Accommodating Heterogeneity in Brand Loyalty Estimation:

Application to the U.S. Beer Retail

Roozbeh Irani-kermani, Pennsylvania State University, Roozbeh@psu.edu

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

Accommodating Heterogeneity in Brand Loyalty Estimation: Application to the U.S. Beer Retail

Roozbeh Irani-kermani

Penn State University, Agricultural Economics, Sociology, and Education, 301 Armsby, UniversityPark, PA 16802. E-mail: Roozbeh@psu.edu

Edward C. Jaenicke

Penn State University, Agricultural Economics, Sociology, and Education, 208C Armsby, UniversityPark, PA 16802. E-mail: tjaenicke@psu.edu

Introduction

Brand loyalty is important not only because of the effect it has on consumers' decision-making processes, but also because firms selling brands with a loyal following enjoy predictable sales and secure consumer demand. In addition, brand loyalty provides an entry barrier that can prevent potential competitors from entering the market (Aaker, 1991). Brand loyalty can also be interpreted as higher willingness to pay, which means that competitors offering products with exact same attributes, characteristics, and price as the original product cannot compete because of the brand-loyalty effect. Brand loyalty of this nature can be achieved by long-term marketing activities and/or personal experiences (Kotler, 2000).

With interest from economists, psychologists, and marketing specialists, a wide range of research exists on how brand loyalty affects choices, enters the decision-making process, and can be modeled and parameterized. Most brand-loyalty research focuses on how one brand competes among other consumer choices in a specific product category (Bentz & Merunka, 2000) .

According to Ballantyne, Warren, & Nobbs (2006), brands are no longer just a representative of products with specific characteristics, but they are now regarded as embodying the individualities, personalities, and lifestyle symbols of customers and their environments. For this reason, behavioral economists have shown increased interest in brand choice investigations during recent decades. For a long time, a brand-choice researcher's focus was testing the effects of marketing mix variables in a consumer's decision-making process (Bentz & Merunka, 2000). However, these marketing-mix variables were product related, but not customer related. Because customer behavior matters, this omission motivated brand-choice modelers to add household characteristics to accommodate consumer heterogeneity as well as state dependent variables, such brand loyalty, to fully account for a wide range of elements in the consumer's decision.

As (Keane, 1997) mentions, distinguishing between consumer heterogeneity and state dependence is of fundamental importance in marketing because both could explain observed persistence in consumer brand choices. For example, the consequences of a price promotion could be viewed two different ways. On one hand, if consumer behavior is truly heterogeneity related and there is no state dependency, a price promotion will not be expected to have a long-lasting affect when it is turned off. On the other hand, if state dependency is important, we would expect to observe some long-lasting effects even after the promotion stops. In other words, with state dependency, a marketer can change a consumer's brand loyalty over time via price promotions.

Brand loyalty has its roots in the consumer's optimization problem. According to Howard and Sheth (1969), households might save time by dismissing the decision-making process and routinizing their purchases by rebuying the same brand repeatedly over time. This outcome happens more often for frequently purchased items, and especially for inexpensive items. In this way, a brand has a higher chance of being chosen again when it was previously preferred. Such a behavior can truly be treated as positive state dependence, or as it has been labeled in these circumstances, brand inertia. Solving the purchasing problems in such a manner is based on situational factors (Labeaga, Lado, & Martos, 2009). Researchers have suggested that this type of state dependence has subtle, but important variations. If a consumer repeatedly purchases a product because of high switching costs and brands availability, even in the long-run, it cannot be labeled as brand loyalty (Jacoby & Chestnut, 1978; Dick & Basu, 1994).

Brand loyalty is defined as a favorable attitude toward a brand (Labeaga, Lado, & Martos, 2009), while inertia is a passive acceptance of a brand, which happens due to high switching costs, brands availability and lack of importance.

As with other products, households make a number of simultaneous choices while purchasing beer, including both the brand as well as the type or style of beer. The main two beer manufacturers in the U.S. are Anheuser-Busch InBev and MillerCoors. Their combined market share of more than 70% dominates the market. However, both manufacturers have focused on brands throughout their histories, introducing many brands both in the past century and very recently. Both supply-side and demand-side factors have motivated brand strategies in the beer market. Substantial changes in both consumption and production have occurred due to mergers on the supply side. On the demand side, we have seen significant variations because of changes in tastes, the popularity of 'real' ale and finally the rise in popularity of premium beers (Pinkse & Slade, 2004).

Accurately modeling and measuring a brand loyalty parameter can drastically affect the measurement of consumer impacts and welfare in brand-choice models. It has been an important concept in both economics and marketing literature for the past few decades (Howard & Sheth, 1969). The first attempt at incorporating brand loyalty as a part of a discrete-choice model was done by Guadagni and Little (1983). From then on, researchers have recurrently used variations of this new state dependent variable in their discrete-choice models by including a lagged purchase variable (Lattin & Bucklin, Reference effects of price and promotion on brand choice behavior, 1989), or by creating an explanatory variable based on past purchase history (Guadagni & Little, 1983; Lattin & Bucklin, Reference effects of price and promotion on brand choice behavior, 1989; Keane, 1997; Ortmeyer, Lattin, & Montgomery, 1991).

Aaker (1991) mentions brand loyalty as one of the basic elements in brand equity measurement, and in his book, he reveals how brand loyalty can cause a reduction in marketing costs, more customers, and higher trade leverage. Later, Dick and Basu (1994) elaborate on brand loyalty by describing how loyal customers might be resistant to competitive strategies and also how loyalty can reduce searching motivation, increase the acceptance of pre-selections and reassuring word of mouth. Keane (1997) clarifies the importance of household demographics and state dependent variables in brand-choice models. Chaudhuri and Holbrook (2001) examine the influences of loyalty on market share and relative price.

