非線形階層ベイズモデルによる顧客满意指数の
評価と予測

<table>
<thead>
<tr>
<th>著者</th>
<th>丁晶 良晴</th>
</tr>
</thead>
<tbody>
<tr>
<td>学位授与機関</td>
<td>東北大学工学系</td>
</tr>
<tr>
<td>学位授与番号</td>
<td>修士(工)第1515号</td>
</tr>
<tr>
<td>URL</td>
<td>hdl.handle.net/10097/00121281</td>
</tr>
</tbody>
</table>
Doctor Thesis

Nonlinear Hierarchical Bayes Modeling of Customer Satisfaction Index

By:

Xing Aijing

Under Supervision of:

Prof. Nobuhiko Terui

International Program of Economics and Management
Graduate School of Economics and Management, Tohoku University

June 2016
Abstract

This paper investigates the nonlinear relationship between customer satisfaction and loyalty, using Hierarchical Bayes modeling and structural heterogeneous functional forms. We extend the relationship proposed by the customer satisfaction index (CSI), and examine different possible functional forms in which satisfaction may affect loyalty. Besides, we propose models that reflect intrinsic characteristics of nonlinear effects, such as saturation-attainable limit of effectiveness, non-constant marginal return, asymmetric response between satisfied and dissatisfied customers, and zone of tolerance, in a parsimonious way.

Two quantitative methods are employed in the research. The first one uses nonlinear hierarchical Bayes model to accommodate structural homogeneity across companies. The empirical analysis with survey data shows that (1) hierarchical Bayes models estimated by borrowing other companies’ data are better than the independent model, (2) nonlinear models perform better than linear models, (3) nonlinear model with asymmetric marginal returns and attainable limits is found to be the best model among all.

The second method uses finite mixture model to accommodate structural heterogeneity across companies. Analysis shows that under the structural heterogeneity assumption, (1) the finite mixture model performs best, (2) companies have various structural forms, (3) about half companies are better accommodated with threshold linear model, in which the consumption tolerance interval can be measured.

Key contributions of the paper include a nonlinear structural equation model that contains nonlinear term of endogenous latent variable, the combination between finite mixture and structural equation model, and an efficient algorithm of MCMC in terms of multi-move sampler by using Gibbs sampling.
Abstract ................................................................................................................................. 1
1. Introduction ........................................................................................................................ 4
2. Review of Related Literature ............................................................................................. 7
   2.1 Nonlinear Relationship between Customer Satisfaction and Loyalty ................................ 7
       2.1.1 Nonlinear Relations of Satisfaction ......................................................................... 7
       2.1.2 Non-linearity and Moderating Effects .................................................................. 12
   2.2 Customer Satisfaction Index (CSI) Model ....................................................................... 14
       2.2.1 Aim and Definition of CSI Model ......................................................................... 14
       2.2.2 Classical Estimation of Linear CSI Model ................................................................. 16
3. Nonlinear Hierarchical Bayesian CSI Model and Empirical Results ............................... 18
   3.1 Nonlinear Model on the Satisfaction to Loyalty ............................................................. 18
   3.2 Structural Equation Modeling with Bayesian Approach ............................................... 21
   3.3 Hierarchical Bayes Modeling for Stable Estimation ....................................................... 23
   3.4 Model Selection ............................................................................................................... 24
   3.5 Data and Data Processing .............................................................................................. 26
       3.5.1 Data Description ..................................................................................................... 26
       3.5.2 Data Transformation ............................................................................................. 26
       3.5.3 Full Conditional Posterior Density and Multi-Move Gibbs Sampler ....................... 27
   3.6 Empirical Results of HB Nonlinear CSI Model ............................................................... 29
       3.6.1 Model Comparison ................................................................................................. 29
       3.6.2 Parameter Estimates of Nonlinear Term from CS to LOY ..................................... 31
       3.6.3 Estimated Functional Form .................................................................................... 34
       3.6.4 Distribution of Satisfaction Score ......................................................................... 34
       3.6.5 Marginal Effects and Indirect Effect of Satisfaction ................................................. 35
       3.6.6 Managerial Implications ......................................................................................... 39
4. Finite Mixture Model of Structural Heterogeneity and Empirical Results ....................... 43
   4.1 Structural Heterogeneity .................................................................................................. 43
   4.2 Finite Mixture Model ...................................................................................................... 43
   4.3 Empirical Results of Structural Heterogeneous Model .................................................. 45
       4.3.1 Model comparison .................................................................................................. 45
       4.3.2 Parameter Estimates of Nonlinear Term from CS to LOY ..................................... 46
       4.3.3 Measure of Tolerance Interval .............................................................................. 51
       4.3.4 Managerial Implications and Suggestions ............................................................... 53
5. Concluding Remarks .......................................................................................................... 57
References ............................................................................................................................... 60
Appendix A: Full Description of Model and Inference Procedure ............................... 64
Appendix B: MCMC Algorithms...........................................................................................................68
Appendix C: Figures of CSI Scores......................................................................................................77
Appendix D: Shapes of Nonlinear Functions.......................................................................................81
1. Introduction

Over the past several decades, the importance of customer loyalty has been largely realized by managers, because high level of loyalty can lead to several advantages, such as stronger competitive position leading to higher marketing share, lower price elasticity, lower business cost, reduced failure cost, and sustainable consumers (Erkan et al 2012). To predict and explain customer loyalty, customer satisfaction has been applied in large number of empirical studies (Kuma et al 2013).

