Estimating Heterogeneous Price Thresholds

Nobuhiko Terui*
and
Wirawan Dony Dahana

Graduate School of Economics and Management
Tohoku University
Sendai 980-8576, Japan
*E-mail:terui@econ.tohoku.ac.jp
*Voice & Fax:+81-22-217-6311

November, 2004
April, 2005, This version

Abstract

This study addresses price thresholds and their induced latitude of price acceptance under consumer heterogeneity. By incorporating price thresholds into a three-regime piecewise-linear stochastic utility function, we propose a brand choice model to estimate heterogeneous price thresholds using scanner panel data. Furthermore, we explore the structures that associate price thresholds with consumer specific information using hierarchical Bayes modeling with Markov chain Monte Carlo (MCMC) method. This study is intended to contribute to the modeling literature on discontinuous likelihoods in choice models. The empirical application using our 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 using price thresholds than by an aggregate (homogeneity) model without price thresholds.

As an application, we demonstrate that a customized pricing strategy based on estimated individual price thresholds can be more efficient than a flat pricing strategy.

Key Words and Phrases
Discontinuous Likelihoods, Reference Effect, Price Threshold, Latitude of Price Acceptance, Brand Choice, Bayesian MCMC, Heterogeneity, Scanner Panel Data, Customized Pricing

*Terui acknowledges the financial support by the Japanese Ministry of Education Scientific Research Grant No.(C)15530137.
The authors acknowledge helpful comments from Area editor and three anonymous reviewers.
Estimating Heterogeneous Price Thresholds

Abstract

This study addresses price thresholds and their induced latitude of price acceptance under consumer heterogeneity. By incorporating price thresholds into a three-regime piecewise-linear stochastic utility function, we propose a brand choice model to estimate heterogeneous price thresholds using scanner panel data. Furthermore, we explore the structures that associate price thresholds with consumer specific information using hierarchical Bayes modeling with Markov chain Monte Carlo (MCMC) method. This study is intended to contribute to the modeling literature on discontinuous likelihoods in choice models. The empirical application using our 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 using price thresholds than by an aggregate (homogeneity) model without price thresholds.

As an application, we demonstrate that a customized pricing strategy based on estimated individual price thresholds can be more efficient than a flat pricing strategy.

1. Introduction

This study specifically addresses the sub-area of price thresholds in the literature of consumer’s brand choice. Under consumer heterogeneity, we estimate the price thresholds and their induced latitude of price acceptance. Related recent studies include Gupta and Cooper (1992), Kalwani and Yim (1992), Kalyanaram and Little (1994), and Han, Gupta and Lehman (2001). Kalwani and Yim (1992) incorporate deterministic thresholds for gain and loss; Han et al. (2001) extend them to stochastic thresholds in logit choice models. The latter study, under the assumption of homogeneous consumers, concludes that the model with price threshold empirically improves on a standard logit model without a price threshold. Further stochastic price thresholds are a more flexible and effective approach. In particular, Kalyanaram and Little (1994), K-L hereafter, propose a three-regime piecewise-linear utility function. Using the logit choice model and scanner panel data, they examine several hypotheses of nonlinear aspects of price response and measure the latitude of price acceptance. Along their line of modeling, the present study introduces a three-regime piecewise-linear stochastic utility function with two price thresholds. We propose a brand-choice model with consumer heterogeneity. We apply
hierarchical Bayes modeling by Markov chain Monte Carlo (MCMC) method to conduct an inference on heterogeneous price thresholds using scanner panel data. Our modeling extends that of K-L in several ways. First, we assume heterogeneity of consumers’ preferences. Second, we do not always assume symmetric price thresholds around their reference prices. Moreover, we use hierarchical modeling to associate price thresholds with household-specific information and thereby explore their underlying structures. On the other hand, Han et al. (2001) do not always assume symmetric latitude price acceptance. Their model has a hierarchical structure to explain the price thresholds. Nevertheless, that model is also estimated under the assumption of consumer homogeneity.

In addition to the points outlined above, our study is intended to contribute to the modeling literature on discontinuous likelihoods in marketing models. Our empirical application adds to the literature on the presence of reference prices and loss aversion that initially addressed effects uncovered in scanner data, for example, as indicated by Winer (1986), Putler (1992), and Mayhew and Winer (1992). Subsequently, Chang, Siddarth and Weinberg (1999) and Bell and Lattin (2000) showed that reference prices’ 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 than in results obtained using the homogeneity model without price thresholds. This result is robust relative to the choice of reference price.

To derive managerial implications from our heterogeneous price threshold model, we investigate the possibility of a customized pricing strategy by evaluating expected incremental profits. We show, under a limited simulation study, that optimal levels could occur near price
thresholds and that a customization strategy on pricing could provide managers with larger
profits than a “non-customized (flat)” pricing strategy.

