Reference Price Formation Model for Heterogeneous Consumers

Wirawan Dony Dahana, Nobuhiko Terui

We propose a hierarchical Bayesian model that accounts for consumer heterogeneity of reference price formation. Two types of reference prices, internal and external, are considered as components of the formation. We model a reference price as a weighted average of these types of reference price. The weight parameter, which indicates the relative importance of one type of reference price, is assumed to vary among consumers and across brands. We apply a hierarchical Bayesian framework that links the weights to some consumers' characteristics to explain the variation. The proposed model is calibrated on a scanner panel data of curry roux and instant coffee category using MCMC method. We observed that more than 50 percent of consumers form their reference prices based either on an internal or external reference price.

1. Introduction

Reference prices are a well established concept in consumer behavior and marketing research. Numerous theoretical and empirical studies have examined this area. An empirical generalization by Kalyanaram and Winer (1995) summarized the effects of reference prices on consumers' demand. They highlighted some theories on which the existence of reference price is based, empirical studies supporting its effects, and how consumers react to price changes relative to reference prices.

Although it is a widely recognized concept, no consensus exists among researchers of how consumers form reference prices. Some researchers argue that a reference price is determined by past prices (Lattin and Bucklin 1989, Kalyanaram and Little 1992, Greenleaf 1995). This type of reference price is called an internal reference price (IRP). This finding is based on Helson's Adaptation-Level Theory that assumes that consumers judge stimuli with respect to past and present stimulation. Therefore, past prices can be considered to serve as an adaptation level, or a reference point relative to an actual price. In addition, consumers' use of past prices as a reference point has been justified in laboratory studies in consumers' price judgments (Della Bitta and Monroe 1974, Kalwani and Yim 1992). An empirical study by Winer (1986) of past prices' effects using coffee scanner panel data showed that a brand choice model fit the data better when the effects were incorporated.

However, some doubt remains about consumers' recall of past price knowledge. A study by Dickson and Sawyer (1990) showed that almost half of consumers demonstrated a limited ability to recall prices they paid for goods (only 47% to 55% of the respondents could accurately recall the prices they had just paid). These facts therefore raised the issue of whether reference price models based on past prices can predict consumer behavior (e.g. brand choice) adequately (Urbany and Dickson 1991). Some researchers have proposed alternative formulations of reference prices based on current prices of some reference brands (Hardie et al. 1993, Mazumdar and Papatla 1995). This type of reference price is called the external reference price (ERP). The underlying arguments are, first, that the most recent prices are more salient in a consumer's memory than past prices (Rajendran and Tellis 1994). Secondly, the most recent brand purchased seems likely to serve as a reference brand; its price is used as reference point (Hardie et al. 1993). In an empirical study applied to refrigerated orange juice data, Hardie et al. (1993) found that the ERP (current price of the last purchased brand) choice model fitted in-sample and holdout data better than an internal reference price (IRP) (exponentially smoothed prices) choice model.

The evidence indicates that both IRP and ERP effects are supported theoretically and empirically. Whereas other formulations of reference price exist such as the reservation price (Hauser and Wernerfelt 1988, Srivastava et al. 2000) and expected future price (Khirsna 1992, Kalwani et al. 1990, Jacobson and Obermiller 1990), IRP and ERP can be conceptualized as encompassing the most relevant prices that consumers use (Rajendran and Tellis 1994). Therefore, it is natural to infer that several types of consumer exist in the market. They form their reference prices based either on IRP or ERP, or both IRP and ERP combined. A study by Mazumdar and Papatla (2000) found that more than 50 percent of the consumers in the liquid detergent and ketchup markets belong to the IRP or ERP segment. That consumers do have multiple reference prices has also been asserted by Mayhew and Winer (1992), who found that both IRP and ERP have considerable effects on purchase probability.

Understanding how reference prices are formed is important because that process determines the optimal pricing policy that should be taken by marketing manager to maximize profits (Kopalle et al. 1996). For example, given some conditions, cyclical (Hi-lo) pricing is the optimal policy for a duopolist when consumers use past prices as reference prices in contrast to constant pricing (e.g. EDLP) when consumers' reference prices are based on the current price of reference brand. Consequently, marketing managers would have an interest in acquiring information of how customers form their reference prices.

Rajendran and Tellis (1994) propose a reference price formation model that incorporates both IRP and ERP to answer the dualism of reference price conceptualizations. They argue that reference prices of target brands are formed as a weighted average of IRP and ERP. The weight parameter is bounded by 0 and 1 and is homogeneous among consumers. Mazumdar and Papatla (2000) extended this specification by assuming that the weight parameters might vary across segments in the market. They apply a finite mixture model to account for weight parameter heterogeneity and estimated parameters at the segment level.

Our study extends the specification of reference prices proposed by Mazumdar and Papatla (2000) by assuming that the weight parameters are heterogeneous among consumers (i.e. each consumer has a different weight). Furthermore, following the work by Briesch and colleagues (1997), which showed that the "best" reference price is the brand specific one, we also assume that the weight parameters vary across brands. This is of particular interest for manufacturers who wish to know the price competitiveness of their brands.

We use a continuous mixture model, instead of a finite mixture model, to capture parameter heterogeneity at the consumer level. This model has several advantages compared to its finite mixture counterpart, which will be mentioned in a subsequent section. We adopt a hierarchical Bayesian model and regress the weight parameters to some consumer characteristics to explain what factors that produce different parameters among consumers.

The remaining sections are organized as follows. In section 2, we describe the basic concepts of the proposed model. The model itself is discussed in section 3. Section 4 summarizes data and variables' operationalization. Empirical results and discussion are presented in section 5. Section 6 concludes this paper.

2. Basic concepts

Some terms are used in the literature to describe IRP and ERP. The IRP and ERP have been called a temporal reference price and a contextual reference price (Rajendran and Tellis 1994), which are memory-based and stimulus-based reference prices Briesch et al. (1997). In this section, we will discuss the process of reference price formation based on these two types of reference price and factors that might influence their relative importance.

How consumers form their reference prices depends on the accessibility of a price in memory (e.g. Biehal and Chakravarti 1983). Consumers who have a better memory of previous prices paid will use that information to construct a reference price. Then IRP will dominate the formation reference price of such a consumer in making price judgment. On the other hand, consumers who have limited ability to access past prices stored in memory will rely greatly on ERP in making such a judgment. Which type of reference price dominates consumer judgment is of particular interest for a retailer and manufacturer who want to identify their customers as belonging to an IRP or ERP segment so that they can make a decision about the best promotion strategy. For example, if the market is dominated by IRP consumers, a frequent price promotion will result in a sales trough after the deal (Blattberg et al. 1981, Neslin et al. 1985, Jain and Vilcassim 1991) because of the change of reference prices (Lattin and Bucklin 1989, Kalwani et al. 1990, Kalwani and Yim 1992).

However, it is well documented that the ability to recall past prices varies among consumers (Dickson and Sawyer 1990, Urbany and Dickson 1991, Vanhuele and Dreze 2002). Some researchers have investigated the extent to which this ability varies and the factors that can be used to explain the variation. According to the studies, variation of consumers' ability to recall past prices arises from the distinction in consumer characteristics (Rajendran and Tellis 1994, Vanhuele and Dreze 2002, Mazumdar and Papatla 2000), brand characteristics (Mazumdar and Papatla 2000), and household inventory position (Kumar, Karande, and Reinartz 1988). We will restrict our discussion on the relation between consumer characteristics and the ability to recall past prices (a summary relating this issue is apparent in Mazumdar, Raj, and Sinha 2005). Characteristics are purchase frequency, interpurchase times, the number of brands sampled, deal proneness, purchase quantity, price volatility, and brand preference. We summarize the relation as follows.

1. *Purchase Frequency*. Consumers who purchase particular brands in the category more frequently are likely to have better memory concerning the past price of those brands. Therefore, we can expect that such consumers tend to use IRP than ERP (Rajendran and Tellis 1994).
2. *Interpurchase times*. A second characteristic that can influence the relative use of IRP and ERP is the interpurchase time (Mazumdar and Papatla 2000). Consumers who leave the market for a long time, i.e. who have longer interpurchase times, are likely to have difficulty recalling past prices. Such consumers would therefore use ERP rather than IRP to evaluate brands' prices.
3. *Brand sampled*. Consumers who tend to switch across numerous different brands in the category will encounter various prices (Rajendran and Tellis 1994, Mazumdar and Papatla

2000). It would be difficult for consumers to recall those prices if the prices of the brands that were purchased varied with a large range.

4. *Deal Proneness*. If a consumer purchases the category mostly when it is on promotion, then ERP will be stronger than IRP (Mazumdar and Papatla 2000). Promotion-sensitive consumers are more likely to notice in-store price information and related cues and instead form a reference price at the point of purchase.
5. *Purchase quantity*. Vanhuele and Dreze (2002) posit that the consumer's budget for the category can improve consumer attention toward prices. Therefore, a consumer who buys a larger size is likely to have better ability to recall past prices and therefore to be an IRP shopper.
6. *Price volatility*. Price volatility is also a potential factor of a consumer's ability to recall past prices (Vanhuele and Dreze 2002). It would be easier for a consumer to recall if a price were relatively stable. On the contrary, if prices fluctuate intensely with a large range, consumers would find it harder to recall past prices. Consequently, ERP is expected to be more important relative to IRP for a consumer who encountered volatile prices.
7. *Brand preference*. The last factor that might engender a greater use of one type of reference price is brand preference (Mazumdar and Papatla 1995). A consumer who has a higher preference to a specific brand would have higher probability to purchase the brand. In turn, that consumer would have better knowledge of the brand characteristics, including price. Such a consumer is more likely to rely on IRP than on ERP.

