Journal of Business Economics and Management
ISSN 1611-1699 print / ISSN 2029-4433 online
2013 Volume 14(1): 182–199
doi:10.3846/16111699.2012.707151

TOWARDS A DYNAMIC ANALYSIS OF MULTIPLE-STORE SHOPPING: EVIDENCE FROM SPANISH PANEL DATA

Noemí Martínez-Caraballo[1], Manuel Salvador[2], Carmen Berné[3], Pilar Gargallo[4]

[1]Dirección y Organización de Empresas, Centro Universitario de la Defensa, Academia General Militar,University of Zaragoza, Carretera de Huesca, s/n., 50090, Zaragoza, Spain
[2,4]Estructura e Historia Económica y Economía Pública, Facultad de Economía y Empresa,University of Zaragoza, C/ Gran Vía, 2, 50005, Zaragoza, Spain
[3]Dirección de Marketing e Investigación de Mercados, Facultad de Economía y Empresa,University of Zaragoza, C/ Gran Vía, 2, 50005, Zaragoza, Spain
E-mails: [1]noemar@unizar.es (corresponding author); [2]salvador@unizar.es; [3]cberne@unizar.es; [4]pigarga@unizar.es

Received 11 April 2012; accepted 25 June 2012

Abstract. This paper aims to determine why consumer purchasing of fast moving consumer goods varies over time in Spain. More specifically, our objective is to explain multiple-store shopping in the households belonging to the Spanish Nielsen Homescan consumer panel that provides information about household shopping decisions between April 2003 and April 2004. In order to achieve this purpose, a Bayesian Dynamic Tobit model is used. The results allow us to confirm the influence of several demographical and geographical variables on household multiple-store shopping during the sample period.

Keywords: Bayesian inference, fast moving consumer goods (FMCG), MCMC, multiple-store shopping, purchasing patterns, store choice, Tobit model.

Reference to this paper should be made as follows: Martínez-Caraballo, N.; Salvador, M.; Berné, C.; Gargallo, P. 2013. Towards a dynamic analysis of multiple-store shopping: evidence from Spanish panel data, *Journal of Business Economics and Management* 14(1): 182–199.

JEL Classification: C11, D12, O24.

1. Introduction

The literature on consumer behaviour in the retail market is skewed towards studying consumer loyalty and retail patronage behaviour. Most of the empirical contributions to this field of study have focused on brand choice and, to a lesser extent, on analysing store choice. Nevertheless, over the last three decades, variety-seeking behaviour and multiple-store shopping have attracted the attention of researchers in the consumer behaviour area.

Variety-seeking attempts to stimulate purchasing behaviour by alternating between objects of choice. For example, a situation of boredom caused by a non-optimum level of stimulation from purchasing can lead to multiple-store shopping.

When there are alternatives available, households complement their purchases at their first-choice store with purchases at other stores (Kahn, McAlister 1997; Rhee, Bell 2002). In highly competitive markets, such as the fast moving consumer goods (FMCG) markets, multiple-store shopping is ever more widespread (McGoldrick, Andre 1997; Gijsbrechts *et al.* 2008).

Multiple-store shopping could be considered as observable either at a given time or over the course of time (Pessemier 1985). Both types of variety offer a wide range of research possibilities that may have implications for business management. In this paper, we are going to focus on studying dynamic variation.

A regular store set is defined as the stores in which households regularly make purchases. These stores complement each other and may even belong to the same retail chain. The budget of the household is allocated among the different stores in the regular store set. Within this set, one store will typically capture the greatest proportion of expenditure, i.e., it is the first-choice store (Rhee, Bell 2002). It is worth analysing the composition of the household store set and its variation over time. This variation in household store set should be analysed and taken into account by retailers. As the variation increases, the portion of the overall budget not allocated to expenditures at the first-choice store will be increasingly important.

The specific objective of this paper is to study, from a dynamic perspective, the determinants of the variation in the regular store set at which households do their FMCG shopping. For this purpose, a Bayesian statistical model is built (Rossi, Allenby 2003; Rossi *et al.* 2005).

From a revision of the specialised literature, we consider different variables that may be having an effect on the dynamic variation of the regular store set. These underlying factors are: (1) shopping pattern variables; (2) demographic variables of households; (3) demographic characteristics of the shopper and (4) geographical characteristics.

The statistical analysis has been carried out from a dynamic viewpoint using Tobit models. The adopted approach is Bayesian because it allows more flexibility and realism in the modelling process, making inferences that are conditional on the data and that do not depend on asymptotic results (Rossi *et al.* 2005). Due to the complexity of the analysis, the estimation of the parameters uses Markov Chain Monte Carlo (MCMC) methods and the data augmentation technique.

The primary motivation for this study is to generate new insights into the nature of dynamic effects that characterize household store choice behaviour. From an academic perspective, we analyse differences across households in store choice in a study of the dynamic behaviour of consumers and we empirically estimate the model on scanner panel data. From a managerial perspective, we provide several managerial guidelines for retailers interested in maintaining their market share, taking into account the profile of multiple-store shoppers.

The remainder of the article is organized as follows. In Section 2, a review of the relevant literature is carried out and hypotheses are formulated. Section 3 describes the database and the statistical methodology used in the paper. Section 4 presents the empirical results and a discussion of substantive insights that can be obtained from this study and, finally, in Section 5, conclusions and managerial implications are drawn, along with suggestions for future research on this topic.

2. Background and hypotheses

The multiple-store shopping approach is considered here from a dynamic perspective. Variation can be observed in the number of stores that belong to the store set and in the different percentage of household budget spent in each store. The greater the variation, the less loyalty there will be to any one store.

In the literature, several relationships between intrinsic and extrinsic motivations for store switching and multiple-store shopping have been posited (e.g., Berné *et al.* 2005; Berné, Martínez-Caraballo 2009). However, there is little research that analyses demographic characteristics and their effect on multiple-store shopping (Popkowski-Leszczyc, Timmermans 1997) and, in some cases, the results are not conclusive. There are several works that have not found any relationship between demographic characteristics and consumer decisions about product categories (Gupta, Chintagunta 1994; Fox *et al.* 2004).

