How Do Experience and Shopping Frequency Affect Consumers' Brand Choice?

Tatiana Farina*
Boston College

October 17, 2011

Abstract

Consumers behave very differently when they do not have perfect information about all brands available on a shelf. In this paper I extend the benchmark discrete choice model of consumer demand to capture two distinct features of experience-goods markets: prior brand experience and shopping frequency. Although the current literature incorporates habit formation in consumer demand models, it has not considered a more fundamental question: how the first experience with a brand affects the consumer's choice. I estimate the model using data on purchases of packaged orange juice, which comes from a new consumer-level panel provided by a large supermarket chain in Brazil. I find that for this product prior experience of a brand is as important for a consumer's choice as price. Furthermore, own- and cross-price elasticities change significantly when experience and shopping frequency are taken into account. My findings have implications for both firms' strategies and for antitrust analysis related to experience-goods markets.

*Department of Economics, Boston College. Email: tatiana.farina@BC.edu.
I am very thankful to Frank Gollop for his advice and mentorship. I thank the other members of my thesis committee: Arthur Lewbel, Christopher Baum and Hideo Konishi. I'm indebted to Samson Alva, Jim Campbell, Donald Cox, Gautam Gowrisankaran, Martin Goetz, Juan Carlos Gozzi, Luigi Pascali, Marc Rysman and Nathan Schiff for helpful discussions. Finally, I would like to thank seminar participants at the Boston College Dissertation Workshop, University of Sao Paulo, Insper and Universite de Cergy-Pontoise for useful comments.

1 Introduction

Consumers behave very differently when they do not have perfect information about all brands available on a shelf. In this paper I extend the benchmark discrete choice model of consumer demand to capture two distinct features of experience goods markets: prior brand experience and shopping frequency. Although the current literature incorporates habit formation in consumer demand models, it has not considered a more fundamental question: how the first experience with a brand affects the consumer's choice. I address this question by using packaged orange juice as an example of an experience good to estimate a consumer demand model.

Consumers choose among brands of experience goods (e.g., orange juice or cereals) differently than they do brands of cameras or DVD players, because the only way to evaluate the brand is to try it. Experience goods are those for which the consumer needs to pay an inspection cost to learn about how much she likes a brand, and for which the cheapest search mechanism is experience. In other words, she needs to consume a brand to learn about her own utility derived from that brand. Foods and beverages are examples of experience goods (Nelson, 1970).

The main hypothesis of this paper is that ignorance about brands generates an experience cost. Many theoretical and empirical papers, to be discussed below, present models in which utility of a brand can be a function of past consumption. None, however, examines the effect of first-time experience with a brand on consumer demand, the topic of this paper.

I am interested in experience costs, i.e. that part of switching costs that is related to learning. In this paper I therefore extend the benchmark discrete choice model of demand to include prior experience, while also accounting for habit formation. The objective is to estimate the magnitude of experience costs and evaluate how they affect consumers' choices. Furthermore, there is evidence that the number of experiments with brands of a product is closely related to the frequency of purchases of that product. Therefore, I incorporate shopping frequency in the model, and estimate how it interacts with experience cost and brand loyalty.

This paper uses a novel consumer-level data set to estimate the model empirically. The data come from a scanner-data panel of 3,000 consumers provided by a large supermarket chain in Brazil. The panel is composed of two years of purchases recorded in consumers' fidelity card accounts, and include detailed data of each shopping trip, as well as demographic characteristics. This type of data is not easily available to academic researchers, and this particular dataset has not been used in any other work. This is the first time, to my knowledge, that consumer-level data is used to analyze consumer behavior in Brazil. The disadvantage of this dataset is that it has only information about one retail chain.

The results of the model imply that, holding all else constant, firms would have to give their product to consumers for free to compensate for their experience cost. When prior experience is taken into account, own-price elasticities decrease, which suggests that standard models over-estimate demand elasticities when applied to experience goods. Moreover, the own-price elasticity is larger when consumers have not experienced a brand before. Estimated cross-price elasticities also change significantly. Finally, frequent shoppers are more loyal to brands.

The findings of this paper are important for a better understanding of consumer behavior in markets of experience goods. As such the paper contributes to both the industrial organization literature and the marketing literature. Firm strategies are different in the presence of experience and other switching costs. When consumers face experience costs, firms with large market shares could charge higher prices since consumers must incur an experience cost to switch to new brands. Smaller competitors on the other hand might be forced to price more aggressively to stimulate experimentation. Experimentation also makes other strategic variables very important for sales performance, like package sizes and free-samples.

The paper is structured as follows. Section 2 reviews the literature, and Section 3 presents the model. Section 4 describes the data. In Section 5 I present the results, followed by a discussion of possible implications and extensions in Section 6. Section 7 concludes.

2 Literature review

2.1 Search and Experience Problems

The problem of imperfect information for consumer choice has been widely explored in the theoretical industrial organization literature. There is a variety of sources of imperfect information one can imagine. There are also a variety of search mechanisms.

Some papers, such as Diamond (1971), Rothschild (1974) and Rosenfield and Shapiro (1981), stress the problem of asymmetry of information in price. That is, consumers have to pay some search cost in order to learn about the price of the good that they want to buy. In Diamond's model, firms charge a monopoly price in equilibrium even if the good is completely homogeneous as long as there is a positive search cost, even if very small.

Nelson (1970, 1974, 1976) and Wilde (1980) focus on the problem of unknown valuation or tastes of the goods. Now consumers have full information about prices, but products are heterogeneous and they do not know their valuation for each brand. Customers must pay a cost to learn about how much they like each brand. Nelson (1970) argues that he chooses to focus on lack of information about quality because that is more costly to acquire than information about price. In this paper this is the information problem of interest.[1]

The method for acquiring information is one of the key features of models with imperfect information problems about brands. Nelson (1970) makes a very clear distinction between search and experience problems, which I follow in this paper.[2] The consumer's information problem is only to evaluate the utility of each option. He defines search as any type of evaluating options subject to two restrictions: i) the consumer must inspect the option; and ii) inspection must occur prior to the purchase of the brand. The second restriction is what makes search different from experimenting. An experience good is one for which the least costly evaluating mechanism is the purchase itself. Many non-durable goods can be classified as experience goods. Food, for instance, is the typical example of an experience good. If one needs to learn how much he or she likes a certain brand, the easiest and cheapest way to do it is to eat it. Therefore, one needs to buy it first. An example of a non-durable good that should not be classified as an experience good is clothing. One could try a dress before buying it. Trying the dress is a less expensive evaluation mechanism than purchasing the dress. The search cost involves time spent traveling to and from the store as well es experimenting at the store. The experience cost is the utility foregone for the best brand that known the consumer. Moreover Nelson (1970) shows that the number of experiments with brands of a product is positively related to the frequency of purchases of that product. This motivates the inclusion of an interaction term between shopping frequency and experience in the demand model of this paper.

[1] Kohn and Shavell (1974), Wolinsky (1983, 1984), and Wolinsky (1986) model both sources of imperfect information.

[2] Nelson (1970) uses the term search more narrowly than Stigler (1961) and Stigler (1962).

While Nelson (1970, 1974, 1976) focuses on experience problems, he does not explicitly solve the optimization problem of the consumer. Villas-Boas (2006) designs a model in which there is a firm duopoly, and goods are heterogeneous and non-durable[3]. The outcome of the model captures some important and interesting features of experience good markets. First, after experiencing a good fit, consumers find it too costly to experiment further. Also, after a large number of consumers have a positive experience with a firm, that firm charges a higher price to exploit those consumers' positive experience. This result is complementary to the literature of switching costs. By introducing experience into the model, one can endogenize part of switching costs. This idea is present in a few recent papers, such as Villas-Boas (2004, 2006), Klemperer (2002), Dubé, Hitsch, and Rossi (2010) and Aghion, Bolton, Harris, and Jullien (1990).

2.2 Discrete choice models of demand estimation

Seminal work by Lancaster (1971) and McFadden (1974) have laid the ground for the use of discrete choice models in demand estimation for differentiated products. In these models, products are considered as a collection of characteristics, and consumers choose the one that maximizes the utility derived from the product's characteristics. With this approach many brands can be projected onto a space of only a few characteristics. Thus, researchers can get around the curse of dimensionality of traditional systems of demand.

In the traditional approach of demand estimation, there are J brands in the model. That means there will be J equations and each equation will have J parameters to estimate. Each equation specifies the demand for a brand

[3]In each period there is a new generation of consumers that enter the market. Consumers live only for two periods and can only consume one unit of the good. In each period, the firms choose the price such as to maximize the expected value of their profits. The equilibrium concept used in the model is the Markov perfect equilibrium Fudenberg and Tirole (1991).

as a function of its own price, the price of other brands, and other variables. That means at least $J \times J + 1$ parameters to be estimated in the system. Even ten brands can make the model intractable. On the other hand, the number of parameters in a discrete choice model does not grow with the number of alternatives but with the number of characteristics. Own- and cross-price elasticities are functions of the estimated parameters, so are not identified directly from the estimates as they do in traditional demand models. This has enabled researchers to include more brands in the demand model. This is particularly beneficial for analysis of differentiated goods.

In the literature of empirical industrial organization Berry (1994) and Berry, Levinsohn, and Pakes (1995) have shown how these models can be altered to be used with aggregate data on market shares. The BLP framework became the most widely used model for demand estimation in industrial organization when only market-level price and quantity data are available. They also have improved on the previous literature by dealing with price endogeneity[4].

There have been few papers that estimate demand models with individual-level data in the empirical industrial organization literature. Ackerberg (2001, 2003) use consumer-level scanner data to model the impact of persuasive and informative advertisements in consumers choice of a brand of yogurt. Hendel and Nevo (2006) use scanner data to measure the impact of sales and consumer behavior on demand of storable goods. On the other hand, in marketing

[4]In discrete choice models of demand price and product characteristics enter demand equations in a non-linear fashion. This frustrates the use of straightforward instrumental variable methods, which rely on a linearity assumption. Most discrete-choice demand models treat price as exogenous, but this can be a even more serious problem when aggregated data is being used. In demand models with consumer-level data the source of price endogeneity is not the traditional one. The problem is the potential unobserved quality that might be part of the stochastic term and be correlated with prices. However, as long as this unobserved quality does not vary over time, it can be picked up by alternative-specific constants.