In the subsequent sections, we will investigate the robustness of the characteristics mentioned above in explaining the heterogeneity of reference price formation. We adopted a hierarchical Bayes structure proposed by Rossi, McCulloch and Allenby (1996) to model the effects of consumer characteristics on weight parameters that represent the domination of IRP toward ERP.

3. The Model

Brand Choice Model. Most studies concerning reference prices have been devoted to investigating its effects on consumers' brand choice behavior (Kalyanaram and Little 1989, Winer 1989, Lattin and Bucklin 1989, Mayhew and Winer 1992). Specifically, the effect is often modeled as a discrepancy of the reference price from the actual price. It is called a "gain effect" if the reference price is greater than the actual price, and a "loss effect" if otherwise. We follow such a manner of modeling reference price effects in our study.

To begin, we first define a random utility function that incorporates reference price effects. Consumer h's utility to brand j at time t is a linear function of the discrepancy of reference price from actual price, marketing mix variables and brand loyalty.

$$
\begin{aligned}
U_{jht} &= \alpha_{jh} + \beta_{Gh} I_{jht}^{G} \left(RP_{jht} - P_{jht} \right) + \beta_{Lh} I_{jht}^{G} \left(P_{jht} - RP_{jht} \right) \\
&\quad + \beta_{1h} X_{jht}^{(1)} + \cdots + \beta_{dh} X_{jht}^{(d)} + \varepsilon_{jht}, \\
\varepsilon_{jht} &\sim N\left(0, \sigma^{2}\right), \quad j = 1, \ldots, J, \quad h = 1, \ldots, H, \quad t = 1, \ldots, T_{h}.
\end{aligned}
\tag{1}
$$

In those equations, RP_{jht} and P_{jht} are the reference price and the actual price of brand j for consumer h at time t. Also, I_{jht}^{G} (I_{jht}^{L}) equals 1 if $RP_{jht} > P_{jht}$ ($P_{jht} > RP_{jht}$), and 0 otherwise. $X_{jht}^{(1)}, \ldots, X_{jht}^{(d)}$ are covariates including the marketing mix variables and brand loyalty. In addition, α_{jh} represents consumer h's intrinsic preference for brand j; $\beta_{Gh}\beta_{Lh}\beta_{1h},\ldots,\beta_{dh}$ denote consumer h's sensitivity to reference price effects and the covariates. Finally, ε_{jht} is an error term.

To model consumers' brand choices based on (1), we employ a multinomial probit model. Using this model, a consumer chooses a brand that has the greatest utility. However, we do not observe the utility U_{jht} directly; instead, we observe the index I_{jht} such that $I_{jht} = 1$ if brand j were chosen by consumer h at time t, and 0 otherwise. The consequence of using the probit model is that the model specified by (1) suffers from identification problems: location invariance and scale invariance. The former can be removed by subtracting the utility function of each brand (i.e. U_{jht}) with respect to the J-th brand. The latter will be examined in the appendix. Thereby, we have a new system:

$$\begin{aligned} u_{jht} &= U_{jht} - U_{Jht} \\ &= \alpha'_{jh} + \beta_{Gh}\left[I^G_{jht}\left(RP_{jht} - P_{jht} \right) - I^G_{Jht}\left(RP_{Jht} - P_{Jht} \right) \right] \\ &\quad + \beta_{Lh}\left[I^L_{jht}\left(P_{jht} - RP_{jht} \right) - I^L_{Jht}\left(P_{Jht} - RP_{Jht} \right) \right] + \beta_{1h} X'_1 + \ldots + \beta_{dh} X'_d + \varepsilon'_{jht}, \end{aligned} \tag{2}$$

where $\alpha'_{jh} = \alpha_{jh} - \alpha_{Jh}$, $X'_{ih} = X^{(i)}_{jht} - X^{(i)}_{Jht}$ $(i = 1, \ldots, d)$, and $\varepsilon'_{jht} = \varepsilon_{jht} - \varepsilon_{Jht}$. It suffices to consider $M = J - 1$ brands' utility functions that we have just defined in (2).

Reference Price Model. We model the reference price as a weighted average of IRP and ERP. We assume that each consumer has their own weight, which might differ among consumers, thereby allowing for heterogeneity. Furthermore, the weights are also assumed to vary across brands because consumers might have different preferences, purchase experiences, etc. against the brands in the category. Consequently, the reference price (RP) of brand j at time t for consumer h is definable as

$$RP_{jht} = \lambda_{jh} IRP_{jht} + \left(1 - \lambda_{jh}\right) ERP_{jht}. \tag{3}$$

The weight parameter λ_{jh} is restricted to take a value in the interval [0, 1]; if λ_{jh} is greater than 0.5, we expect that the consumer is likely to use IRP rather than ERP. We call it the "memory parameter" because it represents a consumer's ability to recall past prices[1].

The model that includes both IRP and ERP as components of reference price formation was introduced for the first time by Rajendran and Tellis (1994). The model is based on the assumption that memory parameter λ is common for all consumers; hence, it does not account for consumer heterogeneity. Rajendran and Tellis (1994) attempted to explain the variation of how IRP consumers might be different from ERP consumers by grouping data based on consumers' characteristics. However, the study addresses limited consumer characteristics. We are not able to test the statistical significance of between-group parameter correlations.

---

1 Kopalle et al. (1996) used this term to denote the weight of exponentially weighted past prices.

Mazumdar and Papatla (2000) extended this specification by introducing consumer heterogeneity into the model. They employed the latent class model (finite mixture model) proposed by Kamakura and Russell (1989) and modeled reference prices as did Rajendran and Tellis (1994), but with segment-specific memory parameters. Consumers and brands' characteristics are used to explain the parametric variation across segments.

Finite mixture models have been used frequently by marketing researchers to model consumer heterogeneity because they are estimated easily and generally yield few segments. However, in spite of their attractive features, these models might not be appropriate in our context for at least two reasons. First, variables such as brand preference, price, and brand loyalty have the greatest impact on segmentation because they explain the greatest variance in brand choices (Keane 1997). Generally, reference price effects are smaller than these variables. Consequently, the resultant segments will be characterized well by parameters of such variables rather than by memory parameters. Secondly, as pointed out by DeSarbo et al. (1997), the derived groups or segments typically relate weakly to consumers' demographic or psychographic data. Furthermore, from a methodological point view, finite mixture models provide worse performance in fitting data than a continuous mixture model (Allenby and Rossi 1999). Therefore, as an alternative in this paper, we adopt a continuous mixture model to capture consumer heterogeneity of relative importance of IRP and ERP.

To account for heterogeneity of the memory parameter $\boldsymbol{\lambda}_h = (\lambda_{1h}, \ldots, \lambda_{Mh})'$, we adopt a flexible hierarchical Bayesian model proposed by Rossi, McCulloch and Allenby (1996). This model is a random coefficient model that regresses memory parameters on some covariates. In our case, we use consumers' characteristics discussed in section 2 as covariates. Because memory parameters are bounded by 0 and 1, it would be tractable to work with their logistic transformation. Therefore, the hierarchical Bayesian model for memory parameters is

$$\boldsymbol{\lambda}_h^* = \boldsymbol{\Theta}\boldsymbol{\delta}_h + \boldsymbol{\omega}_h, \quad \boldsymbol{\omega}_h \sim N(\mathbf{0}, \boldsymbol{\Omega}) \ , \tag{4}$$

where

$$\boldsymbol{\lambda}_h^* = \left(\lambda_{1h}^*, \lambda_{2h}^*, \ldots, \lambda_{Mh}^*\right)', \quad \lambda_{jh}^* = \ln\left(\frac{\lambda_{jh}}{1-\lambda_{jh}}\right).$$

In those equations: $\boldsymbol{\delta}_h$ is a vector of length *m* of consumer *h*'s characteristics; $\boldsymbol{\Theta}$ is the $M \times m$ coefficient matrix; $\boldsymbol{\omega}_h$ is a vector of length M of error terms whose elements are allowed to correlate across brands, but is independent across consumers.

Extension to include no reference price shoppers. To this point, we have assumed in our model that all consumers respond to a reference price effect. However, as pointed out by Erdem, Mayhew, and Sun (2001), the responsiveness to reference price is heterogeneous across consumers, indicating the possible existence of consumers who do not respond to reference prices. Such no-reference price shoppers respond to observed prices without going through any subjective encoding prices when making purchase decisions (Moon, Russell, and Duvvuri 2006). We extend the reference price model (3) by adding the observed price in the third term to include the existence of no-reference price shoppers in our model. Therefore, the reference price and random utility function are defined as

$$RP'_{jht} = \lambda_{1jh} IRP_{jht} + \lambda_{2jh} ERP_{jht} + \lambda_{3jh} P_{jht} \tag{5}$$

$$\begin{aligned} U_{jht} &= \alpha_{jh} + \beta_{Gh} I_{jht}^G \left(RP'_{jht} - P_{jht}\right) + \beta_{Lh} I_{jht}^G \left(P_{jht} - RP'_{jht}\right) \\ &\quad + \beta_{1h} X_{jht}^{(1)} + \cdots + \beta_{dh} X_{jht}^{(d)} + \varepsilon_{jht}, \end{aligned} \tag{6}$$

where we impose a restriction on λ_{1jh}, λ_{2jh} and λ_{3jh} such that they are bounded respectively by 0 and 1 and sum to 1. In addition, λ_{1jh} and λ_{2jh} respectively indicate the relative importance of IRP and ERP; λ_{3jh} is a parameter denoting whether or not a consumer responds to reference price. Therefore, if $\lambda_{3jh} = 1$, then the second and third term of utility function are cancelled out. The utility function reduces to a standard utility function (e.g. Guadagni and Little 1983), which implies that such a consumer does not respond to a reference price.