In the light of this, we attempt to measure the dynamic variation in the store set of each household by analysing the relationship of the degree of variation and (1) shopping pattern variables, (2) demographic variables of households, (3) demographic characteristics of the shopper, and (4) geographical characteristics.

2.1. Shopping pattern variables

The shopping patterns of the households have been analysed in the literature by using the aggregate purchase volume and the interpurchase time.

Regarding the relationship between the purchase volume of the household and its shopping behaviour, the higher the household's total expenditure, the larger the benefits of shopping around for better prices (Mägi 2003) and, consequently, the greater the variation in the store set. In our context, we think that a greater effort devoted to looking for better prices will lead to a store set composed of many more stores. These stores may also experience continuous exchanges of position in relative expenditure and entrances to and exits from the set, reflecting attempts to take advantage of different special offers. Thus, this larger store set may show a bigger variation over time.

However, large purchase volumes are related to household size and it is possible that larger households have less time to go shopping and, thus, tend to concentrate their shopping in just one store. In order to capture the effect of the purchase volume, discounting the effect of the household size variable, we study the impact of the *per capita* purchase volume on the dynamic variation in the store set and we posit that households with large *per capita* basket sizes will exhibit a smaller level of multiple-store shopping.

Consumers' need for variation can be affected by shopping frequency or interpurchase time. The smaller the interpurchase time, or the greater the frequency with which the consumer goes shopping, the sooner the consumer will become satisfied and the need for variety-seeking will lead to boredom or satiation. That is to say, in a brand choice context, boredom or satiation is induced by an accumulated experience of the same brand (Givon 1984). In sequential choice contexts, consumers believed that repetition is associated with boredom and signals closed-mindedness, whereas variety-seeking prevents satiation and signals open-mindedness (Fishbach *et al.* 2011). Although this intrinsic need for stimulus or for innovation can be encouraged by external stimuli, the choice of variety is an internal decision (Kahn 1995). A retailer can protect the principal store by providing variety in complementary ones (Kahn 1995). Inman (2001) demonstrated that switching between flavours and brands in terms of purchase behaviour have been examined, and it appears that consumers do switch more intensively between flavours than between brands for two product categories (tortilla chips and cakes). In the context of other product category (fruits), an experimental research has been conducted and purchase frequency had a significant and positive impact on the consumption variety-seeking behaviour (Berné, Múgica 2010).

Hence, we posit the following hypotheses:

H1: "The higher the aggregate purchase volume per capita, the smaller the dynamic variation in the store set will be".

H2: "The smaller the interpurchase time, the greater the dynamic variation in the store set will be".

2.2. Demographic variables of households

In this subsection, we consider social class and household size.

Social class is determined by a complex set of variables including income, occupation and education. Household social class is an important determinant not only of how much is spent but also of how it is spent. The potential of social class as a marketing segmentation variable was first noted in the 1940s when Warner (in Coleman 1983) found that each of the social class groups that he identified displayed unique purchase motivations and shopping behaviours (Henry 2002). In the marketing literature and, especially in that referring to buyer behaviour, social class has been considered a better variable than income as a predictor of consumer behaviour (e.g., Martineau 1958; Schaninger 1981). Households with higher income levels are usually more loyal to the first-choice store (McGoldrick, Andre 1997; Seetharaman, Chintagunta 1998) and, in our context, we posit that households with higher socioeconomic status may show a smaller variation over time.

Household size may have a positive effect on multiple-store shopping (Seetharaman, Chintagunta 1998). Larger households will be more likely to have different tastes and needs (Seetharaman, Chintagunta 1998) and, so, a higher level of multiple-store shopping will be expected. Conversely, Mägi (2003) maintains that larger households may have more time restrictions and a greater tendency to concentrate their purchases at a single store, so the dynamic variation in the store set will be less.

Therefore, we hypothesise:

H3: "The higher the social class of the household, the smaller the dynamic variation in the store set will be".

H4: "The bigger the household, the smaller the dynamic variation in the store set will be".

2.3. Demographic variables of the shopper

Multiple-store shopping could also be explained by exogenous factors, out of the control of retailers, such as demographic variables (Berné *et al.* 2005). The variables we employ are the age of the shopper, employment status and whether the shopper has young children. We will posit several hypotheses regarding these demographic characteristics.

On the one hand, several studies have described the elderly as more regular customers than younger ones. Botwinick (1978) argued that older consumers could be more likely to choose not to change – that is to say, to remain persistently attached to the same option – due to their cautiousness in decision or to risk aversion. Regarding new car purchases, Lambert-Pandraud *et al.* (2005) found that the consideration of a single model increased markedly with age, whereas the percentage of buyers who considered three or more models dropped sharply with age and conclude that, in this context, older consumers consider fewer brands, fewer dealers, and fewer models, and they choose long-established brands more often. On the other hand, several studies have demonstrated that the age of the shopper is positively related to multiple-store shopping (East *et al.* 1995, 2000; Mägi 2003). One explanation for this relationship is that older consumers, especially those who are retired, have more free time and, thus, they can dedicate more time to shopping, to comparing offers and to using several stores to cover their shopping needs (East *et al.* 2000). Hence, regarding FMCG purchases, a positive relationship is expected between the age of the shopper and the variation in the store set.

Several studies have pointed out that households with greater work commitments and time restrictions avoid variety-seeking. People who have less free time will concentrate their purchases in a more limited number of stores in order to invest less time and effort into making purchases of frequently-used products (McGoldrick, Andre 1997). Furthermore, shoppers who work outside the home will be more loyal to their first-choice store (McGoldrick, Andre 1997; Fox *et al.* 2004). Therefore, the dynamic variation in the store set is expected to be lower.

If the person in charge of shopping for frequently-used products is working full time is between 25 and 40 and in a large household, they are more likely to show loyal behaviour, given that their household commitments and their time restrictions are greater (East *et al.* 1997). Time restrictions also emerge with the presence of children at home (Soberon-Ferrer, Dardis 1991). In fact, we can posit that these households will be prone to concentrate their FMCG purchases in a limited number of stores. As a result, a smaller dynamic variation in the store set could be expected.