Guadagni and Little (1983)'s seminal paper has started a large literature on choice models estimated using scanner data.

Many papers have used past consumption in models of consumer choice. Ackerberg (2001, 2003), Hendel and Nevo (2006), Costantino (2008) and Osborne (2005) are examples in the empirical industrial organization literature. In marketing, models have included variables related to consumption history to incorporate habit formation and brand loyalty like Guadagni and Little (1983), Erdem (1996), Keane (1997), and Shum (2004), among others.

This paper contributes to both the industrial organization and marketing literature by extending the benchmark consumer discrete choice demand model to include features of experience goods markets. Despite the introduction of prior experience and shopping frequency, the model is fairly tractable and straightforward to compare with the benchmark model.

3 Model

In order to model the demand for packaged orange juice, I follow the current literature in using a discrete choice model of consumer demand for differentiated products.

3.1 Benchmark Model of Consumer Demand

First I introduce the benchmark discrete choice model for consumer demand. There are J alternatives available for the consumer.[5] The consumer selects alternative j that gives her the highest indirect utility, U_{ij}. The indirect utility can be decomposed into observable and unobservable parts. For each alterna-

[5]Each alternative in this paper is a module-brand. I define module as a combination of size, container type (bottle or UHT container) and variety (light or regular). In the estimations, $J = 12$.

tive $j \in J$ in the shopping trip t, the indirect utility of individual i can be written as

$$U_{ij} = \xi_j + \alpha P_{ijt} + \beta_k \sum_{k=1}^{K} X_{jkt} + \theta_{jm} \sum_{m=1}^{M} Y_{im} + \epsilon_{ijt} \tag{1}$$

where ξ_j is an alternative-specific constant; P_{ijt} is the price offered to consumer i for the module-brand j at purchase occasion t; X_{jkt} is a vector of alternative-specific characteristics j at occasion t; Y_i is a vector of socio-demographic variables; and ϵ_{ijt} is the random component for consumer i, brand j and time t. The alternative-specific constant captures the average effect on utility of all factors that are not included in the model (e.g. intrinsic quality). [6] It is important to control for that to avoid endogeneity problems in the price variable as quality is expected to be correlated with prices. Social-demographic variables do not vary with j, only with i, thus only differences in θ's can be estimated. That implies that the consumer characteristics are only relevant to the model if they create differences in utility over alternatives. The unobserved part of the utility, ϵ_{ijt}, is assumed to be distributed as type I extreme value. The extreme value distribution leads to the logit formula for the choice probabilities, and the multinomial logit is by far the most widely used discrete choice model. The fact that the choice probabilities have closed form is one of the reasons for its popularity.

The choice probability is $Prob_i(y = j) = Prob(U_{ij} > U_{il})$ for $\veebar j \neq l$. However, the absolute level of utility is irrelevant for the decision maker and for the

[6] In the empirical IO literature, it is common to see a ξ_{jt} term that varies with time in models of consumer demand, as first introduced by Berry (1994). Advertisement that varies over time is the main example used in the literature for this term. Packaged orange juice was intentionally selected as a product that has low intensity advertising to avoid addressing this issue. Since the model is estimated with consumer-level data, the BLP method to correct for the endogeneity caused by the ξ_{jt} term cannot be used.

econometrician. Only differences in utility matter. So the choice probability can be rewritten as $Prob(U_{ij} - U_{il} > 0)$ for $\veebar j \neq l$. As the error term is type I extreme value, then these choice probabilities take the well known logit form:

$$Prob_{it}(y = j) = \frac{exp(\xi_j + \alpha P_{ijt} + \beta_k \sum_{k=1}^{K} X_{jkt} + \theta_{jm} \sum_{m=1}^{M} Y_{im})}{\sum_{l=1}^{J} exp(\xi_l + \alpha P_{ilt} + \beta_k \sum_{k=1}^{K} X_{lkt} + \theta_{lm} \sum_{m=1}^{M} Y_{im})} \tag{2}$$

It can be noticed from Equation 1 that the consumer knows everything there is to know about each module-brand offered in the market. The only unknown part of her utility of a brand is specific to the purchase occasion t.

3.2 Extended Model of Consumer Demand for Experience Goods

I introduce prior experience to the model in order to capture the experience cost incurred due to lack of information about brands in the market. Moreover I include shopping frequency in the model because of its hypothetical impact on the willingness to try new brands. I also control for brand loyalty to ensure that my experience variable does not capture other types of habit formation other than learning. This variable controls for state dependence in the consumer's utility.

Prior experience enters the consumer's utility function as a dummy variable which is equal to one if she has never purchased the brand and is equal to zero if she has purchased that brand before. The variable is coded in this way so that the its estimated coefficient reflects the experience cost of trying a new brand. I follow other papers from the marketing literature for the definition of the brand loyalty variable (BL_{ijt}).

$$Ex_{ijt} = \begin{cases} 1 & \text{if the brand of } j \text{ has not been purchased by consumer } i \text{ yet,} \\ 0 & \text{otherwise.} \end{cases}$$

I estimate the models with three alternative specifications of that variable as a robustness check: (i) the simplest version is a dummy that captures if brand b has been purchased in the past shopping trip; (ii) brand loyalty is equal to the number of times that the consumer has purchased the brand in the past observed shopping trips; and (iii) the brand loyalty variable as specified in Guadagni and Little (1983) as an exponentially smoothed weighted average of lagged purchase indicators, $GL_{ijt} = \delta GLijt - 1 + (1 - \delta)d_{ijt-1}$, where d_{ijt-l1} is an indicator function equal to one if the consumer purchased brand j at $t - 1$, and zero otherwise.

Some alternatives belong to the same brand and therefore are simply different alternatives because they are offered in different packages or types. For those cases the experience and brand loyalty variables are the same. If a consumer buys Xando 300 ml at t, the experience dummy will be switched off for Xando 1L as well from $t + 1$ on. The same approach is taken for the brand loyalty variable.

Shopping frequency is modeled as the number of days since the last shopping trip. It is is denoted by Du_{it} for duration. There is no subscript t on it because it does not vary with alternative j. This is a consumer's shopping trip specific variable. Thus, only the difference in its coefficients can be identified. Two types of duration are calculated: (i) duration between every visit that the consumer makes to the supermarket (Du_{it}), and (ii) duration between trips where orange juice was purchased. I denote the latter one by $DuOJ_{it}$. The model is estimated with both variables, but I prefer to focus on the results with the first one, as it captures better consumer shopping behavior. The inclusion of $DuOJ_{it}$ in the model can also cause concerns about endogeneity. It is easier to argue that Du_{it} is exogenous to the choice of packaged orange juice.

However what I am really interested on is how heterogeneity in shopping behavior affects experience cost and loyalty to brands. Therefore experience and brand loyalty are interacted with median duration between shopping trips for each consumer. That is, consumers are allowed to have different experience cost and loyalty to brands depending on their ẅshopping types:. Previous research shows evidence on the relationship between experience and shopping frequency but it is never formally tested.

Therefore, consumer i's utility for alternative j at shopping trip t is given by

$$U_{ij} = \xi_j + \alpha P_{ijt} + \beta_k \sum_{k=1}^{K} X_{jkt} + \theta_{jm} \sum_{m=1}^{M} Y_{im} + \delta Ex_{ijt} + \lambda BL_{ijt} + \kappa Du_{it} + \varphi Ex_{ijt} \times Du_{it} + \epsilon_{ijt} \tag{3}$$

f Define the deterministic part of the the utility as V_{ijt}[7] :

$$V_{ijt} = \xi_j + \alpha P_{ijt} + \beta_k \sum_{k=1}^{K} X_{jkt} + \theta_{jm} \sum_{m=1}^{M} Y_{im} + \delta Ex_{ijt} + \lambda BL_{ijt} + \kappa_j Du_i + \varphi Ex_{ijt} \times Du_{it} \tag{4}$$

The new logit probabilities are given by Equation 5:

$$Prob_{it}(y = j) = \frac{exp(V_{ijt})}{\sum_{l=1}^{J} exp(V_{ilt})} \tag{5}$$

[7]In the literature V_{ijt} is called representative utility. It is a function of observable variables (the regressors) and parameters that are unknown to the econometrician and therefore must be estimated.

The novelty of this paper is the introduction of experience (Ex_{ijt}) and shopping frequency (Du_{it} and $Ex_{ijt} \times Du_{it}$) in the discrete choice demand model. By doing so I incorporate stylized features of the experience market into the model, thereby eliminating the estimation problems related to the assumption of perfect information for this type of product. Although learning here is myopic, the model is still fairly easy to estimate. It also allows the researcher to make direct comparisons with the results of the benchmark model.

4 Data

4.1 Description

I use consumer-level panel data on grocery purchases provided by a large supermarket chain in Brazil. The dataset consists of purchase information of fidelity card members who scan their card when they pay. This information is paired with the household's demographic information provided when they signed up with the fidelity program.

The dataset contains information on 3,000 consumers from January 2005 to December 2006, inclusive. Every product that the consumer buys is one observation, so the data come in their most disaggregated format. The panel covers food, beverages, and many non-food items commonly found in supermarkets.

The socio-demographic variables provided in the panel are income, gender, age, number of people in a household[8], occupation, schooling and residential

[8] The variable number of people in the household unfortunately does not seem to be reliable, because of the excess of zeros in the data. I believe that the reason why this variable is problematic is the way that they pose the question in the form. They ask how many dependents are there in the household, and that might leave room for misunderstanding. If a housewife is filling the form, she might not consider her husband as her dependent even if he lives in the house. Indeed 60% of the housewives in the data answered that there were 0

zip code. Since consumers fill out the form only once, characteristics are constant within a household.

The purchase specific variables are name of each product, brand, package size, quantity, price, discount value, date of purchase and code of the store.

Women represent 60% of the sample. The median and the mean of age is 50 years. The average monthly income in the sample is R$ 3,145 and the median is R$ 2,400.[9] On average, there are 106 recorded shopping trips per consumer in these 26 months. Consumers spend R$50 on average per trip and buy 14 different products.

For this paper, I focus on purchases of packaged orange juice. I use only consumers that have bought this product at least ten times. Furthermore, I focus only on consumers of the state of São Paulo to minimize issues of product availability[10]. Using these criteria I end up with 221 consumers and 6092 shopping trips during which packaged orange juice was purchased. That represents roughly 20% of the total number of visits to the grocery store recorded by these consumers.

