purchase environments, helping to reinforce the validity of our findings.

Other factors in Table 4 also deserve mention. Notice that the dummy variable distinguishing online and offline purchases is not significant for both the original and generalized Dirichlet models. This indicates that the difference between actual and model-estimated SCR does not depend on the shopping environment. For the multisegment Dirichlet both the niche and change-of-pace effects are significant, with signs in the expected direction. Niche brands have higher-than-expected loyalty, with the reverse effect for change-of-pace brands. Also note that, as found by Bhattacharya (1997), the price effect is significant and negative, but the price-cut effect is not significant.

Table 5 gives the estimated coefficients for the logistic regression model in Equation (4) of the sign of the difference between actual and estimated SCR. The results are consistent with those seen in Table 4, with market share being significant for both offline and online purchases when the single-segment model is used. However, when the generalized multisegment Dirichlet is used as the baseline, only the online brand purchases show excess brand loyalty. The difference between this analysis and Table 4 is that it shows that when in an online setting, high share brands tend to have actual SCR exceeding estimated SCR (excess brand loyalty), while low share brands tend to have actual SCR less than estimated SCR (diminished brand loyalty).

In sum, Tables 4 and 5 give reasonably strong evidence that high share brands in an online environment exhibit loyalty (as measured by SCR) that is greater than would be predicted by a baseline multisegment Dirichlet model. Moreover, the reverse effect is observed for small share brands, which have lower loyalty levels than expected. No such brand size effect is evident for brands purchased offline,

---

[9] We found a number of outliers due to a low number of purchases of some brands, particularly in the offline data, where the panel size was smaller. To eliminate these outliers, we set a threshold of at least 40 purchases for each brand. This reduced the number of online brands to 128 and offline to 96.

Table 4 Regression of (Actual–Estimated) SCR on Brand Characteristics

	One Dirichlet Segment			More than One Dirichlet Segment		
	Coefficient	T-Statistic	p-Value	Coefficient	T-Statistic	p-Value
Intercept	0.091	0.89	0.375	0.124	1.13	0.262
Online dummy	−0.044	−0.47	0.638	−0.022	−0.22	0.829
Offline market share	0.212	2.72	0.007	0.113	1.34	0.181
Online market share	0.395	5.84	0.000	0.316	4.32	0.000
Niche	0.690	5.76	0.000	0.547	4.23	0.000
Change of pace	−0.502	−4.55	0.000	−0.549	−4.61	0.000
Price	0.042	0.82	0.413	−0.178	−3.19	0.002
Price cut	−0.143	−2.73	0.007	−0.058	−1.02	0.309

$R^2 = 46\%$ $R^2 = 35\%$

Adj-$R^2 = 44\%$ Adj-$R^2 = 33\%$

$n = 224$ $n = 224$

F-statistic $= 26.2$ (p-value < 0.0001) F-statistic $= 16.9$ (p-value < 0.0001)

Notes. Dependent and independent variables have been normalized by subtracting the category mean and dividing by the category standard deviation (see Bhattacharya 1997), except the online, niche, and change-of-pace dummies. All variance inflation factors are less than 1.4, indicating that there are no multicollinearity problems.

however. An additional factor which emerges as important is whether a brand is niche or change of pace, with niche brands enjoying excess loyalty in much the same way a large share brand does, while change-of-pace brands have lower brand loyalty than is expected. Finally, of the marketing-mix variables, only price is ever significant (when the multisegment Dirichlet is the baseline), with higher-priced brands having lower loyalty. This is consistent with the results of Bhattacharya et al. (1996) and Bhattacharya (1997), which are justified on the rationale that higher-priced brands gain additional (temporary) buyers when they price promote, thereby lowering the SCR.

The effect is asymmetric, with low price brands not gaining additional new buyers when they price promote.

