MARKETING SCIENCE
Vol. 25, No. 4, July–August 2006, pp. 384–391
ISSN 0732-2399 | EISSN 1526-548X | 06 | 2504 | 0384

DOI 10.1287/mksc.1050.0181

Research Note

Estimating Heterogeneous Price Thresholds

Nobuhiko Terui
Graduate School of Economics and Management, Tohoku University, Sendai 980-8576, Japan, terui@econ.tohoku.ac.jp

Wirawan Dony Dahana
Graduate School of Economics, Osaka University, Osaka 560-0043, Japan, dony@econ.osaka-u.ac.jp

A brand choice model with heterogeneous price-threshold parameters is used to investigate a three-regime piecewise-linear stochastic utility function. The model is used to explore the relationships between aspects of consumer price sensitivity and price thresholds using hierarchical Bayes modeling with the Markov chain Monte Carlo (MCMC) method. This study contributes to the modeling literature on discontinuous likelihoods in choice models. The empirical application using our scanner panel data set shows that the reference effect and loss aversion are more marked after price thresholds are taken into heterogeneous price response models. Furthermore, loss aversion is attenuated by using price thresholds than by an aggregate (homogeneity) model without price thresholds.

Key words: discontinuous likelihoods; reference effect; price threshold; latitude of price acceptance; brand choice; Bayesian MCMC; heterogeneity; scanner panel data
History: This paper was received November 11, 2004, and was with the authors 5 months for 2 revisions; processed by Greg M. Allenby.

1. Introduction

Asymmetric price effects are an active area of research in marketing. Beginning with the work of Helson (1964) who first motivated the usefulness of including reference prices in sales and choice models, researchers have found asymmetric effects and the presence of price thresholds that reflect the presence of discontinuities in consumer behavior. The literature includes findings of asymmetric effect relative to the reference price (Kahneman and Tversky 1979), the disappearance of these effects in models that account for heterogeneity in price sensitivity (Chang et al. 1999, Bell and Lattin 2000), and models in which price has no effect within a price interval, i.e., a latitude of price acceptance (Gupta and Cooper 1992, Kalwani and Yim 1992, Kalyanaram and Little 1994, Han et al. 2001). However, models with threshold effects yielding intervals of price acceptance have yet to be estimated heterogeneously.

This paper introduces a heterogeneous choice model with reference prices, asymmetric responses, and price thresholds. It is characterized as a piecewise linear form so that consumers switch their utility structure accordingly as determined by the relationship between the sticker shock; i.e., the difference between retail price and reference price, and price thresholds. Heterogeneity is incorporated through a random-effects specification that allows determination of the relationship between the threshold values and consumer characteristics. MCMC algorithms are provided for estimation. The model is illustrated using a scanner panel data set of instant-coffee purchases.

Our model contributes to the modeling literature on discontinuous likelihoods. Price thresholds generate unconventional likelihood functions and discreteness at the threshold values that make it impossible to estimate with gradient-based methods such as maximum likelihood. As discussed by Gilbride and Allenby (2004) for models with consideration sets, the Bayesian method of data augmentation is ideally suited for handling discontinuities in the likelihood surface. Here, the thresholds are associated with alternative likelihood functions that result in an apportionment of household data to alternative models. Thus, the model contributes to the marketing literature on structural heterogeneity (Yang and Allenby 2000) by allowing for abrupt changes among alternative models within the unit of analysis.

Our empirical application adds to the literature on the presence of reference prices and loss aversion that initially began with effects uncovered in scanner data, for example, as indicated by Winer (1986), Mayhew and Winer (1992), and Putler (1992), and then Chang et al. (1999) and Bell and Lattin (2000) showed that reference price effects disappear when heterogeneity is incorporated. Our empirical application shows that the reference effect and loss aversion return—at least for the data used in this study—after price thresholds are taken into a heterogeneous model. In doing so, the degree of loss aversion is attenuated relative to results obtained using the homogeneity model without price thresholds.

The remainder of the paper is organized as follows. Section 2 presents our stochastic utility function and its consequent choice model—the threshold Probit model—and we develop a hierarchical Bayes modeling for it. Section 3 describes the application of scanner panel data to our model. Section 4 concludes this paper. The appendix explains details of the model estimation including the algorithm for hierarchical Bayes modeling via MCMC.

2. The Model

2.1. Threshold Probit Model and Hierarchical Bayes Modeling

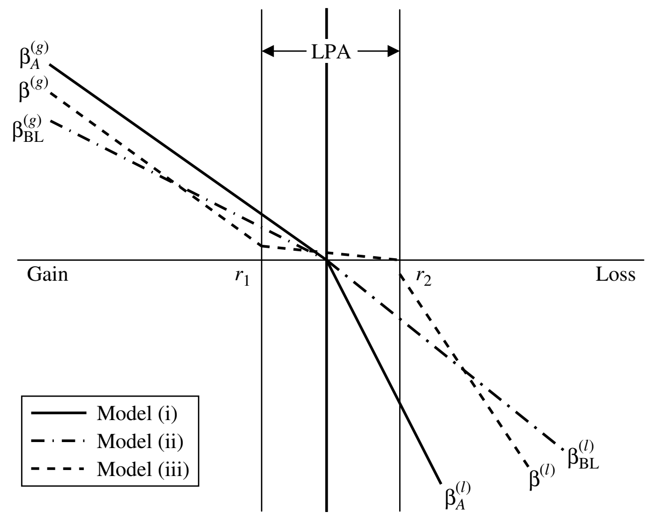
To specify the utility function, we assume that consumer h's utility to brand j at time t of purchase, U_{jht}, reflects a linear function of k kinds of explanatory variables. We also suppose that consumer h has a reference price RP_{jht} for brand j, and two price thresholds r_{1h} and r_{2h} ($r_{1h} < 0 < r_{2h}$). Consequently, we define the three regimes—gain "(g)," price acceptance "(a)," loss "(l)"—utility function to brand j as

