A Proposal to Distinguish State Dependence and Unobserved Heterogeneity in Binary Brand Choice Models

by

Jose Mª Labeaga-Azcona
Mercedes Martos-Partal

DOCUMENTO DE TRABAJO 2007-02

February 2007

We would like to thank Nora Lado and James Nelson for their very helpful comments and ACNielsen Spain for providing the scanner panel data used in the analysis. This research was partially supported by Ministerio de Educación y Ciencia Dir. Gral de Investigación, Grants SEJ2004-00672 and SEJ2005-08793-CO4-04.

* FEDEA and UNED.
** Universidad de Salamanca.

Los Documentos de Trabajo se distribuyen gratuitamente a las Universidades e Instituciones de Investigación que lo solicitan. No obstante están disponibles en texto completo a través de Internet: http://www.fedea.es.

These Working Paper are distributed free of charge to University Department and other Research Centres. They are also available through Internet: http://www.fedea.es.
Abstract

This paper uses binary choice models that specify four possible sources of observed regularity in the consumer brand choice decision over purchase occasion: namely, state dependence, observed and unobserved heterogeneity and correlation effects. The objective is to distinguish correctly among the effects of these four variables. The estimation method proposed is an alternative to the most commonly used estimation methods in marketing choice models. We consider that the alternative method appropriately controls for observed heterogeneity and unobserved heterogeneity correlated with the state dependence variable because of the way the state dependence variable is built. The model is used for the first time in marketing following the methodology proposed by Chamberlain (1984). A relationship for unobserved heterogeneity is specified, taking into account the correlation among unobserved heterogeneity and other choice determinants. In this way, we split the influence of household state dependence and tastes on brand choice. The findings are very conclusive. We find that because the individual effects and the covariates are correlated, traditional estimation methods cannot be used to split state dependence and unobserved heterogeneity. The proposed model is found to yield better measures of predictive performance than the conventional model. The results are found to be robust across categories of laundry detergent and have significant implications for marketing policy.

Key Words: binary brand choice models, state dependence, unobserved heterogeneity, correlated effects, laundry detergent.
1. Introduction

Despite many advances in marketing brand choice models, the Guadagni and Little (1983) model still serves as the benchmark. In this, the key variable that allows the model to fit the choice probability accurately is the state dependence variable (sometimes called purchase feedback or loyalty). Since this seminal study, a rich literature on disaggregate brand choice models has emerged, supporting the existence of state dependence in the household’s brand choice decisions.

The influence of observed past experience (through actual purchase) with a brand on current choice probabilities is often referred to as structural state dependence (Heckman 1981). Put differently, if an identical household with no previously experienced event has different future behavior from a household with a previously experienced event, previous experience is the determinant for temporal unobserved persistence in brand choice (Hsiao 1986).

From a managerial perspective, strong state dependence effects imply a managerial incentive for inducing promotion (e.g., product sampling or price promotion). This is because if state dependence is present in household brand choice behavior, some households who bought the brand through price promotion or product sampling will be persuaded to stay with the brand after the promotion ends.

State dependence in brand choice can be influenced by chance in the market; that is, future event probabilities are disturbed by variables like price or promotions that change future brand choices. For example, low-income households can keep buying the cheapest choice. Therefore, state dependence can be explained as a consequence of environmental effects.

Persistence of brand choice can also be due to unobserved effects, but because these are present in the individual information set, they have a strong impact on event probability in the future. Unobserved heterogeneity refers to (residual) interindividual variations in purchase behavior, which cannot be explained by the observed brand choice experiences. These variations can be intrinsic preferences that consumers have for brands, or different ways in which consumers respond to marketing stimuli. Unobserved effects are also referred to as household tastes or household heterogeneity in preferences and prices. In contemporary households’ scanner panel data, household tastes are not available.

In this case, if the model of consumer behavior of the household includes only unobserved preference heterogeneity, from a managerial perspective, there
will be less incentive for inducing promotion than in the case of state dependence because when the promotion ends, the household will revert to the preferred brand.

State dependence, unobserved heterogeneity and environmental effects are three sources of regularity in households’ decisions over time. However, it can sometimes be difficult to split the relative importance of these effects on household decisions. When the data contain economic and sociodemographic variables, it is also possible to control for environmental effects. However, state dependence and unobserved heterogeneity are often undistinguished. In this way, the managerial conclusions inferred from wrongly identified parameters could be erroneous.

Following Heckman (1981), there is another source of possible persistence in brand choice. This is the influence of prior propensities to select a brand on current selection probabilities or habit persistence. These habitual purchase inclinations are manifested as serial correlations in the random component of the utility functions. If one does not control for this contingency, state dependence may have an inordinately large impact on future choices. This is because previous experience is functioning as a proxy for serial correlation (in the utility-maximizing alternatives) that influences current choices.

It is important to distinguish among the effects of unobserved heterogeneity, state dependence and correlation effects. If heterogeneity is present in the true model (because households included in household scanner panel data are heterogeneous) and we ignore heterogeneity by fitting a model with only permit state dependence, then the state dependence parameter will be overestimated (Heckman 1981). Accordingly, we will incorrectly conclude state dependence (spurious state dependence). If unobserved heterogeneity and state dependence are correlated, and we do not correctly control for correlation in the empirical analysis, then previous experience is the only determinant in future experience because previous experience is a proxy for temporal unobserved persistence. Less problematic is the opposite situation, where if state dependence is present in the true model, and we ignore state dependence fitting a model which only permits heterogeneity, then the value of heterogeneity in the population will be overestimated\(^1\).

Few studies have generally controlled for the four notions of temporal dependence in models of dynamic brand choice behavior: state dependence, environmental effects and unobserved heterogeneity and correlation effects.

\(^1\) In both cases, we assume that the correlation between state dependence and heterogeneity is positive. While this is common in applied terms, theoretically the correlation can also be negative.
This is because previous works have shown that correlation effects matter little in frequently purchased categories after heterogeneity and state dependence have been properly accounted for (Roy et al. 1996; Keane 1997; Seetharaman 2004). In fact, only Erdem and Sun (2001) test whether individual effects are uncorrelated with the covariates. As covariates, they include the price and display of the bought brand. They find individual effects are uncorrelated with the covariates. The outcome of this body of work is that marketing researchers appear to have little interest in building brand choice models that control for serial correlation in the residuals.

On the other hand, it is common in the economic literature to find evidence of correlation between individual effects and covariates. For example, Jones and Labeaga (2003) model current household consumption of tobacco as a function of the current real price of tobacco, some relevant sociodemographic variables of the household and past tobacco consumption. They also allow for the possibility of unobserved individual effects that are correlated with past tobacco consumption (reflecting the individual’s propensity to be addicted). In this case, sociodemographic variables are included as covariates. However, it is uncommon in the marketing literature to include this kind of variable as an explanation for brand choice.

We suggest that in the way in which the state dependence variable is built, the state dependence effect and unobserved heterogeneity are correlated. Abramson et al. (2000, p. 424) argue that the underspecification of serial correlation has serious consequences for the parameters in that the presence of serial correlation caused the loyalty coefficient to inflate. These differing views among econometricians and marketing researchers have brought about the need to undertake research on the ways in which serial correlation in the residuals in brand choice models can be controlled to avoid biased parameters.

In this analysis, we use a richer specification model where economic and sociodemographic variables are included to help control for the environmental effects. We test whether individual effects are correlated with these covariates, and we find that the individual effects are indeed correlated. The novelty of the analysis is that we allow correlation by introducing some assumptions to distinguish between state dependence and unobserved heterogeneity effects.