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. Extending the
framework by Rossi, McCulloch and Allenby (1996), we develop a hierarchical Bayes
modeling for it. Section 3 describes the application of scanner panel data to our model. Section
4 explores a customized pricing strategy using the knowledge of price thresholds for individual
consumers. Section 5 concludes this paper. The appendix explains details of the algorithm for
our 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 that 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

$$
U_{jht} = \begin{cases} 
    u^{(g)}_{jht} + X^{(g)}_{jht} P_{ht}^{(g)} + \varepsilon^{(g)}_{jht} (\equiv U^{(g)}_{jht}) & \text{if } P_{jht} - RP_{jht} \leq r_{1h} \\
    u^{(a)}_{jht} + X^{(a)}_{jht} P_{ht}^{(a)} + \varepsilon^{(a)}_{jht} (\equiv U^{(a)}_{jht}) & \text{if } r_{1h} < P_{jht} - RP_{jht} \leq r_{2h}, \\
    u^{(l)}_{jht} + X^{(l)}_{jht} P_{ht}^{(l)} + \varepsilon^{(l)}_{jht} (\equiv U^{(l)}_{jht}) & \text{if } r_{2h} < P_{jht} - RP_{jht},
\end{cases}
$$

(1)

where $X^{(i)}_{jht}$ is the $k$-dimensional row vector of explanatory variables allocated to regime “$i$”
according to the level of sticker shock \( P_{jht} - RP_{jht} \) (\( P_{jht} \): the retail price exposed to consumer \( h \)) at the occasion, and \( k \)-dimensional column vectors \( \beta^*(i), i = g, a, l \), represent different market responses around the reference price. Finally, \( \epsilon_{jht}, i = g, a, l \), respectively represent stochastic error components in the utility for each regime. We assume that they are independent across regimes. From that definition, the latitude of price acceptance (LPA hereafter) for consumer \( h \) under the conditions discussed below can be expressed as \( L_h = (r_{1h}, r_{2h}) \). Figure 1 portrays the meaning of our stochastic 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, in other words, 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 \( \beta^{(a)}_{hp} \sim N(0, \sigma^{(a)^2}_{hp}) \), which is an element of \( \beta^*_h \). Under the assumption of consumer homogeneity, K-L seek symmetric LPA that have an insignificant price response estimate over this range. On the other hand, Han et al. (2001) have 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.

Figure 1: Model for Price Thresholds and Market Responses

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 difference among latent utilities \( U_{nth}, n = 1, 2, \ldots, m \). Thereby, the probability of choosing brand \( j \) over other brands is defined as
where \( y_{jht} = U_{jht} - U_{nht} \) represents the relative utility from the last brand. 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,

\[
y_{ht}^{(i)} = X_{ht}^{(i)} \beta_h^{(i)} + \varepsilon_{ht}^{(i)} \quad \varepsilon_{ht}^{(i)} \sim N(0, \Lambda^{(i)}) \quad h = 1, \ldots, H, \quad t = 1, \ldots, T_h, \quad i = g, a, l, \tag{3}
\]

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. Regarding price thresholds, we employ the model

\[
r_{1h} = Z_{1h}^\prime \gamma_{1h} + \eta_{1h}; \quad r_{2h} = Z_{2h}^\prime \gamma_{2h} + \eta_{2h}; \quad h = 1, \ldots, H, \tag{4}
\]
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 $\eta_{sh} \sim N(0, \sigma^2_{sh})$ for $s = 1, 2$. We also set a hierarchical structure of “between subjects model for regime $i$” for the market response parameter

$$
\beta_h^{(i)} = \Delta^{(i)} Z^B_h + \nu^{(i)}; \quad \nu^{(i)} \sim N(0, \Sigma^{(i)}), \quad h = 1, \ldots, H, \quad i = g, a, l, \quad (5)
$$

where $Z^B_h$ 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 (5).

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

$$
\begin{align*}
\Pr\left\{ y^{(a)}_{jk} = \max(y^{(a)}_{1h}, \ldots, y^{(a)}_{m-1h}) > 0 \right\} & \quad \text{if } P_{jh} - R_{jkh} \leq r_{1h} \\
\Pr\left\{ y^{(a)}_{jk} = \max(y^{(a)}_{1h}, \ldots, y^{(a)}_{m-1h}) > 0 \bigg| R \right\} & \quad \text{if } r_{1h} < P_{jh} - R_{jkh} \leq r_{2h}, \quad (6) \\
\Pr\left\{ y^{(i)}_{jk} = \max(y^{(i)}_{1h}, \ldots, y^{(i)}_{m-1h}) > 0 \right\} & \quad \text{if } r_{2h} < P_{jh} - R_{jkh},
\end{align*}
$$

where $\Pr\left\{ y^{(a)}_{jk} = \max(y^{(a)}_{1h}, \ldots, y^{(a)}_{m-1h}) > 0 \bigg| R \right\}$ indicates the choice probability under the restriction on price response $\beta_{hp}^{(a)} \sim N(0, \sigma^{(a)}_{hp})$ in the LPA regime.

As for model calibration, extending the framework of Rossi, McCulloch and Allenby (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 $P_{jh} - R_{jkh}$, we first allocate data $\{X_{jh}\}$ of the explanatory variable to make $\{X_{jh}^{(i)}, i = g, a, l\}$ at each purchase occasion. The corresponding latent utility vector $\{y^{(i)}_{jh}, i = g, a, l\}$ is generated based on personal choice data $\{I_{jh}\}$ (the index of observed choices) using the algorithm of the Bayesian
probit model (Rossi et al. (1996)) applied to each regime, and then, except for price threshold “$r_h$ | -”, we can use conditional posterior distributions: $y_h(i) \mid \{ I_{h(i)} \}, \{ X_{h(i)} \}, \beta_h(i), \Lambda(i), r_h$.

$\beta_h(i) \mid \{ y_h(i) \}, \{ X_{h(i)} \}, \Lambda(i), \Delta(i), V_{\beta(i)}, z_h, r_h, \Lambda(i) \mid \{ y_{h(i)} \}, \{ X_{h(i)} \}, \{ \beta_h(i) \}, \{ \beta_h \}$,

$\Delta(i) \mid \{ \beta_h(i) \}, V_{\beta(i)}, \{ z_h \}, \{ r_h \}, \text{and } V_{\beta(i) \Delta} \mid \{ \beta_h(i) \}, \Delta(i), \{ z_h \}, \{ r_h \}$ for $i = g, a, l$.