In short, a greater volume of consumption needs, less time and more commitments can favour a lower level of multiple-store shopping. In particular, it is highly likely that households with time restrictions derived from the presence of children will concentrate their expenditure on groceries and household products at fewer stores, so their store set will probably have less variation.

Consequently, the following working hypotheses are set forth:

H5: "The older the shopper, the greater the dynamic variation in the store set will be".

H6: "If the shopper works outside the home, then the dynamic variation in the store set will be lower".

H7: "If the shopper has young children, then the dynamic variation in the store set will be lower".

2.4. Geographical characteristics

Other potentially relevant influences on the dynamic variation of the store set are the geographical area and the size of the town or village in which the household lives.

In the literature, it has been seen that geographical location exercises a significant effect on the allotment of expenditure on household services (Soberon-Ferrer, Dardis 1991). Companies consider these geographical areas to plan their commercial routes (decisions on sales and distribution of products). This distribution of the market explains a lot of the heterogeneity that exists from the supply side (i.e., number of stores, etc.) and from the demand side (sociological, demographic or economic differences). So, we will include geographical area in our model in order to test the effects of this heterogeneity. We will posit that belonging to a geographical area can determine a bigger or smaller variation in the store set. Moreover, the commercial supply in the geographical areas will differ according to the size of the town or village.

Consequently, we can posit the following hypotheses:

H8: "The degree of dynamic variation in the store set will differ among geographical areas".

H9: "The bigger the town, the greater the dynamic variation in the store set will be".

3. Empirical analysis

3.1. The database

The database has been built from a Nielsen household panel data containing information about the purchases of groceries and household products carried out by 2,016 Spanish households from April 2003 to April 2004. Purchases from all outlets are captured (e.g., grocery stores, mass merchandisers, supermarket, hypermarkets, convenience stores, and so on). Tracking only grocery store purchases might obscure the phenomenon or bias the analysis.

Statistical analysis was performed with statistical software (SPSS 16.0 for Windows). To measure the variation in the store set, we have used the Consumer Behavior Herfindahl-Hirschman Index (HHI) (see Theil, Finke 1983; Ginevičius, Čirba 2009 for a review on measurement of market concentration; see Van Trijp, Steenkamp 1990, 1992; Van Trijp

1995 for a review on measurement of consumer behaviour) calculated in each period "t". This index is given by the following expression:

$$HHI_t = -\sum_{k=1}^{m_t}\left[p_{k,t}\right]^2, \tag{1}$$

where: "$p_{k,t}$" is the percentage of expenditure in store "k" from the initial period of the study until the current period "t" and "m_t" is the total number of stores belonging to the store set of each household in period "t".

The variation in the store set is smaller when the percentage of the budget allocated to the first-choice store is bigger; when the number of stores belonging to the store set is smaller; and when the percentage of budget allocated to the complementary stores is smaller. In Figure 1, the histogram of HHI at the end of the one-year period analysed is shown. It can be seen that HHI is a mixed variable with one discrete mass point and a continuous part. The discrete part has its mass point in HHI = –1 due to the existence of a group of households that only buy in one store. On the other hand, it can be observed that the continuous part is roughly unimodal and left-skewed.

In order to increase the normality degree of the dependent variable, a logarithmic transformation has been used. In particular, the following natural logarithmic transformation was performed to achieve normality: $h_t = -\ln(-IHH_t)$.

Table 1 includes the number of inhabitants, number of households and a brief description of stores (number and retail format) in each Nielsen Area.

Each household will have a certain store set formed by one, two or more regularly used stores, one of which will be the first-choice store and the others will be complementary stores. The regular store set consisted of a maximum of thirteen stores. In Table 2, we can see the percentage of the sample that solved their shopping needs in one store, in two stores, and so on, during the year of the study. In particular, we highlight that 6% of the sample solved their shopping needs at just one store.

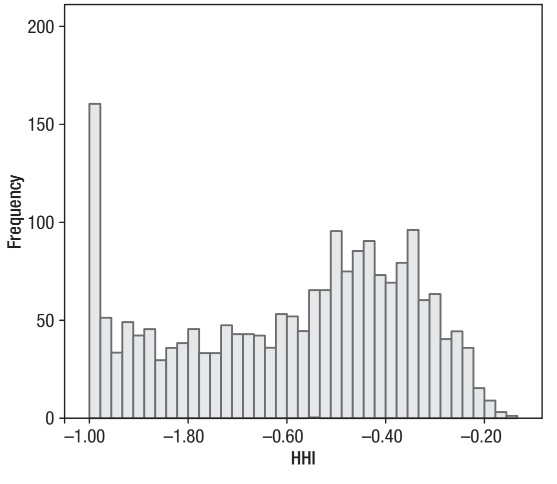

Fig. 1. HHI Histogram

Table 1. Number and retail format of stores in Nielsen areas of the Iberian Peninsula and the Balearic Islands

	Inhabitants	House-holds	Tradi-tional retail stores	Discount & Self-Service Stores	Super-mar-kets	Hyper-markets	Stores (total)
Nielsen Area I	4,914,321	1,750,153	3,706	1,821	1,972	53	7,552
Nielsen Area II	6,131,576	2,163,050	5,526	2,477	1,792	51	9,846
Nielsen Area III	8,201,414	2,906,759	9,960	2,738	3,107	76	15,881
Nielsen Area IV	3,954,904	1,447,663	5,757	1,691	1,389	31	8,868
Nielsen Area V	4,376,585	1,658,511	5,377	2,034	1,608	35	9,054
Nielsen Area VI	4,130,496	1,517,848	4,649	1,441	1,181	46	7,317
Nielsen Area VII (Metropolitan Area of Barcelona)	3,935,604	1,440,723	3,418	1,381	1,282	26	6,107
Nielsen Area VIII (Metropolitan Area of Madrid)	5,356,389	1,883,791	3,381	1,281	1,210	47	5,919
TOTAL	41,001,289	14,768,498	41,774	14,864	13,541	365	70,544

Source: Own elaboration from Nielsen (2005).

