Analyzing the Effects of Past Prices on Reference Price Formation

Rutger van Oest*
Tinbergen Institute
Erasmus University Rotterdam

Richard Paap
Econometric Institute
Erasmus University Rotterdam

Econometric Institute Report EI 2004-36

Abstract

We propose a new reference price framework for brand choice. In this framework, we employ a Markov-switching process with an absorbing state to model unobserved price recall of households. Reference prices result from the prices households are able to remember. Our model can be used to learn how many prices observed in the past are used for reference price formation. Furthermore, we learn to what extent households have sufficient price knowledge to form an internal reference price. For A.C. Nielsen scanner panel data on catsup purchases, we find that the prices observed at the previous purchase occasion have an average recall probability of about 20%. Furthermore, the average probability that a household has sufficient price knowledge to form a reference price is estimated at about 30%. Even though price recall is very limited the impact of reference price formation on brand choice is substantial, and it is stronger than two popular alternative models in the literature suggest. Moreover, contrary to the two alternative models, our model does not suggest asymmetry between price gains and losses.

Keywords

Reference price, brand choice, Markov-switching process, household scanner panel data.

JEL Classification

C25, C51, C53, M31

*Corresponding author: Tinbergen Institute, H16-32, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam (e-mail vanoest@few.eur.nl, tel. +31-10-4088946, FAX +31-10-4089162).

The results in the paper have been obtained using Ox v3.30 (Doornik 1998).

1 Introduction and motivation

The literature provides ample evidence that households do not only consider current prices when deciding which brand to buy from a category, but also take into account past prices via the formation of internal reference points, see Krishnamurthi, Mazumdar and Raj (1992), Putler (1992), Kalyanaram and Little (1994), among many others. A conceptual basis for such reference price formation is provided by the Adaptation-Level Theory of Helson (1964). Kalyanaram and Winer (1995) translate the cumulative evidence into an empirical generalization.

The idea behind reference price effects is as follows. If the price of a brand is below its reference price, the observed price is lower than anticipated, resulting in a perceived gain. This would make the brand more attractive. Similarly, the opposite situation would result in a perceived loss, reducing the probability that the brand is purchased. An important consequence of internal reference price formation is that although frequent price discounts may be beneficial in the short-run, they may damage the brand in the long-run when households get used to these discounts and reference prices drop. The reduced price become anticipated and loses its effectiveness, whereas the non-promoted price becomes unanticipated and would be perceived as a loss.

As the existence of internal reference prices might imply a tradeoff between current and future brand sales, it is important to have a good understanding of how reference prices are formed from past price experience. This includes questions such as (i) what is the size of the reference effect?, (ii) what is the duration of the reference effect?, and (iii) how often are households able to construct a reference price? Still, current brand choice models dealing with reference price effects do not pay much attention to the operationalization of the reference price construct.

In this paper, we focus on internal reference prices, which are constructed from past prices, and we do not consider external reference prices, based on the in-store prices during the purchase occasion. The two most popular specifications for internal reference price are the price observed at the previous purchase occasion and an exponentially smoothed average of previously observed prices. Examples of the former reference price include Krishnamurthi, Mazumdar and Raj (1992), Mayhew and Winer (1992) and Chang, Siddarth and Weinberg (1999), whereas examples of the latter specification include Lattin and Bucklin (1989), Kalyanaram and Little (1994) and Erdem, Mayhew and Sun (2001). By assuming that the reference price of a brand equals its previous price, a reference effect of only one period is imposed, while multi-period dynamics may be more appropriate. An argument which is often made in favor of one-period dynamics is that the literature consistently shows that households have a very limited ability to recall prices, see, for example, the price knowledge surveys by Dickson and Sawyer (1990) and Vanhuele and Drèze (2002). However, this argument would imply that households might not even be able to recall the previously observed price. On the other hand, the exponentially-smoothed-average measure for reference price often results in extremely long price dynamics. For example, Briesch, Krisnamurthi, Mazumdar and Raj (1997) conclude for several product categories that the lag in the formation of reference price is about six periods.

Two limitations of current reference price specifications are that (i) the dynamics underlying the reference price are determined a priori (possibly up to a tuning parameter), and (ii) it is implicitly assumed that households always have sufficient price knowledge to form a reference price. Both might result in underestimation of the size of the actual reference effect. Briesch, Krisnamurthi, Mazumdar and Raj (1997) note that a misspecified reference price model can obscure the reference effect even when it may actually exist. Similarly, by not accounting for the possibility that households forget past prices, one might confound households not *forming* a reference price with households not *reacting* to it, resulting in lower response estimates. A third limitation of current practice is that the assumed processes usually do not account for uncertainty in reference price formation. However, as reference prices are not observed, their existence cannot be inferred without error.

In this paper we propose a new reference price model, which (i) does not a priori impose a rigid dynamic price structure, (ii) accounts for uncertainty in reference price formation, (iii) allows that a household may forget past prices, and (iv) even allows that a household cannot construct a reference price at all. Hence, the model is an attempt to bridge the gap between survey studies, such as Dickson and Sawyer (1990) and Vanhuele and Drèze (2002) which report that households have very limited price knowledge and hence may not be able to construct a reference price, and current reference price models which do not account for this. In the proposed framework, price recall of households is modeled as a hidden Markov-switching process with an absorbing state, and a reference price is constructed from the prices the household is able to recall. Our model can be used to get insights into the extent to which past price information is used in individual brand choice decisions, that is, it can address the question "which past prices do consumers use in forming a reference price, and how many are used?", raised by Kalyanaram and Winer (1995).

The structure of the paper is as follows. In Section 2, we put forward our model. Next, in Section 3, we discuss parameter estimation, and in Section 4, we explain how the estimation results can be used to analyze the dynamic structure underlying reference price formation and to analyze the extent to which internal reference prices are formed. We apply our model to an A.C. Nielsen scanner panel data set on catsup purchases in Section 5. Finally, in Section 6, we conclude with a discussion of the implications of our model and we provide some directions for further research.

2 The model

In this section we develop our reference price model and we show how it can be incorporated in a scanner panel data model for brand choice.

2.1 Reference price

We allow that households may forget prices observed in the past, as suggested by several price knowledge surveys in the literature. To this end, we introduce the unobserved 0/1 price recall variable $S_{i,t}^{\tau}$ such that

$$S_{i,t}^{\tau} = \begin{cases} 1 & \text{if at time } t \text{ household } i \text{ remembers the prices encountered at time } \tau \leq t \\ 0 & \text{otherwise,} \end{cases}$$

where the time indices t and τ correspond to purchase occasions. As $S_{i,t}^{\tau}$ does not contain a brand-index, price recall is the same across brands. Hence, a household is either able to recall the prices of all brands or cannot recall any brand price. To keep implementation of the model feasible, we impose that $S_{i,t}^{\tau} = 0$ if $t - \tau > L$, that is, we assume that households always forget prices which were observed more than L purchase occasions ago. The value of L has to be chosen by the researcher. As it does not affect the number of parameters in the model, a possible selection strategy would be to set $L \in \{1, 2, \ldots\}$ such that it gives the highest maximum likelihood value. The price knowledge of household i at purchase occasion t can be summarized in one *composite* price recall variable $\tilde{S}_{i,t} = (S_{i,t}^{t-L}, \ldots, S_{i,t}^{t-2}, S_{i,t}^{t-1})$, which indicates which of the past L prices are recalled at purchase occasion t.

The second step in the development of our model concerns the description of how price recall of households evolves over time. We assume that, starting from purchase occasion τ when prices were observed, the subsequent price recall variables $\{S_{i,t}^{\tau}\}_{t=\tau,\tau+1,\ldots,\tau+L}$ obey a first-order Markov process with states 0 and 1. So, the probability that prices are recalled at the current purchase occasion depends on whether these prices were still in memory at the previous purchase occasion. Moreover, as it is plausible that prices are never recalled again once they have been forgotten, we let the "forgetting state" 0 be an absorbing state. The Markov transition probabilities are given by

$$\Pr(S_{i,t}^{\tau} = 0 | S_{i,t-1}^{\tau} = 0) = 1, \tag{1}$$

$$\Pr(S_{i,t}^{\tau} = 1 | S_{i,t-1}^{\tau} = 0) = 0, \tag{2}$$

$$\Pr(S_{i,t}^{\tau} = 0 | S_{i,t-1}^{\tau} = 1) = 1 - p_{i,t}^{\tau}, \tag{3}$$

$$\Pr(S_{i,t}^{\tau} = 1 | S_{i,t-1}^{\tau} = 1) = p_{i,t}^{\tau}. \tag{4}$$

The process is initialized by setting $S_{i,\tau}^{\tau} = 1$, meaning that households are aware of the prices of brands at the moment of purchase. The memory processes defined by

(1)−(4) are independent across households i as well as independent across purchase occasions τ during which prices were observed.