The early work by Guadagni and Little (1983) shows that adding a brand-loyalty index to a brand-choice model can improve the model's fit. By their definition, brand loyalty is a state-dependent persistence of preference for a particular brand despite potential price changes. They argue that because past purchase history reveals a lot of information on consumer preferences, a brand-loyalty variable defined as an exponentially weighted sequence of past purchases can add a lot of explanatory power to a model. In their model, brand loyalty could capture two main effects, state dependence and heterogeneity. The purchase feedback or state dependence denotes the impact of past purchases, while heterogeneity refers to the differences across households in brand preference or market response. While their index could encompass the differences in purchase behavior, it could not distinguish between sources of variation. In other words, there was no way to understand what parts of the behavior are due to heterogeneity and what parts are related to changes within a household through time. Other research has explored the possible shortcomings of this well-known index (Ortmeyer, Lattin, & Montgomery, 1991; Fader & Lattin, 1993; Dong & Stewart., 2012). One other main shortcoming of the Guadagni and Little's model was a lack of accuracy in the measurement of the "smoothing" parameter in their model for brand loyalty. Later, Fader and Latin (1992) proposed a method for linear approximation of the loyalty index with Taylor expansion. While this method partly resolves the smoothing parameter measurement issue, it relies on a single smoothing parameter for all households during the whole period of the investigation, and it is static in the sense that neglects the effect of time-related variables on the smoothing parameter. Thus, not only is the smoothing parameter identical for all households, but it is also static.

To overcome the shortcomings of the original model, Fader and Latin (1993) also introduced a new loyalty index based on Dirichlet-Multinomial model that could accept sudden preference changes among households. Later Dong and Stewart (2012), took a step further and improved this model by incorporating heterogeneity in the model as well.

In this paper, we introduce a method that is able to estimate dynamic smoothing parameters for heterogeneous households. Thus, we resolve the main shortcomings of the Guadagni and Little's (1983) model by a simpler method than the Dirichlet-multinomial choice model. The model developed in this study is a new choice model that integrates household-level demographics and purchase history so that the brand-loyalty smoothing parameter reflects household heterogeneity and time variation. Formally, we use a conditional multinomial logit model for the brand choice that incorporates our brand loyalty variable calculated with our new method. We show how the weight of the brand-loyalty smoothing parameter changes through time and across households. We also illustrate this variation is linked to observed household characteristics.

Our new method for estimating the smoothing parameter and brand loyalty has an impact on our brand-choice model estimation. Almost all of the estimated coefficients show an increase in absolute value. This is an expected result if our conditional logit model explains more variance. In our analysis, we estimate the marginal probability of choosing each beer brand after accounting for heterogeneity in the brand-loyalty process and compare our results with those from a model that treats brand loyalty in a more restricted fashion.

According to our results, when income changes, it is primarily the demand for high-end beer brands that is affected. Our results also reveal that price is not always the most effective factor in the household decision-making system and it does not have the highest coefficient in absolute value. Reviewing elasticities uncovers how effective both location and seasonality are in beer brand purchasing probability. Seasonality has a significant and contradictory effect on different brands. Spring increases the probability of purchasing some brands in some regions while it works against some other brands in the same region.

Model:

The application of Logit models to brand-choice analysis is well established and widely accepted. The common multinomial logit (MNL) model assumes that choice probabilities are related to a linear combination of the attributes:

$$P^i_{jt} = \frac{e^{v^i_{jt}}}{\sum_k e^{v^i_{kt}}} \quad , where \qquad (1)$$

P^i_{jt}= The probability that household *i* chooses alternative *j* on purchase occasion *t*

v^i_{jt}= the deterministic part of utility of the brand j to household i at purchase occasion t

$= \sum \beta_r x^i_{jrt}$

x^i_{jrt}= rth explanatory variable of brand j and household *i* on purchase occasion *t*

β_r= Coefficients to be estimated

Here our main concern is household characteristics and among them we are focusing on a special case of state dependency and not demographics or income level. Based on Guadagni and Little (1983), we know that habit formation can have a significant effect on food choices. In their study, which investigates ground-coffee products, they show how households' past purchase histories may affect their decisions in the future in a manner similar to inertia in mechanics. Other research shows that educational level, income, household size, race, and ethnicity can influence the brand a household chooses (Davis, Dong, Blayney, & Owens., 2010). The multinomial logit permits an axiomatic derivation (Erlander, 2010). In our setting, we are using different products as our alternatives while each product has its unique specifications which we simply summarize all this package as one unique characteristic: "Brand".

To use a multinomial logit as our brand-choice model we maintain four main assumptions:

1- Each alternative holds a preference or utility for our consumer.
2- Our consumer chooses the alternative that satisfies him the most, or gives him the highest utility. This assumption enables us to calculate the probability of a product being chosen based on the probability of its utility being the highest in the set.
3- The random components of the consumer's utility are independently distributed random variables with a double exponential distribution.
4- We assume that each household would choose wisely to maximize her utility.

We estimate a brand-choice model that accounts for the main determinants of a household's choice between beer brands. A brand is a product of a manufacturer distinguished by unique attributes. It can be sold in different containers using different UPC codes but it still contains the same product attributes. For example, Bud Light is a brand in our model which is different from Bud Light Platinum, but Bud Light would be treated as the same brand either it is in a bottle or in a can. Due to spurious state dependence and to prevent exaggerating the effect of state dependency in our main model, we follow Keane (1997) to control for the heterogeneity (Keane, 1997). However, we specifically allow for heterogeneity in our brand loyalty index computation. Hence, we define a customer's utility function for our household i consuming brand j at purchase occasion t, as follows:

$$U_{jt}^{i} = \alpha_j + \beta X_{jt} + L_{jt}^{i}\left(\lambda_{it}(Z_i), \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) + \varepsilon_{jt}^{i} \qquad (1)$$

Where:

α_{j} = Beer brand *j*'s specific contribution to household's utility.

X_{jt} = A vector of alternative-variant variables or marketing mix variables such as prices, color, taste and availability. Some of these attributes may vary over time; some may not. For example a brand's style or color may not change over time, but availability or price are more likely to change.

$\varepsilon_{\mathrm{jt}}^{\mathrm{i}}$ = An unobservable random brand- and household-specific error term

$\mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}$ = A $(\mathrm{t}-1) \times 1$ vector of 0's and 1's with the r^{th} element equal to 1 if household *i* purchased product j at time r and equal to zero otherwise.

Z_{i} = A vector of consumer's characteristics that are not related beer brand alternatives. These are not only demographic characteristics but also geographical locations and seasonality variables.