However, the major difficulty to estimate isolatable state dependence and unobserved preference heterogeneity is due to the fact that dependent variable is a latent variable. In the random utility model, the reality is that we observe the event (a household chooses or does not choose a brand), but the choice decision is made after a brand utility comparison among alternatives. Because the decision is to choose a single brand, we only know the chosen brand (the brand
with higher utility). This is very useful in estimating models that assume continuous distribution (logistic or normal) errors, but at the same time, we cannot transform above unobserved preferences. There is not currently model transformation that can remove unobserved components.

In this paper, we do not only attempt to test whether brand choice regularity is also found in Spain but also provide a step forward by proposing a way to control for state dependence, observed heterogeneity (environmental effects) and unobserved heterogeneity. Unobserved heterogeneity correlated with the state dependence variable and other variables is included in the model specification. We follow the methodology proposed by Chamberlain (1984) that can be applied to linear and nonlinear models; in particular, brand choice models. Using this methodology, if state dependence and individual effects are correlated, then a functional relationship can be established where the unobserved effects are split into unobserved and state dependence effects on brand choice. In this way, it is possible to specify a reduced form for the model. The reduced form parameters can be consistently estimated in each time period and, then it is possible to derive, in a second stage, the ‘parameters of interest’ to obtain the correct prediction of the unobserved latent variable. By controlling for unobserved heterogeneity, we obtain structural form parameters such that the state dependence variable becomes less significant.

The aim of this paper is very clear. Starting with a simple binary choice model with state dependence, and in order to account for the heterogeneity of the brand choice process, we propose estimation in some different contexts proposed in the literature. We test for state dependence and unobserved heterogeneity in the consumer brand choice model. We then introduce to the marketing literature a new parameter estimation technique for heterogeneous logit models. We assume that unobserved heterogeneity depends on observed characteristics in a simple way; thus, we try to distinguish between the influences that unobserved heterogeneity and state dependence have on the probability of choosing a brand. To obtain robust results, we use two different product categories. We intuitively test all the model results and formally establish findings about the influence on choice of state dependence, unobserved heterogeneity, environmental and correlation effects. We compare the estimation results, the fit and the predictive ability of the traditional maximum likelihood estimation procedure with the estimation technique put forward here and we find better predictive performance in the method proposed.

The remainder of the paper is organized in four sections. In Section 2, we provide a brief literature review. In Section 3, we specify the proposed model and other nested models specified in the literature. We discuss the data used to estimate the models and the results in Section 4. Finally, in Section 5, we
present the findings and our conclusions and suggest some managerial implications.

2. Literature Review

The control of the previous choice experience can be accomplished in a number of different ways in brand choice models. The key variable that allows the Guadagni and Little (1983) model to fit data accurately is the loyalty variable, which is an exponential smoothing of past purchases. The loyalty variable confounds two effects: state dependence and household heterogeneity. Heterogeneity refers to the differences across households in brand preferences or market responses, and state dependence refers to the impact of past purchases on current preferences.

The Guadagni and Little measure of brand loyalty is able to track the differences in purchase behavior across consumers and over time, but it cannot properly distinguish among sources of variation in utility from heterogeneity (across households) and sources of variation because of nonstationarity (within households over time). Many studies show that the Guadagni and Little loyalty variable does not sufficiently capture consumer heterogeneity (e.g., Ortmeyer et al. 1991; Fader and Lattin 1993). As pointed out by Lattin (1987), by using a single loyalty term, one implicitly assumes that differences across consumers and differences over time contribute equally to the heterogeneity in the base level utility. If such an assumption is inappropriate, it could have a distorting effect on the choice model. To avoid this problem, other measures have been proposed that split the cross-sectional and longitudinal effects. Jones and Landwehr (1988) proposed a discrete choice model that split heterogeneity and state dependence (by measuring the latest purchasing behavior). Modified versions of the Guadagni and Little loyalty variable have also emerged (Krishnamurthi and Raj 1988; Ortmeyer et al. 1991; Erdem 1996 and Keane 1997).

To avoid the critique usually applied to the loyalty variable in Guadagni and Little (1983), we model both sources of the variation in utility separately; that is, we separate the variation associated with nonstationarity by using the latest purchasing behavior and account for heterogeneity by specifying household sociodemographic characteristics and controlling for any unobserved preferences and response heterogeneity.

Unobserved heterogeneity can be included in model specification in a number of different forms (fixed and random effects) and with different
assumptions (random effects correlated with environmental effects, correlated with the other choice’s determinants or without correlation). Approaches to estimating the parameters for unobserved heterogeneity in models of consumer brand choice behavior can be grouped into two broad classes: (1) those that estimate parameters for each household; a fixed effects model (e.g., Jones and Landwehr 1988) and the hierarchical Bayesian approach (e.g., Rossi and Allenby 1993); and (2) those that assume that household parameters are distributed according to a probability distribution and estimate the parameters of that distribution (random effects models). One approach is the finite mixture or latent class model approach, which captures heterogeneity across households in the form of discrete support points (e.g., Kamakura and Russell 1989; Chintagunta et al. 1991; Bucklin and Gupta 1992). The second approach is the continuous mixture approach, which models heterogeneity in the form of continuous mixture distributions (e.g., Gönül and Srinivasan 1993; Erdem 1996; Keane 1997). There has been some debate in the literature as to whether the finite mixture or continuous approach is better. However, research by Andrews et al. (2002) suggests that both approaches are equally good at parameter recovery and predictive validity and that “… whether an analyst prefers to use models with continuous or discrete representations of consumer heterogeneity is a matter of opinion and personal preference”.

In order to account for heterogeneity in the brand choice process, Jones and Landwehr (1988) introduced an extension of a technique developed by Chamberlain (1984). Chamberlain proposed a conditional maximum likelihood estimation technique. In this technique, household-specific parameters are taken out of the likelihood expression by conditioning on their sufficient statistics. One disadvantage of using this technique is that it can only explain the choice behavior of households that change brands, not households that are absolutely brand loyal or never choose the brand. This estimation technique can only explain the variety brand choice behavior. In the conditional heterogeneous model, they assumed that households in the sample that always purchase the brand or never purchase the brand will continue this pattern of behavior for all purchases. Under this assumption, purchases do not add to the likelihood function and so are dropped from the data set in the conditional estimation procedure. Thus, the estimation sample contains only households that ‘switch’ brands. Fixed logit models have rarely been used. One alternative is to assume some probability distribution for the intercept and slope terms using the random effect or hierarchical Bayesian approaches. These approaches eliminate some of the undesirable assumptions of the fixed effects model.

State dependence, unobserved heterogeneity and environmental effects are three sources of regularity on the household’s decisions over time. The state dependence variable usually is based on the household’s previous brand choice;
therefore, unobserved heterogeneity is correlated with the state dependence. Researchers have tried to split the relative importance of these three effects on household decisions. They have often found evidence for true state dependence in the choice process, even after controlling for a rich heterogeneity structure (e.g., Keane 1997; Ailawadi et al. 1999; Varki and Chintagunta 2004). However, there are sufficient empirical tests to support the introduction of observed heterogeneity in the model limits the explanatory power of the state dependence variable in brand choice decision models. It is also clear that controlling for unobserved heterogeneity reduce the explanatory capability of the state dependence variable (Keane 1997). All of these generalizations in brand choice models seek to remove spurious state dependence.

In sum, recent work on state dependence and heterogeneity in the context of disaggregate panel data of consumer brand choices has found evidence of true state dependence in the choice process. On the other hand, we also find empirical evidence of a zero-order choice process in models that use aggregate data. For example, Bass and Wind (1995) argued that zero-order consumer brand-choice behavior is an empirical generalization, and Uncles et al. (1995) concluded that applications of the Dirichlet model have provided strong evidence for zero-order choice process.