The total number of purchases by a consumer of a particular module-brand combination is set to be the number of observed shopping trips during which

dependents in the households; this seems rather odd. Their objective was to get information on the number of people in the household, but they just seemed to ask the wrong question. Therefore I will not use this variable in the analysis.

[9] Income per capita in the state of São Paulo in 2005 was R$ 1,247 per month. The average exchange rate in 2005 was 2.44 R$/US$, and 2.17 R$/US$ in 2006 (source: IPEA).

[10] Availability is a common problem for empirical papers of consumer choice. In most cases, there are no data on what alternatives were available to the decision maker at the time of purchase. In general the assumption is that every alternative is always available. That is what I assume in this paper as well. It does not generate any serious problems for estimations as long as products are randomly unavailable: that is, if availability is not correlated with any of the observables. Restricting the region helps us avoid this issue, as some brands might not be offered everywhere in the country.

the household purchased at least one item of that module-brand. A trip counts as a single purchase regardless of how many units were bought of the same module-brand[11]. A single brand may be sold in multiple modules. I define module as a combination of size, container type (bottle or UHT container) and variety (light or regular). For instance, Maguary, one of the brands, is sold in three different formats: regular juice in a 1 liter UHT package, light juice in a 1 liter UHT package and regular juice in a 200 milliliter UHT container. These three types have different characteristics, including price. Therefore I chose to consider them as different alternatives for the consumer. This approach is very similar to related papers in the literature.

4.2 Additional Sample Characteristics

There are 20 different brands of packaged orange juice in my sample and 28 module-brand combinations. I only use twelve module-brand combinations in the model, which cover 7 brands. This accounts for 83% of the market share of those alternatives in the dataset. The market share of the leading brands are 33%, 22% and 9%, which is very close to the actual market share of those brands in São Paulo in that period. The Herfindahl-Hirschman Index (HHI) for the market is 1759. For the sake of comparison, an individual concentration index is calculated using the same formula as the market HHI, but instead of calculating the level of concentration for a market, the index is calculated for each individual[12]. The average individual consumer's HHI is 5369.[13]This

[11]If two different module-brands are observed in the same shopping trip, then that shopping trip is duplicated. It has exactly the same characteristics as the first one but it will refer to the other alternative. In other words, a shopping trip with two module-brands becomes two observations in my data. The same procedure is followed if more than two module-brands are purchased in a single shopping trip.

[12]To calculate the "consumer's HHI", I calculate the share of each brand in the consumers expenditure with packaged orange juice in the sample period. The index is the sum of squares of the market shares. In the limit, if each consumer only purchased one brand then the individual HHI would be one.

[13]From this point, descriptive statistics or comments always refer to the restricted sample of 221 packaged orange juice consumers.

difference between market concentration and consumer choice concentration reflects brand heterogeneity in this market[14]. The expenditure on orange juice is, as expected, much more concentrated than the market itself. The median consumer buys four brands over the whole sample considering all the twenty brands. In two years, 10% of consumers only buy two. There is no individual that is observed ever spanning all brands over the sample period. Figure 1 in the appendix shows the distribution of consumers for number of brands and module-brands.

The median orange juice consumer shops on average every five days and buys orange juice every 15 days. Figures 3 and 4 in the appendix show the distributions of days between shopping trips. The median consumer spends R$ 67 in each visit, but spends R$ 82.71 in visits in which she purchases orange juice.

Figure 2 shows that more frequent consumers have tried more brands over the sample period. For illustrative purposes, I divided consumers into two groups according to their median duration (i.e., number of days) between orange juice shopping trips. Individuals that shop more than once a month were classified as frequent consumers, and the others as infrequent consumers. The median frequent consumer tries three new brands over the sample, and the median infrequent consumer tries two.

4.3 Data for the Discrete Choice Model

In order to estimate a multinomial model the data must be arranged in long format. In other words, for every purchase occasion the dataset must have J rows of data (one for each alternative, $j = 1, \ldots, J$). A dataset with N

[14]In a market with homogenous goods and no transportation cost, these two indexes should be the same. Differences in market share would only be explained by differences in firms' cost structure.

observations, therefore expands to $N \times J$ observations. [15]

4.3.1 Multiple module-brand purchases in one shopping trip

Multiple module-brands bought in one single shopping trip by a consumer can be problematic for a discrete choice model, because only one alternative can be chosen at a time. Some researchers define criteria to select which products to consider, and drop the secondary ones. However for this paper it is important that I keep track of all choices of orange juice module-brands because consumers' purchase histories are of central interest in my model. In order to avoid loss of information, I use the following method: if two different module-brands are observed in the same shopping trip, then I duplicate that shopping trip. The second trip has exactly the same characteristics as the first one (total expenditure in the shopping trip and number of days since last trip), but it refers to the other alternative. In other words, a shopping trip with two module-brands becomes two observations in my data. One observation has all the characteristics of one brand-module (price, product characteristics, prior experience and habit with the product) and the characteristics of the shopping trip . The characteristics of the brand-module change for the second observation but the shopping trip characteristics are identical. In long format, there will be $2 \times N \times J$ for each shopping trip with purchases of two module-brands of orange juice. The same procedure is followed if more than two module-brands are purchased in a single shopping trip.

4.3.2 Prices

I model the choice of module-brands but not the quantity purchased in a shopping trip. However, the packages come in different sizes. Therefore I calculate the 1L-equivalent price in order to make it comparable across the different alternatives. The average unit price of the 200ml packages is lower than the

[15] Notice that I end up with $N \times J$ observations because of the assumption that individuals always have the option of buying any of the J alternatives

average unit price of the 1L packages, but that does not mean that they are cheaper: see Table 1. It is more reasonable to assume that consumers choose according to weight-equivalent prices instead. Thus the latter is the variable that is used in the analysis.

Only prices paid by consumers are observable in the dataset: a common feature of scanner data. However the model requires data on all alternatives, including those not chosen, to empirically identify the parameters of interest. Therefore I need to impute values for the missing prices. The steps for imputation are the following: i) generate a price imputation dataset with a price for each module-brand for each day during thr sample period; ii) calculate price for alternative j at day t as a weighted average of the prices that consumers paid for it on that day;[16] iii) carry forward prices for the remaining missing values in the price imputation dataset;[17] iv) substitute all the missing prices in the main dataset by the corresponding values from the price imputation dataset. I separate the observations of the city of São Paulo and the other cities of the state for imputing prices. I follow the same steps listed above in each subset.

Value discounts are observed in the dataset. I use only non-discounted prices in the model to avoid endogeneity problems. Erdem, Keane, and Sun (1998) discuss this problem. Coupons are only observed for the chosen alternatives. One cannot assume that all consumers have access to the same coupons as is done with prices. Moreover, consumers might have coupons that they do not use, which the researcher does not observe. But if a coupon value is included as a variable in the model, that value of the unused coupon should have been there. If I simply set coupons to zero for the unchosen alternatives, then

[16] in some cases the prices are all the same, and then the average is just equal to that value.

[17] There is a missing value for a module-brand j at day t if no consumer bought it that day.

the estimated marginal effect of that variable will be over-estimated. Another way to see the endogeneity problem is that if the value of a coupon in positive in the dataset, then it perfectly predicts the consumer's choice. The same problem would appear if I let the price variable be $P^d_{ijt} = P_{ijt} - Coupon_{ijt}$. Although the endogeneity problem is now more subtle, it is still present. The price sensitivity of consumers would be over-estimated. Thus due to these issues I do not use coupon values. The fact that consumers do have coupons is not a problem as long as coupons are independently distributed[18].

4.3.3 Alternative characteristics

I include four characteristics in the consumer's utility in addition to price: i) if the module-brand has added sugar, ii) if it is light or regular, iii) if it is nectar or juice, and iv) size.

Leco and Maguary are the only brands that offer light versions of their juices, and offered in a 1L container. Both also offer the same container size of the regular juice. Light is an indicator variable that is set to 1 if the juice is labelled as light, and 0 otherwise.

Out of the 27 module-brands in the supermarket, only 7 of them are offered in smaller sizes (200 ml or 300 ml), and only 3 brands offer only small containers. Most of packaged orange juices in supermarkets come in 1L containers. Small containers represent 14% of consumers' choices in the sample used for estimation. Size enters the model as a continuous variable.

The distinction between nectar and juice in Brazil is very similar to the one in other countries. Most countries define a standard purity for a beverage to be considered a "fruit juice". This name is generally reserved for beverages

[18] If individuals acquire coupons because of previous purchase of a module-brand, then the effect of those coupons will be absorbed by the brand loyalty variable

that are 100% fruit juice. In Brazil any beverage with 30% concentration of orange juice can be called nectar of orange.[19]. In the United States a nectar must contain between 35% to 50% depending on the fruit.[20] Thus a nectar is more diluted than a juice and may contain artificial sweeteners. Brazilian law limits the addition of non-artificial sugar to only 10% of the weight of juices. Only a few packaged orange juices are classified as nectars in Brazil (3 brands our of 20). In the set of alternatives of my model the only module brand that has added sugar is the Kapo 200 ml. Table 1 shows the average prices of the module-brands that are used for estimation. As expected the average price for nectars is slightly lower than the average price of juices. In the model nectar enters as an indicator variable that is equal to 1 if the module-brand is labeled as nectar, and is 0 otherwise.

These characteristics are displayed on the product's container. The orange juice containers also have information on calories, carbohydrates and amount of vitamin C. But they do not appear to affect significantly the consumers' choice between module-brands once I condition for the four characteristics listed above.

5 Results

Tables 2 through 14 show the main results. The models are estimated using either alternative fixed effects or a set of alternative specific characteristics. They cannot be included simultaneously because the module-brand characteristics described above are constant over time, unlike price. However, the price elasticities and the marginal effects of the main variables of interest do not change significantly between these specifications.

[19] In Brazil Law no. 8.918 of 1994 defines the standards, classification, inspection, registration and production of all beverages.

[20] In the United States the FDA and USDA define the standards for beverages.

The first columns of Tables 2 and 3 show the results of a standard discrete choice demand model. For the former table, alternative-specific constants and price are included. For the latter table, price and alternative characteristics are included. Both models imply very similar price elasticities. The estimated price coefficient ($\hat{\alpha}_{ijt}$) is equal to -0.279 for the first model and -0.305 for the second model. These values for $\hat{\alpha}_{ijt}$ generate very similar marginal effects for prices and price elasticities (see Tables 9 and 12).