$$U_{jht} = \begin{cases} u_{jh}^{(g)} + X_{jht}^{(g)}\beta_h^{*(g)} + \epsilon_{jht}^{(g)} \ (\equiv U_{jht}^{(g)}) \\ \quad \text{if } P_{jht} - \mathrm{RP}_{jht} \le r_{1h}, \\ u_{jh}^{(a)} + X_{jht}^{(a)}\beta_h^{*(a)} + \epsilon_{jht}^{(a)} \ (\equiv U_{jht}^{(a)}) \\ \quad \text{if } r_{1h} < P_{jht} - \mathrm{RP}_{jht} \le r_{2h}, \\ u_{jh}^{(l)} + X_{jht}^{(l)}\beta_h^{*(l)} + \epsilon_{jht}^{(l)} \ (\equiv U_{jht}^{(l)}) \\ \quad \text{if } r_{2h} < P_{jht} - \mathrm{RP}_{jht}, \end{cases} \tag{1}$$

where $X_{jht}^{(i)}$ is the k-dimensional row vector of explanatory variables allocated to regime i according to the level of sticker shock $P_{jht} - \mathrm{RP}_{jht}$ (P_{jht}: the retail price exposed to consumer h) at the occasion, and k-dimensional column vectors $\beta_h^{*(i)}$, $i = g, a, l$, represent different market responses around the reference price. Finally, let $\epsilon_{jht}^{(i)}$, $i = g, a, l$, respectively represent stochastic error components in the utility for each regime, and they are assumed to be independent across regimes.

Equation (1) is capable of simultaneously reflecting the presence of reference prices (Helson 1964), asymmetric effects (Kahneman and Tversky 1979), and price thresholds (Sherif et al. 1958). Asymmetries are present when the β coefficients in the different regions take on different values. The price thresholds, (r_{sh}, $s = 1, 2$), identify different regimes associated with the data. The model captures a latitude of price acceptance (LPA hereafter) $L_h = (r_{1h}, r_{2h}]$ where prices may not affect choice if β_{price} is sufficiently small.

Figure 1 **Model for Price Thresholds and Market Responses**



Notes. Model (i): aggregate (homogeneous) reference price Probit model without a threshold; Model (ii): two-regime heterogeneous reference price Probit model without a threshold (Bell and Lattin 2000); Model (iii): three-regime heterogeneous reference price Probit model with thresholds.

Figure 1 displays our utility function. In order for the second regime of the utility function (1) to be characterized as the LPA, and for r_{1h} and r_{2h} to be interpreted as price thresholds literally, i.e., for our proposed model to be recognized as a price-threshold model, we impose the restriction of insensitiveness on the price-response parameter in the LPA regime as $\beta_{hp}^{(a)} \sim N(0, (\sigma_{hp}^{(a)})^2)$, which is an element of $\beta_h^{*(a)}$. Under the assumption of consumer homogeneity, Kalyanaram and Little (1994) estimated symmetric LPA that had an insignificant price-response estimate over this range. On the other hand, Han et al. (2001) find a significant response parameter estimate in the LPA. We note that their framework differs from ours particularly in the respect that they use the same data set throughout regimes. In contrast, our modeling method allocates data into three regimes.

Following the standard multinomial-brand choice model, we assume that consumer h is observed to make a choice at the period t between m alternatives ($c_{ht} = j$). That choice is driven by the relative utility from the last brand $y_{jht} = U_{jht} - U_{mht}$ among latent utilities U_{nht}, $n = 1, 2, \ldots, m$. We consider a panel of H households observed over T_h periods for each household. Conditional on the vector of threshold $r_h = (r_{1h}, r_{2h})$ for consumer h, the underlying latent structure at period t when the latent utility induced by the choice belongs to the regime i, the so-called "within subject model for regime i" is expressed in the form of $(m-1)$ dimensional multiple regression,

$$y_{ht}^{(i)} = X_{ht}^{(i)}\beta_h^{(i)} + \varepsilon_{ht}^{(i)}; \qquad \varepsilon_{ht}^{(i)} \sim N(0, \Lambda^{(i)}),$$
$$h = 1, \ldots, H,\ t = 1, \ldots, T_h^{(i)},\ i = g, a, l, \tag{2}$$

where $y_{ht}^{(i)}$ is the $(m-1)$ dimensional relative-utility vector, $X_{ht}^{(i)}$ is the $(m-1)\times(k+m-1)$ explanatory variable matrix measured from the last brand, $\beta_h^{(i)}$ is $(k+m-1)$ dimensional coefficient vector, $\varepsilon_{ht}^{(i)}$ is the $(m-1)$ dimensional stochastic-error vector, and we have the relation $T_h^{(g)}+T_h^{(a)}+T_h^{(l)}=T_h$. Consumers' heterogeneity is formulated by way of a hierarchical regression model, as proposed by Rossi et al. (1996). Regarding price thresholds, we employ the model

$$r_{1h}=Z_h^r\phi_1+\eta_{1h};\qquad r_{2h}=Z_h^r\phi_2+\eta_{2h},\qquad h=1,\ldots,H,\quad (3)$$

where Z_h^r is a vector of d kinds of household specific variables. We assume that $r_{1h}<0<r_{2h}$ for identification and $\eta_{sh}\sim N(0,\sigma_{s\eta}^2)$ for $s=1,2$. We also set a hierarchical structure of "between subjects model for regime i" for the market-response parameter

$$\beta_h^{(i)}=\Delta^{(i)}Z_h^\beta+v_h^{(i)};\qquad v_h^{(i)}\overset{\text{i.i.d}}{\sim}N\big(0,V_\beta^{(i)}\big),\qquad h=1,\ldots,H,\ i=g,a,l,\quad (4)$$

where Z_h^β contains another vector of d' kinds of household specific variables. In particular, we note that price response $\beta_{hp}^{(a)}$ in the LPA is assumed a priori to have zero mean in (4).