$L_{jt}^{i}(\lambda_{it}(Z_i), \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}})$ = Our beer brand loyalty variable which is based on our alternative-invariant variables Z_{i} and past purchase history, $\mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}$. In this paper, we specify our loyalty variable for individual *i* for brand *j* at purchase occasion *t* as follows:

$$L_{jt}^{i} = \lambda_{it} L_{j(t-1)}^{i} + (1-\lambda_{it}) Y_{j(t-1)}^{i} \qquad (2)$$

Loyalty in this function is defined as a weighted average of past purchase history. Here in (2), the smoothing parameter λ_{it} shows how a household weighs her past opinion in comparison to her recent experience. We allow this weight to change though time and it also might be different among households. Both Fadder and Latin (1993) and later Dong and Stewart (2012) define their loyalty variables by using a single smoothing parameter common to all households. Finding that single smoothing parameter is easily computable, but causes specific limitations in dealing with heterogeneity and stationarity. To overcome this problem, we are not only calculating the smoothing parameter λ_{it} for households separately, but we are also treating λ_{it} as a dynamic element of the loyalty variable that may change through time due to environmental, demographical, or geographical changes. Hence, we introduce heterogeneity. Using this method and by using a dynamic λ_{it}, we do not need to think of stochastic renewals in the preferences.

We compute the smoothing parameter, (λ_{it}), by relying on a modified version of the nonlinear estimation algorithm first introduced by Fader and Latin (1992). Our modifications accommodate the heterogeneity in the smoothing parameter.

First, we expand $L_{jt}^{i}\left(\lambda, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)$ in a Taylor series around a starting value or initial point λ_{it}^{0} :

$$L\left(\lambda_{it}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) = L\left(\lambda_{it}^{0}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) + \frac{dL\left(\lambda_{it}, Y_{jt-1}^{i}\right)}{d\lambda_{it}} (\lambda_{it} - \lambda_{it}^{0}) + \sum_{n=2}^{\infty} \frac{d^n L\left(\lambda_{it}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)}{d\lambda_{it}{}^{\mathrm{n}}} \frac{\lambda_{it} - \lambda_{it}^{0}}{n!} \quad (4)$$

If $L\left(\lambda_{it}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)$ is a smooth function (e.g., its derivatives with respect to λ_{it} are bounded) in an interval containing both λ_{it}^{0} and the maximum likelihood estimate (MLE) value of λ_{it}, then the second and higher –order terms in (4) will approach 0 as λ_{it}^{0} approaches its MLE value.

If we now rename $\frac{dL\left(\lambda_{it}, Y_{jt-1}^{i}\right)}{d\lambda_{it}}$ as $L'(\lambda_{it}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}})$, we can rewrite our estimation for $L\left(\lambda_{it}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)$ as:

$$L\left(\lambda_{it}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) \cong L\left(\lambda_{it}{}^{0}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) + L'\left(\lambda_{it}{}^{0}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)\left(\lambda_{it} - \lambda_{it}{}^{0}\right) \qquad (5)$$

This approximation will become exact when the second part converges to zero. Using a dynamic smoothing parameter which changes not only through time but also among different households is the main characteristic of our model which makes if different from the previous works that has been done in this field. Similarly, as we are calculating the smoothing parameter for households at different points in time, we do not need to think about stationarity or non-stationarity of this parameter anymore. Controlling for heterogeneity among households and accepting behavioral changes through time, helps us to calculate a more exact brand loyalty index for every household at different points in time.

Because extreme heterogeneity, where every household has a different smoothing parameter, may require too much information for out data to handle, we must make some assumptions about similar households. If we assume that similar households would have similar smoothing parameters, then we just have to find a method for identifying these similarities and categorize them accordingly. By using the K-means method of clustering we can create C clusters and calculate the smoothing parameter for each cluster separately. We can estimate the number of clusters using the elbow method. We use a number of demographic characteristics and time-variant characteristics of our households to do the clustering in each period. In appendix I, we specifically show how clustering is done. For each cluster, similar to each household we can write the loyalty function similar to (5) but we have to replace λ_{it} with λ_{cp} which is the smoothing parameter for the cluster at each period of time.

Using $L\left(\lambda_{cp}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)$ in a conditional multinomial logit brand choice estimation for every cluster separately in a rich data set with frequent buyers, we can get a coefficient β, and if we apply the equality in (5) in our conditional multinomial logit model:

$$\beta L\left(\lambda_{cp}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) \cong \beta L\left(\lambda_{cp}{}^{0}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) + \beta L'\left(\lambda_{cp}{}^{0}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)\left(\lambda_{cp} - \lambda_{cp}{}^{0}\right) \quad (6)$$

Thus, if we use (6) to replace $L\left(\lambda_{ct}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right)$ in our conditional multinomial logit model, we can write down the result as follows:

$$\beta L\left(\lambda_{cp}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) \cong \beta L\left(\lambda_{cp}{}^{0}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) + \beta' L'\left(\lambda_{cp}{}^{0}, \mathrm{Y}_{\mathrm{jt-1}}^{\mathrm{i}}\right) \qquad (7)$$

where:

$$\beta' = \beta\left(\lambda_{cp} - \lambda_{cp}{}^{0}\right) \quad (8)$$

Using (8) we can simply calculate the new smoothing parameter for each cluster as:

$$\lambda_{cp} = {\lambda_{cp}}^{0} + \frac{\beta'}{\beta} \qquad (9)$$

In this way, we can accurately estimate different smoothing parameters for different clusters at any point in time.

To replace $L(\lambda_{cp}, \mathrm{Y}^{\mathrm{i}}_{\mathrm{jt-1}})$ using (6) we need both $L({\lambda_{cp}}^{0}, \mathrm{Y}^{\mathrm{i}}_{\mathrm{jt-1}})$ and $L'({\lambda_{cp}}^{0}, \mathrm{Y}^{\mathrm{i}}_{\mathrm{jt-1}})$. Calculating the loyalty variable based on the initial smoothing parameter is a simple recursive function similar to what we described in (2). Here, we can rewrite our function based on the new parameters as:

$$L^{i}_{jt} = \lambda_{cp} L^{i}_{j(t-1)} + (1 - \lambda_{cp}) Y^{i}_{j(t-1)} \qquad (10)$$

And if we replace $L^{i}_{j(t-1)}$ by its definition, (10) becomes:

$$L^{i}_{jt} = \lambda_{cp}\left(\lambda_{cp} L^{i}_{j(t-2)} + (1 - \lambda_{cp}) Y^{i}_{j(t-2)}\right) + (1 - \lambda_{cp}) Y^{i}_{j(t-1)} \qquad (11)$$

Continuing this process and simplifying the results will lead us to:

$$L^{i}_{jt} = (\lambda_{cp})^{t-1} L^{i}_{j(1)} + (1 - \lambda_{cp})[\sum_{s=0}^{t-2} (\lambda_{cp})^{s} Y^{i}_{j(t-1)}] \qquad (12)$$