The marginal effect tells us by how much the probability of buying a certain module-brand of orange juice changes when the price of module-brand j changes, that is, the change in the probability that decision maker i chooses alternative j given a change in P_{ijt}. The derivative of $Prob_{it}(y = j)$ with respect to P_{ilt} is calculated as $\frac{\delta V_{ijt}}{\delta P_{ilt}} Prob_{it}(y = j) Prob_{it}(y = l)$. If $j = l$, then $\frac{\delta Prob_{it}(y=j)}{\delta P_{ilt}} = \frac{\delta V_{ijt}}{\delta P_{ilt}} Prob_{it}(y = j)(1 - Prob_{it}(y = j))$.[21] When V_{ijt} is linear in prices, then the marginal effect can be written as $P_{ijt} = -\alpha Prob_{it}(y = j) Prob_{it}(y = l)$ [22]. The own-price elasticities are calculated as $E_{jj} = \frac{\delta V_{ijt}}{\delta P_{ijt}} P_{ijt}(1 - Prob_{it}(y = j))$ [23] [24]. The cross-price elasticities are $E_{jl} = \frac{\delta V_{ijt}}{\delta P_{ilt}} P_{ilt} Prob_{it}(y = l)$.

Tables 9 and 12 report the marginal effects and price elasticities for the benchmark demand model. Since the estimated price coefficient ($\hat{\alpha}_{ijt}$) is slightly

[21] Since the marginal effect depends on the choice probabilities, that means that the y also depend on the values of the regressors. Just as in any other non-linear model, the magnitude of the marginal effects always depends upon where you evaluate them at. The derivative is largest at $Prob_{it}(y = j) = 0.5$. As a result of the sigmoid-shaped probability. This implies that the larger the degree of uncertainty regarding a choice, the larger the impact of an observable variable on the change of the choice probability. In tables 9 and 12 those derivatives are evaluated at the mean values of the regressors.

[22] The marginal effects of other continuous variables are computed in exactly the same way.

[23] If V_{ijt} is linear in prices, then $\frac{\delta V_{ijt}}{\delta P_{ijt}} = -\alpha$

[24] Notice that the cross-elasticity is the same for all j. A decrease in price of an alternative reduces the probabilites for all the other alternatives by the same percentage. This is a manifestation of the Independence of Irrelevant Alternatives (IIA) property. Proportional substitution is the main drawback of multinomial logit models.

smaller when the model is estimated with alternative specific coefficients than with alternative characteristics, the own-price elasticities implied by this model are also smaller. Module-brands offered in small containers have own-price elasticites larger than 1, and the 1L module-brands all have estimated elasticities between 0.6 and 1. Therefore consumers are less sensitive to changes in prices of larger containers. The cross-price elasticites of the small module-brands are lower than the ones of the 1L products of the same brand. This implies that a change in prices of orange juices offered in small containers steals less market share from competitors than if the same proportional change in prices of the larger containers is observed. In the model with alternative dummies, the light juices have the lowest price cross-elasticities perhaps due to consumers perceiving these products as more distant from the others. In the model with alternative characteristics cross-price elasticities are more similar to each other.

Columns 2 to 6 of Tables 2 and 3 show results of the demand model with experience. In most discrete choice models, the ratio of coefficients has an economic meaning: the ratio of alternative characteristics' coefficients reflects the marginal rate of substitution between these two characteristics. In particular, all else equal, the ratio between the experience coefficient and the price coefficient shows the consumer's willingness to pay for knowledge about a brand.[25] This reflects the consumer's marginal rate of substitution between these variables. In the third column of table 3 the estimated coefficients for price and experience are -0.210 and -1.909, respectively[26]. The ratio of the two implies

[25] Ignoring the interaction between experience and shopping frequency (i.e. $\varphi = 0$),take the total derivative of Equation 3, holding constant all variables other than price and experience and set it to zero to hold utility constant: $dU = \alpha P + \delta Ex$. The marginal rate of substitution can be calculated by solving for the change in price that would maintain utility constant for switching on the experience variable. $\frac{dP}{dEx} = -\frac{\delta}{\alpha}$. If the interaction between experience and shopping frequency is also included in the model (i.e. φ is allowed to be different than zero), then $\frac{dP}{dEx} = -\frac{\delta + \varphi \times Du_i}{\alpha}$.

[26] Column 3 of table 3 displays the results of the model estimated with experience and brand loyalty (with exponential discounting).

that, all else constant, consumers are willing to pay roughly R$ 10 per liter more for a brand of orange juice that they have already purchased. Given that the most expensive module-brand sells on average for R$ 5.69 per liter, this result implies that firms would have to give away their products for free in order to compensate for the consumer's ignorance about that brand. The marginal rate of substitution between the two implies that firms have to give their products for free in order to compensate for the experience cost. This results suggests free samples can be very important.

The marginal effect of experience is large compared to the marginal effect of price regardless of the brand loyalty variable used. Tables 4 and 5 display the results of the model estimated with the three specifications of the loyalty variable introduced in Section 3.2. This is an important result because it separately identifies two different parts of switching costs: i) experience cost and ii) habit formation or brand loyalty. In the case of brands of packaged orange juice the experience cost seems to be significant. Separating the two is important since each type of switching cost might imply a different strategy for the firm.

In the fifth column of Tables 2 and 3, an interaction between price and experience is added. The own-price elasticities of a brand are roughly twice as large for unexperienced consumers than for experienced consumers. The interaction between price and experience is no longer significant in column 6 when the interaction between loyalty and price is included. Consumers that are loyal to a brand have lower price elasticity to that brand.

This is in line with results from the switching cost literature (e.g. Villas-Boas (2006)) that say that non-loyal consumers have more price sensitive demand than loyal ones.[27] A firm with small market share, that faces a large

[27] In Villas-Boas (2004) and Villas-Boas (2006), consumers are considered loyal to a brand

number of uninformed consumers, is likely to price more aggressively. A market leader, on the other hand, has many informed consumers, so its optimal strategy might be to charge high prices to harvest its loyal customers.

Columns 4 to 6 of Tables 2 and 3 show the results with shopping frequency in the model. The aim of introducing shopping frequency is to investigate the impact of different shopping patterns on experimentation and habit formation. The variable of interest is the interaction between frequency and experience, and frequency and loyalty, not frequency per se. Shopping frequency is expected to have an impact on the decision of purchasing orange juice in a certain shopping trip, but not to affect the choice between the different brand-modules of orange juice. As explained previously the model can only identify the differential impact of a case-specific variable (i.e. a variable that only varies with i but not with j) on the choice of each alternative. Shopping frequency is modeled as the median for each consumer of the number of days between each shopping trip. The marginal effect of experience becomes larger as median duration increases. The results imply that high frequency consumers have higher switching costs than those that go less frequently to the supermarket. This is consistent with papers that show that there is addiction in brand choice. One could interpret these findings as implying that if a consumer stays out of the market long enough his addiction diminishes. Moreover, for packaged orange juice, shopping frequency affects the switching cost related to the loyalty variable but not the first-time experience cost.

5.1 Robustness Checks

Tables 4 and 5 show the results of the model with different specification for the brand loyalty variable. Likelihood ratio tests show that, when that variable is

if they have chosen that brand in the last period. The equilibrium concept used in the paper is Markov Perfect, therefore only the decision of the last period matters. Moreover his consumers only live for two periods, so they can only be considered loyal consumers in the second period. The objective of this strategy is to make the problem more tractable.

set to be equal to the number of times a brand has been purchased, the model has the worse fit. The best fit is achieved with the Guadagni and Little (1983) specification. That is also the specification that generates the largest marginal effect for habit formation.

Nevertheless, estimating the model with any of the three specification of the habit formation variable shows that consumers of packaged orange juice are loyal to brands, as the estimated marginal effect for those variables are positive and significant. The marginal effect of experience is also large and significant no matter which specification of brand loyalty is used.

Because consumers are not observed since the first day of their lives, the model is estimated with different sample sizes. For Tables 6, 7 and 8, the experience and loyalty variable are constructed using the whole sample, but different sample sizes are used for estimation. The model is estimated dropping the first 5, 10 and 15 shopping trips for each consumer. The results are robust to the modifications.

6 Implications and Extensions

In this section I briefly discuss implications for firm strategies of modeling demand of experience goods. Then I examine implications for firm strategies. I also present possible generalizations to my model.

The results of my model show that in the case of packaged orange juice, firms would have to give away their product for free to compensate uninformed consumers for their experience costs. This suggests that free samples are useful for these types of goods.

Furthermore, container sizes can be used strategically by firms. The estimated marginal effects of experience for each module-brand imply lower experience costs for smaller modules (200 and 300 ml containers). This result is intuitive because if one is going to try a new orange juice, she might rather buy the smallest possible package; after all, there is a chance that she might not like it. Thus entrants and fringe competitors should also offer their products in small packages to reduce the cost of experimentation. Some of the smaller competitors do offer smaller sizes, but not all of them.

Advertising is also different for experience goods. The decision of marketers about what to advertise about a product depends on what kind of product it is. If the product is an experience good, then it is difficult to inform the consumer about his own valuation of the brand, as it is hard to describe taste. This type of advertisement is different from advertisements for electronic products. Computers, for instance, are not experience goods. Computer advertisements, in general, describe their characteristics (amount of memory, size of hard drive, type of processor, screen size, etc). Information about those characteristics is normally enough for the individual to rank different computer types and brands, and then make a decision about what to buy. Firms have means of transmitting that information. However, for something like soda or orange juice, advertisement can give a little bit of information (e.g. sugar content, calories, etc), but cannot really say anything about taste. Thus what one observes are advertisements that try to transmit a nice feeling to the consumer in order to try to attract attention to that product.[28]

[28]There is a large literature that discusses distinctions between different types of advertisements, for example Stigler (1961), Butters (1977), and Grossman and Shapiro (1984). Ackerberg (2001) and Ackerberg (2003) empirically distinguish the effect of influential advertisement and of informational advertisement for a brand of yogurt. Ackerberg (2001) mentions a television advertisement for "Molson Ice" beer that portrayed twenty-somethings dressed in hip clothes in a bar drinking the beer. Clearly such advertisements should stimulate demand for the product, otherwise Molson would not pay for them. However it is also clear that they do not provide any information about the characteristics of the good.