To build empirical evidence, we test for state dependence in consumer brand choice. First, we apply a simple test suggested by Chamberlain (1978) to distinguish state dependence from heterogeneity and serial correlation. This approach was subsequently applied in marketing by Erdem and Sun (2001), who found strong evidence of state dependence. Chamberlain’s test consists of including lagged exogenous variables (but not lagged choices) in the utility specification while allowing for unobserved heterogeneity in taste and response parameter. This is because Chamberlain argued that the key distinction between heterogeneity and state dependence is the dynamic response to the exogenous variables. If true state dependence is present, lagged exogenous variables affect current choices because they affect lagged choices. However, if true state dependence is not present, lagged exogenous variables cannot affect current choices. Thus, a test for whether lagged exogenous variables are significant determinants of current choices is also a test for state dependence.

Nonetheless, there are three main drawbacks of Chamberlain’s test. First, the test cannot be used to make further distinctions with regard to state dependence, heterogeneity, and serial correlation. Second, Chamberlain’s test depends on the assumption of individual effects being uncorrelated with the covariates if a random-effects specification is used to implement the test (or the assumption that such correlation is correctly modeled). Third, Chamberlain’s test as a ‘test of state dependence’ requires at least one exogenous variable that
would not have a lag structure in the absence of state dependence. Chamberlain (1984) proposed a new methodology to avoid these drawbacks.

We follow the methodology proposed by Chamberlain (1984) that has never before been applied to marketing. Using this methodology, if state dependence and unobserved heterogeneity are correlated, then a functional relationship can be established on the unobserved effects to separate them from the state dependence effects on brand choice. This methodology is not only a test of state dependence, but also a way to split state dependence and unobserved heterogeneity. Another advantage is that the variables included need not be strictly exogenous.

We use a random effects specification and estimate the structural parameters using the GMM procedure. Working in this way, we avoid a disadvantage of the conditional maximum likelihood technique in that it can only explain the choice behavior of households that change brands, not households that are absolutely loyal or households that never choose the brand. The technique used in the current analysis can explain any type of brand choice behavior, including variety and no-variety behavior.

3. Model Specification and Estimation Methods

3.1 General model

We specify a binary choice model that includes state dependence, environmental effects and unobserved heterogeneity in preferences and the price response. We express the model in the following equations:

\[
y^*_i t = \beta_i p_{it} + \lambda y^*_{it-1} + \tilde{\varepsilon}_i + v_{it} \\
v_{it} = \alpha_i + \mu_i \\
y^*_i t = \alpha_i + \beta_i p_{it} + \lambda y^*_{it-1} + \tilde{\varepsilon}_i + \mu_i \\
y^*_{it} = \{1 \text{ if } y^*_{it} > 0; 0 \text{ whether } y^*_{it} \leq 0\} \\
y^*_{it-1} = \alpha_i + \beta_i p_{it-1} + \lambda y^*_{it-2} + \tilde{\varepsilon}_i + \mu_{it-1}
\]

Where \( y^*_i t \) is the utility of the brand choice for consumer \( i (i = 1, \ldots, N) \) at purchase occasion \( t (t = 1, \ldots, T_i) \) in the product category, \( \beta_i \) is the price parameter for consumer \( i \), and \( p_{it} \) is the price of the chosen alternative for consumer \( i \) on purchase occasion \( t \). In the literature, this is referred to as ‘unobserved price heterogeneity’ because it contains effects unobserved to the researcher. We assume that is a time-invariant effect, \( z_i \) are socio-demographic variables specific to household \( i \) which do not exhibit time variation and \( \delta \) is the
vector of parameters associated with them. Information about \( z_i \) allows knowledge of the households’ observed heterogeneity, \( \lambda \) is the state dependence parameter, \( y_{it-1} \), and is a lagged term of the dependent variable that incorporates purchase-event feedback. \( y_{it-1} \) denotes an indicator about consumer \( i \)’s previous brand choice behavior on purchase occasion \( t-1 \), and \( \alpha_i \) is the household-specific intercept term that characterizes the differences in brand preferences among consumers and is time invariant. In the literature, this is referred to as ‘unobserved preference heterogeneity’ because it also contains effects unobserved by the researcher. The specific household preference effect is obtained when we decompose the random error \( \nu_i \) following equation (2) into a household preference effect (\( \alpha_i \)) and mixed error (\( \mu_i \)) that changes both across time and households.

We further assume that the mixed error components (\( \mu_i \)) are independent and identically Weibull distributed. Thus, the conditional probability of the chosen alternative at time \( t \) by consumer \( i \) is given by the binary logit model.

\[
P(y_{it} = 1) = \frac{\exp(\alpha_i + \beta_i p_{it} + \lambda y_{it-1} + \xi_i)}{1 + \exp(\alpha_i + \beta_i p_{it} + \lambda y_{it-1} + \xi_i)}
\]

and

\[
P(y_{it} = 0) = \frac{1}{1 + \exp(\alpha_i + \beta_i p_{it} + \lambda y_{it-1} + \xi_i)}
\]

In general, we must assume some observability rule linking observed and latent components. The rule is contained in the second part of equation (3), and it establishes the relationship between \( y^*_it \) (the utility of choosing the brand) and \( y_{it} \) (the true choice realization), namely, the variable takes a value of one if the consumer \( i \) chooses the alternative on purchase occasion \( t \). The difficulties in separating the state dependence effect and the unobserved preference effects have been proven sufficiently in the literature, and we reiterate the problem in this paper. However, the contribution of this paper is that we estimate the state dependence effect on brand choice when we control for the unobserved heterogeneity in the correct way. The proposed model is based on the specified general model; we also consider another nested model in the general model that we use to test the proposed model. We follow a sequential estimation process with the aim of differentiating each effect as a determinant on brand choice, and at the same time, we obtain an intuitive test for each group of variables.
3.2 Proposed model

In nonlinear models, like those applied to brand choice behavior, it is not possible to use transformation to first differences or orthogonal deviations to yield estimators of the \( \lambda \)s that are asymptotically independent of the household-specific effects and hence, consistent for all the parameters. A conditional likelihood approach can be followed in order to sweep out the fixed effects (Chamberlain 1984, pp. 1274–1278). However, the dynamic specification of equation (3) adds an additional difficulty in that the presence of lags of the latent endogenous variable induces correlation between this regressor and the effects. This is especially difficult to control for in nonlinear models.

There are some circumstances where \( \lambda \) could not be identified. First, \( \alpha_i \) could be correlated with \( y_{it-1} \), \( E[\alpha_i, y_{it-1}] \neq 0 \) such as shown in equation (4), in a way such as \( y_{it-1} \) is at best predetermined for the mixed error but correlated with the time invariant part of the error. In this situation, \( \lambda \) cannot be separately identified from the correlation effect.

Second, it is possible that in different situations at the household level, the state dependence parameter cannot be separated from the household-specific effect: (i) in case of absolute loyalty, state dependence will be a constant for the household, as it is the household-specific effect, and therefore we cannot separate both effects; (ii) when the alternative is never chosen, it implies the absence of loyalty, then the state dependence for the household takes the value zero for all purchase occasions. Once again, we will have two constants and cannot separate the effects of both choice determinants.

One of the advantages of using panel data is the possibility of accounting for the correlation among the effects and the explanatory variables. Chamberlain’s (1984) suggestion of using a random effects approach and specifying a distribution for the effects conditional on the exogenous variables can be applied to (3).

We follow Chamberlain in assuming:

\[
E(\alpha_i / X_1) = \sum_{r=0}^{r} \pi_r X_{1it} + \pi_r R_{it} + \omega_i \tag{6}
\]

where \( X_{1it} \) are considered exogenous variables and \( R_{it} \) contains nonlinear terms and interactions in \( X_{1it} \). We have to choose the instrument set carefully. Natural choices for instruments are past prices as well as demographic variables not included in the household decision set. However, we do not use past prices to
avoid potential problems with multicollinearity. We expect these instruments to be highly correlated with the unobserved effects and uncorrelated with the mixed error$^2$.