The initial period, or starting point, presents another analytic problem that requires one more assumption: Based on the definition we accepted for the loyalty variable, we assume that at the starting point the consumer has a similar brand loyalty toward all brands or:

$$(L^{i}_{j(1)} = \frac{1}{n}) \qquad (13)$$

where n is the number of brands. While our initialization is different from Guadagni and Little (1983), it is in line with Fader and Lattin (1992). Using (12) as a general form of (2) we can now estimate brand loyalty for any household at any point in time. In this paper, we are using (12) as the base of our calculations while Fader and Latin (1992) has used a more restricted form of it. To estimate $L'({\lambda_{cp}}^{0}, \mathrm{Y}^{\mathrm{i}}_{\mathrm{jt-1}})$ we can now use (12) and find the first derivative of $L({\lambda_{cp}}^{0}, \mathrm{Y}^{\mathrm{i}}_{\mathrm{jt-1}})$ with respect to the smoothing parameter λ_{cp}:

$$DL^{i}_{jt} = \lambda_{cp} DL^{i}_{j(t-1)} + L^{i}_{j(t-1)} - Y^{i}_{j(t-1)} \qquad (14)$$

We can rewrite (14) using the same recursive form as:

$$DL^{i}_{jt} = \sum_{s=0}^{t-2} (\lambda_{cp})^{s} L^{i}_{j(t-1-s)} - \sum_{s=0}^{t-2} (\lambda_{cp})^{s}.\, Y^{i}_{j(t-1-s)} \qquad (15)$$

Using (15), we can calculate the derivative based on all these known elements. Equation (15) is similar to (12), and both are computable recursively based on the initial conditions and the available data. Thus, we can compute a smoothing parameter for each household and each period.

Because we are interested in investigating how the smoothing parameter is related to household characteristics, we specify another equation to investigate this relationship:

$$\lambda_{it} = \gamma_0 + \gamma_1 Z_{it} + \gamma_2 S_{it} + \gamma_3 G_{it} + \gamma_4 X_{it} + \varepsilon\,, \quad (10)$$

where Z_{it} is a vector of household demographic characteristics, S_{it} is a vector of seasonality and environmental variables, G_{it} is another vector controlling with locational variables, and finally we are looking at the effects of time variant variables like the amount of beer consumed per period and the amount of wine consumed at the same period. The smoothing parameter is the weight each household puts on her long-term experience in comparison to the weight the household gives to her recent understanding of a brand.

By having our smoothing parameter calculated, we are now armed with an accurate brand-loyalty index which we can use in a brand-choice multinomial logit. Based on MNL assumptions, we test our alternatives to see if Independence of Irrelevant Alternatives holds.

Data:

We use the Nielsen Homescan panel data on alcohol and cigarette purchases for this study. Our sample period runs from 2009 through 2011. Nielsen picks its sample of households in a way to be geographically and demographically demonstrative of the United States. Households use a home scanner to record purchases of food and related items, including the quantities purchased, the price paid, and the date. Purchases consumed before returning home are generally not scanned and thus missing from the data. Einav, Leibtag, and Nevo (2008) analyze the accuracy of the Homescan data but note that its shortcomings and errors are of the same order of magnitude as many other collected datasets.

We have divided the panel data into eight three-month seasons. During out study period, 24 beer brand names account for the bulk of market sales. Adding an outside option to these 24 we have gathered 25 alternatives in table 1 to create our choice alternatives for households in this study. Using an outside option is classic technic to control for brands which are not in the main alternative list (Erdem & Keane, 1996) Market shares are also presented in table 1. For our conditional logit model, we need the full vector of prices faced by each consumer on each purchase occasion for all brands in the choice set. The problem , however, is that the Homescan dataset records only the price paid by the household for the purchased brand. To overcome this difficulty, we look for a series of matching prices. First, we look for purchases by other households from a similar store in the same region at the exact same week for each specific brand among our alternatives. If we fail to find a match with this method, we have use the week's average state price for that brand for each purchase occasion. For our outside option, we similarly use the week's price weighted average in each state.

The Homescan data provide detailed information on every transaction and a wide array of demographic characteristics for participating households. During 2009-2011, our dataset shows transactions over more than 2700 different brands. Despite the great diversity among brands, 75% of transactions are related to the top 24 brands.

Table 1. Alternatives and Marketing Variables

Brand Names	Average Price/Oz	Discount	Market share	Brands market share by season			
				Spring	Summer	Fall	Winter
BUD LIGHT	0.0639	0.2203	9.96%	9.88%	10.11%	9.93%	9.92%
MILLER LITE	0.0622	0.2281	7.60%	7.78%	7.65%	7.46%	7.46%
BUDWEISER	0.0666	0.1868	6.84%	6.81%	6.67%	6.76%	7.16%
COORS LIGHT	0.0634	0.2140	6.58%	6.34%	7.03%	6.72%	6.18%
NATURAL LIGHT	0.0472	0.1610	4.86%	5.04%	4.61%	4.75%	5.10%
BUSCH LIGHT	0.0473	0.1723	3.53%	3.43%	3.50%	3.60%	3.61%
MILLER HIGH LIFE	0.0489	0.1559	3.32%	3.25%	3.28%	3.31%	3.44%
BUSCH	0.0487	0.1397	3.11%	3.20%	2.96%	3.10%	3.18%
KEYSTONE LIGHT	0.0461	0.1533	2.97%	2.91%	2.85%	3.10%	3.03%
MICHELOB ULTRA LIGHT	0.0756	0.1671	2.69%	2.74%	2.75%	2.57%	2.68%
MILWAUKEE'S BEST LIGHT	0.0432	0.1109	2.27%	2.40%	2.15%	2.11%	2.44%
CORONA EXTRA	0.0956	0.2319	1.94%	2.12%	2.30%	1.55%	1.73%
HEINEKEN	0.1025	0.2208	1.92%	1.86%	1.89%	1.90%	2.03%
MILLER G DRAFT	0.0652	0.2379	1.90%	1.93%	1.98%	1.84%	1.82%
NATURAL ICE	0.0472	0.1218	1.61%	1.73%	1.48%	1.54%	1.72%
MILLER G D LIGHT	0.0642	0.2277	1.46%	1.34%	1.61%	1.64%	1.24%
BUD LIGHT LIME	0.0849	0.1643	1.33%	1.31%	1.71%	1.22%	1.01%
MILWAUKEE'S BEST	0.0454	0.1297	1.30%	1.30%	1.23%	1.34%	1.35%
MILLER HIGH LIFE LIGHT	0.0457	0.1589	1.13%	1.12%	1.17%	1.14%	1.07%
MILWAUKEE'S BEST ICE	0.0431	0.1170	1.04%	1.09%	0.99%	0.99%	1.10%
BUDWEISER SELECT	0.0622	0.3187	1.04%	1.10%	1.02%	0.94%	1.08%
YUENGLING AMBER LAGER	0.0695	0.1294	0.89%	0.89%	0.96%	0.86%	0.84%
COORS BANQUET	0.0652	0.2549	0.83%	0.84%	0.84%	0.84%	0.79%
CORONA LIGHT	0.0945	0.2624	0.79%	0.93%	0.90%	0.67%	0.62%
Outside Option	0.0724	0.1800	29.09%	28.67%	28.35%	30.11%	29.38%