2-2. Discontinuous Likelihoods for Thresholds and Their Modeling

Our model includes threshold variables in the probit model. It brings discontinuous likelihood functions relative to thresholds, as shown in Fig. 1. Conditional posterior density of $r_h$ is unavailable in terms of the known distribution. Conditional on the value of $r_h$, the operation that generates the latent utility above constitutes 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 $h$

$$\prod_{i=g,a,l} \left\{ \prod_{r_h \in R^{(i)}(r_h)} |\Lambda(i)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y_{h(i)} - X_{h(i)} \beta_h(i))' \right| \Lambda(i)^{-1} \left( y_{h(i)} - X_{h(i)} \beta_h(i) \right) \right\}, \quad (7)$$

where $R^{(g)}(r_h) \cap R^{(a)}(r_h) \cap R^{(l)}(r_h) = T_h$. In turn, conditional on $(\beta_h(i), \Lambda(i))$, we take (7) 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

$$L(\{r_h\} \mid \{ I_{h(i)} \}, \{ X_{h(i)} \} \mid \{ \beta_h(i) \}, \{ \Lambda(i) \}) \propto \prod_{h=1}^{H} \left\{ \prod_{i=g,a,l} \left\{ \prod_{r_h \in R^{(i)}(r_h)} |\Lambda(i)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y_{h(i)} - X_{h(i)} \beta_h(i))' \right| \Lambda(i)^{-1} \left( y_{h(i)} - X_{h(i)} \beta_h(i) \right) \right\} \right\}, \quad (8)$$
and jointly with (4) expressed as hierarchical structure, we can apply Metropolis-Hasting sampling algorithm for price thresholds to obtain conditional posterior

\[ r_h \mid \{ I_n \}, \{ X_n \}, \{ \beta_h^{(i)} \}, \Lambda_h^{(i)}, \{ z_h \}, \phi, \Sigma_n \]. Prior distributions and MCMC estimation procedures for these hierarchical Bayes models are described in the appendix. In general, the proposed model includes difficulty of calibration because the estimation of the threshold parameter \( r_h = (r_{h1}, r_{h2}) \) is not well defined for this discontinuity of likelihood function, which is not differentiable in \( r_h \). Conventional maximum likelihood estimation collapses and classical asymptotic distribution theory is not operative on this parameter. Conditional inference is 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. Unfortunately, this approach does not allow statistical inferences such as testing hypotheses and confidence intervals. Another approach is Bayesian inference. Ferreira (1975) provided Bayesian analysis on the timing of structural change of economy by using switching regression models. Geweke and Terui (1993) and Chan and Lee (1995) extended Bayesian switching models including threshold parameter in nonlinear time series models. In particular, the latter two studies prove that the unconventional likelihood function mentioned above is shown to pose no barrier to recovery of the posterior density and statistical inference on the threshold is possible using the posterior density. Their topics differ mutually and with this study, however we retain a statistical structure relative to the threshold parameter in common with our choice model.

In the price-threshold literature, K-L assume symmetric LPA by setting \( r_{h1} = r_{h2} = r \) (homogeneous). They estimate not the threshold, but the length of LPA, where several possible lengths are specified \textit{a priori}; and they seek the LPA that has insignificant
estimate of price response inside this range. Han et al. (2001) also indirectly estimate thresholds through their hierarchical structure. Kalwani and Yim (1992) take a grid search approach. Our proposed method directly models price thresholds in a general way and conducts coherent statistical inference on them under a small-sample situation.

3. Empirical Application to Scanner Panel Data

3-1. Data, Variables and Model Specification

Data and Variables

Video Research Ltd., 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/100 g to equalize quantitative differences for each package of the five brands.

| Table 1 Descriptive Statistics for Data |

Data matrices for our model (3), (4) and (5) include the following variables.

- **Explanatory Variables**: \( X = \{ \text{Constant, Price, Display, Feature, Brand loyalty} \} \), 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

\[
GL_{jht} = \alpha GL_{jht-1} + (1 - \alpha) I_{jht-1},
\]

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: $Z^\prime = \{\text{Constant, Dprone, Pfreq, } RP, BL\}$,
where Dprone is the deal proneness defined as the proportion of purchase (of any of the five brands) made on promotion (Bucklin and Gupta (1992), Han et al. (2001)); Pfreq is the shopping frequency (three categories) ($H_3$ of K-L); $RP$ represents the measure of average reference price level as defined by $\sum_{j=1}^{m} \left( \sum_{t=1}^{T_h} \log\left( \frac{RP_{jht}}{T_h} \right) / m \right)$ and $BL$ represents the brand loyalty measure defined as $BL_h = \max_j \left( \sum_{t=1}^{T_h} GL_{jht} / T_h \right)$, both of which are used by K-L and Han et al. (2001).

We further define $Z^\beta = \{\text{Constant, Hsize, Expend}\}$, where Hsize is 1–6 (number of household members) and Expend is nine categories (shopping expenditure / month) used in Rossi et al. (1996).