We define the transition probability $p_{i,t}^{\tau}$ as

$$p_{i,t}^{\tau} = \frac{1}{1 + \exp(-[\,\gamma_0 + \gamma_1(t-\tau)\,])}\,. \tag{5}$$

This conditional probability of price recall depends on the number of purchase occasions $t-\tau$ which have passed since the prices were observed at purchase occasion τ. Hence, the *unconditional* probability that the prices at purchase occasion τ are still remembered at purchase occasion t is given by

$$\Pr(S_{i,t}^{\tau} = 1) = \prod_{\tilde{\tau}=\tau+1}^{t} p_{i,\tilde{\tau}}^{\tau} = \prod_{\tilde{\tau}=1}^{t-\tau} \frac{1}{1 + \exp(-[\,\gamma_0 + \gamma_1\tilde{\tau}\,])}\,, \tag{6}$$

which is a decreasing function in $t-\tau$. Figure 1 displays some price memory patterns which can be reproduced by tuning the parameters γ_0 and γ_1 in (5). It shows a geometrically decaying pattern, a 1-period full memory pattern which amounts to the previously observed brand price being the reference price, a 2-periods full memory pattern, and a pattern such that prices are sometimes recalled at the next purchase occasion after which they are kept in memory forever. The graph illustrates that our model is able to mimic a wide variety of possible memory structures. Finally, we note that it is straightforward to extend (5) such that the conditional price recall probability $p_{i,t}^{\tau}$ also depends on, for example, the number of days elapsed since the previous purchase occasion $t-1$ and the degree of promotional activity at purchase occasion τ when the prices were observed.

Insert Figure 1 *about here.*

In our model, a reference price is only based on the prices which a household is able to recall, where price recall develops according to a first-order Markov process with absorbing state. We define the reference price $R_{i,j,t}$ of household $i \in \{1,\ldots,N\}$ for brand $j \in \{1,\ldots,J\}$ at purchase occasion $t \in \{1,\ldots,T_i\}$ as

$$R_{i,j,t} = \begin{cases} \frac{\sum_{\tau=t-L}^{t-1} S_{i,t}^{\tau} P_{i,j,\tau}}{\sum_{\tau=t-L}^{t-1} S_{i,t}^{\tau}} & \text{if } \sum_{\tau=t-L}^{t-1} S_{i,t}^{\tau} > 0 \\ P_{i,j,t} & \text{if } \sum_{\tau=t-L}^{t-1} S_{i,t}^{\tau} = 0. \end{cases} \tag{7}$$

Hence, this reference price equals the average of the prices which are recalled provided that at least one price observed in the past is still available in memory. However, it might also be possible that a household has forgotten all past prices. In such a case, the household does not have enough price information to form an internal reference price, and it perceives neither a price gain nor a loss. This is captured by the condition $P_{i,j,t} - R_{i,j,t} = 0$ in case $\sum_{\tau=t-L}^{t-1} S_{i,t}^{\tau} = 0$. The proposed reference price is an unweighted average from the perspective of the household, reflecting that households may have forgotten the order in which the recalled prices were observed. From the perspective of the researcher, it is however a weighted average. The weights are the unobserved 0/1 price recall variables for which the probability distributions have to be inferred from the data. Note that the weights differ across purchase occasions and households.

2.2 Brand choice

Our reference price can be incorporated in a standard model for brand choice. In this paper, we consider a conditional logit model which (i) allows for asymmetric response of households to price gains and losses, and (ii) accounts for heterogeneity across households via the latent segments approach of Kamakura and Russell (1989). By doing so, we follow the mainstream literature on modeling brand choice and reference price effects, see, for example, Bell and Lattin (2000).

Prospect Theory, developed by Kahneman and Tversky (1979), predicts that households react more negatively to price losses than they react positively to price gains of equal size. Indeed, many reference price studies support this hypothesis, see Mayhew and Winer (1992), Putler (1992), Kalyanaram and Little (1994), and Erdem, Mayhew and Sun (2001), among others. Kalyanaram and Winer (1995) propose this gain-loss asymmetry as an empirical generalization. However, the literature also contains some studies which report either a lack of significance or even an opposite effect. Such examples include Briesch, Krishnamurthi, Mazumdar and Raj (1997), Chang, Siddarth and Weinberg (1999), and Bell and Lattin (2000). Besides allowing for separate response parameters for price gains and losses, it is also important to account for household heterogeneity. Several studies show that ignoring heterogeneity might result in biased response estimates, see, for example, Chintagunta, Jain and Vilcassim (1991), Jain, Vilcassim and Chintagunta (1994) and Chang, Siddarth and Weinberg (1999).

We assume that household i with unobserved (composite) memory state $\tilde{s}_{i,t} = (s_{i,t}^{t-L}, \ldots, s_{i,t}^{t-2}, s_{i,t}^{t-1})$, belonging to the unobserved response segment $z \in \{1, \ldots, Z\}$, perceives utility

$$\begin{aligned} U_{i,j,t|z,\tilde{s}_{i,t}} = \alpha_{j|z} &+ \beta_{1|z} BL_{i,j,t} + \beta_{2|z} PM_{i,j,t} + \beta_{3|z} P_{i,j,t} \\ &+ \beta_{4|z} G_{i,j,t|z,\tilde{s}_{i,t}} + \beta_{5|z} L_{i,j,t|z,\tilde{s}_{i,t}} + \varepsilon_{i,j,t} \end{aligned} \tag{8}$$

from purchasing brand j at purchase occasion t, where the random disturbance $\varepsilon_{i,j,t}$ is assumed to be independently and identically distributed obeying a Type-I Extreme Value distribution. In (8), $\alpha_{j|z}$ captures the intrinsic brand preferences of the household, $BL_{i,j,t}$ is the brand loyalty measure of Guadagni and Little (1983), $PM_{i,j,t}$ is a 0/1 promotion indicator (feature or display), $P_{i,j,t}$ is the shelf price, $G_{i,j,t|z,\tilde{s}_{i,t}}$ is the price gain relative to the reference price, and $L_{i,j,t|z,\tilde{s}_{i,t}}$ is the price loss. The gain and loss variables are defined as

$$G_{i,j,t|z,\tilde{s}_{i,t}} = I\{P_{i,j,t} < R_{i,j,t|z,\tilde{s}_{i,t}}\}(R_{i,j,t|z,\tilde{s}_{i,t}} - P_{i,j,t}), \tag{9}$$

$$L_{i,j,t|z,\tilde{s}_{i,t}} = I\{P_{i,j,t} > R_{i,j,t|z,\tilde{s}_{i,t}}\}(P_{i,j,t} - R_{i,j,t|z,\tilde{s}_{i,t}}), \tag{10}$$

where $I\{\cdot\}$ is the 0/1 indicator function. We note that the reference price $R_{i,j,t|z,\tilde{s}_{i,t}}$ is conditional on both the price recall state $\tilde{s}_{i,t}$ and segment membership z, as we allow the price recall parameters γ_0 and γ_1 in (5) to be different across the latent segments.

As suggested by the mental accounting framework of Thaler (1985), total utility $U_{i,j,t|z,\tilde{s}_{i,t}}$ derived from purchasing brand j consists of two components, that is, acquisition utility and transaction utility. The acquisition component corresponds to the "monetary value" of the deal. It is related to the discrepancy between the value of the brand for the household and the brand price. Additionally, the transaction component corresponds to the "psychological value" of the deal, which is determined by the discrepancy between the brand price and the corresponding reference price. In (8), transaction utility is captured by the price gain and price loss components. Brand choice models which do not consider reference effects, such as the model of Guadagni and Little (1983), assume acquisition utility but do not account for transaction utility. In his seminal paper on reference price modeling, Winer (1986) refers to the discrepancy between the observed price and the reference price as a sticker shock effect.