If we substitute (6) in (1), we obtain the reduced form of the general model in the following equation:

$$y^*_i = \pi W_i + \epsilon_i, \quad i = 1, \ldots, N \tag{7}$$

where we can derive the parameters of interest ($\beta_i$ and $\lambda$) from the non-linear relationship between them and the reduced form parameters $\pi$ in (7). Chamberlain (1984) suggested that the parameters in the reduced form (7) could be estimated by fitting a model for each time period where we have information$^3$. We suppose the unobserved effects to be time invariant, but this hypothesis is not very strong because the period used in the current analysis is short.

After having the estimator $\hat{\pi}_t$ ($t = 1, \ldots, T$), we can build predictions for the latent variables, $\hat{y}^*_t = \hat{\pi}_t z$, and derive ‘parameters of interest’ ($\beta_i$ and $\lambda$) in a second step, using the following equation.

$$\hat{y}^*_i = \alpha_i + \beta_i p_u + \lambda y_{it-1} + \delta \epsilon_i + u_{it} \tag{8}$$

Finally, we derive the relevant vector of parameters by applying GMM to (8).

In the discussion in this section, we have shown two assumptions concerning the state dependence effect on brand choice. The first assumption implies that it is the realization of the dependent variable in the previous period, $y_{it-1}$, what influences the choice probability in the current period. In this way, state dependence is in the information set of the analyst in period $t$ and as a result it is an observed variable. The alternative assumption implies that the choice brand probability in period $t$ could be influenced by the previous choice probability; namely, the variable that influences the probability in $t$ is $y_{it-1}^*$. We can postulate it because of missing information to the analyst or because the

---

$^2$ We include in $X_{1it}$ the age of the main buyer of the household, age of the household’s wife, number of children of different ages and number of adults, dummies for sex, dummies indicating the geographic location of the household, and dummies indicating the population size of the geographic location of the household and household expenditure on the category during the analyzed period. We include in $R_{it}$ the age squared and age cubed of the main buyer of the household, the age squared of the household’s wife, and interactions between the sociodemographic variables.

$^3$ In the empirical application, we detail how the time period is defined.
current choice probability is also influenced by the probability of choosing the same alternative in \( t-1 \). If this were the case in (1), we would have \( y^*_{it-1} \) instead of \( y_{it-1} \), and as consequence, in (8) we would also replace the latent explanatory variable by the predicted values using the reduced form predictions for \( \hat{y}^*_{it-1} \).

In both cases, this approach assumes that the conditional expected value of the household preference effect is linear and independent between the variables in \( W \) and the disturbance \( \varepsilon_i \). We assume that the mixed error \( u_{it} \) in (8) does not present autocorrelation of any order. Using these assumptions, we derive the parameters of the structural form by any method that allows for the control of the household preference unobserved effects in equation (8). Chamberlain (1984) proposed deriving parameters using a minimum distance process; but we can also use alternative methods, as least squares or instrumental variables (on the first differences of the model).

Estimators of (8) are sensitive to assumptions about the distribution of \( \varepsilon_i \), the linearity of the expected value (6) and the conditional mean independence assumption implied by (8). However, these hypotheses can be checked by specification tests at the level of the reduced form. In order to estimate the model we make use of the fact that the distribution of \( y^*_{it} \) conditional on the explanatory variables but marginal to the effects, is of the same form as the joint distribution (Chamberlain 1984, section 3.1). This allows us to estimate the reduced form (7) using discrete choice models for each \( t \) and then combining predictions as shown in (8).

### 3.3 Nested models

We compare the performance of the proposed model with the traditional brand choice model proposed by the literature. We compare our model with three models nested in the general one. These models are frequently used in the marketing literature.

**Model 1**

We consider that choice can be influenced by prices, state dependence and sociodemographic variables. We keep the assumption that unobserved household preference and price effects do not exist; therefore \( \alpha_i = \alpha \) and \( \beta_i = \beta \). This assumption implies that all the households have the same preferences toward the brand and the same response to prices. We can rewrite (5) as follows.
This model is very similar in nature to the model in Guadagni and Little (1983). One critique is that under the presence of unobserved effects, the state dependence will be overstated if the heterogeneity is positively correlated with the choice determinants.

**Model 2**

We remove the assumption that unobserved preference and price effects are not present and consider that unobserved effects exist and they follow a random form specification. We keep equation (5).

\[
P(y_{it} = 1) = \frac{\exp(\alpha + \beta p_{it} + \lambda y_{it-1} + \delta \epsilon_i)}{1 + \exp(\alpha + \beta p_{it} + \lambda y_{it-1} + \delta \epsilon_i)}
\]

and

\[
P(y_{it} = 0) = \frac{1}{1 + \exp(\alpha + \beta p_{it} + \lambda y_{it-1} + \delta \epsilon_i)}
\]  

(5')

The specification is estimated using the latent class model (Kamakura and Russell 1989; Wedel and Kamakura 1998). **Model 1** and **Model 2** can be estimated using traditional unconditional maximum likelihood estimation procedures.

**Model 3**

We specify **Model 3** only to allow us to test for state dependence. We suppose price as the lagged covariate to be included in the model. If there is state dependence, a consumer is more likely to buy a brand on the current purchase occasion if they bought it on a previous occasion. Then, even though the lagged price has no effect on the current purchase decision, state dependence induces a negative correlation between the lagged price and the current purchase probability for a brand (Erdem and Sun 2001). Chamberlain’s (1978) approach entails testing for such correlation. We can rewrite in this case (5) as follows.

\[
P(y_{it} = 1) = \frac{\exp(\alpha + \beta_1 p_{it-1} + \delta \epsilon_i)}{1 + \exp(\alpha + \beta_1 p_{it-1} + \delta \epsilon_i)}
\]
and
\[ P(y_{it} = 0) = \frac{1}{1 + \exp(\alpha_i + \beta_i p_{it-1} + \delta_i)} \] (5'')

This last specification is also estimated using the latent class model. We must take into account that Chamberlain’s test depends on the assumption of absence of correlation between the individual effects and the covariates. Therefore, if correlation is present, the test is incorrect.

One difficulty with these nested models is that we cannot transform (1) to remove the unobserved household effects. In fact, transformation is only possible in model (8) where we replace the latent variable using its predicted counterpart. The problem is that in maximum likelihood estimation, it is not possible to split the ‘parameters of interest’ from the unobserved household effects. Because we have few observations for each household, it is not possible to estimate the unobserved household effects consistently, and these inconsistencies are translated to the parameter of interest. This is called the *incidental parameter problem* (for more details, see Chamberlain 1984 or Bover and Arellano 1988). Conversely, these models do not assume correlated effects.

4. Estimation and Empirical Tests

In this section, we present the data and variables. We also present our results for the models specified above. In all cases, we use a binary discrete choice model, in particular the logit model, as a simple way to split state dependence and unobserved heterogeneity. We could also use models with normal errors, but the results are very similar. The third part of this section is used to compare the results.

4.1 Brand choice, data and variables

We define the dependent variable of the logit choice model by taking into account the store brand market share in Spain. Store brands, also known as private labels or retail brands, have enjoyed increased success in recent years. Europe shows a traditional dominance in terms of the market share of store brands. Spain is among the top five markets (ACNielsen 2005) where the market share of store brands reached 26% in 2005, and according to the ACNielsen forecast, sales will increase at double the rate of national brands in 2006. Laundry detergents and non-food groceries is one of the categories in which store brands have had more success in Spain, enjoying a market share approaching 30%. Therefore, the choice indicator takes a value of one when
household $i$ purchases a store brand (SB) on occasion $t$ and zero otherwise (national brands, NBs).