4. Model Application

Data and Variable Operationalization. We calibrate the proposed model on scanner panel data from two categories in the Japanese market: curry roux and instant coffee. Curry roux data are drawn from the Nihon Keizai Shimbun and coffee data are drawn from Video Research Co. Ltd.

Curry Roux. The data are scanner panel records from January 1998 to December 1999. We considered eight brands that account for 86.1% of the purchases of this category. The brand packages are of different sizes. For that reason, we rescaled their prices to prices per 100 g to make them comparable. Households included in these analyses are those that have made at least 10 purchases of the brands. We have 331 qualified households and 4999 observations.

Instant Coffee. The data are scanner panel records from October 1993 to September 1996. Five brands were selected for calibration. They account for 75.6% of the purchases of the category during the calibration period. Their prices are rescaled to prices per 100 g to facilitate their comparison. As in the case of curry roux, we selected households that had made at least 10 purchases of the brands, resulting in 154 qualified households and 2788 observations. Table 1 shows a description of the data.

Insert Table 1 about here

Marketing mix and Brand Loyalty. In model (1), we expressed the latent utility as a linear function of the reference price effect, marketing mix variables, and brand loyalty. We use a log price and promo for marketing mix variables. The promo is a dummy variable that is 1 if the price is discounted and 0 otherwise. For brand loyalty, we use the specification proposed by Guadagni and Little (1983).

Internal and External Reference Price. We use the respective brands' prices on the last purchase occasion (Mayhew and Winer 1992, Krisnamurthi et al. 1992) for IRP. We use the current price of the brand chosen on the last purchase occasion (Hardie et al. 1993) for ERP.

Consumer Characteristics. As mentioned previously, we model the reference price as a weighted average of IRP and ERP, where the weight or memory parameter is assumed to be household-specific and brand-specific. To explain the variation of the memory parameter across households, we set a prior on the memory parameter where the hyper parameter is a function of the consumer characteristics described in section 2. That is, average purchase frequency, average interpurchase times, the number of brands sampled, percentage of purchases at promotion, average purchase quantity, price volatility, and brand preference. However, to avoid a multicollinearity, we dropped purchase frequency from the analyses because it is strongly correlated with interpurchase times (-0.81 in curry roux and -0.87 in instant coffee). Price volatility is the standard deviation of prices paid by household. Brand preference is the maximum value of the average brand loyalty assigned to each brand (Kalyanaram and Little 1994).

5. Empirical Results and Discussion

Model Comparison. In addition to the proposed model, we calibrate six benchmark models for validation: the reference price model that contains IRP only (IRP model), ERP only (ERP model), both IRP and ERP with homogeneous memory parameter (HMP model), the Finite Mixture Model (FM Model) used by Mazumdar and Papatla (2000), the Continuous Mixture Model with constant prior on $\boldsymbol{\lambda}_h$, and the extended model. We include the Continuous Mixture Model with constant prior on $\boldsymbol{\lambda}_h$ to allow comparison of the performance of the finite and continuous model under the same set of information. To test the goodness of fit of the proposed model and its benchmarks, we compute the log of marginal likelihoods (Newton and Raftery 1994) and summarize the results in Table 2.

In both categories, the ERP model fit the data better than the IRP model. This result occurred perhaps because ERP households form a larger segment in the markets. Model fitness was improved when the reference price was modeled as a weighted average of IRP and ERP with a common memory parameter (HMP model). The results suggest that modeling the reference price as IRP or ERP alone will fail to portray the fact that consumers might have multiple reference prices. Furthermore, the performance was improved when heterogeneity was imposed on the memory parameter through the finite mixture model. This fact suggests that, although consumers might use both IRP and ERP, their relative importance might vary among individuals. The FM model therefore captures this heterogeneity at the segment level. Under the same set of information, however, the ability to fit data of FM model is inferior to that of the Continuous Mixture Model with constant prior on $\boldsymbol{\lambda}_h$, which accounts for heterogeneity at the individual level. Finally, as the table shows, the proposed model improved the previous models' performance of fitting the data closely. That improvement might have arisen from two sources of heterogeneity imposed on memory parameters. That is, the heterogeneity of memory parameters across individual households and across brands was not considered in previous models. For the extended model, we found no improvement of goodness of fit, even after introducing no-reference price shoppers. Instead, we observed that its log of marginal likelihood is slightly worse than that of the proposed model.

Insert Table 2 about here

Heterogeneous Distribution of Memory Parameters. The estimates of memory parameters are shown in Table 3, which also shows the mean of memory parameters of all consumers in the respective markets. We observed that the mean memory parameters vary across brands. In the curry roux category, the mean memory parameters for brands A, B, C, and E is less than 0.5, indicating that their typical consumers are likely to use ERP when purchasing them. In contrast, the respective mean memory parameters for brand F and brand G are greater than 0.5, indicating that IRP is important when consumers consider purchasing these brands. The variation of mean memory parameter in the instant coffee category is less noticeable than in curry roux. However, we can see that for brands A, B, and D, the values are less than 0.5, whereas for brand C, the values are greater than 0.5, which implies that the relative importance of memory in the purchase of those brands is different.

Insert Table 3 about here

Information summarized in Table 3 reveals nothing about the distribution of memory parameters. Therefore, we discuss how memory parameters are distributed across consumers. Figure 1.a presents a histogram of memory parameters for curry roux. It is immediately apparent that the distributions of memory parameters are U-shaped or J-shaped (rather than bell-shaped), where the modes are within the interval of either [0, 0.1] or [0.9, 1]. Households that use both IRP and ERP moderately (i.e. whose memory parameters are between 0.4 and 0.6) account for only 11% to 15% of the market. The results reveal that most households have great probability to use either IRP or ERP, but not both. Furthermore, as the figure shows, memory parameters vary not only across households, but also across brands. Households are more likely to use ERP when they evaluate brand A, B, C, D and E. In contrast, when brand F and brand G are considered, ERP is important for most households. Thus, the assumption that a consumer assigns a common weight for all brands will fail to capture this heterogeneity.

The distribution of memory parameters in the instant coffee category is presented in Table 1.b. The results resemble those of curry roux. That is, the market is dominated by households that use IRP or ERP alone. As for curry roux, the distributions vary across brands. Internal reference price households (i.e. whose memory parameters are between 0.9 and 1) account for more than 29% of the total market when they evaluate brand A. On the other hand, households have great probability to use ERP for other brands.

Insert Figure 1.a about here

Insert Figure 1.b about here

Although we observed that the distributions of memory parameters are U-shaped or J-shaped, we found that such is not the case when we used FM model. We present the distributions of memory parameters generated by the FM model in Fig. 2. The figures plot estimates of memory parameters of each segment on the horizontal axis and segment sizes or membership probabilities on the vertical axis. We can see that the distributions are not U-shaped or J-shaped as observed in the proposed model. Similar results of FM model-induced memory parameter distribution have also been observed in a study by Mazumdar and Papatla (2001) using liquid detergent and Yogurt data.

Insert Figure 2 about here

The question now is why the two models can suggest different distributions of memory parameters. As noted by Keane (1997), marketing mix variables such as price and promotion more strongly affect brand choice probability than reference price. Therefore, in the FM model, the derived segments do not represent a group of households having similar memory parameters because they are characterized mostly by marketing responses rather than by reference price responses.

On the other hand, unlike the FM model, our model allows the estimation of memory parameters at a household level. Therefore, we can use these household level estimates to infer their distribution without having to risk misclassifying the distributions. For this reason, we suggest that the FM model might be inappropriate in this analysis: the model incorrectly describes the distribution of memory parameters.

Choice Model Parameters. Table 4 presents the average of estimates of β. Estimates should be interpreted as relative to the last brand. That is, if the sign of any brand constant is positive, it implies that the brand is preferred to the last brand. Marketing responses (price and promo) and brand loyalty have the expected signs. The sign is positive for gain and negative for loss, as expected. For both categories, the loss effect is greater than the gain effect (i.e. $|\beta_L|>|\beta_G|$), indicating that loss aversion was observed in our data. However, this result does not mean that all households exhibit loss aversion.

Insert Table 3 about here

Hierarchical Structure for Memory Parameters. Now we discuss the robustness of consumer characteristics in explaining the variation of memory parameters. Table 5.a provides estimates of Θ in the curry roux category. Average interpurchase times have negative signs for all brands, as expected, indicating a negative relation between interpurchase times and memory parameters, which is consistent with previous works. The number of sampled brands has significant negative signs for brand A, D, F and G. Therefore, the result partially supported our hypothesis that consumers who sampled a larger number of brands will have higher probability to use ERP. Except for brand E, deal proneness has negative signs and is significant for brands A, B, D, and F, indicating that consumers with higher propensity to purchase on deal will tend to use ERP. For purchase quantity, the relation is mixed and not significant, except for brand B and brand F. Price volatility exhibits negative signs and is significant for all brands except brand C, which implies that households that confront volatile prices are likely to use ERP instead of IRP. Brand preference has significant positive signs for brands B, D, F and G, which supports our hypothesis, but such is not the case for the other brands.