**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 take up four kinds of reference prices of two categories: memory-based and stimulus-based. That is, [1] Memory-based as (A) $RP_{jht} = P_{jht-1}$, the price at its last purchase and (B)

$$RP_{jht} = \alpha RP_{jht-1} + (1 - \alpha) P_{jht-1},$$

the smoothed price over previous purchases; and [2]

Stimulus-based as (C) $RP_{jkt} = P_{klt}$, where $k$ means the brand at the last purchase and (D)

$$RP_{jhr} = P_{rhr},$$

where $r$ indicates the price of a brand chosen randomly at the time of purchase. (Briesch et al. (1997))

**Model Specification**

We consider three models for comparison with other candidates: (i) aggregate (homogeneous) reference price probit model without a threshold (Winer (1986), Putler (1992) and Mayhew and
Winer (1992)); (ii) two regimes heterogeneous reference price probit model without a threshold (Bell and Lattin (2000), and Chang, Siddarth and Weinberg (1999)); 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 for comparison. Tables 2-1 and 2-2 describe logs of marginal likelihood for model comparison and summary statistics of parameter estimates. To save space, estimates of constant terms are not listed here. The model (C)-(iii): stimulus-based RP (C)-3 regimes heterogeneous probit with thresholds is most supported by marginal likelihood criterion. Consequently, some evidence supports the hypothesis of price threshold existence ($H_1$ of K-L) under models incorporating heterogeneous consumers.

| Tables 2-1 and 2-2: Model Specification and Parameter Estimates |

### 3-2. Reference Effects, Loss Aversion and Price Thresholds

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: -4.722, loss: -7.098) 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). Therefore, we infer that loss aversion revisited once after price thresholds were incorporated in 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) (Fig. 1 of p. 190), we present intuitive reasoning for our results below. First, Bell and Lattin (2000) discussed the relation between models (i) and (ii) relative to loss aversion: model (i) has steeper slopes in the loss regime ($|\beta_A^{(l)}| > |\beta_{BL}^{(l)}|$) as well as the gain regime ($|\beta_A^{(g)}| < |\beta_{BL}^{(g)}|$) than model (ii), however the difference ($|\beta_A^{(l)}| - |\beta_{BL}^{(l)}|$) is larger in the loss regime than that ($|\beta_A^{(g)}| - |\beta_{BL}^{(g)}|$) 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 Fig. 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. On the other hand, model (iii) has three regimes. We suppose that the true response function has LPA (insensitive range) limited by the price thresholds, as shown in Fig. 1, and that the data are observed along with three lines. Therefore, 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

$$\left(\beta_{BL}^{(g)}, \beta_{BL}^{(l)}\right)$$

will be less steep than the slopes ($\beta^{(g)}, \beta^{(l)}$) of the three-regime model in the gain and loss regimes. That is, $|\beta_{BL}^{(g)}| < |\beta^{(g)}|$, $|\beta_{BL}^{(l)}| < |\beta^{(l)}|$ 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. Tables 2-1 and 2-2 present 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
\( \beta^{(g)} - \beta^{(l)} > 0 \) (we define it 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 \( \Pr \left\{ \beta^{(Gain)} - \beta^{(Loss)} \mid \text{data} \right\} \) at the aggregated level for models (i), (ii) and (iii). Graphs in Fig. 2 show frequency distributions and their empirical distribution functions of random draws of \( \beta^{(Gain)} - \beta^{(Loss)} \) through MCMC iterations for the respective models, where \( \beta^{(Gain)} - \beta^{(Loss)} \) is defined formally as \( \beta^{(g)} - \beta^{(l)} \) for model (i), and

\[
\frac{1}{H} \sum_{h=1}^{H} \left( \beta^{(g)}_{h,BL} - \beta^{(l)}_{h,BL} \right) \text{ for (ii) and } \frac{1}{H} \sum_{h=1}^{H} \left( \beta^{(g)}_{h} - \beta^{(l)}_{h} \right) \text{ for (iii).}
\]

These graphs show that loss aversion \( \beta^{(Gain)} - \beta^{(Loss)} > 0 \) is observed mostly for these models. Furthermore, the degree of loss aversion is strongest for model (i) in the sense that the posterior probability of loss aversion \( \Pr \left\{ \beta^{(Gain)} - \beta^{(Loss)} > 0 \mid \text{data} \right\} \) 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 aversion would increase after incorporation of price thresholds into the models. These results are robust to the selection of reference price, except for (B).

Figure 2: Posterior Distribution of Loss Aversion

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

Table 3 shows that all hierarchical Bayes estimates of regression coefficients \( \phi_1 \) and \( \phi_2 \) of (4) on household specific variable are significant in the sense that 95% HPD region does not include zero. The graphs in Fig. 4 show the relations between LPA and respective household variables. For “Pfreq”, according to shopping frequency, we grouped the households into three segments and averaged \( \{ \hat{\theta}_{1h} \} \) and \( \{ \hat{\theta}_{2h} \} \) inside respective segments. We connected those averages as their LPAs in the graph. The same operation was applied to other variables.

Throughout the graphs, we first note that LPAs relative to segments are not symmetric around the zero levels of their respective sticker shocks. We also observe the following: (1) For Pfreq vs. LPA, the mean level of LPA increases concomitant with the purchase frequency, which is consistent with previous empirical findings, e.g. on the hypothesis \( H_3 \) of K-L. (2) For Dprone vs. LPA, the mean level of LPA decreases concomitant with the increase in deal proneness, which is compatible with Han et al. (2001). (3) For RP vs. LPA, according to the level of average reference price, the lower and upper 30% households are extracted as two segments and are plotted as the relation between LPAs and RP. That plot shows that the segment with high average reference price has wider LPA, which is consistent with the result to the hypothesis \( H_2 \) of K-L. (4) For BL vs. LPA, according to the level of household-specific brand loyalty measure,
two segments were generated similarly as (3) and we observe that the high brand loyalty segment has slightly wider LPA, which is consistent with results by hypothesis \( H_4 \) of K-L.