6. Discussion and Conclusion

In this study we compare the brand loyalty of grocery products for closely matched samples of brands purchased either online or offline. The panels on which the purchase history is observed have identical geographic location, time period, and product categories and are a very near match on brands. To avoid the obvious limitation of differing demographic

Table 5 Logistic Regression of the Sign of (Actual–Estimated) SCR on Brand Characteristics

	One Dirichlet Segment			More than One Dirichlet Segment		
	Coefficient	T-Statistic	p-Value	Coefficient	T-Statistic	p-Value
Intercept	0.198	0.57	0.570	0.221	0.65	0.518
Online dummy	−0.068	−0.19	0.849	−0.006	−0.02	0.986
Offline market share	0.678	2.22	0.026	0.205	0.74	0.458
Online market share	0.936	3.32	0.001	1.125	3.83	0.000
Niche	1.580	3.49	0.001	1.233	2.85	0.004
Change of pace	−1.827	−4.55	0.000	−1.781	−4.48	0.000
Price	−0.098	−0.51	0.610	−0.452	−2.38	0.018
Price cut	−0.190	−0.96	0.339	−0.153	−0.78	0.433

Note. The independent variables have been normalized by subtracting the category mean and dividing by the category standard deviation (see Bhattacharya 1997), except the online, niche, and change-of-pace dummies.

composition of online and offline buyers, we develop a segmented Dirichlet to act as a benchmark for brand loyalty for the respective online and offline panels. In particular, the generalized Dirichlet does not suffer from the known problem of understating brand loyalty for high share brands, which has already been noted by Bhattacharya (1997) and Fader and Schmittlein (1993) for offline purchases. Furthermore, when we compare brand loyalty deviations for purchases made by the whole offline panel and a subset of the offline panel with the demographic characteristics of online shoppers, there are no significant differences. This gives us reasonable confidence that the segmented Dirichlet model accurately represents purchase behavior irrespective of the underlying demographic group.

Our results are clear and straightforward. For purchases made offline, brand market share is not related to the difference between actual and model-estimated brand loyalty, as measured by share of category requirements. However, for online purchases a comparison of actual with model-estimated brand loyalty shows that Fader's and Schmittlein's (1993) "excess brand loyalty" still persists. In particular, greater brand loyalty is observed for brands with high market share, and vice versa for low share brands. Because the segmented Dirichlet model is an accurate benchmark for purchase behavior and we continue to observe excess brand loyalty only in the online environment, we have strong evidence of higher brand loyalty for online purchases compared with offline.

Our findings are consistent with those of Degeratu et al. (2000). They found that brand name was important in the sense that a "strong" brand did better in an online environment compared with a "weak" brand. For their data, "strong" and "weak" are synonymous with large and small share, respectively. Some possible explanations for this phenomenon are:

(i) online shoppers may infer product quality from the brand name and the greater relative salience of the brand name online compared with offline means that consumers are likely to place more emphasis on the importance of brand name when grocery shopping online (Moore and Andradi 1996);

(ii) buying a well-known rather than a lesser-known brand online has less perceived risk (Ernst and Young 1999);

(iii) after initial use of this particular online grocery service, a shopper is able to select subsequent purchases from a checklist of previous purchases. This helps build inertia into the buying process by making it easy for brands to be "rolled over" from one purchase occasion to the next with little consideration given to changing brands.[10]

Any or all of these reasons are plausible explanations for our finding that online brand loyalty for high market share brands exceeds that of a traditional shopping environment, with the reverse effect for low share brands. A limitation of this study is that with our data we are not able to pinpoint which reason(s) might be driving our results, so we leave this as an area for future research. Some promising areas for examination include the area of consumer learning, where choice of a particular brand tends to induce further downstream selection and enhance quality perceptions of that brand, even in the presence of price promotions (Akcura et al. 2004). Furthermore, Moshkin and Shachar (2002) develop a theoretical model that differentiates between future choice based on past choice or based on the information set built up from previous choices, with the information set dominating. Applying the Moshkin and Shachar (2002) finding to the Internet, where brand information and brand perception are key (Alba et al. 1997), it is likely that an even higher proportion of people (than the 71% observed for TV-viewing choices) base their future choices on information sets evolved from previous choices. High share, and hence well-known brands, have an advantage when developing an information set due to greater levels of national advertising, for example.