Results for Instant Coffee are reported in Table 5.b. As in the case of curry roux, the results are consistent with previous works for interpurchase times, brand sampled, deal proneness and price volatility. However, in the case of purchase quantity and brand preference, the results are mixed or not significant otherwise.

Insert Table 5.a about here

Insert Table 5.b about here

Results of the extended model. Although the extended model is inferior to the proposed model in terms of the log of marginal likelihood, it is useful to see the distribution of $\boldsymbol{\lambda}$ when no reference price shoppers are included in the model. In Figs. 3.a and 3.b, we present joint distributions of λ_{1jh},

λ_{2jh} and λ_{3jh}. Interestingly, the conditional distributions of λ_{mjh} given λ_{njh} and λ_{ojh} $(m \neq n \neq o)$ reveal U-shaped distributions, even after no reference price shoppers are included. This fact indicates that consumers who use solely IRP and ERP, and those who do not respond to reference price, account for a large fraction of the market.

The fact that no-reference price shoppers account for a large fraction of the whole market is important from a managerial perspective. Their existence is useful to explain why post promotion sales are not zero. This kind of consumer would not regard a reversion to a normal price as a price increase because they do not compare past prices as IRP consumers do. Therefore, targeted coupon distribution to this segment will not cause a trough after the deal.

Insert Figure 3.a about here

Insert Figure 3.b about here

Beside the existence of no-reference price shoppers, a retailer also has to consider price sensitivity of targeted segment to conduct a successful pricing strategy. Therefore it is necessary to make a comparison about price sensitivity between IRP, ERP, and no-reference price consumers. In a simulation by Moon, Russell, and Duvvuri (2006), it is reported that IRP consumers are more price sensitive than other consumers. In our study, however, the results are inconclusive. Figure 4.a and 4.b show the scatter plots of the estimates of weight parameters and price responses for curry roux and instant coffee respectively. In the case of curry roux, the results are consistent with those of Moon, Russell, and Duvvuri (2006). The average of the estimates of price responses of IRP consumers (those whose $\lambda_{1jh} \in [0.9,1]$) is equal to -3.310 which is greater in absolute value than those of ERP and no-reference price consumers (those whose $\lambda_{2jh}, \lambda_{3jh} \in [0.9,1]$) which are -3.204 and -3.253. On the other hand, the results in instant coffee reveal that ERP consumers are more price sensitive than other consumers. The average of the estimates of price responses of IRP, ERP, and no-reference price consumers are -1.272, -1.444 and -0.428 respectively.

Insert Figure 4.a about here

Insert Figure 4.b about here
----

6. Concluding Remarks

A recent popular issue in marketing studies is how to build models that can capture consumer heterogeneity. Consumers have different preferences and sensitivity to marketing variables. Modeling consumer behavior at an aggregate level merely provides summary statistics of the entire market. Representation of parameters in the aggregate model is limited to an average group of consumers. For that reason, the information contained would not be adequate for firms wishing to pursue their marketing activities because there might be numerous consumers who are not well represented. An effective marketing strategy can be attained only if a firm offers a product that is priced, promoted, and placed in a manner such that they meet the targeted group's needs. Therefore, models that provide individual level estimates are required for it to be possible.

Our proposed model of reference price formation allows weight parameters to vary across consumers and brands. Application of Bayesian analyses of MCMC method made it feasible to infer individual level parameters. The model provides information related to how individual consumers form their reference prices, which might be useful for a manufacturer and retailer who want to identify their customers. For example, a retailer wishing to remove a post-promotion dip in sales might target coupons to ERP consumers because they are likely to be unaffected by past prices.

The hierarchical Bayesian model of memory parameters also provided theoretical insights for researchers interested in investigating the process of reference price formation. Our study only investigated effects of consumer characteristics on the memory parameter. However, as noted by

Dickson and Sawyer (1990), variables such as product characteristics might influence consumers' ability to recall past prices. In addition, consumer demographics can also be considered to influence the process. Incorporating the effects of such variables on reference price formation has now become straightforward because we can treat them as explanatory variables in the proposed hierarchical framework.

However, the proposed reference price model leaves room for improvement. For example, if forward-looking consumers account for a large share of the market, we should consider expected future prices as a component of reference prices in addition to IRP and ERP. It would therefore be appropriate to model reference prices as a function of past, present, and expected future price. This would be particularly important in consumer durable markets such as that for automobiles, in which consumers tend to predict future prices when considering purchases.

Appendix (Estimation of the Model)

Posterior Distribution on Memory Parameter. First, for convenience, we denote the utility system (2) in matrix notation. Thus we have

$$\mathbf{U}_{ht} = \mathbf{X}_{ht}\boldsymbol{\beta}_h + \boldsymbol{\varepsilon}_{ht}, \quad \boldsymbol{\varepsilon}_{ht} \sim N\left(\mathbf{0}, \boldsymbol{\Lambda}\right), \tag{A.1}$$

where $\mathbf{U}_{ht} = (u_{1ht}, \ldots, u_{Mht})'$, $\boldsymbol{\beta}_h = (\alpha'_{1h}, \ldots, \alpha'_{Mh}, \beta_{Gh}, \beta_{Lh}, \beta_{1h}, \ldots, \beta_{dh})'$, $\boldsymbol{\varepsilon}_{ht} = (\varepsilon'_{1ht}, \ldots, \varepsilon'_{Mht})'$. In that equation, $\mathbf{X}_{ht} = (I_M, \mathbf{G}_{ht}, \mathbf{L}_{ht}, \mathbf{X}_{ht}^{(1)}, \ldots, \mathbf{X}_{ht}^{(d)})$ is a matrix including an intercept term of the M brands and covariates, where I_M is an M-dimensional identity matrix, $\mathbf{G}_{ht}$ and $\mathbf{L}_{ht}$ are vectors of length M of gain and loss effects, and $\mathbf{X}_{ht}^{(1)}, \ldots, \mathbf{X}_{ht}^{(d)}$ are covariate vectors of length M, including marketing mix variables and brand loyalty.

In our model, the set of parameters to be estimated is $(\{\boldsymbol{\beta}_h\}, \boldsymbol{\Lambda}, \boldsymbol{\Theta}, \boldsymbol{\Omega}, \{\boldsymbol{\lambda}_h\})$. The parameters are estimated using MCMC method. We generate random draws using the Gibbs sampler from a set of full conditional distributions. To proceed, we must define the conditional posterior distribution of $\boldsymbol{\lambda}_h^*$ from which the sample of logistic transformed memory parameters are drawn. The distribution is defined as follows.

$$\pi\left(\boldsymbol{\lambda}_h^* \mid \mathbf{U}_{ht}, \mathbf{X}_{ht}, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right) \propto l\left(\boldsymbol{\lambda}_h^* \mid \mathbf{U}_{ht}, \mathbf{X}_{ht}, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}\right) P\left(\boldsymbol{\lambda}_h^* \mid \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right)$$

$$\propto \exp\left(-0.5\left[\left(\sum_{t=1}^{T_h}\left(\mathbf{U}_{ht} - \mathbf{X}_{ht}\boldsymbol{\beta}_h\right)' \boldsymbol{\Lambda}^{-1}\left(\mathbf{U}_{ht} - \mathbf{X}_{ht}\boldsymbol{\beta}_h\right)\right) + \left(\boldsymbol{\lambda}_h^* - \boldsymbol{\Theta}\boldsymbol{\delta}_h\right)' \boldsymbol{\Omega}^{-1}\left(\boldsymbol{\lambda}_h^* - \boldsymbol{\Theta}\boldsymbol{\delta}_h\right)\right]\right) \quad \text{(A.2)}$$

This distribution contains information of each consumer along with market level information summarized in the prior. The likelihood implies the probability that consumers achieve utility $\mathbf{U}_{ht}$ given $\boldsymbol{\lambda}_h^*$, which is expressed implicitly in $\mathbf{X}_{ht}$. We set a conjugate prior on $\boldsymbol{\lambda}_h^*$ based on consumers' characteristics so that it will be determined not only by the individual utility function but also by consumers' characteristics. However, the distribution is not proper (i.e. its integral does not equal 1); it has no closed form. Therefore, we use the Metropolis-Hastings algorithm to determine the distribution.

The Priors. We set four priors for the hierarchical model: (1) the prior on the error variances in random utility model $\boldsymbol{\Lambda}$, (2) the prior on the coefficient of random utility model $\boldsymbol{\beta}_h$, (3) the prior on $\boldsymbol{\Omega}$, the covariance matrix of $\boldsymbol{\lambda}_h^*$, and (4) the prior on the matrix of regression coefficients in the memory-parameter model of heterogeneity $\boldsymbol{\Theta}$.