Table 2. Number of stores in the household store set

# of stores	# of households	%
1	121	6.0
2	262	13.0
3	230	11.4
4	417	20.7
5	321	15.9
6	248	12.3
7	135	6.7
>8	135	7.3
Total	2,016	100.0

Source: Own elaboration.

Table 3 displays the choice of retail format by households in the sample and the composition of their regular store sets. The supermarket is the most used as first-choice store while the least used is the traditional retail store, which serves as complementary store.

In Table 4, we can see the descriptive analysis of the independent variables considered in the study.

Table 3. Choice of retail format by households

	Traditional retail store	Discount & Self-Service Store	Supermarket	Hypermarket
First-choice store	1.40%	21.40%	57.50%	19.70%
First complementary store	3.37%	23.81%	49.42%	23.40%
Second complementary store	5.73%	21.22%	52.74%	20.31%
Third complementary store	9.54%	19.05%	52.36%	19.05%
Fourth complementary store	13.83%	19.69%	51.82%	14.66%
Fifth complementary store	15.76%	17.45%	55.91%	10.88%
Sixth complementary store	20.35%	15.44%	55.09%	9.12%
Seventh complementary store	22.15%	11.41%	61.07%	5.37%
Eighth complementary store	36.37%	9.09%	43.93%	10.61%
Ninth complementary store	37.04%	3.70%	51.85%	7.41%
Tenth complementary store	46.66%	6.67%	40.00%	6.67%

Source: Own elaboration.

Table 4. Descriptive analysis of the independent variables[+]

	Mean	Standard Deviation
Expenditure *per capita* (€)	11.70	13.42
Interpurchase time (days)	6.90	7.97

	Lower	Middle*	Higher
Social Class	25.9%	52.6%	21.5%

	1	2	3	4	>4*
Household Size	6.0%	21.8%	24.1%	30.9%	17.2%

	<35	Between 35 and 54*	>55
Age (years)	11.6%	51.9%	36.5%

	No*	Yes
Employment	81.0%	19.0%
Young Children	64.8%	35.2%

	I*	II	III	IV	V	VI	VII	VIII
Nielsen Area	12.1%	15.4%	17.6%	9%	11.9%	10.1%	11.2%	12.7%

	<10	Between 10 and 50	Between 50 and 200	>200*
Town Size (thousands of inhabitants)	15.7%	22.5%	27.6%	34.3%

Note: [+]For each categorical variable, the reference category is marked with*.
Source: Own elaboration.

In order to measure the aggregate *per capita* purchase volume of a household, we have considered the mean of all the shopping tickets accumulated until the considered date. The categories of the social class variable are based on the classification used by Nielsen. In our database, we consider three different classes: Lower, Middle and Higher.

Work commitments are measured through two variables: the employment status and family situation of the shopper, the latter depending on whether there are any children under 12 in the household.

For the analysis of the influence of the geographical area in which the household lives, we have considered the Nielsen area. Nielsen divides Spain into 8 areas depending on their geographical situation (see Figure 2).

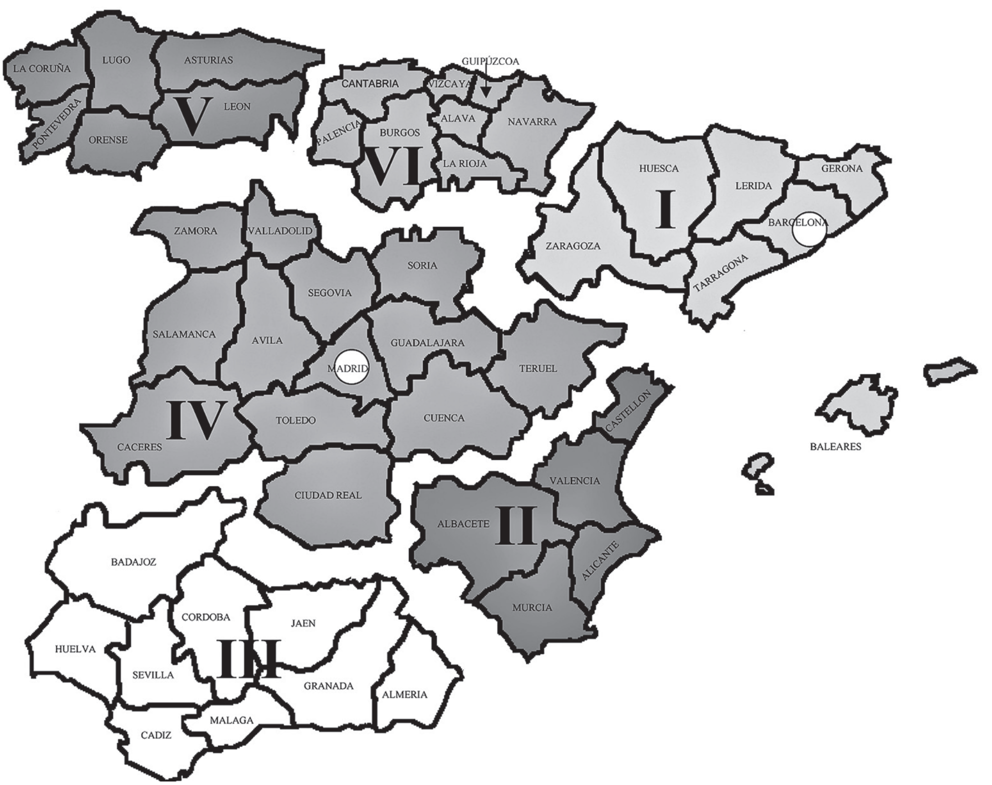

Fig. 2. Map of Spanish Nielsen geographical areas.
Source: Own elaboration.

3.2. Specification of the model

In this section, the statistical model used to test the hypotheses posited in Section 2 is described. A Bayesian hierarchical Tobit model is used. All simulations were performed in MatLab 6.5.