Following the utility function defined as (1), consumer h's probability of choosing brand j is written as

$$\Pr\{c_h=j\}=\begin{cases}\Pr\{y_{jh}^{(g)}=\max\big(y_{1h}^{(g)},\ldots,y_{m-1h}^{(g)}\big)>0\}\\ \quad\text{if } P_{jht}-\text{RP}_{jht}\le r_{1h},\\ \Pr\{y_{jh}^{(a)}=\max\big(y_{1h}^{(a)},\ldots,y_{m-1h}^{(a)}\big)>0\mid R\}\\ \quad\text{if } r_{1h}<P_{jht}-\text{RP}_{jht}\le r_{2h},\\ \Pr\{y_{jh}^{(l)}=\max\big(y_{1h}^{(l)},\ldots,y_{m-1h}^{(l)}\big)>0\}\\ \quad\text{if } r_{2h}<P_{jht}-\text{RP}_{jht},\end{cases}\quad (5)$$

where $\Pr\{y_{jh}^{(a)}=\max(y_{1h}^{(a)},\ldots,y_{m-1h}^{(a)})>0\mid R\}$ indicates the choice probability under the restriction on price response $\beta_{hp}^{(a)}\sim N(0,(\sigma_{hp}^{(a)})^2)$ in the LPA regime.

2.2. Discontinuous Likelihoods for Thresholds and Their Modeling

Our model includes a threshold variable of sticker shock $P_{jht}-\text{RP}_{jht}$ in the model, and it induces discontinuity in the likelihoods, as shown in Figure 1. Conditional on the value of $r_h=(r_{1h},r_{2h})$, the operation generates the latent utility with three different likelihood functions for consumer h. The independence assumption of stochastic errors across regimes generates the likelihood function of $(\beta_h^{(i)},\Lambda^{(i)})$ for consumer h:

$$\prod_{i=g,a,l}\Big\{\prod_{t\in R^{(i)}(r_h)}|\Lambda^{(i)}|^{-1/2}\exp\{-\tfrac{1}{2}\big(y_{ht}^{(i)}-X_{ht}^{(i)}\beta_h^{(i)}\big)'\cdot(\Lambda^{(i)})^{-1}\big(y_{ht}^{(i)}-X_{ht}^{(i)}\beta_h^{(i)}\big)\}\Big\},$$

where each datum is assigned to one regime $R^{(i)}(r_h)$, and it holds that $R^{(g)}(r_h)\cup R^{(a)}(r_h)\cup R^{(l)}(r_h)=T_h$.

In turn, conditional on $(\beta_h^{(i)},\Lambda^{(i)})$, we take the above as a function of r_h to compose the likelihood function of price thresholds. Under the assumption of independent choice behavior across consumers, we have the conditional likelihood function of $\{r_h\}$ by taking products over respective consumers as

$$L(\{r_h\};\{I_{ht}\},\{X_{ht}\}\mid\{\beta_h^{(i)}\},\{\Lambda^{(i)}\})\propto\prod_{h=1}^{H}\Big\{\prod_{i=g,a,l}\Big\{\prod_{t\in R^{(i)}(r_h)}|\Lambda^{(i)}|^{-1/2}\exp\{-\tfrac{1}{2}\big(y_{ht}^{(i)}-X_{ht}^{(i)}\beta_h^{(i)}\big)'\cdot(\Lambda^{(i)})^{-1}\big(y_{ht}^{(i)}-X_{ht}^{(i)}\beta_h^{(i)}\big)\}\Big\}\Big\},\quad (6)$$

where $\{I_{ht}\}$ means the personal observed choice data. Coupled with the prior (3) expressed as hierarchical structure, we can apply the Metropolis-Hasting sampling algorithm for price thresholds to obtain conditional posterior $r_h\mid\{I_{ht}\},\{X_{ht}\},\{\beta_h^{(i)}\},\Lambda^{(i)},\{Z_h\},\phi,\Sigma_\eta$. Prior distributions and MCMC estimation procedures for these hierarchical Bayes models are described in the appendix.

The proposed model is difficult to estimate with conventional methods because the likelihood is not differentiable in r_h. Conventional maximum likelihood estimation collapses and classical asymptotic distribution theory is not operative on this parameter, and it must therefore rely on an adaptive approach: conditioned on some specific value of r_h, estimation is conducted through extensive use of goodness-of-fit criterion, like AIC, in the way of a grid search to find the estimate that minimizes the criterion.

In the price-threshold literature, Kalyanaram and Little (1994) assume symmetric LPA by setting $r_{h1}=r_{h2}=r$ (homogeneous), and they estimate not the threshold but the length of LPA, where several possible lengths are specified a priori, and they seek the LPA that has an insignificant estimate of price response inside this range. Kalwani and Yim (1992) take a grid search approach. This approach does not allow statistical inferences such as testing hypotheses and confidence intervals. Another approach is Bayesian inference. For example, Ferreira (1975), Geweke and Terui (1993), and Chen and Lee (1995) provided Bayesian approaches to deal with discontinuous likelihoods in econometrics and nonlinear time-series models. Our proposed method, using more updated tools, directly models price thresholds in choice models in a general way and conducts coherent statistical inference on them under a small-sample situation. The advantage of our approach is that it avoids such specification search and can accommodate heterogeneity.

3. Empirical Application to Scanner Panel Data

3.1. Data, Variables, and Model Specification

Data and Variables. Video Research Inc., Japan, supplied scanner panel data for the instant-coffee category. In all, 2,840 records for 197 panels during 1990–1992 were available. We assume that five national brands existed in the market during the tracking period. We deal with five primary brands in the market: A, B, C, D, and E. Table 1 provides descriptive information about the data. Brand B has the maximum share—over 48.03%; the minimum share—approximately 5.74%—is for brand E. We rescale all prices as yen/100g to equalize quantitative differences for each package of the five brands.