(1) Prior on $\boldsymbol{\Lambda}$

$$\boldsymbol{\Lambda}^{-1} \sim \text{Wishart}\left(\nu_0, V_0\right), \text{ where } \nu_0 = M+2,\ V_0 = \nu_0 I_M$$

(2) Prior on $\boldsymbol{\beta}_h$

$$\boldsymbol{\beta}_h \sim \text{N}(\mathbf{0}, \mathbf{V}_{\boldsymbol{\beta}}), \text{ where } \mathbf{V}_{\boldsymbol{\beta}} = 0.1 I_{rank(\mathbf{X}_h)}$$

(3) Prior on $\boldsymbol{\Omega}$

$$\boldsymbol{\Omega}_{\boldsymbol{\beta}}^{-1} \sim \text{Wishart}(p_0, P_0),\ \ p_0 = M+2,\ P_0 = p_0 I_M$$

(4) Prior on $\boldsymbol{\Theta}$

$$\text{vec}(\boldsymbol{\Theta}) \sim \text{N}\left(\bar{\mathbf{d}}, \left(\boldsymbol{\Omega} \otimes \mathbf{C}^{-1}\right)\right),\ \ \bar{\mathbf{d}} = \mathbf{0},\ \mathbf{C} = 0.1 I_M$$

Gibbs Sampler.

(1) $\mathbf{U}_{ht} \mid I_{ht}, \mathbf{X}_{ht}, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\lambda}_h$

Generate $\mathbf{U}_{ht}$ from truncated $\mathrm{N}(\mathbf{X}_{ht}\boldsymbol{\beta}_h, \boldsymbol{\Lambda})$.

(2) $\boldsymbol{\beta}_h \mid \mathbf{U}_{ht}, \mathbf{X}_{ht}, \boldsymbol{\Lambda}, \mathbf{V}_\beta, \boldsymbol{\lambda}_h$

Generate $\boldsymbol{\beta}_h$ from $\mathrm{N}(\overline{\boldsymbol{\beta}}_h, (\mathbf{X}_h'\mathbf{X}_h + \mathbf{V}_\beta^{-1})^{-1})$ where $\overline{\boldsymbol{\beta}}_h = (\mathbf{X}_h'\mathbf{X}_h + \mathbf{V}_\beta^{-1})^{-1}\mathbf{X}_h'\mathbf{U}_h$.

(3) $\boldsymbol{\Lambda}^{-1} \mid \{\mathbf{U}_{ht}\}, \{\mathbf{X}_{ht}\}, \{\boldsymbol{\beta}_h\}, \{\boldsymbol{\lambda}_h\}$

Generate $\boldsymbol{\Lambda}^{-1}$ from $\mathrm{Wishart}(\nu_0 + HT_h, V_0 + S_1)$,

where $S_1 = \sum_{h,t}(\mathbf{U}_{ht} - \mathbf{X}_{ht}\boldsymbol{\beta}_h)(\mathbf{U}_{ht} - \mathbf{X}_{ht}\boldsymbol{\beta}_h)'$. Note that, for identification, the $M \times M$-th

element of $\boldsymbol{\Lambda}$ is fixed to 1.

(4) $\boldsymbol{\lambda}_h^* \mid \mathbf{U}_{ht}, \mathbf{X}_{ht}, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\Theta}, \boldsymbol{\Omega}, \boldsymbol{\delta}_h$

First, generate $\boldsymbol{\lambda}_h^*$ from its conditional posterior distribution $\pi\left(\boldsymbol{\lambda}_h^* \mid \mathbf{U}_{ht}, \mathbf{X}_{ht}, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right)$. However, it does not constitute a conjugate family, so we use Metropolis-Hastings with a random walk algorithm. The procedure is described as follows.

For $j = 1, \ldots, M$

Set the starting point $\lambda_{jh}^{*(0)}$.

Generate γ_{jh} from $\mathrm{N}(0,1)$ and renew $\lambda_{jh}^{*(k)} = \lambda_{jh}^{*(k-1)} + \gamma_{jh}$ at the *k*-th iteration.

Generate u_{jh} from $\mathrm{U}(0,1)$.

Set $\lambda_{jh}^* = \lambda_{jh}^{*(k)}$ if $u_{jh} \le \alpha\left(\lambda_{jh}^{*(k)}, \lambda_{jh}^{*(k-1)} \mid \mathbf{U}_h, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right)$, and $\lambda_{jh}^* = \lambda_{jh}^{*(k-1)}$ otherwise.

$\alpha\left(\lambda_{jh}^{*(k)}, \lambda_{jh}^{*(k-1)} \mid \mathbf{U}_h, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right)$ is the acceptance probability defined as

$$\alpha\left(\lambda_{jh}^{*(k)}, \lambda_{jh}^{*(k-1)} \mid \mathbf{U}_h, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right) = \min\left(\frac{\pi_{jh}^{k}\left(\boldsymbol{\lambda}_h^* \mid \mathbf{U}_h, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right)}{\pi_{jh}^{(k-1)}\left(\boldsymbol{\lambda}_h^* \mid \mathbf{U}_h, \boldsymbol{\beta}_h, \boldsymbol{\Lambda}, \boldsymbol{\delta}_h, \boldsymbol{\Omega}\right)}, 1\right),$$

where π_{jh}^{k} is posterior density of $\boldsymbol{\lambda}_h^*$ evaluated at $\lambda_{jh}^* = \lambda_{jh}^{*(k)}$.

Finally, substitute λ_{jh}^{*} to obtain λ_{jh}.

(5) $\boldsymbol{\Theta} \mid \boldsymbol{\Omega}, \{\boldsymbol{\delta}_h\}, \{\boldsymbol{\lambda}_h\}$

Generate $\text{vec}(\boldsymbol{\Theta})$ from $\mathrm{N}\left(\tilde{d}, \left(\boldsymbol{\Omega} \otimes \boldsymbol{\delta}'\boldsymbol{\delta} + \mathbf{C}\right)^{-1}\right)$,

where $\tilde{d} = vec\left(\tilde{D}\right)$, $\tilde{D} = \left(\boldsymbol{\delta}'\boldsymbol{\delta} + \mathbf{C}\right)^{-1}\left(\boldsymbol{\delta}'\boldsymbol{\delta}\hat{D} + \mathbf{C}\bar{D}\right)^{-1}$, and $\hat{D} = \left(\boldsymbol{\delta}'\boldsymbol{\delta}\right)^{-1}\boldsymbol{\delta}'\boldsymbol{\lambda}^{**}$.

$\boldsymbol{\lambda}^{**}$ is the $H \times M$ matrix with each $\left(\boldsymbol{\lambda}_h^{**}\right)'$ as a row.

$\boldsymbol{\delta}$ is the $H \times m$ matrix with each $\boldsymbol{\delta}_h'$ as a row.

$\bar{D}$ = stack$(\bar{\mathbf{d}})$, the $m \times M$ matrix formed column by column from the elements of $\bar{\mathbf{d}}$.

(6) $\boldsymbol{\Omega}^{-1} \mid \{\boldsymbol{\lambda}_h\}, \boldsymbol{\Theta}$

Generate $\boldsymbol{\Omega}^{-1}$ from $\text{Wishart}(p_0 + H, P_0 + S_2)$, where $S_2 = \sum_h (\boldsymbol{\lambda}_h^{*} - \boldsymbol{\Theta}\boldsymbol{\delta}_h)(\boldsymbol{\lambda}_h^{*} - \boldsymbol{\Theta}\boldsymbol{\delta}_h)'$.

References

Allenby, G. M. and P. E. Rossi (1999), “Marketing Models of Consumer Heterogeneity,” *Journal of Econometrics*, 89, 57-78.

Blattberg, R. C., G. D. Eppen and J. Lieberman (1981), “A Theoretical and Empirical Evaluation of Price Deals for Consumer Nondurables,” *Journal of Marketing*, 45, 1, 116-129.

Briesch R. A., L. Krisnamurthi, T. Mazumdar and S. P. Raj (1997), “A Comparative Analysis of Reference Price Models,” *Journal of Consumer Research*, 24, 2, 202-214.

Della Bitta, A. J. and K. B. Monroe (1974), “The influence of Adaptation Levels on Subjective Price Perceptions,” in Scott Ward and Peter Wright (Eds.), *Advances in Consumer Research*, Vol. 1, Urbana, IL: Association for Consumer Research, 359-269.

DeSarbo, W. S., A. Ansari, P. Chintagunta, C. Himmelberg, K. Jedidi, R. Johnson, W. Kamakura, P. Lenk, K. Srinivasan and M. Wedel (1997), “Representing Heterogeneity in Consumer Response Models,” *Marketing Letters*, 8, 3, 335-348.

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

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

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.

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

Hardie B. G. S., E. J. Johnson and P. S. Fader (1993), "Modeling Loss Aversion and Reference Dependent Effects on Brand Choice," *Marketing Science*, 12, 378-394.

Hauser, J. R. and B. Wernerfelt 1988, "Existence and Uniqueness of Price Equilibria in Defender," *Marketing Science*, 7, 1, 92-93.

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

Jacobson, R. and C. Obermiller (1990), "The formation of Expected Future Prices: A Reference Price for Forward-Looking Consumers," *Journal of Consumer Research*, 16, 420-432.

Jain, D. C. and N. J. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10, 1, 23.

Kalwani, M. U., H. J. Rinne, Y. Sugita, and C. K. Yim (1990), "A Price Expectation Model of Customer Brand Choice," *Journal of Marketing Research*, 27, 251-262.

Kalwani, M. U. and C. K. Yim (1992), "Consumer Price and Promotion Expectations: An Experimental Study," *Journal of Marketing Research*, 29, 90-100.