We estimate our models using scanner panel data, supplied by ACNielsen Spain, on household purchases in two product categories: fine laundry detergent and non-fine laundry detergent. SBs are present in both categories. The Spanish data set includes a representative sample of households across the country, rather than households in specific cities. Their purchase activities are recorded from January 1999 to December 2000. We have 1,107 households accounting for 5,347 purchases of fine laundry detergent, and 1,557 households accounting for 33,246 purchases of non-fine detergent. To obtain robust results, we estimate the models in these two detergent categories because store brand market shares differ.

Households without information during the estimation period or during the prediction period were dropped from the analysis. We use the first-year period to estimate the models, while the second-year period is used to predict the choice market share. In the fine laundry detergent category, we have 622 households with 4,366 purchase occasions (2,172 in the estimation period and 2,194 in the prediction period) and 1,499 households with 30,050 purchase occasions for non-fine laundry detergent (15,402 in the estimation period and 14,648 in the prediction period).

To analyze the relative importance of state dependence compared with other variables in the model, we include some explanatory variables describing brands and consumers. The variables included in the model specifications are grouped into the following three categories.

**Purchase-occasion-specific variable:**

‘Price$_{it}$’ is the price/weight of the bought brand on purchase occasion $t$ of household $i$. Weight is equal to one kilogram and price corresponds to the sales price$^4$.

**State dependence variable:**

‘Last$_{it}$’ is a dummy variable that reflects the relative impact of recent choice behavior, measured by whether household $i$ purchased the brand on occasion $t-1$ (Jones and Landwehr 1988, variable of purchase-event feedback).

**Household-specific sociodemographic variables:**

---

$^4$ We only have shelf prices for the bought brands. Information about features and displays is not available; nor do we have data about the size discounts associated with price promotion.
‘Sizei’ is a count variable ranging from 1 to 5 representing the size of household i, where value 5 identifies households with five or more members.

‘Social class 1i’ is a dummy variable that takes a value of one for high and high-medium class households and zero otherwise.

‘Social class 2i’ is a dummy variable that takes a value of one for medium class households and zero otherwise.

‘Social class 3i’ is a dummy variable that takes a value of one for medium-low and low class households and zero otherwise.

‘Workeri’ is a dummy variable, where a value of one identifies households with a working housewife.

Table A.1 of the Appendix provides descriptive statistics for the dependent and explanatory variables for the two categories of detergent. When we compare the fine and the non-fine laundry detergent markets, we find that the former had a larger store brand market share, and the price gap between store brands and national brands is large. Households choose store brands on 37 percent of purchase occasions for fine laundry detergent and on 19 percent of purchase occasions for non-fine laundry detergent. The fine laundry detergent category is a less-frequently bought category.

Market concentration is higher in the fine laundry detergent category. In this market there are 61 bought brands: 23 SBs and 38 NBs. However, in the non-fine laundry detergent category, there are 116 bought brands: 39 SBs and 77 NBs.\(^5\)

### 4.2 Estimation model, results and tests

We estimate and discuss the models specified in Section 3. The estimation results are shown in Table 1 for the case of fine laundry detergent and in Table 2 for the case of non-fine laundry detergent. All models are estimated in the first-year period of purchase occasion by the household in both categories, and we keep the second-year period as the prediction sample.

To estimate the proposed model, we need to define a period where we can get a consistent estimator. We defined as the time unit the month of purchase for the non-fine laundry detergent and the quarter of purchase for the fine laundry detergent because fine laundry detergent is bought less frequently. We could

---

\(^5\) According to Dhar and Hoch (1997) and Ailawadi and Keller (2004), SBs appear to enjoy a higher share in large, less-promoted categories with high market concentration when the price gaps between national brands and store brands are large. We found evidence of this in the differing success of SBs in the Spanish laundry detergent market.
choose other time units, but our choice is based on the interpurchase time in each category\(^6\).

As remarked in Sections 1 and 3, the major problem in estimating isolatable state dependence and unobserved preference heterogeneity is that the dependent variable is a latent variable. Therefore, the model does not accept any transformation that can remove the unobserved components, which are very useful when using scanner or panel data. For this reason, after we specify a relationship between the individual preference effects and observed exogenous variables, such as that described in equation (6); we fit a binary logit model in each time period and obtain the predictions. At this point, it is still possible at the second stage to derive the structural parameter using methods where transformations are possible. In the second stage, the estimation method is random errors GMM because these methods allow us to model the random household effects from the model specification. In this way, coefficient estimates, including the corresponding to the state dependence variable, will be bias free.

We assume that there is the same heterogeneity structure in price and in preferences as in Model 2. We also adjust the probability of membership to each particular segment following Kamakura and Russell (1989) and Gupta and Chintagunta (1994); using an assignment rule such as ‘membership in the segment with highest probability’, it is possible to assign households uniquely to segments with differential preference and price sensitivity.

We analyze the results shown in Tables 1 and 2. We make comparisons of the t-statistics from the variables in Model 1 and Model 2 in each category. Because the t-statistic for the last variable is large, we conclude that state dependence is an important choice determinant for these two models in both categories. The results for Model 1 show evidence that in the brand choice decision the household’s previous choice (state dependence) increases the probability of choosing the same brand. In this model, state dependence is the most important brand choice determinant. However, if unobserved effects exist and are positively correlated with choice determinants, the effect of state dependence is overestimated. We check whether there is any unobserved heterogeneity. We use a latent class model with heterogeneity in preferences and prices. We use the four information criteria suggested by Elrod and Keane (1995) to choose the optimal latent class model. They are the Akaike Information Criterion (AIC), the Hannan–Quinn (HQ) Criterion, the Bayesian

\(^6\) In the full time period of 727 days under study, the median interpurchase occasions of fine laundry detergent is 8 purchases and 26 purchases of non-fine laundry detergent; therefore, we can approximate a time unit of one month for non-fine laundry detergent and one quarter for fine laundry detergent.
Information Criterion (BIC), and the Consistent Akaike Information Criterion (CAIC). The information criteria are computed as $AIC = -2\log L + 2K$, $HQ = -2\log L + 2K \ln(\ln(N))$, $BIC = -2\log L + K\ln(N)$ and $CAIC = -2\log L + K(\ln(N)+1)$, where $\log L$ is the value of the log-likelihood function for each model, $K$ is the number of parameters estimated, and $N$ is the sample size. We prefer those models with higher values of the log-likelihood and smaller values of AIC, HQ, BIC and CAIC. The values of $\log L$, $K$, $N$, AIC, HQ, BIC and CAIC for the optimal model in each category are reported as measures of fit in Tables 1 and 2.
Table 1. Estimation results for the logit model in the fine laundry detergent category