An important feature of the data, which influences the modelling strategy, is the mixed character of the dependent variable, HHI. The most parsimonious model for dealing with its mixed character is the Tobit model. These models are often conceptualized in a hierarchical manner, where movement from one model component to the next proceeds in a logical manner. Hierarchical Bayes methods have recently become available to marketing researchers, and there is ample evidence of their superiority for estimation of this kind of models (see Gelman *et al.* 2004). An advantage of estimating hierarchical Bayes models with MCMC methods is that it yields estimates of all model parameters, including estimates of model parameters associated with specific respondents.

Several factors favour our approach. First, the Bayesian hierarchy captures systematic (as well as random) sources of heterogeneity in multiple-store shopping, so we can assess the predictive contribution of three different types of variables: purchase histories, demographics, and geographic variables. Also, because the Gibbs Sampler allows us to sample from the posterior distribution of any function of model parameters, we can construct Bayesian prediction intervals for multiple-store shopping. Finally, our approach requires only households' data, which most grocery retailers already gather, and multi-outlet panel data.

3.2.1. The dataset

We consider a sample of N households and we analyse the shopping trips of each household during a fixed period T (one year).

Our data set is given by $\mathbf{D} = \{(\mathbf{x}_{ij}, t_{ij}, h_{ij}); j = 1,\ldots,n_i; i = 1,\ldots,N\}$ where:

- $\mathbf{x}_{ij} = (x_{i,j,1},\ldots,x_{i,j,p})'$ are the covariates corresponding to the i-th household in the j-th trip;
- $0 \leq t_{i1} < t_{i2} < \ldots < t_{i,n_i} \leq T$ are the days on which the i-th household goes shopping;
- $h_{ij} = -\ln\left(-IHH_{i,t_{i,j}}\right)$ with $IHH_{i,t_{ij}} = -\sum_{k=1}^{m_{t_{ij}}} p^2_{i,k,t_{i,j}}$ where: $p_{i,k,t_{ij}}$ is the percentage of expenditure of the i-th household in store k on the j-th shopping trip.

3.2.2. The model

Taking into account the mixed character of the dependent variable HHI, we consider the Dynamic Tobit model in a Bayesian framework, given by:

$$h_{ij} = \begin{cases} 0 & \text{with probability } 1 - \Phi\left(\tau_i^{\frac{1}{2}}\left(\boldsymbol{\beta}'\mathbf{x}_{ij}\right)\right), \\ \boldsymbol{\beta}'\mathbf{x}_{ij} + \varepsilon_{ij} & \text{with probability } \Phi\left(\tau_i^{\frac{1}{2}}\left(\boldsymbol{\beta}'\mathbf{x}_{ij}\right)\right), \end{cases} \tag{2}$$

where: $\varepsilon_{ij} \sim N\left(0, \tau_i^{-1}\right)$ and $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)'$ is the vector of the regression coefficients that determine the sign and the intensity of the influence of the independent covariates on the multiple-store shopping of a household.

The model is a multivariate system of hierarchical Bayesian Tobit censored regressions, which is estimated using the Gibbs Sampler. Every predictor variable specified in the model is found in panel data. Although retailers may not currently gather every predictor variable, they could; and retailer decisions about gathering additional variables could be informed by our evaluation of the variables' predictive contributions.

3.2.3. Prior distribution

Given that we adopt a Bayesian approach to the problem, we need to specify a prior distribution on the parameters of the model. In our case, we have adopted the usual fully conjugate prior distributions given by:

$$\boldsymbol{\beta} \sim N_p(\mathbf{0}, \mathbf{S}_\beta), \tag{3}$$

$$\tau_i \sim \text{Gamma}\left(\frac{n_\tau}{2}, \frac{n_\tau s_{0\tau}}{2}\right); \; i = 1, \ldots, N, \tag{4}$$

With: known constants $n_\tau > 0$, $s_{0\tau} > 0$ and a known symmetrical definite positive matrix $\mathbf{S}_\beta$ (pxp) and all the distributions (3)–(4) mutually independent.

3.2.4. Posterior distribution

In order to calculate the posterior distribution, we use the data augmentation technique (Tanner, Wong 1987) and we introduce the non-positive latent variables $\boldsymbol{\lambda} = \{\lambda_{ij}; j \in \{1, \ldots, n_i\}: h_{ij} = 0; I = 1, \ldots, N\}$. We also define $\lambda_{ij} = h_{ij}$ for $j \in \{1, \ldots, n_i\}: h_{ij} > 0$.
Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\lambda})$ the vector of parameters where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_N)$.
We consider the probability distribution given by:

$$[\boldsymbol{\theta}|\mathbf{D}] \propto \prod_{i=1}^{N}\prod_{j=1}^{n_i}\left[\lambda_{ij}|\boldsymbol{\beta}, \tau_i, \mathbf{x}_{ij}\right][\boldsymbol{\beta}|\mathbf{S}_\beta]\prod_{i=1}^{N}\left[\tau_i \mid n_\tau, s_{0\tau}\right] \propto,$$

$$\prod_{i=1}^{N}\prod_{j=1}^{n_i}\tau_i^{\frac{1}{2}}\exp\left[-\frac{\tau_i}{2}\left(\lambda_{ij} - \boldsymbol{\beta}'\mathbf{x}_{ij}\right)^2\right]\prod_{(i,j):h_{ij}=0} I_{(-\infty,0)}\left(\lambda_{ij}\right),$$

$$\exp\left[-\frac{1}{2}\boldsymbol{\beta}'\mathbf{S}_\beta^{-1}\boldsymbol{\beta}\right]\prod_{i=1}^{N}\tau_i^{\frac{n_\tau}{2}-1}\exp\left[-\frac{n_\tau s_{0\tau}}{2}\tau_i\right]I_{(0,\infty)}\left(\tau_i\right), \tag{5}$$