4. A Managerial Implications

Based on knowledge of heterogeneous price thresholds for respective consumers obtained using the proposed model, we consider customized pricing and explore possible efficient pricing in this section.

**Incremental Profit**

Conditional on the draw \( \left\{ (r_{ih}, <0), \beta_h^{(p)}, \beta_h^{(a)}) \right\} \) of MCMC, we set the discount level of \( (r_{ih} + \alpha) >0 \), \( \alpha = 0, \pm 1\%, \pm 2\%, \ldots \) for \( h = 1,\ldots,H \). We define the expected incremental profit for “customized discounting” averaged over households as

\[
IP_{\text{C}}^{-j}(\alpha \mid (r_{ih}, \beta_h^{(p)}, \beta_h^{(a)}), h = 1,\ldots,H) = \\
\sum_{h=1}^{H} \left[ \Pr \left( \beta_h^{(p)}, P_{j0} \left( 1 - (r_{ih} + \alpha) \right) \right) - \Pr \left( \beta_h^{(a)}, P_{j0} \right) \right] \left( M - (r_{ih} + \alpha) \right) \text{ if } \alpha \geq 0: \text{(Price Gain)} ; \\
\sum_{h=1}^{H} \left[ \Pr \left( \beta_h^{(a)}, P_{j0} \left( 1 - (r_{ih} + \alpha) \right) \right) - \Pr \left( \beta_h^{(a)}, P_{j0} \right) \right] \left( M - (r_{ih} + \alpha) \right) \text{ if } \alpha < 0: \text{(LPA)}
\]

where \( P_{j0} \) is brand \( j \)’s price that belongs to LPA and \( M \% \) represents the margin. Taking the expectation of (9) with respect to posterior distribution of \( (r_{ih}, \beta_h^{(p)}, \beta_h^{(a)}) \) generates unconditional incremental profit

\[
IP_{\text{C}}^{-j}(\alpha) = E_{(r_{ih}, \beta_h^{(p)}, \beta_h^{(a)})} \left[ IP_{\text{C}}^{-j}(\alpha \mid (r_{ih}, \beta_h^{(p)}, \beta_h^{(a)}), h = 1,\ldots,H) \right]. \tag{10}
\]

These values are calculated as by-products of sampling through MCMC iterations.
Figure 5: Customized Pricing for Brand A

The negative part of the left picture in Fig. 5 shows these expected incremental profits for brand A. The largest profit is obtained at boundary \( r_{1h} \), where the margin was set as \( M = 30\% \). As for a price hike strategy, we set the hike rate \((r_{2h} + \alpha)\% (\geq 0)\) for \( h = 1, \ldots, H \) (\( \alpha = 0, \pm 1\%, \pm 2\%, \ldots \)). In that case, the unconditional expected incremental profits are defined similarly as \( IP^{(+)}_j(\alpha) = E \left( \{ r_{2h}, \beta_j^{(+)}(r_{2h}), M = 30\% \} \right) \). In that case, the unconditional expected incremental profits are defined similarly as

\[
IP^{(+)}_j(\alpha) = E \left( \{ r_{2h}, \beta_j^{(+)}(r_{2h}), M = 30\% \} \right).
\]

The positive region of the graph shows that the maximum profit is obtained at the upper threshold \( r_{2h} \).

**Comparison with Non-Customized Pricing**

Aggregated difference between an optimal customized discounting at the level \( r_{1h} \) and a non-customized (flat) discounting at the level \( d^\star \) is denoted as

\[
DIF^{(c)}_j(d^\star \mid \{ r_{1h}, \beta_j^{(c)}, h = 1, \ldots, H \})
\]

\[
= \frac{1}{H} \sum_{h=1}^{H} \left( Pr \left( \beta_j^{(c)}(r_{1h}) \right) - Pr \left( \beta_j^{(c)}(d^\star) \right) \right) (M - r_{1h})
\]

\[
- \frac{1}{H} \sum_{h=1}^{H} \left( Pr \left( \beta_j^{(c)}(r_{1h}) \right) - Pr \left( \beta_j^{(c)}(d^\star) \right) \right) (M - d^\star),
\]

where market response \( \beta_j^{(c)} \) for non-customized pricing depends on the regime determined by discount level \( d^\star \). An unconditional estimate is obtained by taking expectation

\[
DIF^{(c)}_j(d^\star) = E \left( \{ r_{1h}, \beta_j^{(c)} \} \right) DIF^{(c)}_j(d^\star \mid \{ r_{1h}, \beta_j^{(c)}, h = 1, \ldots, H \})
\].

The same operation is applied to a price hike strategy at an optimal customized price hike at level \( r_{2h} \), and compared with non-customized discounting at the level \( d^\star \) to obtain
We note that a manager taking a non-customized pricing strategy is oblivious to respective consumers’ price thresholds, and thus the manager applies constant (flat) pricing to every consumer. The negative and positive regions of the graph on the right side of Fig. 5 respectively show plots of \( \text{DIF}_j^{(-)}(d^*) \) and \( \text{DIF}_j^{(+)}(d^*) \) for \( d^* = 1, 2, \ldots, 15\% \). Different margins produce slightly different graph shapes, but the findings described below do not change. First, this graph shows that two kinds of optimal customized pricing – discount at \( r_{1h} \) and hike at \( r_{2h} \) – dominate every level of non-customized pricing \( d^* = 1, 2, \ldots, 15\% \). In the negative LPA region, non-customized pricing does not bring as great a sales increase as expected because of price change insensitivity over this region. In contrast, customized pricing does not take a discount strategy over this region. Identical logic applies for a price hike shown in the positive part of LPA. Those gains for customized pricing arise from price threshold information. In the price gain regime as well as in the price loss regime, it would be reasonable to consider that efficiency can occur at those limits. Almost identical situations apply to other brands, but the change is largest for the highest-priced brand E. These results are based on limited simulation study under a simple framework for illustrating possible applications of our model. Full discussion, including recent related work on pricing by Van Heerde et al. (2004) and others which explores factor decomposition of price promotion and its long term effects, is left for future investigation.