One possible extension of the model presented in this paper is to include outside options. I do not model a consumer's decision to buy something else other than orange juice. For instance, another fruit juice that might be a substitute for orange juice. One problem of not modeling this is that if all brands increase their prices by say 20%, nothing happens to the probabilities of consumers buying each brand. This result is due to the fact that, in multinomial logit models, only relative changes in variables matter. However, price for the outside option could be normalized to 0, and therefore I would be able to introduce changes in the relative prices, even if we observe an increase of 20% in prices for all module-brands of orange juice[29] In that situation, the outside good would steal market share from every brand in the market. This prediction sounds more realistic. The problem is that it is not clear what the outside option should be. I could define it as every other purchase of fruit juice of the consumer or I could define it as every visit that the consumer does to the grocery store and does not buy orange juice, among other options. It is easier to define the outside option when the researcher is working with aggregate data instead of consumer-level data. In that case he needs only to assume something about the consumption pattern of the average consumer for that product category. Nevo (2000) presents an example for the industry of ready-to-eat cereals.

This problem can be serious if the econometrician is trying to use the model for the Test of the Hypothetical Monopolist that appears in the Merger Guidelines of the U.S. Department of Justice and the Federal Trade Commission. This test is used as a reference for the definition of relevant market in antitrust cases. It works as follows. Suppose all products in the relevant market are produced by a hypothetical monopolist. Assume that it raises the prices of all products by a small but significant amount (i.e. by the same percentage). If consumers do not substitute to other products, then the market is defined.

[29]The outside option is included in a multinomial model as an additional alternative. One can normalize the value of all product characteristics to zero. It is common to allow for the utility of the outside good to vary only with socio-demographic characteristics.

However, if one uses the multinomial logit without an outside option for this exercise, then it will be concluded that no other product needs to be included in the relevant market. The introduction of the outside option solves this caveat.

Another extension would be to add random coefficients into the model. This would allow for a more realistic pattern of substitution among brands, therefore addressing the problem generated by the problem of independence of irrelevant alternatives. When this property holds, a change in price of an alternative generates a proportional change in the probability of choosing all the other alternatives. It would be more realistic to predict that consumers will switch mostly to the alternative that is most similar to the one that changed its price.

7 Conclusions

In this paper I argue that introducing prior experience and shopping frequency is important to model consumers' demand for experience goods. By estimating the extended model for packaged orange juice using consumer-level scanner data, I show that experience costs are significant for this type of product, even when controlling for brand loyalty.

The estimated marginal effects of price and experience imply that firms would need to give away their products for free to compensate uninformed consumers for their experience cost. Moreover, the results show that frequent shoppers are on average more loyal to brands. Furthermore, implied price elasticities change significantly when experience, brand loyalty and frequency are included. Without taking prior experience and loyalty into account, consumers seem to be more willing to substitute among brands. This means that the modifications of the model should be taken into account if one estimates demand for experience goods.

The paper contributes to a large body of literature of demand estimation of differentiated goods, both in industrial organization and in marketing, by introducing the main features of experience goods markets in a tractable way. The results offer important insights for academic research, policy research and for the private sector.

References

Ackerberg, D. A. (2001): "Empirically Distinguishing Informative and Prestige Effects of Advertising," *RAND Journal of Economics*, 32(2), 316–33.

Ackerberg, D. A. (2003): "Advertising, learning, and consumer choice in experience good markets: an empirical examination," *International Economic Review*, 44(3), 1007–1040.

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1990): "Optimal Learning By Experimentation," DELTA Working Papers 90-10, DELTA Ecole Normale Superiore.

Berry, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *The Rand Journal of Economics*, 25(2), 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Butters, G. R. (1977): "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies*, 44(3), 465–91.

Costantino, C. (2008): "Gone In Thirteen Seconds: Advertising and Search in the Supermarket," Ph.D. thesis, University of Maryland.

Diamond, P. A. (1971): "A model of price adjustment," *Journal of Economic Theory*, 3(2), 156–168.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010): "State dependence and alternative explanations for consumer inertia," *RAND Journal of Economics*, 41(3), 417–445.

Erdem, T. (1996): "A Dynamic Analysis of Market Structure Based on Panel Data," *Marketing Science*, 15(4), 359–378.

Erdem, T., M. P. Keane, and B. Sun (1998): "Missing price and coupon availability data in scanner panels: Correcting for the self-selection bias in choice model parameters," *Journal of Econometrics*, 89(1-2), 177–196.

Fudenberg, D., and J. Tirole (1991): *Game Theory*. MIT Press.

Grossman, G. M., and C. Shapiro (1984): "Informative Advertising with Differentiated Products," *Review of Economic Studies*, 51(1), 63–81.

Guadagni, P. M., and J. D. C. Little (1983): "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2, 203–238.

Hendel, I., and A. Nevo (2006): "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 74(6), 1637–1673.

Keane, M. P. (1997): "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *Journal of Business & Economic Statistics*, 15(3), 310–27.

Klemperer, P. (2002): "Markets with Consumer Switching Costs," *The Quarterly Journal of Economics*, 102(2), 375–94.

Kohn, M. G., and S. Shavell (1974): "The Theory of Search," *Journal of Economic Theory*, pp. 93–123.

Lancaster, K. (1971): "Consumer Demand: A New Approach," *Review of Economic Studies*, 74, 132–157.

McFadden, D. (1974): *Conditional Logit Analysis of Qualitative Choice Behavior*. New York: Academic Press.

Nelson, P. (1970): "Information and Consumer Behavior," *Journal of Political Economy*, 78(2), 311–29.

Nelson, P. (1974): "Advertising as Information," *Journal of Political Economy*, 82(4), 729–54.

——— (1976): "Political Informantion," *Journal of Law and Economics*, p. 315.

Nevo, A. (2000): "A Practitioner's Guide to Estimation fo Random-Coefficients Logit Models of Demand"," *Journal of Economics & Management Strategy*, 9(4), 513–548.

Osborne, M. (2005): "Consumer Learning, Habit Formation, and Heterogeity: A Structural Examination," .

Rosenfield, D. B., and R. D. Shapiro (1981): "Optimal Adaptive Price Search," *Journal of Economic Theory*, 25, 1–20.

Rothschild, M. (1974): "Searching for the lowest price when the distribution of prices is unknown," *Journal of Political Economy*, 82, 689–712.

Shum, M. (2004): "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market," *Journal of Economics & Management Strategy*, 13(2), 241–272.

Stigler, G. J. (1961): "The Economics of Information," *The Journal of Political Economics*, 69, 212–225.

——— (1962): "Information in the Labor Market," *Journal of Political Economy*, 70, 94–105.

Villas-Boas, J. M. (2004): "Consumer Learning, Brand Loyalty, and Competition," *Marketing Science*, 23, 134–145.

——— (2006): "Dynamic Competition with Experience Goods," *Journal of Economics & Management Strategy*, 15(1), 37–66.

Wilde, L. L. (1980): "The Economics of Consumer Information Acquisition," *Journal of Business*, 53(3), S143–58.

Wolinsky, A. (1983): “Retail Trade Concentration Due to Consumers’ Imperfect Information,” *Bell Journal of Economics*, 14(1), 275–282.

——— (1984): “Product Differentiation with Imperfect Information,” *The Review of Economic Studies*, 51(1), 53–61.

——— (1986): “True Monopolistic Competition as a Result of Imperfect Information,” *The Quarterly Journal of Economics*, 101(3), 493–512.