It immediately follows from the distributional assumptions on $\varepsilon_{i,j,t}$ that brand j is chosen, providing maximum utility, with probability

$$\Pr(B_{i,t} = j | Z_i = z, \tilde{S}_{i,t} = \tilde{s}_{i,t}) = \frac{\exp(U_{i,j,t|z,\tilde{s}_{i,t}})}{\sum_{k=1}^{J} \exp(U_{i,k,t|z,\tilde{s}_{i,t}})}, \tag{11}$$

where Z_i describes the unobserved segment membership variable for household i, see McFadden (1974).

3 Parameter estimation

In this section we discuss how the parameters of our model can be estimated. Parameter estimation is not straightforward, as the model contains two kinds of unobserved variables, that is, segment membership and the price recall states. The price recall variables obey a first-order Markov process, and hence are not independent. We use the EM algorithm of Dempster, Laird and Rubin (1977) to deal with the unobserved response segments, see McLachlan and Krishnan (1997) for a textbook discussion. Mixture models, such as our model, provide a natural application area for this algorithm, see Wedel, DeSarbo, Bult and Ramaswamy (1993), Ramaswamy, Anderson and DeSarbo (1994) and Böckenholt (1999), among others. Within the EM algorithm, we apply an iterative filter, put forward in Hamilton (1989) and also described in Hamilton (1994, p. 692−693), to sum out the unobserved price memory states $\tilde{s}_{i,t}$.

Using the shorthand notation $B_{i,1:\tilde{t}}$ to denote the sequence of brand choices

$B_{i,1}, B_{i,2}, \ldots, B_{i,\tilde{t}}$, the (unconditional) likelihood function is given by

$$\begin{aligned}\mathcal{L} &= \prod_{i=1}^{N} \Pr(B_{i,1:T_i} = b_{i,1:T_i}) \\ &= \prod_{i=1}^{N} \sum_{z=1}^{Z} \pi_z \Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = z),\end{aligned} \tag{12}$$

where $b_{i,1:T_i}$ corresponds to realized brand choice and $\pi_z = \Pr(Z_i = z)$ is the size of segment z.

3.1 Applying the Hamilton filter

In order to evaluate the likelihood function (12), one needs to compute

$$\begin{aligned}&\Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = z) \\ &= \sum_{\tilde{s}_{i,1}} \cdots \sum_{\tilde{s}_{i,T_i}} \Pr(B_{i,1:T_i} = b_{i,1:T_i}, \tilde{S}_{i,1} = \tilde{s}_{i,1}, \ldots, \tilde{S}_{i,T_i} = \tilde{s}_{i,T_i} | Z_i = z),\end{aligned} \tag{13}$$

which involves 2^{LT_i} summations to get rid of all unobserved memory states. We note that the Markov property of $\tilde{S}_{i,1}, \ldots, \tilde{S}_{i,T_i}$ implies that all possible price recall paths have to be considered. For example, even if price recall is restricted to $L = 4$ periods and household i has only made $T_i = 6$ purchases in the category, (13) would already consist of more than 16 million components. Fortunately, the Hamilton (1989) filter turns out to be a very useful tool to evaluate this objective function, as it avoids such infeasible summations.

Hamilton (1989) originally developed his nonlinear filter to make inference on changes in economic regimes using time series. He applied the filter to establish the dates of historical business cycles, assuming that the underlying recession-expansion regimes follow a first-order Markov process. Econometric models with latent Markov-switching processes were first introduced by Goldfeld and Quandt (1973), and our reference price framework can be considered a member of this class of models as well.

In our application of the Hamilton filter, the unobserved components are the *composite* price recall variables $\tilde{S}_{i,t}$, $t = 1, \ldots, T_i$, for which the transition probabilities still have to be derived. It follows from the independence assumptions made in

Subsection 2.1 and the assumption of complete price information at the moment of purchase that

$$\begin{aligned}
&\Pr\Big(\tilde{S}_{i,t} = \tilde{s}_{i,t}\Big|\tilde{S}_{i,t-1} = \tilde{s}_{i,t-1}\Big) \\
&= \prod_{\tau=t-L}^{t-1} \Pr(S^{\tau}_{i,t} = s^{\tau}_{i,t}|S^{\tau}_{i,t-1} = s^{\tau}_{i,t-1}) \qquad \text{with} \qquad S^{t-1}_{i,t-1} = 1,
\end{aligned} \tag{14}$$

where the transition probabilities $\Pr(S^{\tau}_{i,t} = s^{\tau}_{i,t}|S^{\tau}_{i,t-1} = s^{\tau}_{i,t-1})$ are defined by (1)–(5). We initialize the process at the first purchase occasion $t = 1$ by setting

$$\Pr\Big(\tilde{S}_{i,1} = \tilde{s}_{i,1}\Big) = \prod_{\tau=1-L}^{0} \Pr\Big(S^{\tau}_{i,1} = s^{\tau}_{i,1}\Big), \tag{15}$$

where, analogous to (6),

$$\Pr(S^{\tau}_{i,1} = 1) = \prod_{\tilde{\tau}=1}^{1-\tau} \frac{1}{1 + \exp(-[\gamma_0 + \gamma_1 \tilde{\tau}])}\,. \tag{16}$$

We note that, for notational convenience, we suppress the dependence on segment membership z.

The Hamilton filter allows for inference on the unobserved memory state $\tilde{s}_{i,t}$ of household i at purchase occasion t, while taking into account the household's purchase history $b_{i,1:t}$ up to purchase occasion t. As a by-product, it also provides an evaluation of the brand choice probability $\Pr(B_{i,t} = b_{i,t}|B_{i,1:t-1} = b_{i,1:t-1})$ *unconditional* on $\tilde{s}_{i,t}$, which has been summed out. We note that current brand choice may depend on previous brand choices via the brand loyalty variable $BL_{i,j,t}$. By iteratively applying the filter, one can obtain the *unconditional* likelihood value for each household.

In the sequel, we use the shorthand notation $\xi_{i,t|\tilde{t}} = \overline{\Pr}(\tilde{S}_{i,t} = \tilde{s}_{i,t}|B_{i,1:\tilde{t}} = b_{i,1:\tilde{t}})$ to denote the $(2^L \times 1)$ vector containing the probabilities of *all* possible states $\tilde{s}_{i,t}$, given the household's observed purchase history $b_{i,1:\tilde{t}}$. Basically, an iteration of the Hamilton (1989) filter can be split up into two steps. Starting from $\xi_{i,t|t-1}$, the first step consists of updating $\xi_{i,t|t-1}$ to $\xi_{i,t|t}$, that is, updating the inference on $\tilde{s}_{i,t}$ by including the information contained in the most recent purchase observation $b_{i,t}$. It immediately follows from Bayes' theorem that

$$\xi_{i,t|t} = \frac{\xi_{i,t|t-1} \odot \overline{\Pr}(B_{i,t} = b_{i,t}|\tilde{S}_{i,t} = \tilde{s}_{i,t}, B_{i,1:t-1} = b_{i,1:t-1})}{\iota' [\, \xi_{i,t|t-1} \odot \overline{\Pr}(B_{i,t} = b_{i,t}|\tilde{S}_{i,t} = \tilde{s}_{i,t}, B_{i,1:t-1} = b_{i,1:t-1})\,]}\,, \tag{17}$$

where ι is the $(2^L \times 1)$ vector consisting of ones, and $\odot$ denotes element-by-element multiplication. We note that the vector of probabilities $\overline{\Pr}(B_{i,t} = b_{i,t} | \tilde{S}_{i,t} = \tilde{s}_{i,t}, B_{i,1:t-1} = b_{i,1:t-1})$ is conditional on the memory states $\tilde{s}_{i,t}$, so that it can be evaluated immediately using (7)$-$(11). Furthermore, it is crucial to note that the denominator in (17) amounts to the *unconditional* probability $\Pr(B_{i,t} = b_{i,t} | B_{i,1:t-1} = b_{i,1:t-1})$. This is the by-product, which allows us to evaluate (13). In the second step of the iteration, $\xi_{i,t|t}$ is updated to $\xi_{i,t+1|t}$, that is, the next memory state $\tilde{s}_{i,t+1}$ is inferred given the currently available purchase observations $b_{i,1:t}$. As price recall obeys a first-order Markov process, the basic Markov identity

$$\xi_{i,t+1|t} = \Lambda_{i,t+1|t}{}' \xi_{i,t|t} \tag{18}$$

holds. Here, $\Lambda_{i,t+1|t}$ is the row-conditional $(2^L \times 2^L)$ transition probability matrix for $\tilde{S}_{i,t}$. The elements $\Pr(\tilde{S}_{i,t+1} = \tilde{s}_{i,t+1} | \tilde{S}_{i,t} = \tilde{s}_{i,t})$ of $\Lambda_{i,t+1|t}$ are defined by (14) and hence (1)$-$(5). The output $\xi_{i,t+1|t}$ of (18) can be used as the input for (17) in the next iteration. This iterative procedure is initialized by setting $\xi_{i,1|0} = \overline{\Pr}(\tilde{S}_{i,1} = \tilde{s}_{i,1})$ in accordance with (15) and (16).