5. Concluding Remarks

This study, while specifically addressing sub-areas of price thresholds and latitude price acceptance in consumer behavior studies, introduced a three-regime piecewise-linear stochastic utility function and their induced choice models – threshold probit models – under a framework of continuous mixture modeling of heterogeneous consumers. Price thresholds generate
unconventional discontinuous likelihood function in the analysis, creating 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. Through estimation using MCMC, and model specification in terms of the criterion of marginal likelihood, we obtained evidence demonstrating that our proposed model is superior to a conventional linear-utility based probit model without price thresholds. Our empirical application, as far as our data set is concerned, showed that 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 than by an aggregate (homogeneous) model without price thresholds. Through hierarchical modeling, we investigated explanatory factors of heterogeneous price thresholds. Some are consistent with previous studies that presumed consumer homogeneity. Finally, as an application of our model, we explored a customized pricing strategy based on the knowledge of price thresholds for respective households. We demonstrated that optimal customized pricing levels are provided at the lower price thresholds for discounting, and at the upper price thresholds for a price hike. Moreover, our performance-comparison exercises with possible non-customized pricing strategies showed that customized pricing could yield greater profits than flat pricing. These results are limited to our simulation study under simple conditions.

Our analysis can be generalized further to incorporate heterogeneity in several aspects. For example, heterogeneous modeling could be applied to smoothing parameters of the reference price derived from exponentially weighted past prices, and a purchase carryover parameter in the state dependent variable. Together with a comprehensive discussion of customized pricing, we leave these problems for future research.
References


Appendix

Markov chain Monte Carlo Algorithm for Hierarchical Bayes Modeling of Threshold Probit Model

[1] Gibbs Sampler for Within Subject and Between Subjects Parameters

Conditional on the thresholds \( r_h = (r_{h1}, r_{h2}) \), the conditional posterior distributions for “within subject” model \( y_{ht}^{(i)} = X_{ht}^{(i)} \beta_h^{(i)} + \epsilon_{ht}^{(i)} \sim \mathcal{N}\left(0, \Lambda^{(i)}\right) \) and “between subjects” model for regime \( i \) \( \beta_h^{(i)} = \Delta^{(i)} Z_h^{(i)} + u_h^{(i)} \); \( u_h^{(i)} \sim \mathcal{N}\left(0, V^{(i)}_\beta\right) \) are respectively described as

\[
\begin{align*}
\begin{cases}
y_{ht}^{(i)} | \{ I_{ht} \}, \{ X_{ht} \}, \beta_h^{(i)}, \Lambda^{(i)}, r_h : & \text{Truncated normal} \\
\beta_h^{(i)} | \{ y_{ht}^{(i)} \}, \{ X_{ht}^{(i)} \}, \Lambda^{(i)}, \Delta^{(i)}, V^{(i)}_\beta, z_h, r_h : & \text{Normal} \\
\Lambda^{(i)} | \{ y_{ht}^{(i)} \}, \{ X_{ht}^{(i)} \}, \beta_h^{(i)}, \Delta^{(i)}, V^{(i)}_\beta, z_h, r_h : & \text{Wishart} \\
V^{(i)}_\beta | \{ \beta_h^{(i)} \}, \Delta^{(i)}, \{ z_h \}, \{ r_h \} : & \text{Wishart}
\end{cases}
\end{align*}
\]

(A-1) and their specifications are available in the appendix of Rossi et al.(1996; pp.338-339).


As for hierarchical modeling for the price threshold, the lower (negative) threshold is modeled as \( r_{h1} = z_h \phi_1 + \eta_h ; \quad \eta_h \sim \mathcal{N}\left(0, \sigma_n^2\right) \) and stacking over all households \( h = 1, ..., H \), leads to matrix notation \( r_i = Z \phi_i + \eta_i ; \quad \eta_i \sim \mathcal{N}\left(0, \Sigma_n\right) \). We first set prior distributions on \((\phi, \Sigma_n)\) as

\[
\begin{align*}
\phi_1 | r_1, \Sigma_{1n} & \sim \mathcal{N}(\phi_{10}, \Sigma_{10}) ; \quad \phi_{10} = 0, \quad \Sigma_{10} = 0.01 I_d, \\
\Sigma_{1n}^{-1} & \sim \text{Wishart}\left(v_{1n0}, V_{1n0}\right) ; \quad v_{1n0} = d^* + 4, \quad V_{1n0} = V_{1n0} I_H,
\end{align*}
\]

then we have conditional posterior distributions

\[
\begin{align*}
\phi_1 | \{ r_{h1} \}, \{ z_h \}, \Sigma_{1n} & \sim \mathcal{N}\left(\bar{\phi}_1, (Z'Z + \Sigma_{1n}^{-1})^{-1}\right), \\
\Sigma_{1n}^{-1} | \{ r_{h1} \}, \{ z_h \}, \phi_1 & \sim \text{Wishart}\left(V_{1n0} + H, V_{1n0} + \sum_i (r_{hi} - Z \bar{\phi}_i) (r_{hi} - Z \bar{\phi}_i)'\right).
\end{align*}
\]