In addition to our market share finding, we also find that niche brands have higher loyalty than estimated by the segmented Dirichlet benchmark model, while change-of-pace brands have lower loyalty than expected. Our result for niche brands supports the finding of Bhattacharya (1997), while the change-of-pace result is novel, although not unexpected. Of the marketing-mix factors, only a brand's relative base

---

[10] While this is a simple explanation for repeat purchases, it does not help explain the selection of a high share brand at initial use of the online service, where explanations (i) and (ii) are more relevant.

price emerged as important, with high-priced brands exhibiting lower loyalty.

The managerial implications of our results are sobering for fledgling brands hoping to gain high loyalty and penetration on the Internet. What we observe is that high share, and therefore better-known, brands have greater-than-expected brand loyalty, with small share (relatively unknown) brands having lower-than-expected loyalty. When e-commerce first became headline news, some business press commentators predicted "vanishing brand loyalty" (Kuttner 1998). In fact, our results show that purchase behavior on the Internet tends to be more conservative than in traditional stores. In terms of brand loyalty, already-familiar brands, with a strong offline presence, do even better in the Internet environment than offline.

Acknowledgments

References

Aaker, D. A. 1991. *Managing Brand Equity: Capitalizing on the Value of a Brand Name*. Free Press, New York.

Akcura, M. Tolaga, Fusan F. Gonul, Elina Petrova. 2004. Consumer learning and brand valuation: An application of over-the-counter drugs. *Marketing Sci.* Forthcoming.

Alba, J., J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, S. Wood. 1997. Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *J. Marketing* **61**(3) 35–53.

Allenby, Greg M., Peter E. Rossi. 1991. Quality perceptions and asymmetric switching between brands. *Marketing Sci.* **10**(Summer) 185–204.

Andrews, Rick L., Imran S. Currim. 2000. Behavioral differences between consumers attracted to shopping online vs. traditional supermarkets: Implications for enterprise design and strategy. Working paper, University of Delaware, Newark, DE.

Bakos, Nickolay V., Erik Brynjolfsson. 2000. Bundling and competition on the Internet. *Marketing Sci.* **19**(1) 63–82.

Bell, David R., James M. Lattin. 1998. Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. *Marketing Sci.* **17**(1) 66–88.

Bhattacharya, C. B. 1997. Is your brand's loyalty too much, too little, or just right?: Explaining deviations in loyalty from the dirichlet norm. *Internat. J. Res. Marketing* **14**(5) 421–435.

——, P. S. Fader, L. M. Lodish, W. S. DeSarbo. 1996. The relationship between marketing mix and share of category requirements. *Marketing Lett.* **7**(1) 5–18.

Brynjolfsson, Erik, Michael D. Smith. 2000. Frictionless commerce? A comparison of Internet and conventional retailers. *Management Sci.* **46**(4) 563–585.

Bucklin, Randolph E., Sunil Gupta. 1992. Brand choice, purchase incidence and segmentation: An integrated modeling approach. *J. Marketing Res.* **29**(May) 201–215.

Burke, R. R., B. A. Harlam, B. E. Kahn, L. M. Lodish. 1992. Comparing dynamic consumer choice in real and computer-simulated environments. *J. Consumer Res.* **19**(1) 71–82.

Degeratu, A., Arvind Rangaswamy, J. Wu. 2000. Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes. *Internat. J. Res. in Marketing* **17**(1) 55–78.

Dillon, William R., Sunil Gupta. 1996. A segment-level model of category volume and brand choice. *Marketing Sci.* **15**(1) 38–59.

East, Robert. 1997. *Consumer Behavior: Advances and Applications in Marketing*. Prentice Hall, Hertfordshire, U.K.

Emmanouilides, Chris J., Kathy A. Hammond. 2000. Internet usage: Predictors of active users and frequency of use. *J. Interactive Marketing* **14**(2) 17–32.