In sum, the following algorithm results in an evaluation of the *unconditional* probability that the sequence of brand choices $b_{i,1:T_i}$ is observed for household i, that is,

> Initialize the unconditional probability: $\Pr_i = 1$.
>
> Initialize $\xi_{i,1|0}$ using (15) and (16).
>
> Do for purchase occasion $t = 1, \ldots, T_i$:
>
>> Compute $\overline{\Pr}(B_{i,t} = b_{i,t} | \tilde{S}_{i,t} = \tilde{s}_{i,t}, B_{i,1:t-1} = b_{i,1:t-1})$ using (7)$-$(11).
>>
>> Compute $\xi_{i,t|t}$ using (17).
>>
>> Update the unconditional probability: $\Pr_i = \Pr_i \cdot [\,\text{denominator of (17)}\,]$.
>>
>> Compute row-conditional $\Lambda_{i,t+1|t}$ using (14) and hence (1)$-$(5).
>>
>> Compute $\xi_{i,t+1|t}$ using (18).
>
> Return $\Pr_i$.

3.2 Applying the EM algorithm

In order to obtain the parameter estimates for our model we maximize the likelihood function (12), which can be evaluated *for given parameter values* using the Hamilton filter. Although numerical techniques such as the BFGS algorithm can be applied directly to find the optimal parameter values, we consider the EM algorithm to take advantage of the specific structure of the problem. The EM algorithm quickly moves to reasonable (but not yet optimal) parameter values and it is quite robust with respect to the starting values, see, for example, Hamilton (1990). As final convergence of the algorithm is usually slow, one can decide to first apply the EM algorithm until the steps in the parameter space become rather small, and next one can do direct optimization of (12) starting from the parameter location obtained from the EM stage.

The EM algorithm considers the *complete* data likelihood of our model, that is, the *joint* likelihood of both observed brand choices and unobserved segment membership. The complete data likelihood is given by

$$\mathcal{L}_c = \prod_{i=1}^{N} \prod_{z=1}^{Z} \Big(\pi_z \Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = z) \Big)^{I\{Z_i = z\}} . \tag{19}$$

After taking logarithms we obtain

$$\ln \mathcal{L}_c = \sum_{i=1}^{N} \sum_{z=1}^{Z} I\{Z_i = z\} \ln(\pi_z) + \sum_{i=1}^{N} \sum_{z=1}^{Z} I\{Z_i = z\} \ln(\Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = z)). \tag{20}$$

The EM algorithm contains an Expectation step and a Maximization step, which are performed iteratively. In the E-step, the expectation of the log complete data likelihood (20) is taken with respect to the unobserved segment membership variables Z_i, $i = 1, \ldots, N$, *given* the household's observed purchase history $b_{i,1:T_i}$, $i = 1, \ldots, N$, and *given* the current parameter values. This results in

$$E[\ln \mathcal{L}_c] = \sum_{i=1}^{N} \sum_{z=1}^{Z} \Pi_{i,z} \ln(\pi_z) + \sum_{i=1}^{N} \sum_{z=1}^{Z} \Pi_{i,z} \ln(\Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = z)), \tag{21}$$

where, using Bayes' theorem,

$$\Pi_{i,z} \equiv \Pr(Z_i = z | B_{i,1:T_i} = b_{i,1:T_i}) = \frac{\pi_z \Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = z)}{\sum_{\tilde{z}=1}^{Z} \pi_{\tilde{z}} \Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = \tilde{z})} . \tag{22}$$

Next, in the M-step, (21) is maximized over all segment sizes $\pi = (\pi_1, \ldots, \pi_Z)$ and all model parameters denoted by $\theta = (\theta_1, \ldots, \theta_Z)$, *given* the posterior segment probabilities $\Pi_{i,z}$ computed in the E-step. This amounts to $Z+1$ *separate* subproblems, that is,

$$\max_{\pi} \quad \sum_{i=1}^{N} \sum_{z=1}^{Z} \Pi_{i,z} \ln(\pi_z), \tag{23}$$

$$\max_{\theta_z} \quad \sum_{i=1}^{N} \Pi_{i,z} \ln(\Pr(B_{i,1:T_i} = b_{i,1:T_i} | Z_i = z)), \qquad z = 1, \ldots, Z. \tag{24}$$

The resulting parameter updates can again be used as input for the E-step, see (22), after which a new iteration starts. It can be shown that subproblem (23) has a closed-form solution, which is given by

$$\pi_z = \frac{1}{N} \sum_{i=1}^{N} \Pi_{i,z}, \qquad z = 1, \ldots, Z. \tag{25}$$

In sum, an iteration of the EM algorithm consists of the follows steps, that is,

Compute the posterior segment probabilities $\Pi_{i,z}$ using (22).

Update the segment sizes $\pi = (\pi_1, \ldots, \pi_Z)$ using (25).

Update the model parameters $\theta = (\theta_1, \ldots, \theta_Z)$ by solving (24).

This iterative procedure converges and the resulting parameter values correspond to the Maximum Likelihood estimates of (12), see Dempster, Laird and Rubin (1977). In the first iteration, we initialize the EM algorithm by randomizing the $\Pi_{i,z}$ under the restriction that $\sum_z \Pi_{i,z} = 1$. By considering multiple runs, this is an effective and easily implemented approach to try out different starting points.

4 Interpretation

In this section we discuss how the estimation results for our model can be used to analyze (i) the dynamic structure underlying reference price formation, and to analyze (ii) the extent to which households are able to form an internal reference price.

Ideally, for given segment membership z, inference on the unobserved memory state $\tilde{s}_{i,t}$ should be based on *all* available information $b_{i,1:T_i}$ in the data set. This would give the most reliable results. However, the Hamilton filter described in the previous section only allows for inference on $\tilde{s}_{i,t}$ using the information $b_{i,1:t}$ available up to purchase occasion $t \leq T_i$. So, again using the notation $\xi^z_{i,t|\tilde{t}} = \overline{\Pr}(\tilde{S}_{i,t} = \tilde{s}_{i,t}|B_{i,1:\tilde{t}} = b_{i,1:\tilde{t}}, Z_i = z)$ but now with explicit segment membership index z, we can only estimate $\xi^z_{i,t|t}$, while we are actually more interested in $\xi^z_{i,t|T_i}$. Fortunately, the smoothing algorithm of Kim (1994), also described in Hamilton (1994, p. 694), allows us to compute $\xi^z_{i,t|T_i}$ from $\xi^z_{i,t|t}$. This algorithm is characterized by the recursion

$$\xi^z_{i,t|T_i} = \xi^z_{i,t|t} \odot [\,\Lambda^z_{i,t+1|t}(\,\xi^z_{i,t+1|T_i} \oslash \xi^z_{i,t+1|t}\,)\,]\,, \tag{26}$$

where $\oslash$ denotes element-by-element division. Using that $\xi^z_{i,T_i|T_i}$ is already available from the Hamilton filter and starting from $t = T_i - 1$, the recursion (26) is iterated backward to obtain all $\xi^z_{i,t|T_i}$, $t = 1, \ldots, T_i$.