The same formulations apply to upper (positive) price threshold \( r_{h2} \).
Next, as for conditional posterior \( \{r_h\} | \{I_{ht}\}, \{X_{ht}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \{z_h\}, \Phi, \Sigma_\eta \), we use the likelihood function \( L(\{r_h\}; \{I_{ht}\}, \{X_{ht}\}; \{\beta_h^{(i)}\}, \{\Lambda^{(i)}\}) \) of (8) in section 2 jointly with the prior \( p(\{r_h\}; \Phi, \Sigma_\eta) \sim N(\Phi, \Sigma_\eta) \) 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. \( i > 1 \); \( z \sim N(0, \sigma^2_{\text{RW}} I_{ht}) \) and set \( r = r^{(i-1)} + z \)
3. \( \alpha(r^{(i-1)}; r) = \min \left\{ \frac{p(r | \Phi, \Sigma_\eta) L(r; \{I_{ht}\}, \{X_{ht}\})}{p(r^{(i-1)} | \Phi, \Sigma_\eta) L(r^{(i-1)}; \{I_{ht}\}, \{X_{ht}\})}, 1 \right\} \)
4. Sample \( u \sim U_{[0,1]} \),
   if \( u \leq \alpha (r^{(i-1)}; r) \) then \( r^{(i)} = r \), otherwise \( r^{(i)} = r^{(i-1)} \).

Thus we have necessary conditional posterior distributions

\[
\begin{align*}
\Phi | \{r_h\}, \{z_h\}, \Sigma_\eta : & \text{ Normal distribution} \\
\Sigma_\eta^{-1} | \{r_h\}, \{z_h\}, \Phi : & \text{ Inverted Wishart distribution} \\
\{r_h\} | \{I_{ht}\}, \{X_{ht}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \{z_h\}, \Phi, \Sigma_\eta : & \text{ Metropolis-Hasting sampling}
\end{align*}
\]

Finally, we denote by \( f_{(i)} \left( \{y_h^{(i)}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \Delta^{(i)}, V^{(i)}_{\beta}, \{r_h\}, \Phi, \Sigma_\eta | \{I_{ht}\}, \{X_{ht}\}, \{z_h\} \right) \) 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

\[
(A-3) \quad \prod_{i \in S, a \in A} f_{(i)} \left( \{y_h^{(i)}\}, \{\beta_h^{(i)}\}, \Lambda^{(i)}, \Delta^{(i)}, V^{(i)}_{\beta}, \{r_h\}, \Phi, \Sigma_\eta | \{I_{ht}\}, \{X_{ht}\}, \{z_h\} \right).
\]

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

<table>
<thead>
<tr>
<th>Alternative</th>
<th>Choice Share</th>
<th>Average Price</th>
<th>% of Time Displayed</th>
<th>% of Time Featured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand A</td>
<td>0.138</td>
<td>623.5</td>
<td>0.264</td>
<td>0.423</td>
</tr>
<tr>
<td>Brand B</td>
<td>0.480</td>
<td>632.9</td>
<td>0.135</td>
<td>0.294</td>
</tr>
<tr>
<td>Brand C</td>
<td>0.099</td>
<td>601.3</td>
<td>0.317</td>
<td>0.405</td>
</tr>
<tr>
<td>Brand D</td>
<td>0.225</td>
<td>693.2</td>
<td>0.182</td>
<td>0.344</td>
</tr>
<tr>
<td>Brand E</td>
<td>0.057</td>
<td>902.4</td>
<td>0.191</td>
<td>0.286</td>
</tr>
</tbody>
</table>

Figure 1: Model for Price Thresholds and Market Responses

Model (i): aggregate (homogeneous) reference price probit model without a threshold
Model (ii): two regimes heterogeneous reference price probit model without a threshold (Bell and Lattin (2000));
Model (iii): three-regime heterogeneous reference-price probit model with thresholds
Table 2-1. Model Specification and Parameter Estimates

<table>
<thead>
<tr>
<th>[1] Memory based</th>
<th>(A) $RP_{jkt} = P_{jkt-1}$</th>
<th>(B) $RP_{jkt} - \alpha RP_{jkt-1} + (1-\alpha)P_{jkt-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(i) Without heterogeneity</td>
<td>With heterogeneity</td>
</tr>
<tr>
<td></td>
<td>(ii) Without threshold</td>
<td>(iii) With thresholds</td>
</tr>
<tr>
<td>Gain Regime</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-4.737* [0.797] -1.277 (-3.179) -2.676 (-4.538) -4.815* [0.762] -1.326 (-3.044) -5.219 (-8.184)</td>
<td></td>
</tr>
<tr>
<td>Display</td>
<td>0.474* [0.156] 0.841 (4.179) 0.935 (2.464) 0.495* [0.165] 0.520 (3.278) 0.908 (5.044)</td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>1.381* [0.114] 1.782 (9.592) 1.919 (5.701) 1.377* [0.103] 1.665 (9.835) 1.589 (8.458)</td>
<td></td>
</tr>
<tr>
<td>Brand Loyalty</td>
<td>0.769* [0.151] 1.364 (6.366) 0.964 (2.222) 0.740* [0.159] 1.717 (9.222) 1.150 (5.671)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td></td>
<td>-0.916 (-1.744)</td>
</tr>
<tr>
<td>Display</td>
<td></td>
<td>0.757 (2.559)</td>
</tr>
<tr>
<td>Feature</td>
<td></td>
<td>1.605 (5.613)</td>
</tr>
<tr>
<td>Brand Loyalty</td>
<td></td>
<td>2.998 (9.126)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss Regime</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-6.958* [1.184] -1.122 (-3.883) -4.615 (-4.774) -5.973* [0.992] -0.784 (-1.886) -1.409 (-2.656)</td>
<td></td>
</tr>
<tr>
<td>Display</td>
<td>0.711* [0.168] 0.622 (2.368) 0.867 (2.568) 0.672* [0.152] -1.013 (-3.549) 0.793 (3.010)</td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>1.444* [0.118] 1.281 (6.736) 1.838 (5.567) 1.450* [0.201] 0.100 (0.304) 1.176 (4.827)</td>
<td></td>
</tr>
<tr>
<td>Brand Loyalty</td>
<td>1.829* [0.157] 2.078 (9.909) 3.432 (9.100) 1.837* [0.260] 0.806 (3.799) 2.136 (7.018)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thresholds</td>
<td></td>
<td>(-0.102, 0.105)</td>
</tr>
<tr>
<td>LML</td>
<td>-36,709.047 -34,981.349 -33,829.653 -36,716.499 -35,836.148 -34,176.325</td>
<td></td>
</tr>
</tbody>
</table>