In addition to a number of household characteristics, we construct a number of transaction variables based on the transactional data. As noted above, we record the price for each brand and construct variables related to beer type (or flavor), purchased volume, amount purchased on promotion, product container information, and ultimately a brand loyalty variable.

For our study, we focus on households with at least XXX transactions over the 8 seasons, and after dropping infrequent purchasers, we are left with 200 frequent beer-purchasing households. To calculate the brand loyalty using a dynamic smoothing parameter we create nine different household clusters in each season. Table A.1 in Appendix A provides information on these clusters and how we have categorized households in these different clusters. We allow cluster membership changes through time so one household is not always assigned to the same cluster.

Table 3 shows household characteristics that we use for clustering. Euclidean distance between clusters is used in k-means clustering. As the scale of our characteristics are different, we normalize our data not to have different weighs among our characteristics. These variables are the same variables we have used to control for the heterogeneity of the smoothing parameter.

Table 2. Household Variables

Variable	Mean	Std. Dev.	Min	Max
Household size	2.20	1.07	1	7
Household income ($)	44331	27550	0	200000
Age of female head	39.60	23.97	0	65
Age of Male head	48.93	17.73	0	70
Female head employment (hours per week)	14.29	18.02	0	40
Male head employment (hours per week)	23.79	18.74	0	40
Mail Head Education (if college education gaind =1 if not=0)	0.59	0.49	0	1
Female Head Education (if college education gaind =1 if not=0)	0.42	0.49	0	1
Marital Status (Married=1 not married=0)	0.31	0.46	0	1
Hisponic (Hispanis =1 not Hispanic=0)	0.06	0.24	0	1
Amount of beer per season (Floz per season)	7943	5744	72	40992
Amount of wine per season (Floz per season)	136.5	492.2	0	4844
Cigarets per season (Total count per season)	892.5	1525.4	0	8000

It worth mentioning that to use time-variang, household-specfic variables for the clustering algorithm , we have to construct some because many households' characteristics are essentially constant (such as race or ethnicity) and a lot of others would not change drastically among all the households during a period of two years (variables like income, education or employment). So, to have a better understanding of behavioral changes among households, we construct household-specific, time-varying variables reflecting the amount of beer purchased per season, the amount of wine (as a substitute) purchased per season, and finally the amount of cigarettes (as a complement) purchased per period for each household.

To make sure that this forced displacement among different clusters is not the source of variation for the smoothing parameter, we did a test to show this is not the case. To overcome this challenge, we picked some of the clusters with varieties of smoothing parameters and looked at them through time without changing the clusters. For example, we took cluster seven from season one and looked at it through time in the next three seasons without changing the cluster members, and the results are the same as before.

Results:

The model shown above was estimated using Matlab software. In the brand choice model estimated in this study, aside from household demographic charachtersitics, seasonality and marketing mix variables, we have also used the brand loyalty measure based on equation (2). This brand loyalty index, accounts for changes in sensitivities across households and tracks changes in sensitivities through time. When first introduced by Guadagni and Little (1983), a complex method was used to estimate the value of the smoothing parameter. They chose an initial value of 0.75 and by adding dummy variables; they captured the carry over effects for the last ten purchases of each household. To sum up, they fit an exponential decay curve to these dummy variable coefficients to derive estimates of $\lambda_B = 0.875$ for brand loyalty and $\lambda_S = 0.812$ for package size loyalty. Although this complex method was accurate, they were still using the same smoothing parameter for all the household neglecting heterogeneity. Here we estimated our smoothing parameter (λ_{it}) based on equation (9). To calculate this smoothing parameter through time, we divided our whole-time period into eight seasons. As the methodology shows, we could calculate smoothing parameter for each household separately, but to make this estimation more practical, and by using K-means clustering, we divided our households into nine different clusters, which was the optimized number of clusters based on the elbow method. . Clusters in this estimation are not the same and we have done the clustering for each season separately. This way we let new clusters to emerge based on changes in household characteristics and we have made mobility possible for different households, in other words, if there are sudden changes in household characteristics they might move from one cluster into another cluster through time. Then, brand loyalty (L_{ijt}) was estimated using a time variant smoothing parameter(λ_{it}), which we calculated in the first step. Table 4 shows different smoothing parameters for nine random households with different characteristics in different seasons.

Table 4. Random households smoothing parameters through time

Household Id	S1	S2	S3	S4	S5	S6	S7	S8
2016156	0.67	0.74	0.71	0.76	0.73	0.74	0.71	0.70
30074365	0.7	0.73	0.66	0.73	0.74	0.78	0.73	0.69
8085592	0.66	0.73	0.69	0.75	0.67	0.73	0.72	0.70
2007345	0.66	0.73	0.69	0.76	0.67	0.73	0.72	0.70
2022433	0.81	0.74	0.55	0.61	0.69	0.73	0.70	0.63
30451196	0.74	0.81	0.72	0.80	0.75	0.75	0.75	0.71
8284863	0.31	0.73	0.30	0.73	0.69	0.77	0.90	0.69
30044347	0.71	0.67	0.72	0.76	0.74	0.71	0.65	0.74
30469226	0.62	0.73	0.69	0.76	0.67	0.73	0.72	0.70

From table 4, we see how different the smoothing parameter can be, not only among different households but also for a single household through time. Both these sample results and the more complete tables we have in the appendix, shows how people's behavior might change through time and how they might weight their own opinion differently through time. Lower smoothing parameters shows how a household might weight her recent experience more than her previous opinions. It is interesting to see how a household might change her behavior toward her own opinion even in just on season. Although the average of 0.7 looks a respectable average to use in these cases, there are strong deviations from this average among households. Using these dynamic smoothing parameters in equation (2) and calculating brand loyalties, equipped us with a more precise index to calculate the purchasing probabilities for each brand. To investigate the smoothing parameters relation to household characteristics, we regressed equation (10) and the results are presented in table 5. As predicted, the smoothing parameter is significantly correlated with household demographics. According to results shown in Table 4 and 5, we accept smoothing parameter as a varying parameter, related to different demographic, seasonal and locational factors. To evaluate the influences of a varying smoothing parameter in comparison with common constant smoothing parameters, we defined two separate models on our beer market data of 2009-2011. First, we used our newly estimated smoothing parameter to calculate brand loyalty of every household for each household in every purchase date. Using this new loyalty index, we have estimated a brand choice model using conditional multinomial logit model explained in the model section.