Kalyanaram, G. and J. D. C. Little (1989), "An Empirical Analysis of Latitude of Price Acceptance in Consumer Packaged Goods," *Journal of Consumer Research*, 21, 408-418.

Kalyanaram, G. and J. D. C. Little (1989), "A Price Response Model Developed from Perceptual Theories," Working Paper, MIT.

Kalyanaram, G. and R. S. Winer (1995), "Empirical Generalization from Reference Price Research," *Marketing Science*, 14, 3, 161-169.

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

Keane, M. P. (1993), "Simulation Estimation for Panel Data Models with Limited Dependent Variables," in Handbook of Statistics, G. S. Maddala, C. R. Rao, H. D. Vinod, eds., New York: Elsevier Science Publishers, 545-571.

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.

Krishna, A. (1992), "The Normative Impact of Consumer Price Expectations for Multiple Brands on Consumer Purchase Behavior," *Marketing Science*, 11, 3, 266-286.

Krisnamurthi, 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, 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.

Mazumdar, T. and P. Papatla (1995), "Loyalty Differences in the Use of Internal and External Reference Prices," *Marketing Letters*, 6, 111-122.

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

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.

Moon, S., G. J. Russell, and S. D. Duvvuri (2006), "Profiling the Reference Price Consumer," *Journal of Retailing*, 82, 1-11.

Neslin, S. A., C. Henderson, and J. Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4, 2, 147-165.

Newton, M. and A. Raftery (1994), "Approximate Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of Royal Statistical Society*, 56, 3-48.

Rajendran, K. N. and G. J. Tellis (1994). "Contextual and Temporal Components of Reference Price," *Journal of Marketing*, 58 (January), 22-34.

Rossi, P. E., R. E. McCulloch, and G. M. Allenby (1996), "The value of Purchase History Data in Target Marketing," *Marketing Science*, 15, 4, 321-340.

Srivastava, J., D. Chakravarti, and A. Rapoport (2000), "Price and Margin Negotiations in Marketing Channels: An Experimental Study of Sequential Bargaining under One-Sided Uncertainty and Opportunity Cost of Delay," *Marketing Science*, 19, 2, 163-184.

Urbany, J. E. and P. R. Dickson (1991), "Consumer Normal Price Estimation: Market versus Personal Standards," *Journal of Consumer Research*, 18, 1, 45-51.

Winer, R. S. (1986), "Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, 4, 74-90.

Winer, R. S. (1989), "A Multi-Stage Model of Choice Incorporating Reference Prices," *Marketing Letters*, 1, 27-36.

Vanhuele, M. and X. Dreze (2002), "Measuring the Price Knowledge Shoppers Bring to the Store," *Journal of Marketing*, 66, 72-85.

Table 1. Description of the Data

Brand	Choice Share	Average Price	Percentage of Purchases on Discount
Curry Roux			
Brand A	0.191	73.674	11.082
Brand B	0.255	79.425	11.084
Brand C	0.253	79.869	13.338
Brand D	0.106	90.457	12.560
Brand E	0.060	108.712	11.775
Brand F	0.057	159.226	11.522
Brand G	0.049	188.380	12.717
Brand H	0.029	57.202	7.233
Instant Coffee			
Brand A	0.480	473.628	8.967
Brand B	0.099	412.663	10.133
Brand C	0.225	477.061	4.874
Brand D	0.057	419.477	11.110
Brand E	0.138	398.495	7.250

Table 2. Log of Marginal Likelihood

	Curry Roux	*Instant Coffee*
IRP Model	-166,354.201	-37,629.538
ERP Model	-166,281.449	-37,540.172
HMP Model	-165,992.520	-37,284.931
FM Model	-165,169.397	-36,571.210
CM Model 1*	-163,273.013	-36,296.178
Proposed Model	**-161,257.878**	**-36,029.202**
Extended Model	-161,636.7793	-361,33.272

* Continuous Mixture Model with constant prior on $\boldsymbol{\lambda}_h$

Table 3. HB Estimates of Memory Parameter

Variable	*Curry Roux*		*Instant Coffee*	
	Post Mean	Post STD	Post Mean	Post STD
Brand A	0.481	0.124	0.474	0.132
Brand B	0.354	0.105	0.446	0.129
Brand C	0.487	0.124	0.513	0.132
Brand D	0.519	0.128	0.368	0.124
Brand E	0.437	0.119	—	—
Brand F	0.658	0.138	—	—
Brand G	0.720	0.142	—	—

Table 4. HB Estimates of β

Variable	*Curry Roux*		*Instant Coffee*	
	Post Mean	Post STD	Post Mean	Post STD
Brand Constants				
Brand A	-0.082	0.065	1.134	0.223
Brand B	0.245	0.108	-0.545	0.195
Brand C	0.172	0.076	0.506	0.206
Brand D	-0.087	0.134	-0.465	0.209
Brand E	-0.121	0.097	—	—
Brand F	-0.244	0.092	—	—
Brand G	-0.300	0.112	—	—
Responses Parameters				
Price	-1.216	0.226	-3.288	0.346
Promo	3.617	0.189	1.693	0.268
Loyalty	0.783	0.204	0.976	0.244
Reference Price Parameters				
Gain	0.232	0.107	0.208	0.102
Loss	-0.275	0.116	-0.224	0.106

Table 5.a. HB Estimates of Θ (*Curry Roux*)

	Brand A	Brand B	Brand C	Brand D	Brand E	Brand F	Brand G
Constant	**3.774** (3.194)	**2.257** (2.082)	0.031 (0.017)	-0.862 (-0.483)	-1.074 (-0.573)	-1.603 (-0.960)	**7.096** (5.507)
Interpurchase times	**-0.166** (-4.536)	**-0.145** (-3.993)	-0.004 (-0.109)	**-0.099** (-2.705)	-0.035 (-0.905)	**-0.229** (-6.440)	**-0.175** (-5.096)
Brand Sampled	**-0.122** (-2.055)	-0.078 (-0.923)	0.033 (0.388)	**-0.770** (-9.060)	-0.042 (-0.493)	**-0.214** (-2.543)	**-0.446** (-5.304)
Deal Proneness	**-1.864** (-2.505)	**-5.390** (-7.252)	-0.422 (-0.560)	**-1.737** (-2.345)	**1.923** (2.551)	**-2.393** (-3.251)	0.834 (1.135)
Purchase Quantity	-0.411 (-0.816)	**-1.385** (-2.388)	0.729 (1.249)	-0.109 (-0.188)	1.010 (1.905)	**1.273** (2.132)	-0.827 (-1.405)
Price Volatility	**-1.946** (-2.929)	**-1.422** (-2.177)	-0.299 (-0.493)	**-7.193** (-10.981)	**-1.481** (-2.263)	**-2.502** (-4.536)	**-2.996** (-5.624)
Brand Preference	0.330 (0.531)	**4.348** (7.151)	-0.058 (-0.089)	**1.398** (2.357)	**-1.526** (-2.302)	**1.272** (2.088)	**2.642** (5.407)

t-value in parentheses. Bold indicates significance at 0.05 level.

Table 5.b. HB Estimates of Θ (*Instant Coffee*)

	Brand A	Brand B	Brand C	Brand D
Constant	**4.294** (2.297)	1.282 (0.683)	-1.517 (-0.767)	**-3.977** (-2.139)
Interpurchase times	**-0.112** (-2.071)	0.006 (0.085)	**-0.122** (-2.026)	**-0.159** (-2.256)
Brand Sampled	**-0.192** (-2.962)	**-0.343** (-3.294)	**-0.410** (-3.474)	**-0.307** (-2.948)
Deal Proneness	**-0.405** (-2.009)	**-0.610** (-2.541)	-0.038 (-0.300)	**-0.588** (-2.712)
Purchase Quantity	**1.717** (4.149)	-0.795 (-1.620)	0.038 (0.173)	-0.386 (-0.937)
Price Volatility	0.793 (0.812)	**-2.535** (-2.423)	**-5.678** (-5.477)	**-1.617** (-1.989)
Brand Preference	-0.048 (-0.040)	**2.541** (2.161)	**-1.027** (-0.806)	**2.811** (2.248)

t-value in parentheses. Bold indicates significance at 0.05 level.