Finally, one can obtain the posterior distribution of $\tilde{S}_{i,t} = (S^{t-L}_{i,t}, \ldots, S^{t-2}_{i,t}, S^{t-1}_{i,t})$ *unconditional* on segment membership z by weighting the segment-conditional $\xi^z_{i,t|T_i}$, $z = 1, \ldots, Z$, with the corresponding posterior segment membership probabilities $\Pi_{i,z}$, $z = 1, \ldots, Z$. This amounts to computing

$$\Pr(\tilde{S}_{i,t} = \tilde{s}_{i,t}|B_{i,1:T_i} = b_{i,1:T_i}) = \sum_{z=1}^{Z} \Pi_{i,z}\, \xi^z_{i,t|T_i}, \tag{27}$$

where $\Pi_{i,z}$ is defined by (22).

In order to increase the interpretability of (27), one can derive the probability that household i is still able to recall at purchase occasion t the prices observed at purchase occasion τ as

$$\Pr(S^\tau_{i,t} = 1|B_{i,1:T_i} = b_{i,1:T_i}) = \sum_{\tilde{s}_{i,t}:s^\tau_{i,t}=1} \Pr(\tilde{S}_{i,t} = \tilde{s}_{i,t}|B_{i,1:T_i} = b_{i,1:T_i}), \tag{28}$$

and one can compute the probability that household i has enough price information to form an internal reference price at purchase occasion t as

$$\Pr(\sum_{\tau=t-L}^{t-1} S^\tau_{i,t} > 0|B_{i,1:T_i} = b_{i,1:T_i}) = 1 - \Pr(\tilde{S}_{i,t} = (0, \ldots, 0, 0)|B_{i,1:T_i} = b_{i,1:T_i}). \tag{29}$$

We emphasize that (28) and (29) take into account *all* information contained in the household's purchase history.

5 Empirical analysis

In this section we apply our reference price model to an A.C. Nielsen scanner panel data set on catsup purchases in the Sioux Falls SD market. The considered period consists of 114 weeks from June 1986 to August 1988. The first 57 weeks are used for initialization purposes and are discarded in the log-likelihood evaluation, while the remaining 57 weeks are used for either parameter estimation or out-of-sample model validation. Only households which made at least three purchases in both periods are considered in the analysis. The estimation sample contains 80% of these households. The remaining 20% is assigned to a hold-out sample. The total sample consists of 619 households who made together 9416 purchases in the catsup category. The brands in our data set are Heinz, Hunts, Del Monte and Private Label.

5.1 Parameter estimates

We consider two variants of our model, that is, the model without unobserved heterogeneity and the model with an unknown number of latent response segments to capture unobserved heterogeneity. For the latter variant, we use the AIC-3 criterion, proposed by Bozdogan (1994), to determine the optimal number of segments. This measure is defined as $-2\ln\mathcal{L}+3K$ with $\mathcal{L}$ being the maximum likelihood value and K being the number of parameters to be estimated. The model corresponding to the smallest AIC-3 value is selected. The AIC-3 criterion suggests more parsimonious models than the standard AIC criterion, but if the number of observations exceeds 20 it is less parsimonious than BIC, the other popular measure in the literature. An extensive simulation study by Andrews and Currim (2003) indicates that AIC-3, applied to multinomial choice data, performs better than several other criteria including AIC and BIC. For our model, after setting $L = 4$ (this lag in price recall gives the highest maximum likelihood value), we find two response segments in the data set. Our parameter estimates are reported in Table 1. For the two-segments model, we show both the segment-specific parameters and the segment-averaged parameters for which the standard errors have been obtained using the delta method.[1]

Insert Table 1 *about here.*

For both the model without unobserved heterogeneity and the two-segments model, all response parameters have the expected sign. Furthermore, all parameters are significant at a 1% level in the first model, whereas in the second model this holds for all parameters except the price parameter. However, the parameter estimates do not indicate a strong response asymmetry concerning price gains and price losses. In the two-segments model, the first segment contains about 45% of the households and the second segment contains the remaining 55%. Comparison of the segment-specific parameter estimates suggests that households in segment 2 react stronger to promotional activities of brands, while households in segment 1 are more driven by state dependence, which is represented by the brand loyalty variable. For both variants of our model, the estimates of the price recall parameters γ_0 and γ_1 in (5) indicate that price memory is very limited, as $\gamma_0 + \gamma_1$ is much smaller than zero, and hence even price recall related to the previous purchase occasion is already far below 50%.

5.2 Impact of purchase timing and promotion on price recall

In order to see whether interpurchase times and (non-price) promotional activity affect price recall, we also estimate our model after including these variables in the price recall probability (5). Here, we define promotional activity as the average promotion rate in the category at purchase occasion τ when the prices were observed. Hence, promotional activity is measured as $\frac{1}{J}\sum_{k=1}^{J} PM_{i,k,\tau}$. We note that it is conceptually straightforward to include such time-varying variables in the memory

[1]The segment-averaged parameter values should not be interpreted as the parameter values for "the average household", as any household belongs to exactly one of the unobserved segments and hence cannot have preferences and response parameters corresponding to some convex combination. For an appropriate meaning, we have to reformulate our model as a random effects model in which the parameters for each household are drawn from some continuous heterogeneity distribution. In a mixture model, the heterogeneity distribution is represented by a discrete distribution. This subtle difference is pointed out by Jain, Vilcassim and Chintagunta (1994). However, as the difference is only conceptual and not methodological, we continue to interpret the segment-averaged parameter values as aggregated estimates.

process, but a practical drawback is that the transition probability matrix $\Lambda_{i,t+1|t}$ has to be computed for each purchase occasion in order to evaluate the likelihood function (12). This makes parameter estimation very time-consuming. In contrast, for the original model in which (5) just depends on $t - \tau$, $\Lambda_{i,t+1|t}$ only needs the be evaluated once (for each response segment) per likelihood evaluation.

It turns out that interpurchase time does not have a significant impact on price recall. For the model without unobserved heterogeneity, the z-score is -0.668. For the two-segments model, the z-score becomes 0.019. This lack of significance is consistent with the price knowledge study by Vanhuele and Drèze (2002), which reports that purchase recency does not affect price knowledge. On the other hand, the effect of promotional activity on price recall is positive and significant at a 1% level for both models, with z-scores of 3.011 and 2.699, respectively. So, promotional activities such as features and displays substantially increase the probability that corresponding prices are kept in memory, and hence affect both current and future brand choices. This result is, in a sense, consistent with the study by Lattin and Bucklin (1989), which also demonstrates that promotional activity has significant reference effects. Still, there is an important difference. We model the reference effect of promotion on brand choice as an *indirect* effect via the reference effect of price, whereas Lattin and Bucklin (1989) model the reference effect of promotion as an effect in itself. However, it is not clear why such promotional activities, which only aim at drawing the attention of households, should directly affect future brand choice utility.

5.3 Model performance

In order to see how well our model performs in-sample and out-of-sample, we compare it with three popular alternatives in the literature. The competing models are a brand choice model which does not account for any reference price, a model in which the reference price of a brand is the price observed at the previous purchase occasion, and a model in which the reference price is defined as an exponentially smoothed average of previously observed prices.

Insert Table 2 *about here.*

Table 2 contains the results of the empirical comparison. The in-sample log-likelihood values without unobserved heterogeneity show that our reference price is the most flexible one. Interestingly, after accounting for heterogeneity, the AIC-3 criterion identifies three segments for the three competing models, whereas it suggests only two segments for our model. Hence, our model has one response segment less. This is probably due to its richer dynamical structure, allowing for additional household heterogeneity to be picked up. The AIC-3 value of our model is slightly lower than it is for the "previous price model", but it is clearly lower than the AIC-3 values of the remaining two models. Our model also has the highest out-of-sample log-likelihood, although the difference with the "previous price model" is again quite small. The out-of-sample comparison would even become more favorable if we set the number of response segments for all models equal to two, as in our model. Overall, we conclude that our model has a good performance relative to three popular alternative models.

5.4 Size and symmetry of the reference effect

The parameter estimates in Table 1 suggest that the effect of reference price on brand choice is substantial, but without a clear asymmetry between price gains and losses. It would however be interesting to consider these results in more detail, and to compare them with the sticker shock effects resulting from the two other reference price specifications that we have considered in the empirical comparison.