Table 5. Smoothing parameters and household characteristics corrolation

lambdah	Whole period	2009	2010
Household Siz	0.0044991* (0.0000375)	0.0053111* (0.0000529)	0.0033672* (0.0000535)
Log of household average annual income	-0.0033116* (0.0000735)	-0.0034694* (0.0001117)	-0.0038154* (0.0000793)
Male head age	0.0001664* (0.00000363)	-0.0000259* (0.00000612)	0.0003815* (0.00000283)
Female head age	0.0002299* (0.00000207)	0.00041* (0.00000312)	0.0000853* (0.00000243)
Male head education	-0.005301* (0.0000732)	-0.0086477* (0.0000991)	-0.0031696* (0.000103)
Female head education	0.00535* (0.0000869)	0.0091585* (0.0001307)	-0.000442* (0.0001018)
Female head working hours per week	-0.0000879* (0.00000223)	-0.0001583* (0.00000361)	-0.0000082* (0.00000249)
Male head working hours per week	0.0005003* (0.00000251)	0.000507* (0.00000383)	0.0005123* (0.0000026)
Beer amount purchased per season	0.000000681* (0.0000000104)	0.00000119* (0.000000015)	-0.000000265* (0.0000000117)
Wine amount purchased per season	-0.00000329* (0.0001013)	-0.00000477* (0.0000000792)	0.000000282* (0.000000119)
Ciggaretts count purchased per season	0.00000125* (0.0000000259)	-0.00000075* (0.000000034)	0.00000404* (0.0000000384)
Black/ African American	-0.0087799* (0.0001013)	-0.0128583* (0.0001589)	-0.0045512* (0.0001072)
Asian	0.010502* (0.0003249)	0.0042997* (0.000458)	0.0165922* (0.0003847)
Other races	-0.0543196* (0.0006921)	-0.1123275* (0.0009598)	0.0230712* (0.0004953)
seasdum1	0.0326982* (0.0000896)	0.0451638* (0.0001574)	0.0288903* (0.0000918)
seasdum3	0.0133337* (0.0001041)	0.0318551* (0.000166)	0.0005599* (0.0001215)
divdum1	0.0094139* (0.0012)	0.0014139* (0.001662)	0.0183865* (0.0014844)
divdum2	0.0023992* (0.0006324)	-0.0032059* (0.0009727)	0.008978* (0.0006803)
divdum3	0.0024304* (0.0006083)	-0.0035027* (0.0009382)	0.0092693* (0.0006545)
divdum4	0.0041525* (0.0015145)	-0.0252131* (0.0028251)	0.020627* (0.001245)
divdum5	0.0023534* (0.0006054)	-0.0034516* (0.000934)	0.0090907* 0.001245)
divdum7	0.0022371* (0.0006363)	-0.0039275* (0.0009796)	0.0088738* (0.0006515)
divdum8	-0.0003822* (0.0008991)	-0.0092617* (0.0014449)	0.0090028* (0.0010109)
divdum9	0.0026025* (0.0006139)	-0.0032916* (0.0009463)	0.0093889* (0.0006603)
_cons	0.697619* (0.0009452)	0.6909325* (0.001482)	0.703504* (0.0009976)
R-squared	0.229	0.28	0.225

Therefore, to compare our results with previous models, in the second step, we have used a constant smoothing parameter equal to 0.7, which was the constant in our regression, and the average of all smoothing parameters calculated through the whole two years. Using a constant smoothing parameter with no modification. This is in line with what most researchers used to do previously. As Fadder and lattin(1992) mentioned, simply choosing a suitable value of λ (usually between 0.7 and 0.9) and not to attempt to refine it has been a common procedure in this line of study. Examples include Lattin (1987), Gupta (1988), Kalwani et al. (1990) and Ortmeyer, Lattin and Montgomery (1991) .

Table 6 shows the results of the first model. It shows the parameter estimates and the value of the log likelihood function. There we have also shown the predicted purchasing probability of each brand. As it is shown in table 5, household variables indirectly effect brand choices through the brand loyalty index. When controlling for these variables, we see that many of them are significant at 5% and some are significant even at 1% level. Looking at the marketing mix variables also reveals how important these variables are in the decision making process for all brands. As we are using a conditional logit model, we had to have a base line for normalization. In this study, we have used Bud Light as our baseline brand so all the results for household characteristics, seasonality and locations in the brand choice model are estimated relative to the value of the Bud Light constant. From Table 6, we calculated the average elasticities of beer brand choice probabilities for household related variables. As Bud light is our base line in the conditional logit model, we have not put it in Table 6.

According to our results, creating the predicted probability variable would help us to have an estimation on the models fit. The model can predict with 82% accuracy, what the next purchase would be. In other words when we fit the probability prediction with the actual purchases we find out that 82% of the times the most probable choice is the one that the target household would actually pick.