Figure 1.a Distribution of memory parameters (*Curry Roux*)

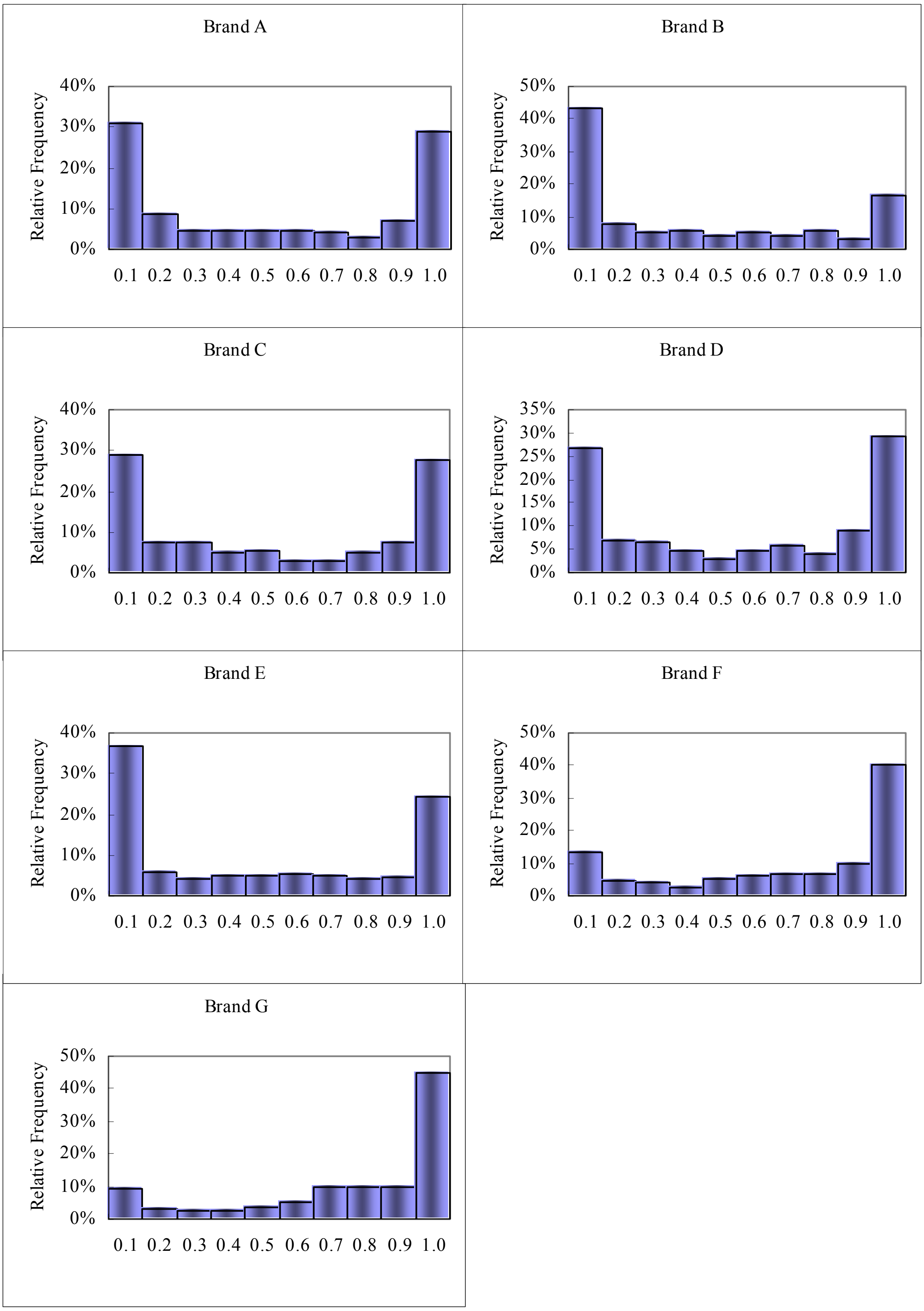


Figure 1.b Distribution of memory parameters (*Instant Coffee*)

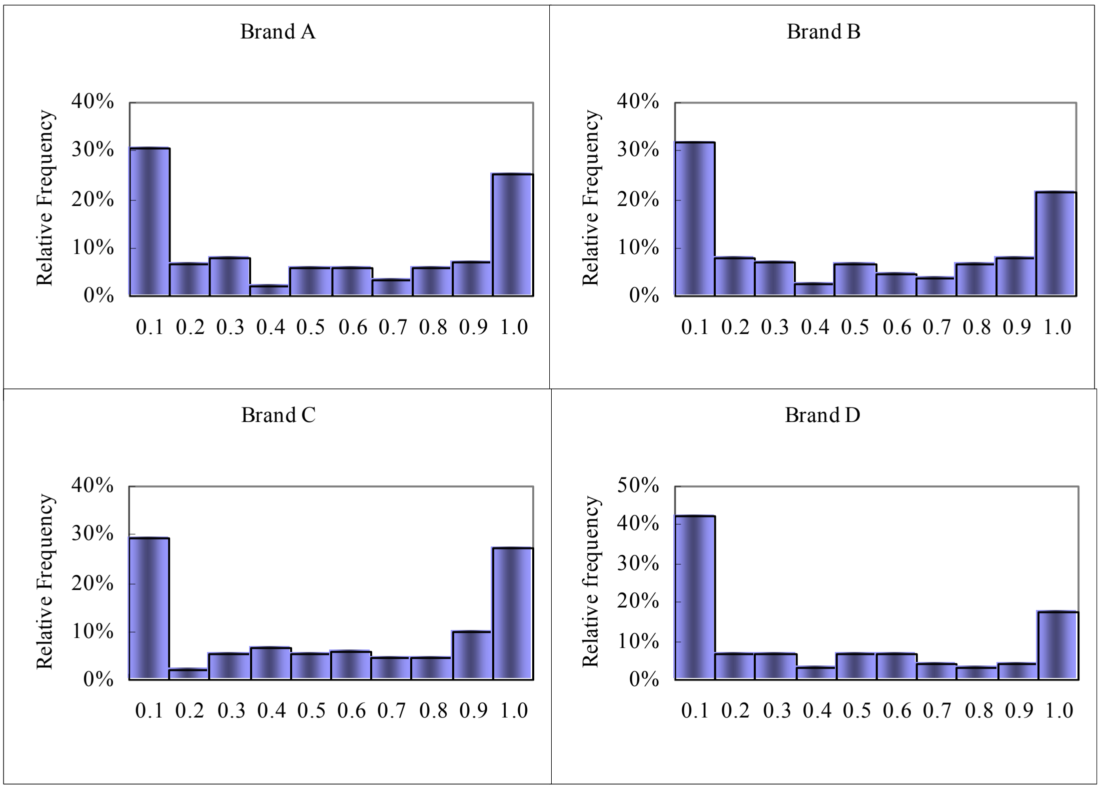


Figure 2 Distribution of Memory Parameters of Finite Mixture Model

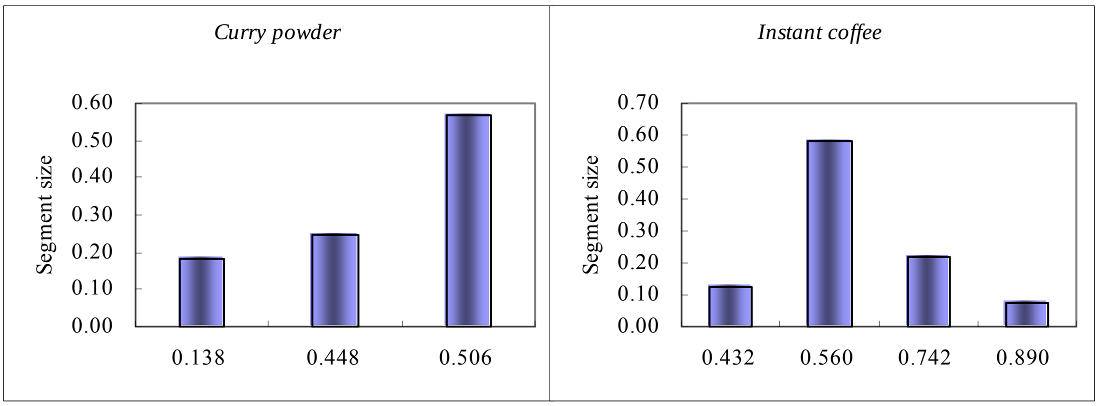
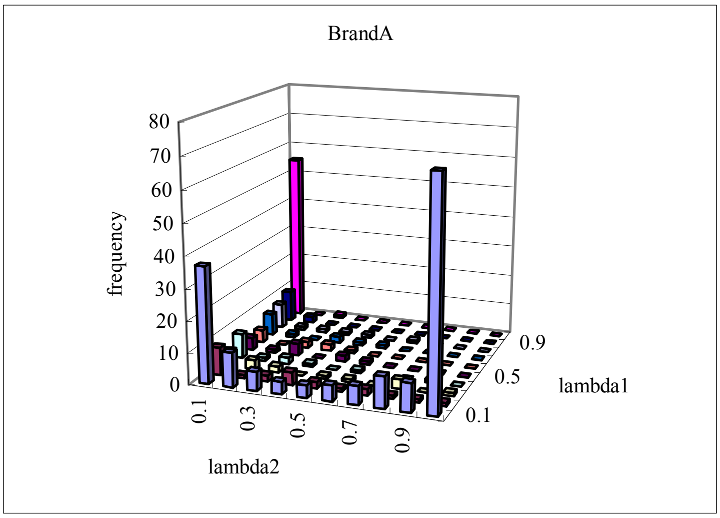
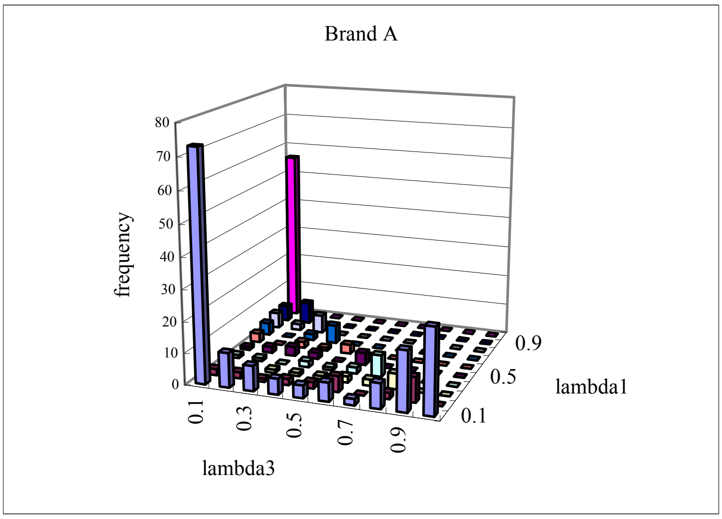