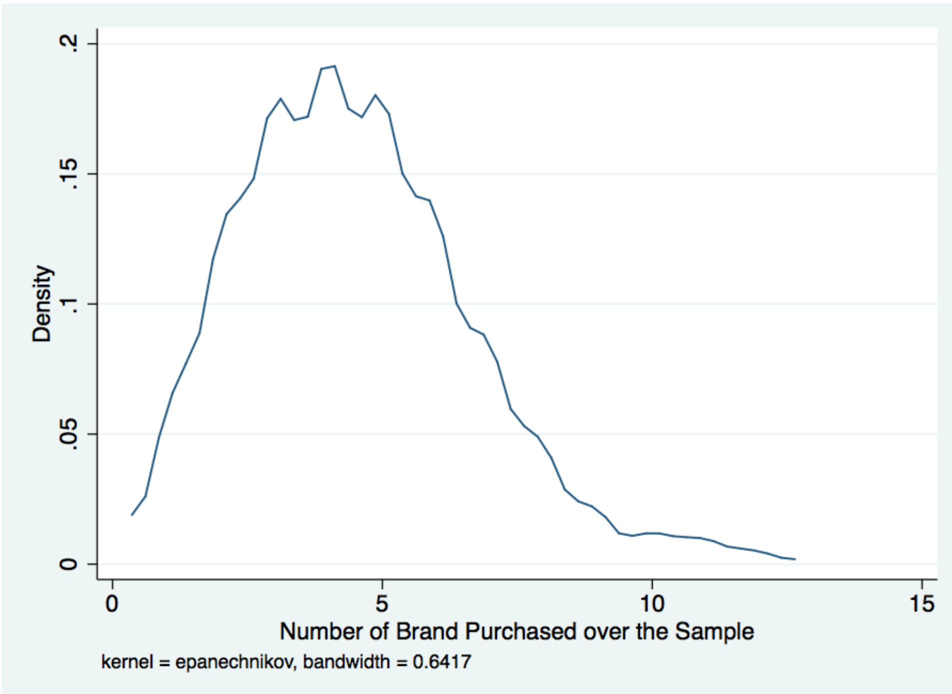

Figure 1: Number of brands purchased over the sample

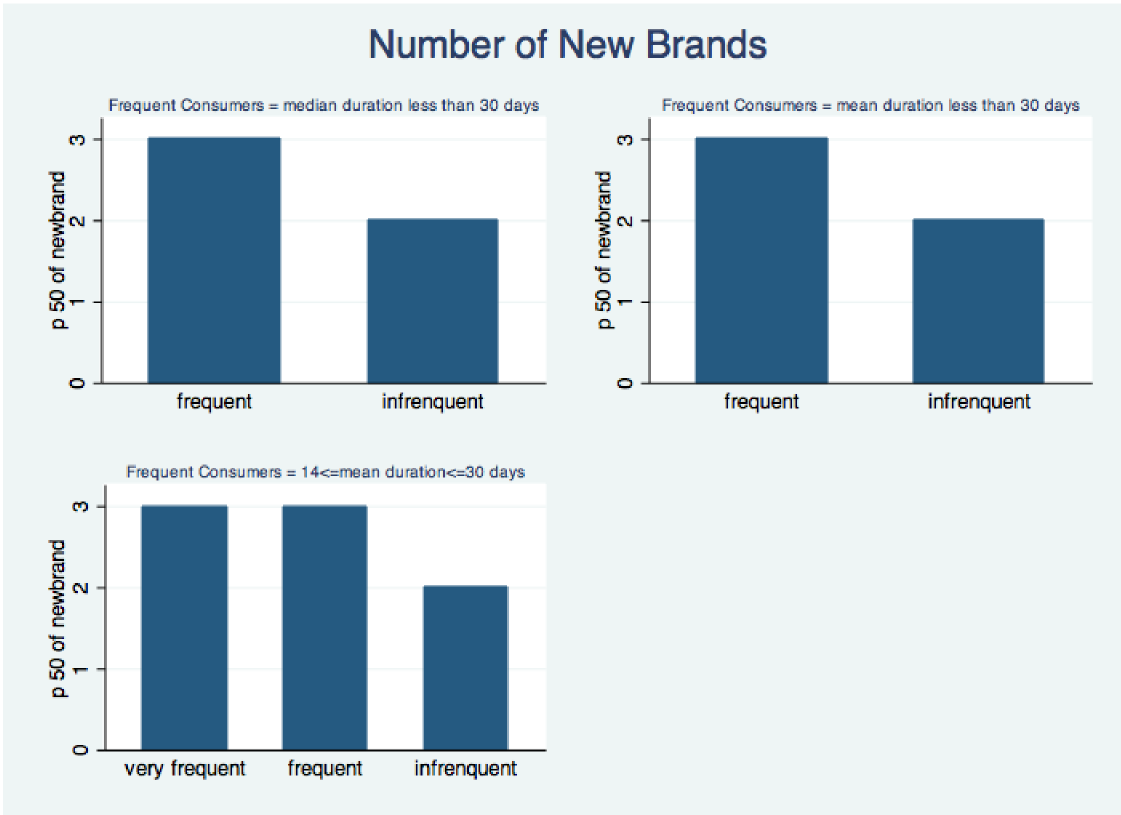

Figure 2: Number of new brands experimented by consumers

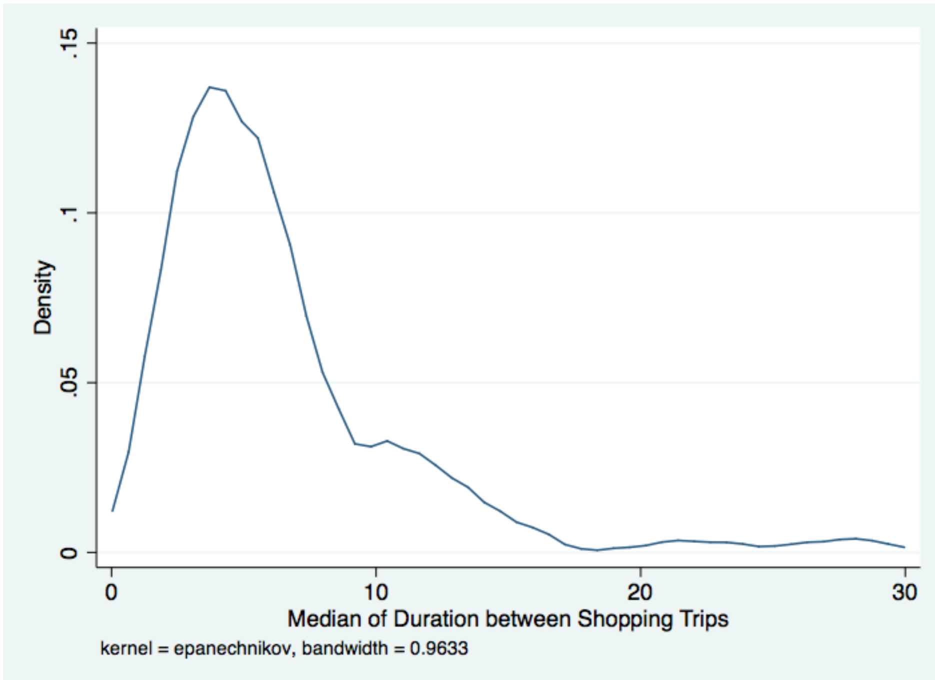

Figure 3: Distribution of median duration between consumers' shopping trips

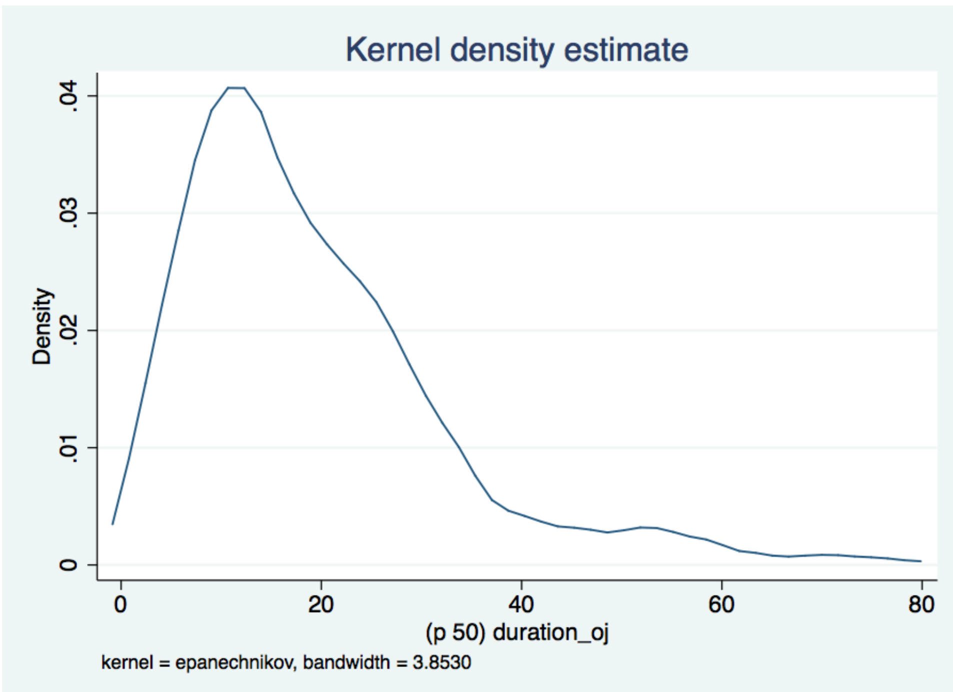

Figure 4: Distribution of median duration between shopping trips on which orange juice was purchased

Table 1: Summary statistics

Prices per 1L	**Average**	**Std. Dev.**
Orange Juice		
Bela Vista 1L	3.405	0.281
Bela Vista 200 ml	5.698	1.613
Xando 1L	3.035	0.23
Xando 300 ml	4.223	0.713
Kapo 200 ml	3.642	0.343
Orange Nectar		
Del Vale 1L	2.668	0.251
Leco 1L	3.413	0.396
Leco 1L light	3.704	0.509
Maguary 1L	2.487	0.3
Maguary 1L light	3.047	0.389
Maguary 200 ml	5.023	0.899
Sufresh 1L	2.375	0.233

Table 2: Estimated Coefficients - Models with Alternative Specific Constants

	(1)	(2)	(3)	(4)	(5)	(6)
Price	-0.279***	-0.257***	-0.210***	-0.211***	-0.184***	-0.407***
	(0.038)	(0.041)	(0.044)	(0.044)	(0.045)	(0.071)
Experience		-3.325***	-1.909***	-2.077***	-1.402***	-1.833***
		(0.052)	(0.064)	(0.099)	(0.262)	(0.302)
Loyalty GL			2.410***	2.403***	2.400***	1.526***
			(0.063)	(0.063)	(0.063)	(0.274)
Price × Experience					-0.215***	-0.042
					(0.078)	(0.089)
Price × Loyalty GL						0.327***
						(0.077)
Median Duration × ...						
... Experience				0.030**	0.030**	0.009
				(0.013)	(0.013)	(0.018)
... Loyalty GL						-0.034*
						(0.020)
... Bela Vista 200 ml				0.091***	0.095***	0.107***
				(0.032)	(0.032)	(0.033)
... Coca Cola 200 ml				0.054**	0.052**	0.059**
				(0.023)	(0.023)	(0.023)
... Del Valle 1L				0.060***	0.058***	0.060***
				(0.020)	(0.020)	(0.020)
... Leco 1L				0.017	0.017	0.022
				(0.032)	(0.032)	(0.032)
... Leco 1L light				-0.014	-0.016	-0.007
				(0.034)	(0.034)	(0.034)
... Maguary 1L				0.065***	0.063***	0.062***
				(0.024)	(0.024)	(0.024)
... Maguary 1L light				0.001	-0.001	0.009
				(0.040)	(0.040)	(0.040)
... Maguary 200 ml				0.074**	0.076**	0.076**
				(0.030)	(0.030)	(0.031)
... Sufresh 1L				0.066***	0.063***	0.066***
				(0.022)	(0.022)	(0.022)
... Xando 1L				0.029	0.028	0.033*
				(0.018)	(0.018)	(0.018)
... Xando 300 ml				-0.022	-0.023	-0.016
				(0.030)	(0.030)	(0.030)
Alternative Specific Effects	✓	✓	✓	✓	✓	✓
Observations	5,213	5,213	5,028	5,028	5,028	5,028

This table reports the results of the mixed logit model with alternative specific fixed effects.
Significance stars are: * p<0.10, ** p<0.05, *** p<0.01.