Table 6. Brand choice model 1 (Varying smoothing parameter)

Variable	Natural light	Budweiser	Miller lite	Busch busch	Keystone light	Miller life	Natural ice	Milwaukee's light	Heineken heineken	Busch light	Coors light	Outside option
Probability	8.443%	7.682%	6.373%	6.013%	3.690%	3.347%	3.122%	2.974%	2.895%	2.673%	1.790%	42.602%
Household Variables:												
Log income	-.3845* (.08)	-.2304* (.0844)	-.0672 (.0797)	.020* (.0058)	-.1850* (.08778)	-.6773* (.0852)	4.9* (1.033)	-.3687* (.0908)	.28 (1.180)	-.2346* (.081)	.01224 (.1057)	.0948 (.0634)
Max age	.0139* (.0051)	.04002* (.0052)	.0003 (.0057)	.0211* (.00388)	.0007 (.0057)	.0057 (.0071)	-.6163* (.0947)	.04016* (.0073)	.0085 (.1023)	.0045 (.0055)	.01202 (.0062)	.014* (.0043)
Max Emp	.0065* (.0031)	.0167* (.0033)	-.0036 (.0032)	.2198* (.1071)	-.0008 (.0038)	.0209* (.0039)	.0256* (.006)	.0203* (.004)	-.0146 (.008)	.0076* (.00376)	.02114* (.0046)	.00021 (.00267)
Married	.0544 (.0993)	.2052* (.0945)	-.0720 (.0966)	-.0302 (.1581)	.46266* (.1053)	-.042 (.11)	.0014 (.0036)	.26747* (.1252)	.0155* (.0051)	.157 (.1073)	-.0366 (.1221)	.2886* (.08)
Black/ African American	-.1297 (.1473)	.1322 (.1468)	-.4844* (.1595)	19.02 (4245.2)	-.4034* (.1504)	-.1571 (.1546)	-.0265 (.1167)	-2.123* (.4291)	-.0393 (.13)	-.4236* (.185)	.2350 (.1565)	.0437 (.130)
Asian	18.51 (4245.2)	.726 (5003)	.4489 (5351)	18.21 (1998.4)	.698 (5078)	.29363 (6435.8)	-2.788* (.5168)	17.38 (4245)	.2648 (.180)	16.66 (4245)	.3492 (5606)	18.268 (4245)
Other Races	17.82 (1998.4)	15.605 (1998)	16.26 (1998)	.00002* (.00001)	17.41 (1998.4)	18.54 (1998.4)	17.33 (4245)	17.09 (1998)	.7372 (5707)	.4314 (2619)	.2771 (2958)	18.48 (1998)
Beer per season	.00003* (1e-05)	0 (0)	-.00001 (1e-05)	.0002 (.00013)	.00005* (1e-05)	-3e-05* (1e-05)	15.08 (1998)	6.94e-07 (1e-05)	17.64 (1998)	3e-06 (1e-05)	-3e-06 (1e-05)	-.00001* (8.21e-06)
wine per season	.0001 (.0001)	.0002* (.0001)	.0002 (.0001)	-.0001* (.00003)	-.0014* (.0003)	-.0007* (.0003)	-1.3e-06 (1e-05)	-.00018 (.00018)	-.0001* (1e-05)	-.00004 (.0001)	.0003* (.00013)	.0001 (.0001)
Cigar per season	-.0002* (4e-05)	-.0001* (3e-05)	-.00002 (3e-05)	.0191 (.171)	-.0004* (4e-05)	-.0003* (6e-05)	.0005* (.00012)	-.000014 (3e-05)	.0003* (.0001)	-.0002* (3e-05)	-.0001* (4e-05)	-.00008* (2e-05)
Marketing Mix Variables												
Constant	3.134* (.8666)	-.3670 (.8848)	1.0965 (.8930)	-.6006* (.0851)	1.447 (.9394)	6.454* (.9205)	1.8e-06 (3e-05)	.8429 (1.034)	.00003 (3e-05)	2.061* (.8731)	-1.704 (1.152)	-1.32 (.696)
Spring	-.0544 (.1533)	.0796 (.1507)	-.0374 (.1590)	-.1341 (.1721)	.1002 (.1645)	-.2888 (.16964)	-.0722 (.18)	-.14767 (.1897)	.1172 (.1981)	.0887 (.1661)	.1327 (.1844)	-.01942 (.127)
Summer	.0632 (.1516)	-.04213 (.1466)	.0117 (.1576)	.01302 (.1638)	.09615 (.16268)	-.2146 (.1668)	-.0640 (.18)	-.2621 (.1885)	.0696 (.1966)	-.06987 (.16731)	.0808 (.1818)	-.0086 (.1248)
Fall	-.1078 (.1468)	.0235 (.1419)	.1092 (.1519)	4.4776* (.9137)	.084 (.1568)	-.0986 (.1593)	.0216 (.1733)	-.106 (.1815)	.1057 (.1882)	.0176 (.16152)	-.2179 (.1769)	.12017 (.1204)
Price						.0946 (.1287)						
Loyalty						5.303* (.0196)						
Discount						-.3833* (.0204)						
Volume						.05383* (.0016)						
Likelihood						-42913						

Investigating the household variables in table 7, reveals that race is an important factor in brand choices. While it is more likely for African Americans to purchase Natural light or Budweiser, at the same time, it is less likely for Asians or Other races to pick these brands. On the other end, being part of any of these three races would increase the likelihood of being interested in the outside option significantly. While higher income would result in higher likelihood for choosing outside option, it decreases the likelihood of picking all other top brands except for Miller lite, Coors light and the popular imported brand, Heineken. Married couples are more likely to choose the outside option. It is interesting to see how household's seasonal consumptions have opposite effects on the outside option and keystone light. Any increase seasonal wine or cigarettes consumption will decrease the likelihood of choosing Keystone light while these increases would increase the likelihood of picking the outside option. It is exactly the opposite for seasonal beer consumption.