where: I_A denotes the indicator function of A and R_{p+1} = {symmetrical definite positive matrices (p+1)x(p+1)}. The posterior distribution of $(\boldsymbol{\beta}, \boldsymbol{\tau})$ is the corresponding marginal distribution of (5). This is not a standard distribution and we use MCMC methods (Rossi *et al.* 2005) to calculate it. In particular, we use the Gibbs sampling algorithm employing the full conditional distributions of (5) which will be available upon request:

$$\{\boldsymbol{\theta}^{(j)} = \left(\boldsymbol{\beta}^{(j)}, \tau_1^{(j)}, \ldots, \tau_N^{(j)}, \boldsymbol{\lambda}^{(j)}; \; j = s_0, s_0 + L, \ldots, s_0 + \left(S-1)L\right)\right), \tag{6}$$

where: s_0 is the "burn-in" number of iterations necessary to achieve convergence, L is the number of estimated steps needed to obtain an approximate uncorrelated sample and S is the sample size. Using sample (6), it is possible to make inferences about the parameters of model (2)–(3) calculating medians and quantiles that let us to obtain point estimations and Bayesian credibility intervals of the parameters of the model.

4. Results

We have taken p = 23 independent variables, namely, the constant, the 20 variables that come from adopting the indicator codification of the categorical variables listed in Table 4, the interpurchase time and the natural logarithm of *per capita* expenditure. We take $n_{0\tau} = 0.1$, $s_{0\tau} = 1$ and $\mathbf{S}_\beta = 100\mathbf{I}$, which constitutes a flat prior distribution of the parameters of the model[1].

Gibbs sampling was run for 10,000 iterations and the convergence was achieved after $s_0 = 1001$ iterations. We took a sample every L = 10 iterations in order to obtain an approximate uncorrelated sample. Therefore, the sample size of (5) was S = 900.

The estimations of the parameters of the model are shown in Table 5. In particular, we have calculated the posterior median and the posterior quantiles 2.5 and 97.5 calculated from sample (6) that constitutes a point estimation and the limits of the 95% Bayesian credibility interval, respectively, of the parameters of the model.

Table 5. Estimation of the parameters of the model[++]

Variable	Quantile 2.5	Median	Quantile 97.5
Constant	0.6062	0.6231	0.6396
ln(Expenditure *per capita*)	–0.0132	–0.0101	–0.0069
Interpurchase Time	–0.0010	–0.0007	–0.0005
Lower Class	0.0657	0.0754	0.0851
Higher Class	0.0314	0.0375	0.0429
Household Size 1	–0.0808	–0.0607	–0.0412
Household Size 2	0.0212	0.0303	0.0395
Household Size 3	0.0679	0.0782	0.0891
Household Size 4	0.0663	0.0736	0.0807
Age < 35	–0.0479	–0.0342	–0.0194
Age > 55	0.0703	0.0778	0.0857
Employment	0.0331	0.0409	0.0476
Children < 12	–0.0254	–0.0170	–0.0091
Nielsen Area II	–0.0725	–0.0583	–0.0457
Nielsen Area III	–0.2030	–0.1889	–0.1750
Nielsen Area IV	–0.0420	–0.0216	–0.0053
Nielsen Area V	0.0214	0.0377	0.0493
Nielsen Area VI	–0.0769	–0.0590	–0.0451
Nielsen Area VII	–0.0427	–0.0212	–0.0032
Nielsen Area VIII	0.0261	0.0369	0.0485
Town Size < 10	–0.0470	–0.0331	–0.0188
10 < Town Size < 50	0.0642	0.0744	0.0846
50 < Town Size < 200	0.0327	0.0434	0.0565

Note: [++]All the coefficients are significant at 95%.
Source: Own elaboration.

[1] A model parameter sensitivity study reveals to have an insignificant effect on results due to sample size.

From the results shown in Table 5, it can be seen that all the independent covariates of the model have a significant influence on household multiple-store shopping, but not all of them are in the expected direction. More specifically:

1) Given that the *per capita* expenditure coefficient is significantly negative, it follows that the higher the aggregate purchase volume, the lower the dynamic variation in the store set. So, hypothesis **H1** must be accepted.
2) The coefficient of interpurchase time is negative. Therefore, the lower the interpurchase time, the higher the variation in the store set over time. So, hypothesis **H2** is accepted and the interpurchase time is directly related to the variation in the store set.
3) The coefficients of the lower and higher categories of the social class variable are positive. This reveals that social class has a non-monotonic effect on the variation in the store set, the middle class households having a lower variation in their store set. Thus, households with higher socioeconomic status show a higher multiple-store shopping than middle ones; in consequence, hypothesis **H3** cannot be accepted.
4) The signs of the coefficients of the household size variables reveal a non-monotonic relation with multiple-store shopping behaviour, the smaller and the larger households being those with less variation in the store set. In particular, households of size 3 and 4 are the ones, which tend to have the greatest variation. Therefore, hypothesis **H4** is rejected.
5) The age of the buyer is directly related to the variation in the store set. Elder people tend to have the greatest dynamic variation and young people the least. Consequently, hypothesis **H5** is accepted. This result is in line with the reasoning and empirical results in East *et al.* (1995, 2000) and Mägi (2003). As an improvement to this study and in order to capture details that shopper age cannot explain, we propose an in-depth study of the family life cycle.
6) The employment commitments of the shoppers have a significant influence on the variation in their store set. However, the relationship operates in the opposite direction to what was expected. The shoppers that work tend to have the greatest dynamic variation, so hypothesis **H6** is rejected.
7) If there are small children in the home, the variation in the household store set seems to be lower. So, hypothesis **H7** can be accepted.
8) The geographical area where the household shops influences the degree of variation. Specifically, households in the south (Nielsen Area III) show the lowest variation over time in their store set. The situation is the opposite in the northwest (Nielsen Area V) and the metropolitan area of Madrid (Nielsen Area VIII). Thus, hypothesis **H8** is accepted.
9) The effect of the town size is non-linear. Households in medium-size towns (between 10,000 and 200,000 inhabitants) tend to have greater dynamic variation than those in smaller and bigger towns. Therefore, hypothesis **H9** must be rejected.