Variables included in the analysis are:

• Explanatory Variables: $X = [\text{Constant}, \text{Price}, \text{Display}, \text{Feature}, \text{Brand Loyalty}]$, where Price is the log(price); Display and Feature are binary values; and Brand Loyalty is a smoothing variable over past purchases proposed by Guadagni and Little (1983) as $\text{GL}_{jht} = \alpha \text{GL}_{jht} + (1-\alpha) I_{jh,\,t-1}$, where a grid search (Keane 1997) is applied to fix the smoothing parameter as 0.75 based on the criterion of minimum marginal likelihood.

• Household Specific Variables: $Z^r = [\text{Constant}, \text{Dprone}, \text{Pfreq}, \text{ARP}, \text{BL}]$, where Dprone is the deal proneness defined as the proportion of purchase (of any five brands) made on promotion (Bucklin and Gupta 1992, Han et al. 2001); Pfreq is the shopping frequency (three categories; H_3 of Kalyanaram and Little 1994: shopping frequency is a factor to affect the price thresholds); ARP represents the measure of average reference-price level as defined by $\text{ARP}_h = \sum_{j=1}^{m} (\sum_{t=1}^{T_h} \log(\text{RP}_{jht})/T_h)/m$; and BL represents the brand loyalty measure defined as $\text{BL}_h = \max_j (\sum_{t=1}^{T_h} \text{GL}_{jht}/T_h)$, both of which are used by Kalyanaram and Little (1994) and Han et al. (2001). We further define

• Household Specific Variables: $Z^\beta = [\text{Constant}, \text{Hsize}, \text{Expend}]$, where Hsize is 1–6 (number of household members) and Expend is nine categories (shopping expenditure/month) used in Rossi et al. (1996).

Table 1 Descriptive Statistics for Data

Alternative	Choice share	Average price	Time displayed (%)	Time featured (%)
Brand A	0.138	623.5	0.264	0.423
Brand B	0.480	632.9	0.135	0.294
Brand C	0.099	601.3	0.317	0.405
Brand D	0.225	693.2	0.182	0.344
Brand E	0.057	902.4	0.191	0.286

Specification of Reference Price. The literature presents some conceptualizations for the reference price RP_{jht}. Following Briesch et al. (1997), we employ a brand-specific reference price. We also consider four kinds of reference prices of two categories: memory-based and stimulus-based. That is,

(1) Memory-based as (A) $\text{RP}_{jht} = P_{jh,\,t-1}$, the price at its last purchase and (B) $\text{RP}_{jht} = \alpha \text{RP}_{jh,\,t-1} + (1-\alpha) \cdot P_{jh,\,t-1}$, the smoothed price over previous purchases;

(2) Stimulus-based as (C) $\text{RP}_{jht} = P_{kht}$, where k means the brand at the last purchase, and (D) $\text{RP}_{jht} = P_{rht}$, where r indicates the price of a brand chosen randomly at the time of purchase.

Model Specification. We consider three models for comparison with other candidates: (i) aggregate (homogeneous) reference-price Probit model without a threshold (Winer 1986, Mayhew and Winer 1992, Putler 1992); (ii) two regimes' heterogeneous reference price Probit model without a threshold (Chang et al. 1999, Bell and Lattin 2000); and our proposed (iii) three-regime heterogeneous reference-price Probit model with thresholds. With four types of reference prices for each model above, summing to 12 models are considered for comparison.

Table 2 describes logs of marginal likelihood for model comparison. The model (C)(iii): stimulus-based RP(C) three regime heterogeneous Probit with thresholds is most supported by marginal likelihood criterion. Consequently, some evidence supports the hypothesis of price-threshold existence that was discussed by Kalyanaram and Little (1994) even by using the models incorporating heterogeneous consumers. The summary statistics of parameter estimates for that model are provided in Table 3. To save space, estimates of constant terms are not listed here.

Our chosen stimulus-based RP is compatible with previous studies of Hardie et al. (1993), stating that it seems likely that characteristics of the brand last purchased will be more available in memory in particular with consumer package goods, and moreover Rajendran and Tellis (1994) discuss the reasons of appropriateness for stimulus-based RP in detail.

Table 2 Log of Marginal Likelihood

	(i)	(ii)	(iii)
Memory-based			
A	−36,700.434	−34,981.349	−33,829.653
B	−36,711.345	−35,836.148	−34,176.325
Stimulus-based			
C	−36,048.875	−34,716.514	**−33,658.890**
D	−36,140.055	−35,372.912	−34,103.315

(i) Conventional Probit without both heterogeneity and thresholds.
(ii) Two-regime Probit with heterogeneity and without thresholds.
(iii) Three-regime Probit with both heterogeneity and thresholds.

Table 3 Model Specification and Parameter Estimates ($RP_{jht} = P_{kht}$)

		With heterogeneity	
	(i) Without heterogeneity	(ii) Without threshold	(iii) With thresholds
Gain regime			
Price	−4.251* [0.556]	−1.637 (−3.875)	−3.142 (−4.749)
Display	0.526* [0.184]	0.744 (2.793)	1.155 (2.251)
Feature	1.370* [0.353]	1.665 (7.830)	1.773 (4.021)
Brand Loyalty	1.112* [0.426]	1.199 (5.161)	0.413 (1.342)
LPA			
Price			−0.907 (−1.711)
Display			0.646 (2.969)
Feature			1.595 (7.361)
Brand Loyalty			3.232 (12.601)
Loss regime			
Price	−7.021* [1.594]	−2.622 (−3.034)	−4.518 (−4.725)
Display	0.688* [0.212]	1.046 (3.135)	1.086 (2.972)
Feature	1.720* [0.389]	3.057 (8.044)	2.679 (7.366)
Brand Loyalty	1.293* [0.460]	3.483 (8.805)	5.427 (12.846)
Thresholds			(−0.113, 0.138]
LML	−36,041.732	−34,716.514	−33,658.890

(i) Conventional Probit without both heterogeneity and thresholds.
(ii) Two-regime Probit with heterogeneity and without thresholds.
(iii) Three-regime Probit with both heterogeneity and thresholds.
*Significant at 0.95 HPD region. The number inside [ ] in the model (i) means standard deviation, and ( ) for (ii) and (iii) means the quasi t-value calculated from the frequency distribution of household's estimates.