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1-step)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2-step)</td>
</tr>
<tr>
<td>Last</td>
<td>2.71 (23.39)***</td>
<td>1.16 (5.93)***</td>
<td>0.58 (2.88)***</td>
<td>57.39 (–70.50)***</td>
</tr>
<tr>
<td>Price</td>
<td>–0.73 (–4.12)***</td>
<td>–77.11 (–6.01)***</td>
<td>–5.39 (–70.50)***</td>
<td>–2.51 (–19.38)***</td>
</tr>
<tr>
<td></td>
<td>5.68 (5.73)***</td>
<td>–3.90 (–6.94)***</td>
<td>–1.96 (–33.26)***</td>
<td>–0.44 (0.79)</td>
</tr>
<tr>
<td>Lagged price</td>
<td>–1.61 (–3.67)***</td>
<td>–1.26 (–1.28)*</td>
<td>–1.02 (–3.15)***</td>
<td>–0.12 (–0.63)</td>
</tr>
<tr>
<td>Worker</td>
<td>–0.26 (–2.31)***</td>
<td>–0.38 (–1.72)***</td>
<td>–0.10 (–3.15)***</td>
<td>–0.12 (–0.63)</td>
</tr>
<tr>
<td>Size</td>
<td>–0.04 (–0.95)</td>
<td>0.18 (1.82)**</td>
<td>0.28 (2.25)**</td>
<td>0.04 (0.43)</td>
</tr>
<tr>
<td>Social Class 1</td>
<td>0.10 (0.64)</td>
<td>–0.61 (–1.83)**</td>
<td>0.38 (0.92)</td>
<td>1.18 (3.84)****</td>
</tr>
<tr>
<td>Social Class 2</td>
<td>0.06 (0.52)</td>
<td>–0.16 (–0.57)</td>
<td>0.83 (2.53)***</td>
<td>0.69 (2.70)*****</td>
</tr>
<tr>
<td>Dummy of segments 1</td>
<td>56.01 (6.01)***</td>
<td>–7.51 (–5.94)***</td>
<td>3.12 (2.49)***</td>
<td>–1.08 (–3.15)****</td>
</tr>
<tr>
<td>Dummy of segments 2</td>
<td>4.04 (4.87)***</td>
<td>–5.50 (–5.66)***</td>
<td>6.67 (13.55)***</td>
<td>–0.41 (–1.32)</td>
</tr>
<tr>
<td>Constant</td>
<td>–0.42 (–1.48)*</td>
<td>0.76 (1.10)</td>
<td>3.92 (2.29)**</td>
<td>6.67 (13.55)***</td>
</tr>
</tbody>
</table>

MEASURES OF FIT

| Log likelihood       | –1,087  | –857  | –662  | –400  |
| # of purchases       | 2,172   | 2,172 | 1,550 | 2,172 |
| # of segments(s)     | S = 1   | S = 1 | S = 1 | S = 1 |
|                      | S1 = 0.46 | S1 = 0.32 | S1 = 0.46 |        |
|                      | S2 = 0.18 | S2 = 0.16 | S2 = 0.18 |        |
|                      | S3 = 0.36 | S3 = 0.52 | S3 = 0.36 |        |
| # of parameters      | 7       | 13    | 12    | 70    |
| AIC                  | 2,188   | 1,740 | 1,348 | 1,360 |
| HQ                   | 2,203   | 1,767 | 1,372 | 1,942 |
| BIC                  | 2,282   | 1,914 | 1,500 | 5,103 |
| CAIC                 | 2,237   | 1,829 | 1,426 | 3,233 |
| Pseudo R²            | 0.70; 0.71; | 0.73; 0.70; |        |        |

Notes:
1. Models:
   Model 1. Logit model with state dependence, observed heterogeneity and without unobserved heterogeneity.
   Model 2. Latent class logit model with state dependence and observed heterogeneity.
   Model 3. Latent class logit model without state dependence and with observed heterogeneity and lagged price.
   Proposed Model. Panel model with state dependence and observed heterogeneity controlling unobserved heterogeneity using random effect. Model with two stages.
2. Table entries are the value coefficient with the t-statistic in parentheses, *p < 0.10, **p < 0.05 and ***p < 0.01.
3. Akaike Information Criterion (AIC = –2LogL + 2K); Hannan–Quinn (HQ = –2LogL + 2Kln(Ln(N))); Bayesian Information Criterion (BIC = –2LogL + Kln(N)); and Consistent Akaike Information Criterion (CAIC = –2LogL + K(ln(N) + 1)). Here, LogL is the value of the log-likelihood function for each model, K is the number of parameters estimated, and N is the sample size.
4. Pseudo-R² = 1 – (LogL/ LogLc), where LogL is the value of the log-likelihood function at the optimum for the complete model and LogLc is the value of the log-likelihood function at the optimum for a restricted model with only a constant as the explanatory variable.
5. Variables used in the first step are the age of the main buyer of the household, age of the wife in the household, number of children of different ages and number of adults, dummies for sex, dummies indicating the geographic location of the household and dummies indicating the population size of the geographic location of the household, and household expenditure on the category during the analyzed period. Age squared and age cubed of the main buyer of the household, age squared of the household wife, and interactions between the sociodemographic variables.
Table 2. Estimation results for the logit model in the non-fine laundry detergent category.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1-step)</td>
<td>(2-step)</td>
<td>(1-step)</td>
<td>(2-step)</td>
</tr>
<tr>
<td>Last</td>
<td>2.95 (59.91)***</td>
<td>1.50 (19.40)***</td>
<td>-3.70 (138.4)***</td>
<td>0.34 (6.68)***</td>
</tr>
<tr>
<td>Price</td>
<td>-1.15 (10.75)***</td>
<td>-33.47 (-19.65)***</td>
<td>-1.89 (62.50)***</td>
<td>-3.70 (-47.85)***</td>
</tr>
<tr>
<td>Lagged price</td>
<td>-0.68 (-1.94)***</td>
<td>-0.33 (-2.04)***</td>
<td>0.65 (0.56)</td>
<td>0.08 (0.13)</td>
</tr>
<tr>
<td>Worker</td>
<td>-0.16 (-3.02)***</td>
<td>0.07 (0.77)</td>
<td>-0.09 (-0.53)</td>
<td>-0.08 (-1.21)</td>
</tr>
<tr>
<td>Size</td>
<td>-0.04 (-1.68)**</td>
<td>-0.03 (-0.81)</td>
<td>-0.01 (-0.16)</td>
<td>-0.14 (-4.49)***</td>
</tr>
<tr>
<td>Social Class 1</td>
<td>-0.22 (-2.80)***</td>
<td>-0.44 (-2.98)***</td>
<td>-0.75 (-2.57)***</td>
<td>-0.36 (-3.36)***</td>
</tr>
<tr>
<td>Social Class 2</td>
<td>-0.06 (-0.94)</td>
<td>-0.40 (-3.44)***</td>
<td>-0.40 (-1.74)**</td>
<td>0.04 (0.49)</td>
</tr>
<tr>
<td>Dummy of segments 1</td>
<td>-0.66 (-4.46)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummy of segments 2</td>
<td>-0.19 (-2.02)**</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MEASURES OF FIT

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log likelihood</td>
<td>-5,340</td>
<td>-3,614</td>
<td>-4,119</td>
<td>-4,254</td>
</tr>
<tr>
<td># of purchases</td>
<td>15,402</td>
<td>15,402</td>
<td>13,903</td>
<td>15,402</td>
</tr>
<tr>
<td># of segments(s)</td>
<td>S = 1</td>
<td>S = 0.62</td>
<td>S = 0.22</td>
<td>S = 0.62</td>
</tr>
<tr>
<td></td>
<td>S = 0.26</td>
<td>S = 0.12</td>
<td>S = 0.60</td>
<td>S = 0.12</td>
</tr>
<tr>
<td># of parameters</td>
<td>7</td>
<td>13</td>
<td>15</td>
<td>70</td>
</tr>
<tr>
<td>AIC</td>
<td>10,694</td>
<td>7,254</td>
<td>8,268</td>
<td>10,188</td>
</tr>
<tr>
<td>HQ</td>
<td>10,712</td>
<td>7,287</td>
<td>8,306</td>
<td>12,315</td>
</tr>
<tr>
<td>BIC</td>
<td>10,815</td>
<td>7,478</td>
<td>8,524</td>
<td>24,707</td>
</tr>
<tr>
<td>CAIC</td>
<td>10,756</td>
<td>7,368</td>
<td>8,398</td>
<td>17,449</td>
</tr>
</tbody>
</table>

Pseudo $R^2$ = 0.36; 0.33; 0.38; 0.39; 0.39; 0.42; 0.43; 0.50; 0.43; 0.42; 0.41; 0.41

Notes:

1. Models:
   - Model 1. Logit model with state dependence, observed heterogeneity and without unobserved heterogeneity.
   - Model 2. Latent class logit model with state dependence and observed heterogeneity.
   - Model 3. Latent class logit model without state dependence and with observed heterogeneity and lagged price.
   - Proposed Model. Panel model with state dependence and observed heterogeneity controlling unobserved heterogeneity using random effect. Model with two stages.