Table 3: Estimated Coefficients - Models with Module-Brand Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Price	-0.305***	-0.292***	-0.238***	-0.221***	-0.205***	-0.316***
	(0.027)	(0.030)	(0.032)	(0.039)	(0.040)	(0.065)
Experience		-3.422***	-1.941***	-2.061***	-1.618***	-1.788***
		(0.054)	(0.065)	(0.099)	(0.252)	(0.289)
Loyalty GL			2.522***	2.443***	2.440***	2.099***
			(0.064)	(0.064)	(0.064)	(0.260)
Price × Experience					-0.144*	-0.051
					(0.076)	(0.087)
Price × Loyalty						0.165**
						(0.072)
Experience × Median Duration				0.028**	0.029**	0.005
				(0.013)	(0.013)	(0.017)
Loyalty × Median Duration						-0.037*
						(0.019)
Mililiters	0.001***	0.001***	0.002***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Light	-0.563***	0.111	-0.213**	0.121	0.143	0.119
	(0.074)	(0.082)	(0.083)	(0.160)	(0.161)	(0.162)
Nectar	-1.141***	-1.005***	-0.812***	-0.774***	-0.779***	-0.771***
	(0.046)	(0.053)	(0.061)	(0.108)	(0.108)	(0.107)
Sugar	-0.523***	-0.174***	0.105*	-0.142	-0.149	-0.167
	(0.039)	(0.052)	(0.062)	(0.107)	(0.107)	(0.107)
Median Duration × Alternative Dummies				✓	✓	✓
Observations	5213	5213	5028	5028	5028	5028

This table reports the results of the mixed logit model with alternative characteristics.
Significance stars are: * p<0.10, ** p<0.05, *** p<0.01.

Table 4: Different Loyalty Specifications - Models with Alternative Specific Constants

	(1)	(2)	(3)	(4)	(5)	(6)
Price	-0.218*** (0.043)	-0.273*** (0.043)	-0.210*** (0.044)	-0.395*** (0.064)	-0.245*** (0.050)	-0.407*** (0.071)
Experience	-2.412*** (0.059)	-2.867*** (0.056)	-1.909*** (0.064)	-2.169*** (0.294)	-2.464*** (0.264)	-1.833*** (0.302)
Loyalty 1	1.383*** (0.041)			0.683*** (0.201)		
Loyalty 2		0.064*** (0.003)			0.008 (0.008)	
Loyalty GL			2.410*** (0.063)			1.526*** (0.274)
Price × Experience				-0.083 (0.088)	-0.211*** (0.079)	-0.042 (0.089)
Price × Loyalty 1				0.285*** (0.059)		
Price × Loyalty 2					0.000 (0.002)	
Price × Loyalty GL						0.327*** (0.077)
Median Duration × ...						
... Experience				0.003 (0.015)	0.069*** (0.014)	0.009 (0.018)
... Loyalty 1				-0.040*** (0.012)		
... Loyalty 2					0.016*** (0.001)	
... Loyalty GL						-0.034* (0.020)
Median Duration × Alternative Dummies				✓	✓	✓
Alternative Specific Effects	✓	✓	✓	✓	✓	✓
Observations	5,028	5,028	5,028	5,028	5,028	5,028

This table reports the results of the mixed logit model with alternative specific fixed effects.
Significance stars are: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 5: Different Loyalty Specifications - Models with Alternative Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Price	-0.251***	-0.249***	-0.238***	-0.339***	-0.193***	-0.316***
	(0.031)	(0.031)	(0.032)	(0.058)	(0.045)	(0.065)
Experience	-2.467***	-2.956***	-1.941***	-2.205***	-2.548***	-1.788***
	(0.060)	(0.057)	(0.065)	(0.282)	(0.254)	(0.289)
Loyalty 1	1.443***			1.027***		
	(0.042)			(0.191)		
Loyalty 2		0.066***			0.015*	
		(0.003)			(0.008)	
Loyalty GL			2.522***			2.099***
			(0.064)			(0.260)
Price × Experience				-0.070	-0.191**	-0.051
				(0.086)	(0.077)	(0.087)
Price × Loyalty 1				0.186***		
				(0.055)		
Price × Loyalty 2					-0.003*	
					(0.002)	
Price × Loyalty GL						0.165**
						(0.072)
Median Duration × ...						
... Experience				0.000	0.070***	0.005
				(0.015)	(0.013)	(0.017)
... Loyalty 1				-0.042***		
				(0.012)		
... Loyalty 2					0.017***	
					(0.001)	
... Loyalty GL						-0.037*
						(0.019)
Mililiter	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Light	-0.074	-0.007	-0.213**	0.242	0.377**	0.119
	(0.084)	(0.084)	(0.083)	(0.162)	(0.160)	(0.162)
Nectar	-0.908***	-0.810***	-0.812***	-0.879***	-0.742***	-0.771***
	(0.060)	(0.056)	(0.061)	(0.105)	(0.101)	(0.107)
Sugar	-0.010	0.124**	0.105*	-0.312***	-0.152	-0.167
	(0.059)	(0.057)	(0.062)	(0.102)	(0.099)	(0.107)
Median Duration × Alternative Dummies				✓	✓	✓
Observations	5028	5028	5028	5028	5028	5028

This table reports the results of the mixed logit model with alternative characteristics.
Significance stars are: * p<0.10, ** p<0.05, *** p<0.01.

Table 6: Different Sample Sizes for Robustness Checks - Model with Alternative Specific Constants and GL 1983 Loyalty Variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price	-0.211***	-0.207***	-0.201***	-0.179***	-0.407***	-0.438***	-0.501***	-0.486***
	(0.044)	(0.047)	(0.052)	(0.058)	(0.071)	(0.077)	(0.090)	(0.104)
Experience	-2.077***	-2.211***	-2.308***	-2.339***	-1.833***	-1.717***	-1.688***	-1.350**
	(0.099)	(0.120)	(0.167)	(0.235)	(0.302)	(0.376)	(0.491)	(0.648)
Loyalty GL	2.403***	2.410***	2.469***	2.480***	1.526***	1.474***	1.361***	1.411***
	(0.063)	(0.065)	(0.073)	(0.083)	(0.274)	(0.292)	(0.338)	(0.401)
Price x Experience					-0.042	-0.108	-0.141	-0.255
					(0.089)	(0.113)	(0.148)	(0.197)
Price x Loyalty GL					0.327***	0.368***	0.448***	0.448***
					(0.077)	(0.083)	(0.095)	(0.110)
Experience x Median Duration	0.030**	0.028	0.012	-0.029	0.009	-0.004	-0.026	-0.077
	(0.013)	(0.017)	(0.028)	(0.047)	(0.018)	(0.021)	(0.031)	(0.050)
Loyalty GL x Median Duration					-0.034*	-0.048**	-0.070**	-0.087**
					(0.020)	(0.021)	(0.027)	(0.038)
Median Duration × Alternative Dummies				✓	✓	✓		
Alternative Specific Effects	✓	✓	✓	✓	✓	✓		
Observations	5,028	4,298	3346	2,591	5,028	4,298	3,346	2,591

This table reports the results of the mixed logit model with alternative specific fixed effects.
Significance stars are: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 7: Different Sample Sizes for Robustness Checks - Model with Alternative Specific Constants and Loyalty Variable 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price	-0.219***	-0.222***	-0.214***	-0.194***	-0.395***	-0.422***	-0.485***	-0.477***
	(0.043)	(0.046)	(0.051)	(0.058)	(0.064)	(0.068)	(0.080)	(0.092)
Experience	-2.599***	-2.740***	-2.855***	-2.856***	-2.169***	-2.081***	-2.031***	-1.701***
	(0.094)	(0.115)	(0.164)	(0.230)	(0.294)	(0.369)	(0.489)	(0.649)
Loyalty 1	1.381***	1.434***	1.521***	1.549***	0.683***	0.667***	0.547**	0.520*
	(0.041)	(0.043)	(0.048)	(0.055)	(0.201)	(0.218)	(0.257)	(0.306)
Price x Experience					-0.083	-0.154	-0.201	-0.315
					(0.088)	(0.113)	(0.148)	(0.198)
Price x Loyalty 1					0.285***	0.314***	0.400***	0.405***
					(0.059)	(0.064)	(0.075)	(0.088)
Experience x Median Duration	0.035***	0.028*	0.022	-0.025	0.003	-0.006	-0.019	-0.066
	(0.012)	(0.017)	(0.028)	(0.047)	(0.015)	(0.019)	(0.030)	(0.048)
Loyalty 1 x Median Duration					-0.040***	-0.046***	-0.063***	-0.062**
					(0.012)	(0.013)	(0.017)	(0.025)
Median Duration × Alternative Dummies				✓	✓	✓		
Alternative Specific Effects	✓	✓	✓	✓	✓	✓		
Observations	5028	4298	3346	2591	5028	4298	3346	2591

This table reports the results of the mixed logit model with alternative specific fixed effects.
Significance stars are: * p<0.10, ** p<0.05, *** p<0.01.

Table 8: Different Sample Sizes for Robustness Checks - Model with Alternative Specific Constants and Loyalty Variable 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price	-0.274***	-0.273***	-0.265***	-0.257***	-0.245***	-0.250***	-0.266***	-0.275***
	(0.043)	(0.046)	(0.051)	(0.057)	(0.050)	(0.055)	(0.064)	(0.075)
Experience	-2.932***	-2.926***	-2.961***	-2.966***	-2.464***	-2.344***	-2.230***	-2.013***
	(0.094)	(0.117)	(0.164)	(0.225)	(0.264)	(0.340)	(0.458)	(0.611)
Loyalty 2	0.065***	0.064***	0.061***	0.056***	0.008	0.004	0.003	0.006
	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)	(0.008)	(0.008)	(0.009)
Price x Experience					-0.211***	-0.290***	-0.345**	-0.441**
					(0.079)	(0.103)	(0.138)	(0.185)
Price x Loyalty 2					0.000	0.001	0.002	0.002
					(0.002)	(0.002)	(0.002)	(0.002)
Experience x Median Duration	0.014	0.002	-0.003	-0.037	0.069***	0.090***	0.093***	0.084*
	(0.013)	(0.017)	(0.027)	(0.044)	(0.014)	(0.018)	(0.028)	(0.045)
Loyalty 2 x Median Duration					0.016***	0.016***	0.015***	0.012***
					(0.001)	(0.001)	(0.001)	(0.002)
Median Duration × Alternative Dummies				✓	✓	✓		
Alternative Specific Effects	✓	✓	✓	✓	✓	✓		
Observations	5028	4298	3346	2591	5028	4298	3346	2591

This table reports the results of the mixed logit model with alternative specific fixed effects.
Significance stars are: * p<0.10, ** p<0.05, *** p<0.01.