3.2. Reference Effects, Loss Aversion, and Price Thresholds

Our analysis indicates the presence of asymmetric reference effects, loss aversion, and heterogeneity. This finding is different from Chang et al. (1999) and Bell and Lattin (2000) where asymmetric effects were negligible when heterogeneity and price thresholds were incorporated into the model.

Corresponding to each type of RP, (A)–(D), we have four sets of models: (A)(i), (A)(ii), (A)(iii) through (D)(i), (D)(ii), (D)(iii). Based on the most supported RP of (C), price-response estimates at the aggregated level show some loss aversion because, in the case of (C)(iii), we have the estimate of regime (g)—gain—which is smaller than that of regime (l)—loss—as (gain: −3.142, loss: −4.518). The same applies to the model (i) (gain: −4.251, loss: −7.021) and (ii) (gain: −1.637, loss: −2.622). We further observe the following: (i) aggregate (homogeneous) Probit without a threshold shows loss aversion most clearly at the aggregated level; (ii) two-regime heterogeneous Probit without a threshold is the most vague; and (iii) our proposed three-regime heterogeneous Probit with thresholds yields performance between (i) and (ii).

We found the presence of asymmetric reference effects and loss aversion after price thresholds were incorporated into the model for this data set. This observation is robust relative to other types of reference price, except for (A)(ii), (B)(ii), and (B)(iii).

Following Bell and Lattin (2000, p. 190, Figure 1), we present intuitive reasoning for our results. First, Bell and Lattin (2000) discuss the relation between models (i) and (ii) relative to loss aversion: model (i) has steeper slopes in the loss regime ($|\beta_A^{(l)}| > |\beta_{\mathrm{BL}}^{(l)}|$) as well as the gain regime ($|\beta_A^{(g)}| > |\beta_{\mathrm{BL}}^{(g)}|$) than model (ii), however, the difference ($|\beta_A^{(l)}| - |\beta_{\mathrm{BL}}^{(l)}|$) is larger in the loss regime than ($|\beta_A^{(g)}| - |\beta_{\mathrm{BL}}^{(g)}|$) is in the gain regime. Therefore, loss aversion would decrease or disappear if heterogeneity were incorporated into the model.

Next, we examine the relation between models (ii) and (iii). The response functions assumed in each model are shown in Figure 1; we note that the response function of model (ii), corresponding to the model of Bell and Lattin (2000), has two regimes without price thresholds. Two response functions are kinked at the zero of sticker shock. Our model (iii), however, has three regimes. We suppose that the true response function has LPA (insensitive range) limited by the price thresholds, as shown in Figure 1, and that the data are observed along with three lines.

If we fit only two response functions, as employed in Bell and Lattin (2000), to data over positive and negative sides of sticker-shock domain separately, the fitted slopes ($\beta_{\mathrm{BL}}^{(g)}, \beta_{\mathrm{BL}}^{(l)}$) will be less steep than the slopes ($\beta^{(g)}, \beta^{(l)}$) of the three-regime model in the gain-and-loss regimes. That is, $|\beta_{\mathrm{BL}}^{(g)}| < |\beta^{(g)}|$, $|\beta_{\mathrm{BL}}^{(l)}| < |\beta^{(l)}|$ because the fitted slopes are calculated so as to catch up with the data inside LPA, to which a flatter line is applied by definition. For that reason, a model with price thresholds is likely to be more responsive than a model without a threshold for both loss and gain regimes. Table 3 presents empirical evidence to support this conjecture, which is robust with respect to the choice of reference price.

On the other hand, that discussion indicates no tendency for loss aversion between models with and without thresholds. However, relying on Kahneman and Tversky (1979), it would be likely for $\beta^{(g)} - \beta^{(l)}$ (we define it as the degree of loss aversion) to be positive in most cases because we are trying to model individual consumers' responses. In fact, we used our data to calculate the posterior probability $\Pr\{\beta^{(\mathrm{Gain})} - \beta^{(\mathrm{Loss})} \mid \text{data}\}$ at the aggregated level for models (i), (ii), and (iii). Figure 2 displays the posterior density of the difference $\beta^{(\mathrm{Gain})} - \beta^{(\mathrm{Loss})}$. The distributions agree with Prospect theory (Kahneman and Tversky 1979), and the degree of loss aversion is strongest for model (i) in the sense that the posterior probability of loss aversion $\Pr\{\beta^{(\mathrm{Gain})} - \beta^{(\mathrm{Loss})} > 0 \mid \text{data}\}$ is largest; it is weakest for model (ii). Model (iii) is located between them. The relation between (i) and (ii) is consistent with the results of Bell and Lattin (2000). The relation between (ii) and (iii) implies that the degree of loss