Insert Table 3 *about here.*

The top part of Table 3 shows the estimated price gain and price loss parameters for the three models, and the middle part of the table provides the corresponding z-scores. Clearly, both in an absolute sense and in terms of significance, the reference effect in our model is largest. This is not surprising, as our reference price is the most flexible one, and only our model is able to distinguish between households not forming a reference price and households not reacting to it. Neglecting this distinction would result in a downward bias of the estimated response to price gains and losses *given* that a reference price is available. As the reference effect resulting from our new model is substantially larger than the size we would obtain from applying current practice, the impact of internal reference price formation on brand choice might be even larger than is currently believed.

Finally, the bottom part of Table 3 shows the p-values of the Likelihood Ratio [LR] test for the null hypothesis "no gain-loss asymmetry", that is, $\beta_4 = -\beta_5$. These results are interesting too. For our model the null hypothesis cannot be rejected at any reasonable significance level, whereas it is strongly rejected at a 5% level for the two other models. So, although applying current practice would result in finding an asymmetric sticker shock effect, as predicted by Prospect Theory, this asymmetry seems to disappear when our more advanced reference price model is considered.

5.5 Analyzing reference price formation

Our model can be used to get insights into (i) the probabilities that each of the past prices is used for reference price formation, and (ii) the probability that households have sufficient price knowledge to form an internal reference price. Current reference price models in the literature do not allow for such inference from scanner panel data.

Insert Figure 2 *about here.*

We discuss the results for our model with two response segments. Figure 2 contains histograms revealing the extent to which households are still able to recall the prices which were observed $t-\tau$ purchase occasions ago, where $t-\tau = 1, 2, 3, 4$. These histograms are obtained by computing (28) for all purchase occasions in the estimation sample. Figure 2 clearly demonstrates that price memory of households is very limited. For example, even the recall probabilities for prices observed at the previous purchase occasion are seldom larger than 50%. The average percentages of price recall are 21.0%, 7.3%, 3.7% and 2.4%, for 1, 2, 3 and 4 purchase occasions ago.

Insert Figure 3 about here.

Figure 3 shows an analogous histogram indicating to what extent households form internal reference prices. This histogram is obtained by computing (29) for all purchase occasions in the estimation sample. It can be seen that the probability of reference price formation is usually smaller than 60%. The average probability is 31.1%. We note that this low percentage is a direct consequence of the low price recall probabilities.

5.6 Price elasticity analysis

The existence of internal reference prices implies a tradeoff between current and future brand sales. Related to this tradeoff, Greenleaf (1995) and Kopalle, Rao and Assuncao (1996) develop dynamic programming models to investigate the impact of reference effects on the profitability of price promotions, and to obtain optimal pricing schemes. They define reference prices at the market level. However, in our model for scanner panel data, we consider reference prices at the household level and we can only analyze the impact on household demand, and not the impact on profitability.

Our model suggests a strong and significant effect of reference price on brand choice, but the managerially relevant implications for current and future brand sales are still unclear. In order to get a better understanding of the effects of a price adjustment, we perform a simulation study. We are not aware of studies in the literature which do something similar, except for a study by Erdem, Imai and Keane (2003) in which the effects of future price expectations on household demand are considered. Households forming price expectations are forward-looking, whereas households forming internal reference prices are backward-looking.

In the simulation study we investigate two scenarios. In the first situation the price of the considered brand is *temporarily* adjusted by 1% at the current purchase occasion, whereas in the second setting this price is kept at its original level. For both scenarios, future prices are unchanged. Moreover, the prices of all other brands remain unaffected. By comparing the current and future brand choice probabilities for the two settings, one can obtain estimates of both the instantaneous sensitivity of own brand choice and the lagged response via the reference price effect. For each purchase occasion in the estimation sample which is followed by at least $L = 4$ future purchase occasions, we compute the current brand choice probabilities and the brand choice probabilities for $L = 4$ periods ahead. Proxies for current and future sales are obtained by adding the corresponding probabilities over different purchase occasions. We repeat this procedure 1000 times, and we add the "sales" values obtained from all runs. For each run, segment membership of each household is drawn from its posterior distribution, and the price recall states are drawn from the corresponding first-order Markov processes.

Table 4 reports the percent change in sales between the two scenarios, where we distinguish between current sales and the sales levels up to $L = 4$ periods ahead. Furthermore, the table shows the "total effect" of the temporary price adjustment. This effect is in percents too. It is computed as the ratio of the net change in sales over time and the current sales level for the scenario without price adjustment. For comparison, Table 4 also contains the results for the exponentially-smoothed-average reference price model. We note that the previous-price model cannot capture dynamic effects lasting for more than one period, while a brand choice model without any reference price component can only capture instantaneous effects.

Insert Table 4 *about here.*

Table 4 illustrates that the impact of internal reference price formation on sales is substantial, even though price recall of households is very limited. Although the instantaneous own price elasticities of all brands are located in between -10 and -4, the net effects after $L = 4$ subsequent purchase occasions are rather close to zero, that is, all net price elasticities are larger than -3. For all brands, the net effect of a price adjustment on current and future sales is less than 40% of its instantaneous effect. A comparison with the results obtained for the exponentially-weighted-average reference price model illustrates that the reference effect implied by our model is much stronger, in an absolute sense as well as in a relative sense.

6 Conclusions

In this paper, we have proposed a new reference price framework for brand choice. Price recall of households is modeled as a hidden Markov-switching process with an absorbing state, and an internal reference price is constructed from the prices the household is able to recall. Features of our reference price model include the following, that is, (i) it does not a priori impose a rigid dynamic price structure, (ii) it accounts for uncertainty in reference price formation, (iii) it allows that households may forget past prices, and (iv) it even allows that a household cannot construct a reference price at all. Our model can be used to analyze how many prices observed in the past are considered for reference price formation, and to what extent households have sufficient price knowledge to form an internal reference price.

Applied to A.C. Nielsen scanner panel data on catsup purchases, our model has a good performance relative to two popular reference price specifications in the literature. Our main findings are as follows. First, a price supported by a feature/display has a higher probability to be recalled than a price which is not supported by such promotional activity. Second, contrary to the two competing reference price models, our model does not indicate an asymmetry between price gains and losses. However, our model suggests a stronger and more significant reference price effect. This is an indication that the impact of internal reference price formation on brand choice might be even larger than is currently believed. Third, we find that the prices observed at the previous purchase occasion have an average recall probability of about 20%. We estimate the average probability that a household has sufficient price knowledge to form an internal reference price at about 30%. Fourth, even though price recall is very limited, the impact of reference price formation on brand choice is very substantial. For all brands in the data set, the total effect of a price discount on current and future sales is less than 40% of its instantaneous effect.

We conclude by mentioning some limitations of our framework and we provide some suggestions for further research. A conceptual remark about our model is that we interpret "recalling a price" as being equivalent to "using this price for reference price formation". We report price recall probabilities, but in fact we can only make inference on the probabilities that prices from the past show up in the reference price. It might be possible that a household is able to recall prices observed in the past, but still does not use this price information to form a reference price. However, this is only a matter of interpretation. A second issue is that households might have different degrees of price knowledge, see, for example, Vanhuele and Drèze (2002) who distinguish between recallable price knowledge, price recognition and deal spotting. Our model does not account for this. Finally, we have focussed on internal reference prices, constructed from past prices, and we have not allowed for an external reference price, based on the in-store prices during the purchase occasion. However, several studies indicate that both are important and can exist at the same time, see Mayhew and Winer (1992), Rajendran and Tellis (1994) and Mazumdar and Papatla (2000).

References

Andrews R.L. and I.S. Currim (2003), "A Comparison of Segment Retention Criteria for Finite Mixture Logit Models", *Journal of Marketing Research*, 40 (May), 235−243.

Bell D.R. and J.M. Lattin (2000), "Looking for Loss Aversion in Scanner Panel Data: The Confounding Effect of Price Response Heterogeneity", *Marketing Science*, 19 (2), 185−200.

Böckenholt U. (1999), "Mixed INAR(1) Poisson Regression Models: Analyzing Heterogeneity and Serial Dependencies in Longitudinal Count Data", *Journal of Econometrics*, 89, 317−338.

Bozdogan H. (1994), "Mixture-Model Cluster Analysis using Model Selection Criteria and a New Informational Measure of Complexity", *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach*, Vol. 2, H. Bozdogan (ed.), Kluwer Academic Publishers, 69−113.

Briesch R.A., L. Krishnamurthi, T. Mazumdar and S.P. Raj (1997), "A Comparative Analysis of Reference Price Models", *Journal of Consumer Research*, 24 (September), 202−214.