(i): Aggregate(homogeneous) Probit without threshold  
(ii): 2 regimes heterogeneous Probit without threshold  
(iii): 3 regimes heterogeneous Probit with thresholds  
* Significant at 0.95 HPD Region. The number inside [  ] in the model (i) means posterior standard deviation, and (  ) for (ii) and (iii) means t-value calculated from the frequency distribution of household’s estimates.  
LML: the log of marginal likelihood (Newton and Raftery(1994))  
Thresholds: Average of heterogeneous price threshold estimates
Table 2-2. Model Specification and Parameter Estimates

<table>
<thead>
<tr>
<th>[1] Stimulus Based</th>
<th>(C) $RP_{jht} = P_{jht}$</th>
<th>(D) $RP_{jht} = P_{jht}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(i) Without heterogeneity</td>
<td>With heterogeneity</td>
</tr>
<tr>
<td></td>
<td>(ii) Without threshold</td>
<td>(iii) With thresholds</td>
</tr>
<tr>
<td><strong>Gain Regime</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-4.722* [0.798]</td>
<td>-1.637 (-3.875)</td>
</tr>
<tr>
<td>Display</td>
<td>0.544* [0.193]</td>
<td>0.744 (2.793)</td>
</tr>
<tr>
<td>Feature</td>
<td>1.430* [0.143]</td>
<td>1.665 (7.830)</td>
</tr>
<tr>
<td>Brand Loyalty</td>
<td>-1.481* [0.288]</td>
<td>1.199 (5.161)</td>
</tr>
<tr>
<td><strong>LPA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Display</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Feature</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Brand Loyalty</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td><strong>Loss Regime</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>-7.098* [1.263]</td>
<td>-2.622 (-3.034)</td>
</tr>
<tr>
<td>Display</td>
<td>0.719* [0.342]</td>
<td>1.046 (3.135)</td>
</tr>
<tr>
<td>Feature</td>
<td>1.862* [0.255]</td>
<td>3.057 (8.044)</td>
</tr>
<tr>
<td>Brand Loyalty</td>
<td>-0.773* [0.323]</td>
<td>3.483 (8.805)</td>
</tr>
<tr>
<td><strong>Thresholds</strong></td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td><strong>LML</strong></td>
<td>-36,053.991</td>
<td>-34,716.514</td>
</tr>
</tbody>
</table>

(i): Aggregate(homogeneous) Probit without threshold
(ii): 2 regimes heterogeneous Probit without threshold
(iii): 3 regimes heterogeneous Probit with thresholds

* Significant at 0.95 HPD Region. The number inside [ ] in the model (i) means posterior standard deviation, and ( ) for (ii) and (iii) means t-value calculated from the frequency distribution of household’s estimates.

LML: the log of marginal likelihood (Newton and Raftery(1994))

Thresholds: Average of heterogeneous price threshold estimates
Figure 2: Posterior Distribution for Loss Aversion

![Figure 2: Posterior Distribution for Loss Aversion](image)

Figure 3: Heterogeneous Distribution of Price Thresholds

![Figure 3: Heterogeneous Distribution of Price Thresholds](image)

Table 3: Hierarchical Regression Coefficients on Household data (Price Thresholds)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$r_1$</th>
<th>$r_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>-0.292*</td>
<td>0.321*</td>
</tr>
<tr>
<td></td>
<td>[0.053]</td>
<td>[0.041]</td>
</tr>
<tr>
<td>Pfreq</td>
<td>0.028*</td>
<td>-0.010*</td>
</tr>
<tr>
<td></td>
<td>[0.004]</td>
<td>[0.005]</td>
</tr>
<tr>
<td>Dprone</td>
<td>0.042*</td>
<td>-0.033*</td>
</tr>
<tr>
<td></td>
<td>[0.006]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>RP</td>
<td>-0.049*</td>
<td>0.026*</td>
</tr>
<tr>
<td></td>
<td>[0.006]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>BL</td>
<td>-0.149*</td>
<td>0.165*</td>
</tr>
<tr>
<td></td>
<td>[0.017]</td>
<td>[0.017]</td>
</tr>
</tbody>
</table>

*Significant at 0.95 HPD Region.

The number inside [ ] means posterior standard deviation.
Figure 4: LPA and Household Specific Information

(1) Preq

(2) Dprone

(3) RP

(4) BL

Figure 5: Customized Pricing for Brand A

(1) Incremental Profit

(2) Difference from Non-customized pricing

The top and bottom of each box show upper and lower 5% percentile posterior density and the bar inside box is Bayes estimate (posterior mean).