Figure 2 Posterior Distribution for Loss Aversion

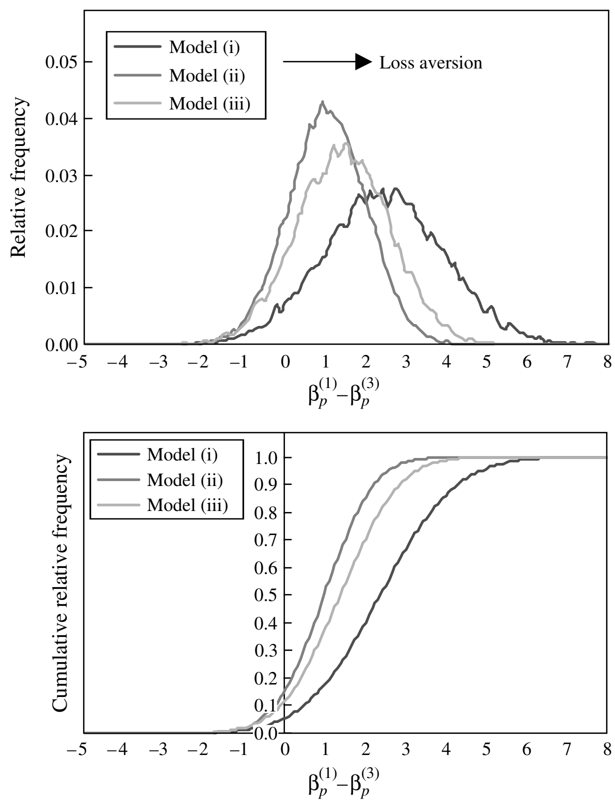
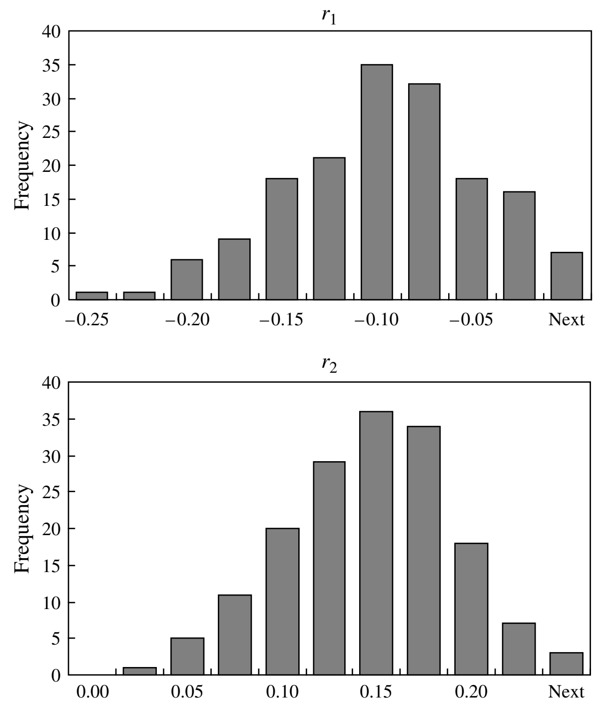


Figure 3 Heterogeneous Distribution of Price Thresholds



aversion would increase after incorporation of heterogeneous price thresholds into the models. These results are robust to the selection of reference price, except for (B).

The relationship between household-specific variables and their market-response parameters was significantly estimated. However, they are not reported here to save space.

3.3. Price Thresholds and Their Hierarchical Structure

Figure 3 shows the frequency distribution of Bayes estimates $\{\hat{r}_{1h}, h = 1, \ldots, H\}$ and $\{\hat{r}_{2h}, h = 1, \ldots, H\}$, where $\hat{r}_{\bullet h}$ is the mean of posterior distribution of threshold parameters for household h. Both distributions exhibit skewness showing distinct features of each other. The average distances from zero are different: −0.113 for the lower threshold $\hat{r}_{1h}$ and 0.138 for the upper threshold $\hat{r}_{2h}$. For that reason, the symmetric LPA around zero could fail to reflect heterogeneity in the brand choice study.

Table 4 shows that all hierarchical Bayes estimates of regression coefficients ϕ_1 and ϕ_2 of (3) on household-specific variables are significant in the sense that the 95% HPD region does not include zero. From these estimates, we first note that LPAs relative to households are not symmetric around the zero levels of their respective sticker shocks. We also observe the following: (1) For Pfreq versus LPA, the mean level of LPA increases with the purchase frequency, which is consistent with previous empirical findings; e.g., on the hypothesis H_3 of Kalyanaram and Little (1994): "Pfreq is a factor to affect price thresholds." (2) For Dprone versus LPA, the mean level of LPAs decreases accordingly as the increase in deal proneness, which is compatible with Han et al. (2001). (3) For ARP versus LPA, the estimates show that the household with high ARP value has wider LPA, which is consistent with the result to the hypothesis H_2 of Kalyanaram and Little (1994): "ARP is a factor to affect price thresholds." (4) For BL versus

Table 4 Hierarchical Regression Coefficients on Household Data (Price Thresholds)

	(C)(iii)	
Variable	r_1	r_2
Const	−0.292*	0.321*
	[0.053]	[0.041]
Pfreq	0.028*	−0.010*
	[0.004]	[0.005]
Dprone	0.042*	−0.033*
	[0.006]	[0.006]
ARP	−0.049*	0.026*
	[0.006]	[0.006]
BL	−0.149*	0.165*
	[0.017]	[0.017]

*Significant at 0.95 HPD region. The number inside [ ] means posterior standard deviation.

LPA, we observe that the high brand-loyalty household has wider LPA, which is consistent with results by hypothesis H_4 of Kalyanaram and Little (1994): "BL is a factor to explain price thresholds."

4. Concluding Remarks

This study introduced a three-regime piecewise-linear stochastic utility function with two price thresholds, and proposed a class of brand-choice model—threshold Probit model—under a framework of a continuous mixture modeling for heterogeneous consumers. Price thresholds generate unconventional discontinuous likelihood functions in the analysis, and they create difficulties in estimation. However, our method directly models thresholds for choice models in a general manner, and coherent statistical inference on the thresholds can be done particularly when the number of samples is scarce.

Our empirical application, as far as our data set is concerned, shows that our proposed model is superior to a conventional linear-utility-based Probit model without price thresholds as well as aggregate Probit models; the reference effect and loss aversion are observed after price thresholds are incorporated into a heterogeneous price-response model, but that loss aversion is attenuated then by an aggregate (homogeneous) model without price thresholds. Moreover, through hierarchical modeling, we investigated explanatory factors of heterogeneous price thresholds. Some are consistent with previous studies that assumed consumer homogeneity.