2. Table entries are the value coefficient with the t-statistic in parentheses, *p < 0.10, **p < 0.05 and ***p < 0.01.

3. Akaike Information Criterion (AIC = –2LogL + 2K); Hannan–Quinn (HQ = –2LogL + 2ln(Ln(N))); Bayesian Information Criterion (BIC = –2LogL + Kln(N(N))); and Consistent Akaike Information Criterion (CAIC = –2LogL + K(ln(N) + 1)). Here, LogL is the value of the log-likelihood function for each model, K is the number of parameters estimated, and N is the sample size.

4. Pseudo-$R^2 = 1 – (LogL/ LogL_{C}),$ where LogL is the value of the log-likelihood function at the optimum for the complete model and LogL_{C} is the value of the log-likelihood function at the optimum for a restricted model with only a constant as the explanatory variable.

5. Variables used in the first step are the age of the main buyer of the household, age of the wife in the household, number of children of different ages and number of adults, dummies for sex, dummies indicating the geographic location of the household and dummies indicating the population size of the geographic location of the household, and household expenditure on the category during the analyzed period. Age squared and age cubed of the main buyer of the household, age squared of the household wife, and interactions between the sociodemographic variables.
The results for Model 2 show that there is unobserved heterogeneity in preferences and prices. Model 1 is a particular case of Model 2 where the number of segments is equal to one. If we compare the measure of fit between both models, we choose Model 2 in both categories because they have higher values of log-likelihood and smaller values of AIC, HQ, BIC and CAIC.

The control of unobserved effects is important for explaining brand choice behavior. For example, in the fine laundry detergent category, we find three consumer segments. In these segments, the state dependence variable, the price and the constant have similar t-statistics; this implies that state dependence and unobserved heterogeneity have a similar impact on brand choice. The state dependence variable losses some capacity to explain choice behavior once we control for unobserved heterogeneity in prices and preferences. The probability to choose today the same brand that in the previous choice is smaller when the probability to choose the brand is conditioned on unobserved heterogeneity. We get similar results in the non-fine laundry detergent category.

We find signs of overstated state dependence in Model 1 not only because the explanatory capability of the last variable is smaller in Model 2 but also because its size is smaller. With this focus, we specified Model 3 with the aim of distinguishing between true and spurious state dependence.

Therefore, Model 3 constitutes a simple test for distinguishing state dependence from heterogeneity and serial correlation. Chamberlain (1978) noted that the key distinction between heterogeneity and state dependence is the dynamic response to exogenous variables. The results indicate that the lagged price coefficients are statistically significant in segments one and two in both categories. Chamberlain’s test results suggest that there are intertemporal dependencies in the deterministic part of the utilities, and hence, choice dynamics. However, this test cannot split state dependence from heterogeneity because of the assumption of absence of correlation between the individual effect and the covariates. We then need to use another method to split state dependence and unobserved heterogeneity correctly in the case that correlation effects exist.

We must not forget in any case that: (i) the results from infraspecified models (without unobserved effects in Model 1) produce bias in the estimates; (ii) the results from the correctly specified model but with random unobserved heterogeneity (Model 2) can also produce bias when the state dependence variable is correlated with the individual effect.

To solve the correlation problem, we show the results from the two-stage estimation process in the proposed model. The determinants of unobserved
heterogeneity in the reduced form model in the first estimation stage are: age of the main buyer in the household, age of the wife in the household, number of children of different ages and the number of adults, dummies for sex, dummies indicating the geographic location of the household and dummies indicating the population size of the geographic location of the household, and household expenditure on the category during the analyzed period. We assume that these variables are strictly exogenous to identify model parameters. Our aim is to use a rich specification to obtain a good model fit. We include all the sociodemographic determinants, along with the squares of the variables, and variable interactions. We need a good model fit because we will use the latent variable prediction as the dependent variable in the second stage. In this way, we avoid unobserved latent variables in the model specification. We fit a logit model in each time period, and the first-stage pseudo-$R^2$ are not less than 0.70 (fine laundry detergent) and 0.33 (non-fine laundry detergent), representing a logit regression of the utility of the choice of brand on all exogenous variables.

We model a reduced logit model, where it is assumed that all the coefficients will be the same during the estimation period. If correlation is not present, then the coefficients will be the same in the reduced model as in the proposed model. If the coefficients are different, we can use this result as a sign of the correlation between the individual preference effects and the covariates. We test whether individual preference effects are uncorrelated with the covariates in the first stage. To do this, we run a Wald test between the model that we propose and the reduced logit model. The Wald statistic is equal to 0.4976. $\chi^2$ for 210 restrictions and significance at the 0.01 level is 156.432 for fine laundry detergent, and the statistic is 0.1947. The $\chi^2$ for 770 restrictions and significance at the 0.01 level is major than 156.432 for non-fine laundry detergent. We cannot reject the hypothesis of different coefficients. Therefore, we interpret these results as driven by correlated effects. We then use a method like that proposed to split state dependence and unobserved heterogeneity when correlation effects are presents to obtain bias-free parameters.

In the second stage of the proposed model, we obtain the parameter of interest. The state dependence variable is less significant than in the other nested models, and there are other brand choice determinants that have more important explanatory power in brand choice than the last variable. With respect to the remainder variables in the proposed model, we point out that price is the most significant variable. Brand intrinsic preference also has an important effect. In the fine laundry detergent category, we find three segments with different characteristics. The biggest is Segment 1. It is very price sensible and has no preference for buying SBs. Segment 2 is the second segment in size and is the lowest in price sensitivity because it has a preference for buying SBs that are cheaper. Segment 3 is the smallest segment and shows median price sensitivity,
and the preference for SBs is similar to that for NBs. In the non-fine laundry detergent category, we find similar results. We find also three segments. The biggest one is Segment 1. It is the most price sensible segment and has no preference for buying SBs. Segment 2 is the second segment in size and is less price sensitive than Segment 1, but neither prefers to buy SBs. Only Segment 3 show high preferences toward SBs. This is the smallest segment and displays the lowest price sensitivity. These results imply that after controlling for unobserved effects in a correct way, the marketing-mix variables have a greater influence on households’ brand choice than state dependence.

To summarize, we can say that at the same time as model specification becomes richer, the state dependence influence on brand choice becomes less significant. In our opinion, these results clarify the way to correctly identify true state against spurious state dependence.

We cannot compare the performance of the proposed model with Model 2 using measures of fit because the former is estimated using maximum likelihood and we use generalized method of moments (GMM). However, in the following section, we compare the predictive ability of the traditional maximum likelihood estimation procedure and the estimation procedure proposed.

4.3 Predictive model performance

A final stage in the model selection procedure is the evaluation of the forecasting performance of one or more selected models. We may consider an ‘out of sample’ interpolation to evaluate the forecasting power of Model 2 because it is the richer traditional maximum likelihood specified model with the proposed model.

One indicator of this forecasting power is the root mean squared error (RMSE) in predictive choice market share. We assume that, on each choice occasion, the brand with the highest predicted choice probability is purchased. The predicted choice market share of each brand is then obtained by aggregating predicted choices across all purchase occasions.

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( \tilde{P}(y_i) - \hat{P}(y_i) \right)^2} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (e_i)^2}
\]

Another indicator of forecasting power, allowing comparison of the performance of different models, is the percentage of correct predictions for each model. This follows directly from a prediction–realization table, where the
value \( \sum_{j=1}^{2} p_{jj} \) can be interpreted as the hit rate. \( P_{jj} \) is the diagonal value in the prediction–realization table, the proportion of times that alternative \( j \) is correctly predicted. Based on simulation experiments, Veal and Zimmermann (1992) recommend the use of the measure suggested by McFadden, Puig and Krischner, which is given by \( F1 = \frac{\sum_{j=1}^{2} p_{jj} - p_{jj}^2}{1 - \sum_{j=1}^{2} p_{jj}^2} \), where \( p_{jj} \) is the proportion of times alternative \( j \) is predicted.