Chang K., S. Siddarth and C.B. Weinberg (1999), "The Impact of Heterogeneity in Purchase Timing and Price Responsiveness on Estimates of Sticker Shock Effects", *Marketing Science*, 18 (2), 178−192.

Chintagunta P.K., D.C. Jain and N.J. Vilcassim (1991), "Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data", *Journal of Marketing Research*, 28 (November), 417−428.

Dempster A.P., N.M. Laird and D.B. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM-algorithm", *Journal of the Royal Statistical Society B*, 39, 1−38.

Dickson P.R. and A.G. Sawyer (1990), "The Price Knowledge and Search of Supermarket Shoppers", *Journal of Marketing*, 54 (July), 42−53.

Doornik J.A. (1998), *Object-Oriented Matrix Programming using Ox 2.0*, Timberlake Consultants, London, `http://www.nuff.ox.ac.uk/Users/Doornik`.

Erdem T., S. Imai and M.P. Keane (2003), "Brand and Quantity Choice Dynamics Under Price Uncertainty", *Quantitative Marketing and Economics*, 1, 5−64.

Erdem T., G. Mayhew and B. Sun (2001), "Understanding Reference-Price Shoppers: A Within- and Cross-Category Analysis", *Journal of Marketing Research*, 38 (November), 445−457.

Goldfeld S.M. and R.E. Quandt (1973), "A Markov Model for Switching Regressions", *Journal of Econometrics*, 1, 3−16.

Greenleaf E.A. (1995), "The Impact of Reference Price Effects on the Profitability of Price Promotions", *Marketing Science*, 14 (1), 82−104.

Guadagni P.M. and J.D.C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data", *Marketing Science*, 2 (3), 203−238.

Hamilton J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57 (March), 357−384.

Hamilton J.D. (1990), "Analysis of Time Series Subject to Changes in Regime", *Journal of Econometrics*, 45, 39−70.

Hamilton J.D. (1994), *Time Series Analysis*, Princeton University Press.

Helson H. (1964), *Adaptation-Level Theory*, Harper and Row Publishers, New York.

Jain D.C., N.J. Vilcassim and P.K. Chintagunta (1994), "A Random-Coefficients Logit Brand-Choice Model Applied to Panel Data", *Journal of Business and Eco-*

nomic Statistics, 12 (3), 317−328.

Kahneman D. an A. Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, 47 (March), 263−291.

Kalyanaram G. and J.D.C. Little (1994), "An Empirical Analysis of Latitude of Price Acceptance in Consumer Package Goods", *Journal of Consumer Research*, 21 (December), 408−418.

Kalyanaram G. and R.S. Winer (1995), "Empirical Generalizations from Reference Price Research", *Marketing Science*, 14 (3), G161−G169.

Kamakura W.A. and G.J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure", *Journal of Marketing Research*, 26 (November), 379−390.

Kim C.-J. (1994), "Dynamic Linear Models with Markov-Switching", *Journal of Econometrics*, 60, 1−22.

Kopalle P.K., A.G. Rao and J.L. Assuncao (1996), "Asymmetric Reference Price Effects and Dynamic Pricing Policies", *Marketing Science*, 15 (1), 60−85.

Krishnamurthi L., T. Mazumdar and S.P. Raj (1992), "Asymmetric Response to Price in Consumer Brand Choice and Purchase Quantity Decisions", *Journal of Consumer Research*, 19 (December), 387−400.

Lattin J.M. and R.E. Bucklin (1989), "Reference Effects of Price and Promotion on Brand Choice Behavior", *Journal of Marketing Research*, 26 (August), 299−310.

Mayhew G.E. and R.S. Winer (1992), "An Empirical Analysis of Internal and External Reference Prices Using Scanner Data", *Journal of Consumer Research*, 19 (June), 62−70.

Mazumdar T. and P. Papatla (2000), "An Investigation of Reference Price Segments", *Journal of Marketing Research*, 37 (May), 246−258.

McFadden D. (1974), “Conditional Logit Analysis of Qualitative Choice Behavior”, in P. Zarembka (ed.), *Frontiers in Econometrics*, Chapter 4, Academic Press.

McLachlan G.J. and T. Krishnan (1997), *The EM Algorithm and Extensions*, Wiley.

Putler D.S. (1992), “Incorporating Reference Price Effects into a Theory of Consumer Choice”, *Marketing Science*, 11 (3), 287−309.

Rajendran K.N. and G.J. Tellis (1994), “Contextual and Temporal Components of Reference Price”, *Journal of Marketing*, 58 (January), 2−34.

Ramaswamy V., E.W. Anderson and W.S. DeSarbo (1994), “A Disaggregate Negative Binomial Regression Procedure for Count Data Analysis”, *Management Science*, 40 (March), 405−417.

Thaler R. (1985), “Mental Accounting and Consumer Choice”, *Marketing Science*, 4 (3), 199−214.

Vanhuele M. and X. Drèze (2002), “Measuring the Price Knowledge Shoppers Bring to the Store”, *Journal of Marketing*, 66 (October), 72−85.

Wedel M., W.S. DeSarbo, J.R. Bult and V. Ramaswamy (1993), “A Latent Class Poisson Regression Model For Heterogeneous Count Data”, *Journal of Applied Econometrics*, 8, 397−411.

Winer R.S. (1986), “A Reference Price Model of Brand Choice for Frequently Purchased Products”, *Journal of Consumer Research*, 13 (September), 250−256.

Table 1: Parameter estimates for our model. The estimated standard errors are given in parentheses.

	no heterogeneity	2-segments model		
		averaged	segment 1	segment 2
brand loyalty	0.232^{***}	0.200^{***}	0.299^{***}	0.120^{***}
	(0.010)	(0.013)	(0.037)	(0.040)
promotion	0.456^{***}	0.469^{***}	0.372^{***}	0.548^{***}
	(0.034)	(0.037)	(0.076)	(0.093)
price	-1.130^{***}	-1.077^{*}	-0.777	-1.321
	(0.421)	(0.654)	(1.375)	(0.893)
price gain	0.949^{***}	0.913^{***}	0.203	1.494^{***}
	(0.255)	(0.330)	(0.277)	(0.578)
price loss	-1.092^{***}	-1.010^{***}	-1.001	-1.017^{***}
	(0.187)	(0.275)	(0.612)	(0.322)
p: intercept	-2.105^{***}	-2.007^{***}	-2.142	-1.897^{*}
	(0.509)	(0.696)	(1.971)	(1.006)
p: $t - \tau$	0.653^{*}	0.686	0.688	0.684
	(0.347)	(0.634)	(1.816)	(0.687)
segment size			0.449	0.551

* significant at 10%.
** significant at 5%.
*** significant at 1%.

Table 2: Empirical comparison of our model and three competing models. From left to right, we report the in-sample log-likelihood value without accounting for unobserved heterogeneity, the number of segments identified by the AIC-3 criterion, the corresponding number of parameters, the in-sample log-likelihood, the AIC-3 values, and the out-of-sample log-likelihood.

	LL no het.	# seg.	# par.	LL in	AIC-3	LL out[a]
our reference price	−2992.3	2	23	−2961.0	5991.0	−716.2
exp. weighted average	−3009.6	3	32	−2951.5	5999.0	−725.8
previous price	−3011.8	3	29	−2952.7	5992.3	−718.7
no reference price	−3019.0	3	23	−2973.1	6015.2	−722.4

a: The LL-out values for the final three models with two segments are −726.4, −723.1, and −723.7, respectively.