5. Conclusions, managerial guidelines and further research

In this research, we have carried out a dynamic analysis of multiple-store shopping using a panel of Spanish households over a one-year period. The focus of our paper is on understanding multiple dimensions of household store choice behaviour. The results confirm that multiple-store shopping is widespread for FMCG and that several variables (interpurchase time, social class, age, whether the shopper has young children and geographical area) have the expected influence on it.

Our research contributes to the consumer behaviour literature since, to our knowledge, no previous work has examined the drivers of multiple-store shopping in Spain in the context of a store set used for the purchase of FMCG. The results reported here are consistent with those obtained in previous studies in a brand choice context and using store scanner databases, in which it has been shown that there is a considerable variation across retailers, across product categories, and within a product category for a given retailer.

Moreover, those households living on the south coast and east coast of the peninsula show a less varied purchasing environment. This result could lead to an interesting, in-depth study on the heterogeneity that this variable could reflect, regarding both the supply side and the characteristics of the households that make up the demand side in the various geographic zones.

In short, belonging to a higher social class, having a smaller household size, a greater volume in the shopping basket and a lower interpurchase time are characteristics of households whose store set varies less over the period analysed.

Due to the high level of disaggregation of our data, the variables related to household and town size and social class have been found to show a non-linear effect. Smaller towns offer fewer shopping alternatives, but larger towns involve significantly higher travel costs for shoppers. So, managers will have more possibilities to influence household multiple-store shopping behaviour in medium size towns.

Work commitments and volume of the shopping basket have shown the opposite relation to what was expected.

Besides, we can affirm that managers taking several decisions, such as the location of a new store and the implementation of retention strategies mixed with variety and multi-format strategies, must consider the demographic and geographical characteristics of their customers. They have to make an effort to study the profiles of households by geographical area before taking their final decisions. For this purpose, advanced Geographical Information Systems may be useful.

Among the limitations of this study is the fact that we have used a secondary data source. Moreover, the Tobit model assumes that the effect of the independent variables is homogeneous for all the households in the sample. This is an aspect that needs to be addressed in the future. Bayesian hierarchical modelling may be a useful tool for performing analyses like these.

Another limitation of this study comes from the limited external validity of the analysis reported here. Hence, it is necessary to replicate the study by using different databases. Besides, variables' categorisation used in the *Homescan* Nielsen panel data for 'Shopper age', 'Employment', 'Nielsen Area' or 'Town Size' do not allow reflecting a high variability. In the future, our intention is to consider a wider time range.

This research may be broadened and the managerial implications enriched through the analysis of the synergies between the defensive strategies, variety strategies and multiformat strategies of retail companies. So, it would be interesting to carry out an in-depth study using a mixed supply and demand database.

Acknowledgements

This research has been undertaken within the framework of the Projects MTM 2008-00625 and ECO2010-20880/ECON, financed by the Research Agency of the Spanish government. The authors also wish to thank for the financial support received under the following research groups: "CREVALOR: Creación de valor en las organizaciones", "IMPROVE" and "Métodos estadísticos no paramétricos sobre datos sesgados en el mercado laboral", financed by DGA-FSE. A previous version of this paper was published as working paper nº 594/2011 in the Working Paper series edited by Fundación de las Cajas de Ahorros (FUNCAS).

References

Berné, C.; Martínez-Caraballo, N. 2009. Determinantes del comportamiento variado del consumidor en el escenario de compra, *Revista Europea de Dirección y Economía de la Empresa* 18(3): 99–113.

Berné, C.; Múgica, J. M.; Rivera, P. 2005. Managerial ability to control the varied behaviour of regular customers in retailing: interformat differences, *Journal of Retailing and Consumer Services* 12(3): 151–164. http://dx.doi.org/10.1016/j.jretconser.2004.05.003

Berné, C.; Múgica, J. M. 2010. Identification of variety cycles in food choices: its' impact on state-dependence coefficients, *Food Quality and Preference* 21(6): 639–647. http://dx.doi.org/10.1016/j.foodqual.2010.04.001

Botwinick, J. 1978. Cautiousness in decision, in J. Botwinick (Ed.). *Aging and Behavior: A Comprehensive Integration of Research Findings*. New York: Springer, 128–141.

Coleman, R. P. 1983. The continuing significance of social class to marketing, *Journal of Consumer Research* 10(3): 265–280. http://dx.doi.org/10.1086/208966

East, R.; Harris, P.; Willson, G.; Lomax, W. 1995. Loyalty to supermarkets, *International Review of Retail, Distribution and Consumer Research* 5(1): 99–109. http://dx.doi.org/10.1080/09593969500000006

East, R.; Hammond, K.; Harris, P.; Lomax, W. 2000. First-store loyalty and retention, *Journal of Marketing Management* 16(4): 307–325. http://dx.doi.org/10.1362/026725700784772907

Fishbach, A.; Ratner, R. K.; Zhang, Y. 2011. Inherently loyal or easily bored? Nonconscious activation of consistency *versus* variety-seeking behaviour, *Journal of Consumer Psychology* 21(1): 38–48. http://dx.doi.org/10.1016/j.jcps.2010.09.006

Fox, E. J.; Montgomery, A. L. A.; Lodish, L. 2004. Consumer shopping and spending across retail formats, *Journal of Business* 77(2): S25–S60. http://dx.doi.org/10.1086/381518

Gelman, A.; Carlin, J. B.; Stern, H. S.; Rubin, D. B. 2004. *Bayesian Data Analysis*. 4th ed. Boca Raton, FL: Chapman & Hall/CRC.