Individual heterogeneity makes it possible to design a customized strategy, which has been demanded recently as we can see, for example, in Gal-Or and Gal-Or (2005) for customized advertising. As an application of our model, we explored a possibility of customized pricing strategy based on the knowledge of price thresholds for respective households. Our limited simulation study showed that customized pricing could yield greater profits than flat pricing, and optimal pricing levels are provided at the lower price thresholds for discounting, and at the upper price thresholds for a price hike. However, this is not reported here to save space.

Our analysis can be generalized further in several aspects. For example, heterogeneous modeling could be applied to smoothing parameters of the reference price derived from exponentially weighted past prices. Moreover, it can be extended to accommodate the dynamic relation on a utility function, for example, such as Leichty et al. (2005) recently proposed. Together with a comprehensive discussion of customized pricing mentioned above, we leave these problems for future research.

Acknowledgments

Nobuhiko Terui acknowledges the financial support from the Japanese Ministry of Education Scientific Research Grant (C)15530137. The authors appreciate helpful comments from the area editor and three anonymous reviewers.

Appendix. Estimation of the Model: Markov Chain Monte Carlo Algorithm for Hierarchical Bayes Modeling of Threshold Probit Model

As for model calibration, extending the framework of Rossi et al. (1996), we employ hierarchical Bayes modeling to implement the threshold Probit model (6). Given the value of r_h, according to the level of consumer h's sticker shock $P_{jht} - \mathrm{RP}_{jht}$, we first allocate data $\{X_{ht}\}$ of the explanatory variable to make $\{X_{jht}^{(i)},\ i = g, a, l\}$ at each purchase occasion. The corresponding latent utility vector $\{y_{ht}^{(i)},\ i = g, a, l\}$ is generated based on personal choice data $\{I_{ht}\}$ (the index of observed choices) using the algorithm of the Bayesian Probit model (Rossi et al. 1996, pp. 338–339) applied to each regime. Then, except for price threshold $r_h \mid -$, we can use conditional posterior distributions: $y_{ht}^{(i)} \mid \{I_{ht}\}, \{X_{ht}\}, \beta_h^{(i)}, \Lambda^{(i)}, r_h$, $\beta_h^{(i)} \mid \{y_{ht}^{(i)}\}, \{X_{ht}^{(i)}\}, \Lambda^{(i)}, \Delta^{(i)}, V_\beta^{(i)}, z_h, r_h$, $(\Lambda^{(i)})^{-1} \mid \{y_{ht}^{(i)}\}, \{X_{ht}^{(i)}\}, \{\beta_h^{(i)}\}, \{r_h\}$, $\Delta^{(i)} \mid \{\beta_h^{(i)}\}, V_\beta^{(i)}, \{z_h\}, \{r_h\}$, and $(V_\beta^{(i)})^{-1} \mid \{\beta_h^{(i)}\}, \Delta^{(i)}, \{z_h\}, \{r_h\}$ for $i = g, a, l$. The variance $(\sigma_{hp}^{(a)})^2$ for $\beta_{hp}^{(a)} \sim N(0, (\sigma_{hp}^{(a)})^2)$ in the LPA was fixed as 0.01.

As for hierarchical modeling for the price threshold, the lower (negative) threshold is modeled as $r_{1h} = z_h \phi_1 + \eta_{1h}$; $\eta_{1h} \sim N(0, \sigma_{1\eta}^2)$ and stacking over all households $h = 1, \ldots, H$, leads to matrix notation $r_1 = Z\phi_1 + \eta_1$; $\eta_1 \sim N_H(0, \Sigma_{1\eta})$. We first set prior distributions on (ϕ, Σ_η) as

$$\phi_1 \mid r_1, \Sigma_{1\eta} \sim N(\phi_{10}, \Sigma_{1\phi}); \quad \phi_{10} = 0,\ \Sigma_{1\phi} = 0.01 I_{d'}$$

$$\Sigma_{1\eta}^{-1} \sim \text{Wishart}(\nu_{1\eta 0}, V_{1\eta 0}); \quad \nu_{1\eta 0} = d' + 4,\ V_{1\eta 0} = \nu_{1\eta 0} I_H.$$

Then we have conditional posterior distributions

$$\phi_1 \mid \{r_{1h}\}, \{z_h\}, \Sigma_{1\eta} \sim N(\bar{\phi}, (Z'Z + \Sigma_{1\phi}^{-1})^{-1}),$$

where

$$\bar{\phi}_1 = (Z'Z + \Sigma_{1\phi}^{-1})^{-1}[Z'Z\hat{\phi}_1 + \Sigma_{1\phi}^{-1}\phi_{10}] \quad \text{and} \quad \hat{\phi}_1 = (Z'Z)^{-1}Z'r_1.$$

$$\Sigma_{1\eta}^{-1} \mid \{r_{1h}\}, \{z_h\}, \phi_1 \sim \text{Wishart}\left(\nu_{1\eta 0} + H, V_{1\eta 0} + \sum_t (r_{1t} - Z_t\hat{\phi}_1)(r_t - Z_t\hat{\phi}_1)'\right).$$

The same formulations apply to upper (positive) price threshold r_{2h}.

Next, as for conditional posterior $\{r_h\} \mid \{I_{ht}\}, \{X_{ht}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \{z_h\}, \phi, \Sigma_\eta$, we use the likelihood function $L(\{r_h\}; \{I_{ht}\}, \{X_{ht}\} \mid \{\beta_h^{(i)}\}, \{\Lambda^{(i)}\})$ of (8) in §2 jointly with the prior $p(\{r_h\} \mid \phi, \Sigma_\eta) \sim N(Z\phi, \Sigma_\eta)$ defined above, and we employ Metropolis-Hastings sampling with random walk algorithm as the following steps:

Denote $\{r_h\} = r$; then

(1) $i = 0$. Set $r^{(0)}$.

(2) $i > 1$; $z \sim N(0, \sigma_{RW}^2 I_H)$ and set $r = r^{(i-1)} + z$.