Table 7. Elasticities of Choice Probability with Respect to Household Variables

Variable	Natural light	Budweiser	Miller lite	Busch busch	Keystone light	Miller life	Natural ice	Milwaukee's light	Heineken heineken	Busch light	Coors light
Logincome	-7.1E-03	-7.1E-03	4.5E-03	-2.8E-02	-1.8E-03	-1.8E-02	-1.5E-02	-6.9E-03	4.2E-03	-2.6E-03	2.7E-03
Max age	1.9E-03	1.9E-03	-9.6E-04	2.8E-04	-5.4E-04	-3.2E-04	3.2E-04	7.4E-04	-8.7E-04	-2.9E-04	-6.0E-05
Max emp	7.9E-04	7.9E-04	-6.4E-04	8.8E-04	-2.7E-04	4.8E-04	-1.6E-04	4.1E-04	2.6E-04	3.1E-05	2.6E-04
Married	1.4E-03	1.4E-03	-1.6E-02	2.0E-03	1.1E-02	-7.4E-03	-6.4E-03	2.4E-03	-6.3E-03	-8.0E-04	-3.8E-03
Black/ African American	2.5E-02	2.5E-02	-1.7E-02	8.8E-03	-7.3E-03	6.6E-04	-4.1E-02	-3.2E-02	1.4E-02	-5.7E-03	8.1E-03
Asian	-8.1E-02	-8.1E-02	-6.7E-02	1.2E-01	-3.9E-02	-3.5E-02	-1.3E-02	-1.1E-02	-3.1E-02	-1.8E-02	-1.9E-02
Other Races	-6.9E-02	-6.9E-02	-5.0E-02	2.1E-02	-1.3E-02	2.8E-02	-3.0E-02	-1.5E-02	-5.8E-03	-4.7E-02	-3.2E-02
Beer per season	1.0E-06	1.0E-06	-3.5E-07	2.3E-06	2.5E-06	-6.7E-07	3.4E-07	3.8E-07	-2.3E-06	4.2E-07	1.7E-07
wine per season	1.6E-05	1.6E-05	8.9E-06	1.1E-05	-5.4E-05	-2.5E-05	1.4E-05	-7.0E-06	9.0E-06	-2.5E-06	4.6E-06
Cigar per season	2.5E-06	2.5E-06	5.7E-06	-2.0E-06	-1.1E-05	-8.3E-06	3.6E-06	3.0E-06	4.2E-06	-3.1E-06	-5.4E-08

Conclusion:

As it has been discovered through other researches and as we have seen in this study, household brand choices depends on household characteristics, marketing mix variables of alternatives, seasonality and last but not least household past purchase history. This study estimated a dynamic smoothing parameter to make the brand loyalty index more accurate. Precisely, we expand on Fader and Lattin's (1992) to create a model with more flexibility through time, which covers household heterogeneity for brand loyalty estimation. Creating a dynamic smoothing parameter has made it possible to accept heterogeneous behaviors among households. This way we have a parameter to explain sudden changes in household behavior and explain how households would weigh their past choices in comparison to their recent experiences.

When it comes to the application of the model in the US beer market, we first see how the classic smoothing parameter of the Guadagni and Little's (1983) should not be considered as a constant through time and then we discover how to treat the smoothing parameter as a household dependent variable. As brand loyalty affects beer brand choices strongly, having an accurate measure of brand loyalty and knowing how it changes through time would help them optimizing their segmentation and targeting plans. In line with previous researches, we find the strong impact of marketing mix variables like price, discount and volume.

Briefly, our study offers introductory evidence about the correlation between the way household weigh their experiences and their characteristics. It also tries to open a discussion on how we can measure these changes through time.

The results for the US beer market shows how using such a dynamic model might be an improvement in using discrete choice models of this kind and how this new method of estimating brand loyalty can create a more effective index. Furthermore, the proposed model in this study enables us to measure the smoothing parameter of each household at any point in time separately. As a market scientist, an economist or a data analysist, measuring brand loyalty accurately, is an important tool in market segmentation, targeting and policymaking.

References

Aaker, D. (1991). *Managing brand equity. Capitalizing on the Value of a Brand Name.* New York: The Free Press.

Ballantyne, R., Warren, A., & Nobbs, K. (2006). The evolution of brand choice. *Journal of Brand Management* , 339-352.

Bentz, Y., & Merunka, D. (2000). Neural networks and the multinomial logit for brand choice modelling: a hybrid approach. *Journal of Forecasting 19.3*, 177-200.

Chaudhuri, A., & Holbrook, M. B. (2001). The chain of effects from brand trust and brand affect to brand performance: the role of brand loyalty. *Journal of marketing*, 81-93.

Davis, C. G., Dong, D., Blayney, D. P., & Owens., A. (2010). *An Analysis of U.S. Household Dairy Demand.* U.S. Dept. of Agr. Retrieved from Http://www.ers.usda.gov/publications/tb1928/tb1928.pdf

Dick, A., & Basu, K. (1994). Customer loyalty: toward an integrated conceptual framework. *Journal of the academy of marketing science 22.2* , 99-113.

Dong, D., & Stewart., H. (2012). Modeling A Household's Choice among Food Store Types. *American journal of agricultural economics 94.3*, 702-717.

Einav, L., Leibtag, E., & Nevo, A. (2008). *On the accuracy of Nielsen Homescan data.* USDA, Economic Research Service, .

Erdem, T., & Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing science 15.1*, 1-20.

Erlander, S. B. (2010). *Cost-minimizing choice behavior in transportation planning: a theoretical framework for logit models.* Springer Science & Business Media.

Fader, P. S., & Lattin, J. M. (1993). Accounting for heterogeneity and nonstationarity in a cross-sectional model of consumer purchase behavior. *Marketing Science 12.3*, 304-317.

Fader, P., & Lattin, J. (1992). Estimating nonlinear parameters in the multinomial logit model. *Marketing Science 11.4*, 372-385.

Guadagni, P. M., & Little, J. D. (1983). A logit model of brand choice calibrated on scanner data. *Marketing science 2.3*, 203-238.

Gupta, S. (1988). Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing research*, 342-355.

Howard, J., & Sheth, J. N. (1969). *The theory of buyer behavior.* New York: Wiley.

Jacoby, J., & Chestnut, R. W. (1978). *Brand Loyalty Measurement and Management.* New York: Wiley.

Kalwani, M. U., Yim, C. K., Rinne, H. J., & Sugita., Y. (1990). A price expectations model of customer brand choice. *Journal of Marketing research*, 251-262.

Keane, M. P. (1997). Modeling heterogeneity and state dependence in consumer choice behavior. *Journal of Business & Economic Statistics 15.3* , 310-327.

Kotler, P. (2000). *Marketing Management, Millennium Edition.* Prentice-Hall.

Labeaga, J. M., Lado, N., & Martos, M. (2009). Brand loyalty and heterogeneity in the Spanish market of non-fine laundry detergents. *capítulo 10*, 99-116.

Lattin, J. M. (1987). A model of balanced choice behavior. *Marketing Science 6.1*, 48-65.

Lattin, J. M., & Bucklin, R. E. (1989). Reference effects of price and promotion on brand choice behavior. *Journal of Marketing research*, 299-310.

Ortmeyer, G., Lattin, J. M., & Montgomery, D. B. (1991). Individual differences in response to consumer promotions. *International Journal of Research in Marketing 8.3*, 169-186.

Pinkse, J., & Slade, M. E. (2004). Mergers, brand competition, and the price of a pint. *European Economic Review 48.3*, 617-643.