Table 3: The top part shows the estimated response parameters of the price gain and price loss variables. For the multi-segment models, these response parameters amount to segment-weighted averages. The middle part reports the corresponding z-scores. The bottom part contains the p-values of the LR test for the null hypothesis "no gain-loss asymmetry".

		no heterogeneity		multi-segment	
		gain	loss	gain	loss
param.	our reference price	0.949	−1.092	0.913	−1.010
	exp. weighted average	0.045	−0.215	0.083	−0.269
	previous price	−0.050	−0.115	0.046	−0.111
z-score	our reference price	3.722	−5.852	2.772	−3.674
	exp. weighted average	0.646	−3.899	0.781	−2.238
	previous price	−1.008	−3.419	0.586	−2.631
p-value	our reference price	0.642		0.740	
	exp. weighted average	0.013		0.000	
	previous price	0.003		0.000	

Table 4: Estimated price elasticities indicating how current and future "sales" are affected by a temporary own price adjustment.

		t	$t+1$	$t+2$	$t+3$	$t+4$	"total"
				our reference price			
	Heinz	−4.45	1.84	0.59	0.28	0.17	−1.65
t	Hunts	−6.07	2.56	0.80	0.40	0.26	−1.86
	Del Monte	−8.54	4.49	1.47	0.74	0.39	−1.37
	Rest	−9.08	3.70	1.41	0.77	0.53	−2.67
				exp. weighted average			
	Heinz	−5.28	1.58	0.75	0.33	0.15	−2.55
t	Hunts	−7.45	2.17	1.02	0.50	0.26	−3.31
	Del Monte	−8.77	2.57	1.27	0.65	0.29	−3.92
	Rest	−8.33	1.19	0.69	0.42	0.26	−5.71

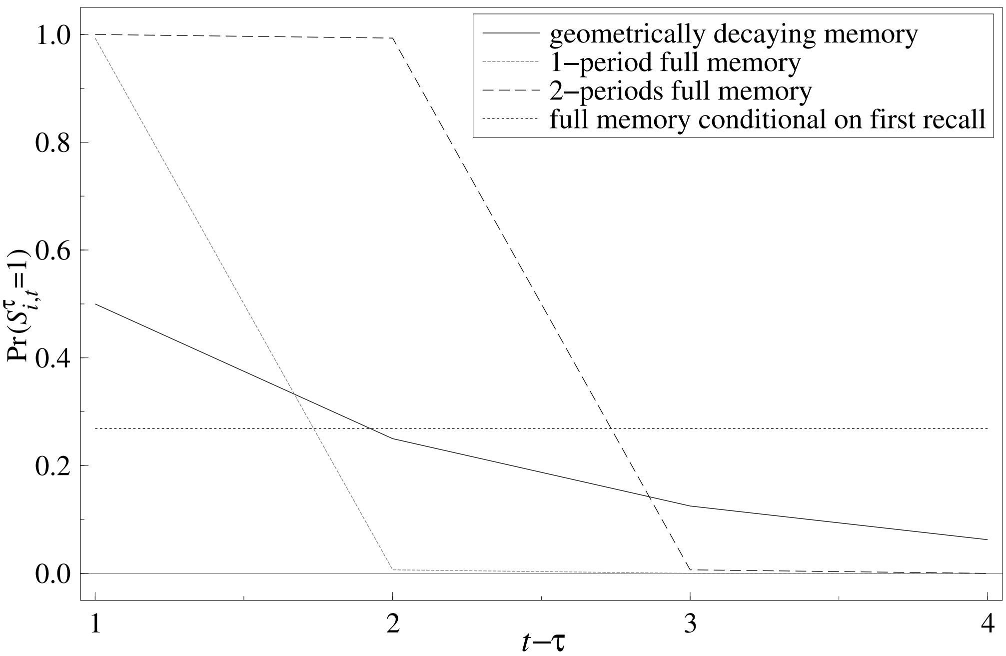

Figure 1: Illustration of some price memory patterns which can be reproduced by our model. The unconditional probability of price recall is plotted against the number of purchase occasions which have passed since the prices were observed.

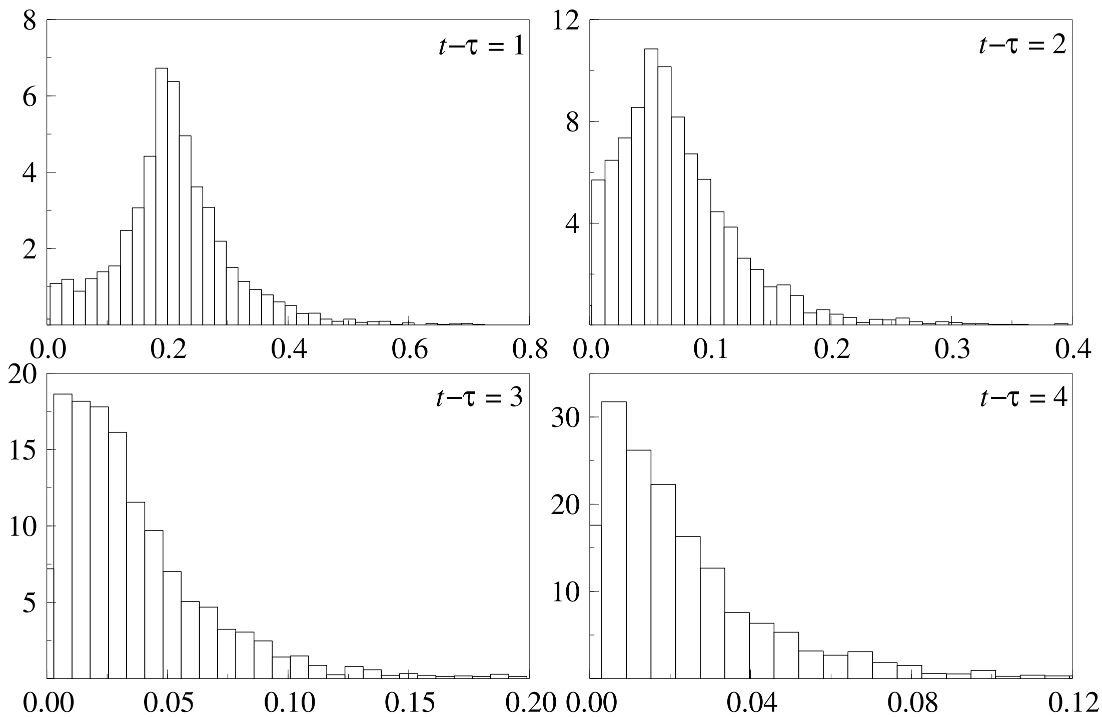

Figure 2: The probability that a household is able to recall the prices observed $t-\tau$ purchase occasions ago, $t-\tau = 1, 2, 3, 4$. The histograms are based on all purchase occasions in the estimation sample.

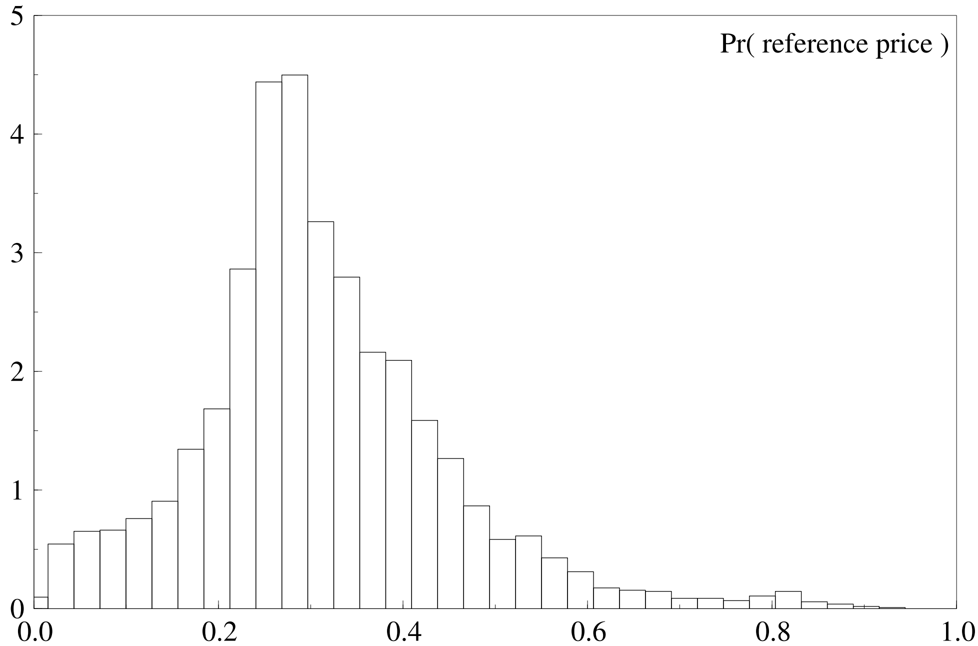

Figure 3: The probability that a household has enough price knowledge to form an internal reference price. The histogram is based on all purchase occasions in the estimation sample.