Figure 3.a Joint Distribution of lambda λ (Curry Roux)





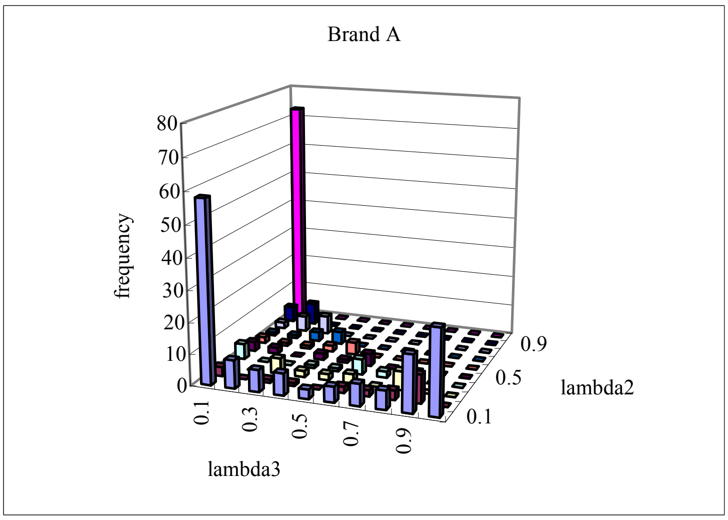
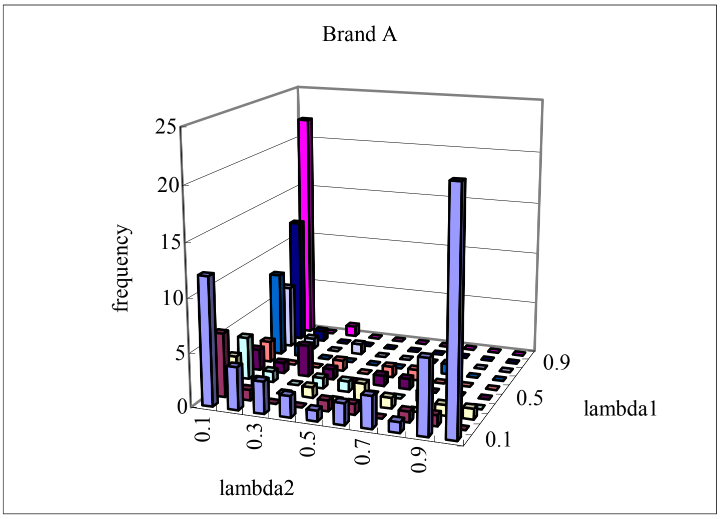
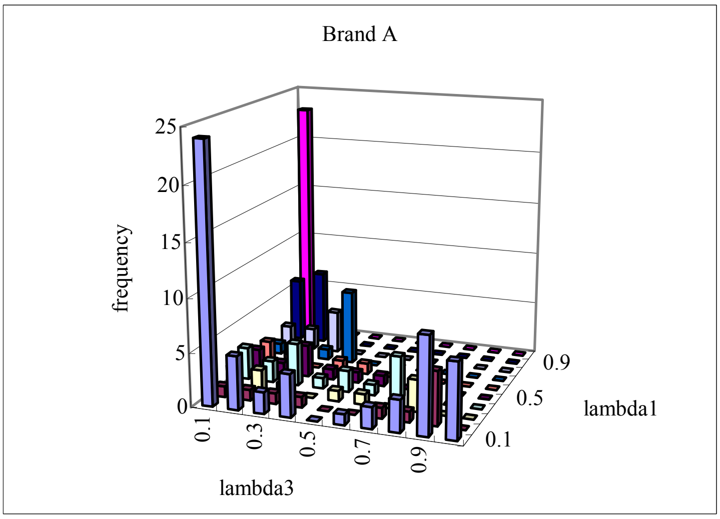


Figure 3.b Joint Distribution of λ (Instant Coffee)





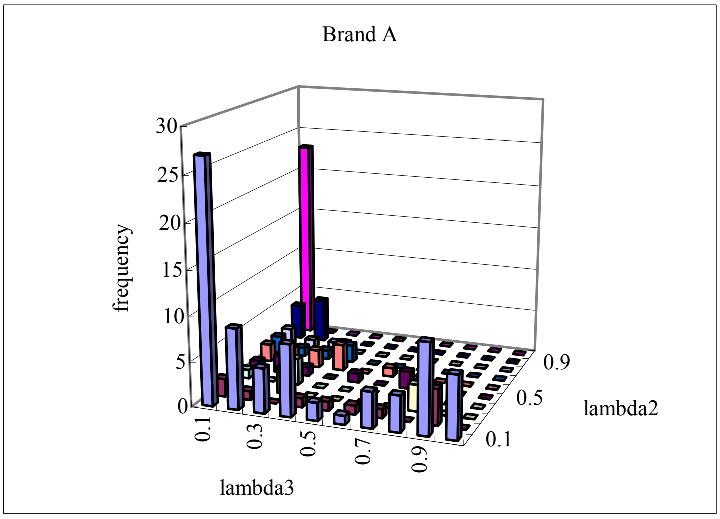
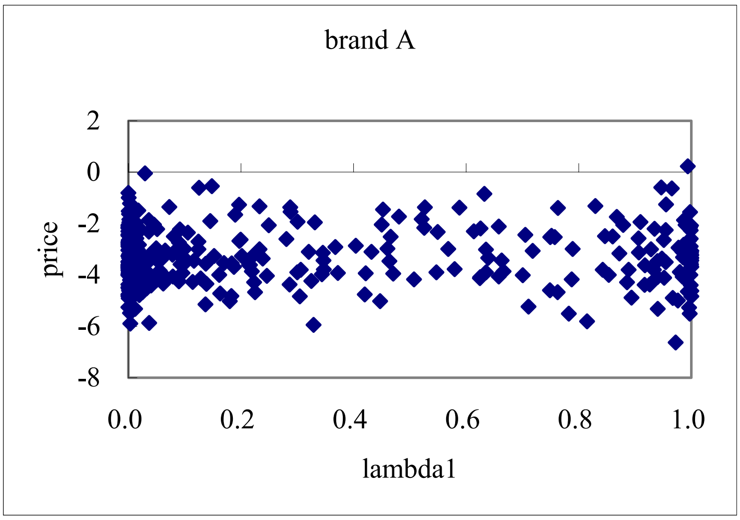
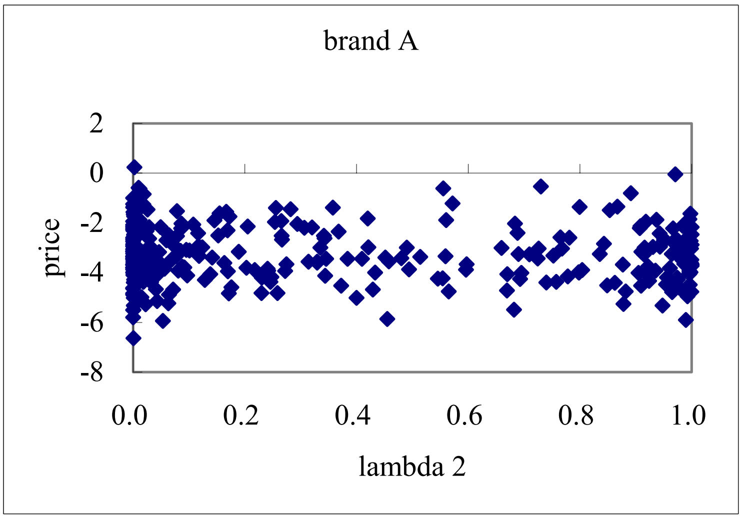


Figure 4.a Relation between lambda and price sensitivity (Curry Roux)





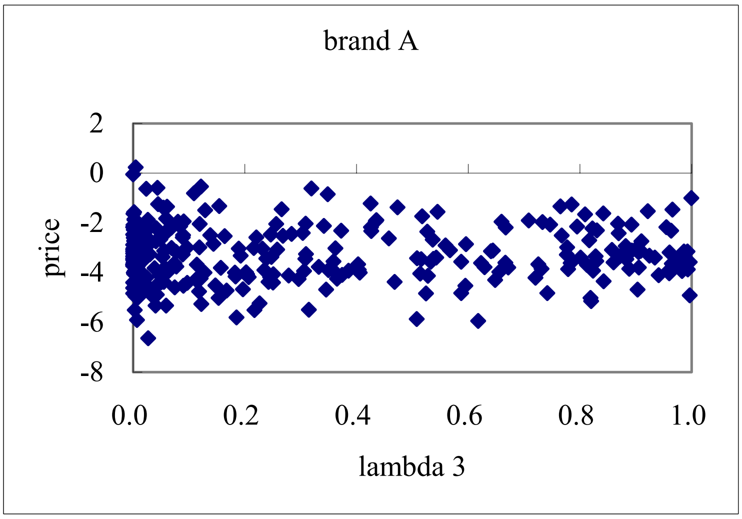


Figure 4.b Relation between lambda and price sensitivity (Instant Coffee)