Ernst and Young. 1999. *The Second Annual Ernst and Young Internet Shopping Study: The Digital Channel Continues to Gather Steam*. Ernst and Young. Available at http://www.ey.com/publicate/consumer/pdf/internetshopping.pdf (18 October).

Fader, Peter S., Leonard M. Lodish. 1990. A cross-category analysis of category structure and promotional activity for grocery products. *J. Marketing* **54**(4) 52–65.

——, David C. Schmittlein. 1993. Excess behavioral loyalty for high-share brands: Deviations from the Dirichlet model for repeat purchasing. *J. Marketing Res.* **30**(4) 478–493.

Farris, Paul, James Olver, Cornelius de Kluyver. 1989. The relationship between distribution and market share. *Marketing Sci.* **8**(Spring) 107–127.

Goodhardt, Gerald J., Andrew S. C. Ehrenberg, Christopher Chatfield. 1984. The Dirichlet: A comprehensive model of buying behavior. *J. Roy. Statist. Soc. A* **147**(5) 621–655.

Häubl, Gerald, Valerie Trifts. 2000. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Sci.* **19**(1) 4–21.

Kahn, Barbara E., Manohar U. Kalwani, Donald G. Morrison. 1988. Nicheing versus change-of-pace brands: Using purchase frequencies and penetration rates to infer brand positionings. *J. Marketing Res.* **25**(November) 384–390.

Kamakura, Wagner A., Gary J. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* **26**(November) 379–390.

Kapferer, J-N. 1998. *Strategic Brand Management: Creating and Sustaining Brand Equity Long Term*, 2nd ed. Kogan Page, Dover, NH.

Keller, Kevin L. 1998. *Strategic Brand Management: Building, Measuring, and Managing Brand Equity, 1/e*. Prentice Hall Business Publishing,

Kuttner, R. 1998. The net: A market too perfect for profits. *Bus. Week* (May 11) 20.

Lynch, John G., Dan Ariely. 2000. Wine online: Search costs affect competition on price, quality and distribution. *Marketing Sci.* **19**(1) 83–103.

Moore, K., B. Andradi. 1996. Who will be the winners on the Internet? *J. Brand Management* **4**(1) 57–64.

Moshkin, Nickolay, Ron Shachar. 2002. The asymmetric information model of state dependence. *Marketing Sci.* **21**(4) 435–454.

Nelson, P. 1974. Advertising as information. *J. Political Econom.* **83**(July–August) 729–754.

Pauwels, Koen, Enrique Dans. 2001. Internet marketing the news: Leveraging brand equity from marketplace to marketplace. *J. Brand Management* **8**(4–5) 303–314.

Shankar, Venkatesh, Amy K. Smith, Arvind Rangaswamy. 2002. Customer satisfaction and loyalty in online and offline environments. *Internat. J. Res. Marketing* **20**(2) 153–175.

——, Arvind Rangaswamy, Michael Pusateri. 2003. The online medium and customer price sensitivity. Working paper, University of Maryland, College Park, MD.

Uncles, Mark D., Andrew S. C. Ehrenberg. 1990a. The buying of packaged goods at US retail chains. *J. Retailing* **66**(3) 278–296.

——, ——. 1990b. Industrial buying behavior: Aviation fuel contracts. *Internat. J. Res. Marketing* **7**(1) 57–68.

——, ——, Kathy Hammond. 1995. Patterns of buyer behavior: Regularities, models, and extensions. *Marketing Sci.* **14**(3) G71–G78.

——, Kathy A. Hammond, Andrew S. C. Ehrenberg, R. E. Davis. 1994. A replication study of two brand-loyalty measures. *Eur. J. Oper. Res.* **76**(2) 378–384.

Ward, Michael R., Michael J. Lee. 2000. Internet shopping, consumer search and product branding. *J. Product Brand Management* **9**(1) 6–12.

Zinkhan, G. M., R. T. Watson. 1998. Electronic commerce: A marriage of management information systems and marketing. *J. Market-Focused Management* **3**(1) 5–22.