(3)

$$\alpha(r^{(i-1)}; r) = \min\left\{\frac{p(r \mid \phi, \Sigma_\eta)L(r; \{I_{ht}\}, \{X_{ht}\})}{p(r^{(i-1)} \mid \phi, \Sigma_\eta)L(r^{(i-1)}; \{I_{ht}\}, \{X_{ht}\})}, 1\right\}.$$

(4) Sample $u \sim U_{[0,1]}$,

$$\text{if } u \le \alpha(r^{(i-1)}; r) \quad \text{then } r^{(i)} = r, \quad \text{otherwise } r^{(i)} = r^{(i-1)}.$$

Thus, we have necessary conditional posterior distributions

$$\text{(A-1)} \quad \begin{cases} \phi \mid \{r_h\}, \{z_h\}, \Sigma_\eta\text{: Normal distribution,} \\ \Sigma_\eta^{-1} \mid \{r_h\}, \{z_h\}, \phi\text{: Inverted Wishart distribution,} \\ \{r_h\} \mid \{I_{ht}\}, \{X_{ht}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \{z_h\}, \phi, \Sigma_\eta\text{:} \\ \quad \text{Metropolis-Hasting sampling.} \end{cases}$$

Finally, we denote by $f_{(i)}(\{y_{ht}^{(i)}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \Delta^{(i)}, V_\beta^{(i)}, \{r_h\}, \phi, \Sigma_\eta \mid \{I_{ht}\}, \{X_{ht}\}, \{z_h\})$ the joint posterior density for the regime i, and under the assumption of uncorrelated errors for latent utility equations of each regime, overall joint posterior density across regimes can be expressed as

$$\text{(A-2)} \quad \prod_{i=g,a,l} f_{(i)}\big(\{y_{ht}^{(i)}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \Delta^{(i)}, V_\beta^{(i)}, r_h, \phi, \Sigma_\eta \mid \{I_{ht}\}, \{X_{ht}\}, \{z_h\}\big).$$

In terms of sampling algorithms (A-1) and (A-2) for Markov chain Monte Carlo, we can constitute the posterior distribution of each regime, respectively, to get overall joint posterior density across regimes.

References

Bell, D. R., J. M. Lattin. 2000. Looking for loss aversion in scanner data: The confounding effect of price response heterogeneity. *Marketing Sci.* **19** 185–200.

Briesch, R. A., L. Krishnamurthi, T. Mazumdar, S. P. Raj. 1997. A comparative analysis of reference price models. *J. Consumer Res.* **24** 202–214.

Bucklin, R. E., S. Gupta. 1992. Brand choice, purchase incidence, and segmentation: An integrated modeling approach. *J. Marketing Res.* **29** 201–215.

Chang, K., S. Siddarth, C. B. Weinberg. 1999. The impact of heterogeneity in purchase timing and price responsiveness on estimates of sticker shock effects. *Marketing Sci.* **18** 178–192.

Chen, C. W. S., J. C. Lee. 1995. Bayesian inference of threshold autoregressive models. *J. Time Ser. Anal.* **16** 483–492.

Ferreira, P. E. 1975. A Bayesian analyis of a switching regression models: Known number of regimes. *J. Amer. Statist. Assoc.* **70** 370–374.

Gal-Or, E., M. Gal-Or. 2005. Customized advertising via a common media distributer. *Marketing Sci.* **24** 241–253.

Geweke, J., N. Terui. 1993. Bayesian threshold autoregressive models for nonlinear times series. *J. Time Ser. Anal.* **14** 441–454.

Gilbride, T., G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Sci.* **21** 391–404.

Guadagni, P. M., J. D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2** 203–238.

Gupta, S., L. G. Cooper. 1992. Discounting of discounts and promotion thresholds. *J. Consumer Res.* **19** 401–411.

Han, S. M., S. Gupta, D. R. Lehmann. 2001. Consumer price sensitivity and price thresholds. *J. Retailing* **77** 435–456.

Hardie, B. G. S., E. J. Johnson, P. S. Fader. 1993. Modeling loss aversion and reference dependence effects on brand choice. *Marketing Sci.* **12** 378–394.

Helson, H. 1964. *Adaption-Level Theory*. Harper & Row, New York.

Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47** 263–291.

Kalwani, M. U., C. K. Yim. 1992. Consumer price and promotion expectations: An experimental study. *J. Marketing Res.* **29** 90–100.

Kalyanaram, G., J. D. C. Little. 1994. An empirical analysis of latitude of price acceptance in consumer package goods. *J. Consumer Res.* **21** 408–418.

Keane, M. P. 1997. Modeling heterogeneity and state dependence in consumer choice behavior. *J. Bus. Econom. Statist.* **15** 310–327.

Leichty, J. C., D. K. H. Fong, W. S. DeSarbo. 2005. Dynamic models incorporating individual heterogeneity: Utility evolution in conjoint analysis. *Marketing Sci.* **24** 285–293.

Mayhew, G. E., R. S. Winer. 1992. An empirical analysis of internal and external reference price using scanner data. *J. Consumer Res.* **19** 62–70.

Newton, M. A., A. E. Raftery. 1994. Approximating Bayesian inference with the weighted likelihood bootstrap. *J. Roy. Statist. Soc.* **B56** 3–48.

Putler, D. 1992. Incorporating reference price effects into a theory of consumer choice. *Marketing Sci.* **11** 287–309.

Rajendran, K. N., G. J. Tellis. 1994. Contexual and temporal components of reference price. *J. Marketing* **58** 22–34.

Rossi, P. E., R. McCulloch, G. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Sci.* **15** 321–340.

Sherif, M., D. Taub, C. I. Hovland. 1958. Assimilation and contrast effects of anchoring stimuli on judgements. *J. Experiment. Psych.* **55** 150–155.

Winer, R. 1986. A reference price model of brand choice for frequently purchased products. *J. Consumer Res.* **13** 250–256.

Yang, S., G. M. Allenby. 2000. A model for observation, structural, and household heterogeneity in panel data. *Marketing Lett.* **11** 137–149.