Table 9: Marginal Effects of Price and Elasticities from the Benchmark Demand Model

Alternative	Marginal Effect	Elasticities	Cross Elasticities
Bela Vista 1L	-0.0234774	-0.9521453	0.0872594
Bela Vista 200ml	-0.0122761	-1.659868	0.0727238
Cocal Cola 200ml	-0.0220274	-1.022482	0.0868402
Del Valle 1L	-0.0287384	-0.7281231	0.0856086
Leco 1L	-0.0234141	-0.9551234	0.0872493
Leco 1L light	-0.0215047	-1.048941	0.0865923
Maguary 1L	-0.0301613	-0.6744017	0.0843357
Maguary 1L light	-0.0258931	-0.8437988	0.0871072
Maguary 200ml	-0.0147961	-1.457867	0.0784692
Sufresh 1L	-0.0310553	-0.6419026	0.0833591
Xand 1L	-0.0260069	-0.8389393	0.0870743
Xando 300ml	-0.0187816	-1.197682	0.0844396

This table displays marginal effects for prices, own-price elasticities and cross-price elasticities. The marginal effects are evaluated at the mean value of the regressors. The model was estimated with alternative characteristics.

Table 10: Marginal Effects and Elasticities from the Demand Model with Experience

Alternative	Marginal Effect	Elasticities	Cross Elasticities	Marginal Effect: Milliliter	Marginal Effect: Experience
Bela Vista 1L	-0.054928	-0.7457911	0.2496056	0.0002618	-0.5409826
Bela Vista 200ml	-0.0118577	-1.588962	0.0702729	0.0000565	-0.0913703
Cocal Cola 200ml	-0.0064768	-1.038273	0.0240808	0.0000309	-0.2634202
Del Valle 1L	-0.0332955	-0.6771435	0.1021352	0.0001587	-0.5804425
Leco 1L	-0.0072828	-0.97272	0.0255191	0.0000347	-0.2977938
Leco 1L light	-0.0059863	-1.064713	0.0227421	0.0000285	-0.3011962
Maguary 1L	-0.0149396	-0.6873624	0.0392504	0.0000712	-0.3412787
Maguary 1L light	-0.0064306	-0.8714314	0.0200604	0.0000306	-0.3421805
Maguary 200ml	-0.0024395	-1.458908	0.0123809	0.0000116	-0.0531647
Sufresh 1L	-0.0101339	-0.6695813	0.0249731	0.0000483	-0.3636482
Xand 1L	-0.0618835	-0.6170213	0.2697853	0.0002949	-0.528895
Xando 300ml	-0.0219016	-1.127687	0.100151	0.0001044	-0.1418059

This table reports the marginal effect for prices and container size, own-price elasticities, cross-price elasticities and marginal effect of prior experience. The marginal effects for prices and size are evaluated at the mean value of the regressors. The marginal effect for experience is an average of the marginal effect calculated for each observation. The model was estimated with alternative characteristics.

Table 11: Marginal Effects and Elasticities from the Demand Model with Experience

Alternative	Marginal Effect	Elasticities	Cross Elasticities	Marginal Effect: Milliliter	Marginal Effect: Experience
Bela Vista 1L	-0.0708241	-0.9650018	0.3214644	0.0002622	-0.5059307
Bela Vista 200ml	-0.0140316	-2.061587	0.0828339	0.0000519	-0.0782086
Cocal Cola 200ml	-0.0080379	-1.343145	0.029857	0.0000298	-0.2081254
Del Valle 1L	-0.0438709	-0.8721063	0.1350456	0.0001624	-0.5279136
Leco 1L	-0.0092262	-1.257828	0.0323114	0.0000342	-0.2405716
Leco 1L light	-0.0074788	-1.377053		0.0000277	-0.2374997
Maguary 1L	-0.0199192	-0.8866516	0.0524341	0.0000737	-0.3033866
Maguary 1L light	-0.0084413	-1.125835	0.0263426	0.0000312	-0.2810884
Maguary 200ml	-0.0028042	-1.887298	0.0142183	0.0000104	-0.040629
Sufresh 1L	-0.0137453	-0.8637149	0.0339379	0.0000509	-0.3146421
Xand 1L	-0.080461	-0.7936843	0.3524383	0.0002979	-0.4997225
Xando 300ml	-0.0274676	-1.461634	0.1252425	0.0001017	-0.1282583

This table reports the marginal effect for prices and container size, own-price elasticities, cross-price elasticities and marginal effect of prior experience. The marginal effects for prices and size are evaluated at the mean value of the regressors. The marginal effect for experience is an average of the marginal effect calculated for each observation. Marginal effects are computed differently due to the interaction term between experience and price (the model is no longer linear in prices and experience). The model was estimated with alternative characteristics.

Table 12: Marginal Effects and Elasticities from the Benchmark Demand Model

Alternative	Marginal Effect	Elasticities	Cross Elasticities	Marginal Effect: Experience	Marginal Effect: Duration 0	Marginal Effect: Duration 1	Marginal Effect: Duration × Experience
Bela Vista 1L	-.045692	-.5240811	.2212425	-.4888151	-.0105493	-.0078962	.0140761
Bela Vista 200ml	-.0042096	-1.218017	.024369	-.0323	.0010174	.0011927	-.0018931
Cocal Cola 200ml	-.0118055	-.7499579	.045501	-.3004909	.0018771	.0003821	-.0014634
Del Valle 1L	-.0190725	-.5272379	.0562629	-.3570154	.0036386	.0045004	-.0025315
Leco 1L	-.0059092	-.7266991	.0207527	-.1812471	-.000199	-.0009247	.0046118
Leco 1L light	-.0056584	-.7926331	.0216213	-.1993032	-.0010817	-.0017757	.0101722
Maguary 1L	-.0100677	-.5177736	.0262925	-.1942248	.0023479	.0010848	-.0030885
Maguary 1L light	-.0053192	-.650888	.0166345	-.2035794	-.0003734	-.0010247	.0062572
Maguary 200ml	-.0045618	-1.078201	.0234567	-.0855745	.0012136	.0006571	-.0020806
Sufresh 1L	-.0101558	-.4946975	.0253642	-.2716918	.0019754	.0007006	-.0029173
Xando 1L	-.0455903	-.4676652	.1963493	-.464644	.0018299	.0044728	.0030601
Xando 300ml	-.0077174	-.8857262	.0336422	-.0574995	-.0016963	-.0013692	.0034

This table reports the marginal effect for prices and container size, own-price elasticities, cross-price elasticities and marginal effect of prior experience. The marginal effects for prices and size are evaluated at the mean value of the regressors. The marginal effect for experience is an average of the marginal effect calculated for each observation. The model was estimated with alternative specific effects.

Table 13: Marginal Effects and Elasticities from the Demand Model with Experience

Alternative	Marginal Effect: Price	Elasticities	Cross Elasticities	Marginal Effect: Experience	Marginal Effect: Experience × Median Duration
Bela Vista 1L	-0.0547334	-0.6043128	0.2694393	-0.6198342	0.0133804
Bela Vista 200ml	-0.0050843	-1.427014	0.0294507	-0.040644	-0.0024017
Cocal Cola 200ml	-0.0135294	-0.8804563	0.05207	-0.472287	-0.0048462
Del Valle 1L	-0.0213173	-0.6215062	0.0625389	-0.5090587	-0.0059155
Leco 1L	-0.0056291	-0.8565856	0.0196613	-0.2802856	0.0076672
Leco 1L light	-0.0058168	-0.9324103	0.0221502	-0.3331852	0.0105295
Maguary 1L	-0.010756	-0.6098574	0.027958	-0.2817594	-0.0060621
Maguary 1L light	-0.0055512	-0.7652347	0.0173102	-0.3492905	0.0062137
Maguary 200ml	-0.0048634	-1.266529	0.0249575	-0.1242159	-0.0028399
Sufresh 1L	-0.010233	-0.5843089	0.0253659	-0.407165	-0.0053055
Xando 1L	-0.0551084	-0.5353889	0.2430435	-0.5881708	-0.0008081
Xando 300ml	-0.0097121	-1.035325	0.0424616	-0.074217	0.0037615

This table reports the marginal effect for prices and container size, own-price elasticities, cross-price elasticities and marginal effect of prior experience. The marginal effects for prices and size are evaluated at the mean value of the regressors. The marginal effect for experience is an average of the marginal effect calculated for each observation. Marginal effects are computed differently due to the interaction term between experience and price (the model is no longer linear in prices and experience). The model was estimated with alternative specific fixed effects.

Table 14: Marginal Effects and Elasticities from the Demand Model with Experience

Alternative	Marginal Effect: Price	Elasticities	Cross Elasticities	Marginal Effect: Milliliter	Marginal Effect: Experience	Marginal Effect: Experience × Median Duration
Bela Vista 1L	-0.0521087	-0.6139091	0.2495276	0.0002294	-0.6044369	-0.0046797
Bela Vista 200ml	-0.0054783	-1.407476	0.0317934	0.0000241	-0.0462017	0.0093301
Cocal Cola 200ml	-0.0100534	-0.8834096	0.0381074	0.0000443	-0.4019735	-0.0214953
Del Valle 1L	-0.0239755	-0.6045082	0.0714613	0.0001055	-0.5425628	0.0029883
Leco 1L	-0.0054171	-0.8469927	0.0189094	0.0000238	-0.2748571	0.0036906
Leco 1L light	-0.0055837	-0.9220434	0.0212477	0.0000246	-0.3267133	0.0062312
Maguary 1L	-0.011763	-0.5995517	0.0307338	0.0000518	-0.3092445	0.0026081
Maguary 1L light	-0.0056509	-0.7556729	0.0176334	0.0000249	-0.3559125	0.0080758
Maguary 200ml	-0.0040996	-1.255264	0.0209757	0.000018	-0.1042336	-0.0080728
Sufresh 1L	-0.0096789	-0.5785311	0.0239459	0.0000426	-0.396493	-0.0090152
Xando 1L	-0.0558237	-0.5175794	0.251663	0.0002457	-0.5934752	-0.0063747
Xando 300ml	-0.0103388	-1.019711	0.0453517	0.0000455	-0.077244	0.0133014

This table reports the marginal effect for prices and container size, own-price elasticities, cross-price elasticities and marginal effect of prior experience. The marginal effects for prices and size are evaluated at the mean value of the regressors. The marginal effect for experience is an average of the marginal effect calculated for each observation. Marginal effects are computed differently due to the interaction term between experience and price (the model is no longer linear in prices and experience). The model was estimated with alternative specific effects.