Gijsbrechts, E.; Campo, K.; Nisol, P. 2008. Beyond promotion-based store switching: antecedents and patterns of systematic multiple-store shopping, *International Journal of Research in Marketing* 25(1): 5–21. http://dx.doi.org/10.1016/j.ijresmar.2007.07.001

Ginevičius, R; Čirba, S. 2009. Additive measurement of market concentration, *Journal of Business Economics and Management* 10(3): 191–198. http://dx.doi.org/10.3846/1611-1699.2009.10.191-198

Givon, M. 1984. Variety-seeking through brand switching, *Marketing Science* 3(1): 1–22. http://dx.doi.org/10.1287/mksc.3.1.1

Gupta, S.; Chintagunta, P. K. 1994. On using demographic variables to determine segment membership in logit mixture models, *Journal of Marketing Research* 31(1): 128–139. http://dx.doi.org/10.2307/3151952

Henry, P. 2002. Systematic variation in purchase orientations across social classes, *Journal of Consumer Marketing* 19(5): 424–438. http://dx.doi.org/10.1108/07363760210437641

Inman, J. J. 2001. The role of sensory-specific satiety in attribute-level variety seeking, *Journal of Consumer Research* 28(1): 105–120. http://dx.doi.org/10.1086/321950

Kahn, B. E. 1995. Consumer variety-seeking among goods and services. An integrative review, *Journal of Retailing and Consumer Services* 2(3): 139–148. http://dx.doi.org/10.1016/0969-6989(95)00038-0

Kahn, B. E.; McAlister, L. 1997. *Grocery Revolution: the New Focus on the Consumer*. Reading, MA: Addison Wesley, Longman.

Lambert-Pandraud, R.; Laurent, G.; Lapersonne, E. 2005. Repeat-purchasing of new automobiles by older consumers: empirical evidence and interpretations, *Journal of Marketing* 69(2): 97–113. http://dx.doi.org/10.1509/jmkg.69.2.97.60757

Lancaster, T. 2004. *An Introduction to Modern Bayesian Econometrics*. Oxford, UK: Blackwell Publishing.

Mägi, A. W. 2003. Share of wallet in retailing: the effects of customer satisfaction, loyalty cards and shopper characteristics, *Journal of Retailing* 79(2): 97–106. http://dx.doi.org/10.1016/S0022-4359(03)00008-3

Martineau, P. 1958. Social classes and spending behaviour, *Journal of Marketing* 23(2): 121–130. http://dx.doi.org/10.2307/1247828

McGoldrick, P. J.; Andre, E. 1997. Consumer misbehaviour: promiscuity or loyalty in grocery shopping, *Journal of Retailing and Consumer Services* 4(2): 73–81. http://dx.doi.org/10.1016/S0969-6989(96)00027-6

Nielsen. 2005. *Anuario 2005. La Guía de los Mercados de Gran Consumo*. Madrid.

Pessemier, E. A. 1985. Varied individual behaviour: some theories, measurement and models, *Multivariate Behavioural Research* 20(1): 69–94. http://dx.doi.org/10.1207/s15327906mbr2001_5

Popkowski-Leszczyc, P. T. L.; Timmermans, H. J. P. 1997. Store switching behaviour, *Marketing Letters* 8(2): 193–204. http://dx.doi.org/10.1023/A:1007910503617

Rhee, H.; Bell, D. R. 2002. The inter-store mobility of supermarket shoppers, *Journal of Retailing* 78(4): 225–237. http://dx.doi.org/10.1016/S0022-4359(02)00099-4

Rossi, P. E.; Allenby, G. M. 2003. Bayesian statistics and marketing, *Marketing Science* 22(3): 304–328. http://dx.doi.org/10.1287/mksc.22.3.304.17739

Rossi, P. E.; Allenby, G. M.; McCulloch, R. 2005. *Bayesian Statistics and Marketing*. Chichester, UK: Wiley Series in Probability and Statistics. http://dx.doi.org/10.1002/0470863692

Schaninger, Ch. M. 1981. Social class versus income revisited: an empirical investigation, *Journal of Marketing Research* 18(2): 192–208. http://dx.doi.org/10.2307/3150954

Seetharaman, P. B.; Chintagunta, P. 1998. A model of inertia and variety-seeking with marketing variables, *International Journal of Research in Marketing* 15(1): 1–17. http://dx.doi.org/10.1016/S0167-8116(97)00015-3

Soberon-Ferrer, H.; Dardis, R. 1991. Determinants of household expenditures for services, *Journal of Consumer Research* 17(4): 385–397. http://dx.doi.org/10.1086/208565

Tanner, M. A.; Wong, W. 1987. The calculation of posterior distributions by data augmentation, *Journal of the American Statistical Association* 82: 528–549. http://dx.doi.org/10.1080/01621459.1987.10478458

Theil, H.; Finke, R. 1983. The consumer's demand for diversity, *European Economic Review* 23(3): 395–400. http://dx.doi.org/10.1016/0014-2921(83)90039-9

Van Trijp, H. C. M. 1995. *Variety Seeking in Product Choice Behaviour: Theory with Applications in the Food Domain*. *Mansholt Series 1*, *Wageningen*. The Netherlands: Wageningen University.

Van Trijp, H. C. M.; Steenkamp, J.-B. E. M. 1990. An investigation into the validity of measures for variation in consumption used in economics and marketing, *European Review of Agricultural Economics* 17(1): 19–41. http://dx.doi.org/10.1093/erae/17.1.19

Van Trijp, H. C. M.; Steenkamp, J.-B. E. M. 1992. Consumers' variety seeking tendency with respect to foods: measurement and managerial implications, *European Review of Agricultural Economics* 19(2): 181–195. http://dx.doi.org/10.1093/erae/19.2.181

Noemí MARTÍNEZ-CARABALLO holds a PhD in Business Administration and is Assistant Professor of Business Management at the Centro Universitario de la Defensa in Zaragoza. Her research focuses on consumer behavior and cross-cultural management, among others.

Manuel SALVADOR is Senior Professor at the Faculty of Economics and Business of the Zaragoza University. He received his PhD in Mathematics in 1987. His current research interests include Bayesian inference in general and dynamic models with applications to economics and business.

Carmen BERNÉ is Senior Professor in Marketing at the Faculty of Economics and Business of the Zaragoza University. She received his PhD in Economics in 1995. Her current research interests include consumer behavior and productivity analysis, among others.

Pilar GARGALLO is Senior Professor at the Faculty of Economics and Business of the Zaragoza University. She received her PhD in Mathematics in 2001. Her current research interests include Bayesian dynamic models with applications to economics and business.