The model with the lower value of RMSE and highest value of the hit rate and F1 may be viewed as the model that has the best forecasting performance. Measures of RMSE, the hit rate and F1 are reported in Table 3.

### Table 3. Forecasting results for the out-of-sample in the fine and non-fine laundry detergent category

<table>
<thead>
<tr>
<th></th>
<th>RMSE(^1)</th>
<th>Hit rate(^2)</th>
<th>F1(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fine laundry detergent category</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.3789</td>
<td>0.8564</td>
<td>0.6941</td>
</tr>
<tr>
<td>Proposed Model</td>
<td>0.3415</td>
<td>0.8833</td>
<td>0.7619</td>
</tr>
<tr>
<td><strong>Non-fine laundry detergent category</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.5032</td>
<td>0.7467</td>
<td>0.4198</td>
</tr>
<tr>
<td>Proposed Model</td>
<td>0.3899</td>
<td>0.8479</td>
<td>0.2849</td>
</tr>
</tbody>
</table>

Notes:
- Model 2. Latent class logit model with state dependence and observed heterogeneity.
- Proposed Model. Panel model with state dependence and observed heterogeneity controlling unobserved heterogeneity using random effect. Model with two stages.
- 1. Root mean squared error (RMSE) in predictive market share. We assume that, on each choice occasion, the brand with the highest predicted choice probability is purchased. Predicted market share of each brand is then obtained by aggregating predicted choices across all purchase occasions (N).

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( P(y_i) - \hat{P}(y_i) \right)^2} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} e_i^2}
\]

2. These follow directly from a prediction–realization table, where the value \( \sum_{j=1}^{2} p_{jj} \) can be interpreted as the hit rate.

3. Measure suggested by McFadden, Puig and Krischner, which is given by \( F1 = \frac{\sum_{j=1}^{2} p_{jj}^2}{1 - \sum_{j=1}^{2} p_{jj}^2} \).
The proposed model performs better out of sample predictions in both categories. We find that the proposed model could be then a correct model for splitting state dependence and unobserved effects, and at the same time it could be a good alternative to make out of sample prediction after estimating the specifications.

5. Conclusions and Implications

The paper’s main aim is to distinguish state dependence from unobserved heterogeneity in the estimation context of a discrete choice model in a binary setting.

The results show evidence of true and spurious state dependence in the choice process when we control for observed heterogeneity and unobserved heterogeneity by methods commonly used in the brand choice literature (Keane 1997). Empirical tests have showed as to control unobserved heterogeneity limit the state dependence explanatory power in both laundry detergent categories used to fit the model.

Advance to split unobserved preference heterogeneity and state dependence do not fully avoid the problem of biased parameters (Abramson et al. 2000); biased parameters would imply wrong policies application by the marketing manager. We find that biased parameters can probably due to the correlation between state dependence and unobserved preference heterogeneity, and the correlation appears to be due to the way in which the state dependence variable is built.

We find that when using traditional maximum likelihood approaches, the individual effects are correlated with the covariates. This novelty of this analysis is that formalizing the correlation structure by imposing some assumptions, it is possible to consider this correlation during the estimation process. It allows derivation of a reduced form from which we obtain reduced form predictions that allow structural parameters to be derived in a simple way when individual effects are uncorrelated with the covariates.

On the other hand, because we have a limited number of observations for each household in consumer panel data, it is not possible to estimate consistently the unobserved household effects using maximum likelihood procedures, and the inconsistent unobserved household effects is translated to the parameter of interest. With the estimation procedure proposed here, we are able obtain consistent estimators.
When we control for unobserved effects in the way proposed by Chamberlain (1984), state dependence becomes a less significant variable. From this result, we conclude that when households make the brand choice decision by conditioning on the environmental effects, observed characteristics and unobserved characteristics without state dependence have an important influence on choice. This result could be the reason for finding mixed evidence about the order of consumer brand choice behavior in disaggregate and aggregate choice models. Because the influence of state dependence is small in a disaggregate setting, this effect could disappear when we model it in an aggregate context and evidence of zero-order brand choice process can be found.

From a managerial perspective, we conclude that there are some variables in the household’s information set that are not available in our sample, and these variables are more important than state dependence for explaining choice behavior. These unavailable variables can cause inertia. For example, the household can choose the most accessible brand to avoid wasting time looking for a cheaper brand because the time wasted cannot compensate for the reduction in price from the process of comparing brands or the lack of information concerning differences among brands. It is possible that the cost of accessing all household information is greater than the benefits in the case of the detergent category. In any case, these variables are individual specific and time invariant and have an important role in explaining brand choice behavior. Therefore, greater household knowledge is necessary for understanding households’ brand choice decisions.

With this comment, we would not like to say that the information contained in household panel data is not sufficient, rather the opposite. Household panel data generally provide richer information about sociodemographics. This can be very useful for predicting market shares by using the correct model that controls for the unobserved variables that are taken into account by the household but that are not available to the researcher. On the other hand, when we model in this way, we can obtain very significant information on the price sensitivity of the consumer segments. With this information, we can apply differentiated strategies to each consumer segment.

Retailers may think about their strategic decisions to hold a store brand portfolio. Store brands sell at a lower price than national brands with good quality levels, so the retailer could launch different store brands positioned with different prices aimed at different customer segments—for example, premium quality store brands that are not priced lower than national brands and are targeted at less price-sensitive consumers.
For example, if a retailer would like to increase the market share of his SB in the categories analyzed in the segment with the biggest size, one way would be to increase preferences for the SB. In both categories, the consumer in these segments does not buy the SB, and a small increase in the price has a big impact in the SB’s market share in this segment. Because SBs are the cheapest alternative, this result leads us to think that this segment is more concerned about quality than price.

As in any research, this investigation has certain limitations that must be considered. First, the proposed model is estimated using two frequently purchased categories. We have used categories where the SB’s market shares are very different and where market conditions and consumer behavior are different. To check the findings and provide more robust results, it would be interesting to apply the model to high-involvement categories and see if the results are maintained. Second, instead of binary choice options, we could further extend the model in future research to the multinomial case to model the level of brand alternatives.
### Appendix.

#### Table A.1 Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Fine laundry detergent category</th>
<th>Non-fine laundry detergent category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SBs</td>
<td>NBs</td>
</tr>
<tr>
<td>Dependent variable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purchase occasions</td>
<td>1,638</td>
<td>2,728</td>
</tr>
<tr>
<td></td>
<td>(37.52%)</td>
<td>(62.48%)</td>
</tr>
<tr>
<td>Explanatory variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price (€)</td>
<td>0.89</td>
<td>2.51</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(1.16)</td>
</tr>
<tr>
<td>Last</td>
<td>0.68</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>Size</td>
<td>3.67</td>
<td>3.62</td>
</tr>
<tr>
<td></td>
<td>(0.92)</td>
<td>(1.04)</td>
</tr>
<tr>
<td>Social class1</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.40)</td>
</tr>
<tr>
<td>Social class2</td>
<td>0.63</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.48)</td>
</tr>
<tr>
<td>Social class3</td>
<td>0.16</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.38)</td>
</tr>
<tr>
<td>Worker</td>
<td>0.25</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.43)</td>
<td>(0.47)</td>
</tr>
<tr>
<td>Number of brands</td>
<td>23</td>
<td>38</td>
</tr>
</tbody>
</table>

Notes:
1. Store brands (SBs) and national brands (NBs)
2. Table entries in the first line is the mean value with the standard deviation in parentheses.
References