The relationship between customer satisfaction and loyalty is one of the hotly discussed relationships in the field of marketing and services (Dong et al 2011; Kumar et al 2013). The premise that customer satisfaction, significantly affects customer loyalty, is widely accepted among both academic researchers and industrial managers. This relationship is believed to form the basis for measuring marketing effectiveness (Fornell 1992, Bolton and Lemon 1999, Anderson et al. 2004), of firms’ market, financial performance, and for firm value (Anderson and Mittal 2000; Gupta, Lehmann and Stuart 2004, Gupta and Zeithaml 2006). The fact that this relationship has been extensively studied in marketing and services over several decades highlights its important and critical role in determining the effectiveness of marketing programs and ensuring the creation of firm value through marketing action.

In this paper, we examine the relationship between customer satisfaction and customer loyalty with the survey data used to develop customer satisfaction index. The framework of analysis uses the customer satisfaction index (CSI) model as the starting point to propose a nonlinear structural equation model which includes a nonlinear function form between satisfaction and loyalty as one equation in the set of equations in this research. We contribute to the existing literature that examines satisfaction and loyalty variables with a comprehensive system of equations (Dong et al 2011; Kumar et al 2014), in contrast to previous studies using
just the metrics of satisfaction in isolation of their context.

As for functional forms of the nonlinear relationship, we consider piece-wise linear and S-shaped functions. The former is motivated by the ease of estimation, being close to linear model, and the latter specification is justified by prospect theory of Kahneman and Tversky (1979) and empirically supported by Ngobo (1999), whose research objective and dependent variable of loyalty are in common with ours.

From the methodological point of view, there are quite a few extant papers on nonlinear structural equation models. For example, Lee (2007) discusses a model with non-linearity only with respect to exogenous latent variables. This article contributes to the modeling literature by modeling nonlinear structural equations that include nonlinear terms of endogenous latent variable. Using the recursive property of system for CSI model, which means that latent variables are determined sequentially, we provide an efficient algorithm of MCMC in terms of multi-move sampler for latent variables by using Gibbs sampling.

In addition, we develop a hierarchical Bayes model with Finite Mixture to deal with structural heterogeneity across individual companies in the survey data. The model connects the structural models for respective company, and it leads to higher reliability of model estimates than the original customer satisfaction index model. This is accomplished by following insights from Terui et al. (2011). Finite Mixture is also employed to estimate the mixing proportions of candidate models, based on the method of Sylvia Fruehwirth-Schnatter (2006).

Chapter 2 discusses extant work on nonlinear relationship between satisfaction and outcomes including loyalty in the literature. We also discuss the perspective for building the framework of our analysis. In chapter 3, we propose the four candidates of HB nonlinear algorithm, the possible alternative specifications and empirical results are also discussed.
Chapter 4 introduces in detail the improved algorithm, which induces structural heterogeneous model- Finite Mixture Model, and the empirical results of this structural heterogeneity are also shown. Chapter 5 ends the conclusions.
2. Review of Related Literature

In the past research efforts, many kinds of nonlinear functional forms have been used to explain how customer satisfaction affects customer loyalty, but some have not reached the anticipated results. One reason is the theoretical limitation, for example, some possible components of satisfaction and loyalty are not considered sufficiently, or the use of some debatable measure indicators of satisfaction and loyalty. Another reason, which is more important, is the measure method of satisfaction and loyalty. Mostly, satisfaction and loyalty variables are directly measured by a survey directed to respondents. The impact of other antecedents on satisfaction is left out of consideration. Fornell (1992) discusses the need to use a comprehensive system of post-purchase outcomes in the way that satisfaction is part of the overall outcome that is measured. This is motivated by the fundamental principles that a variable should take on meaning depending on the context (Fornell, 1982, 1988), the joint analysis of measurement and latent variable modeling (Fornell and Yi, 1992), survey variables containing some degrees of errors (Andrews, 1984), and satisfaction which is not directly observable (Howard and Sheth, 1969, Oliver, 1981, Westbrook and Riley, 1983). In addition, he insists that, if satisfaction variable is measured in isolation of the context and it is used respectively to estimate the relationship, we tend to have results with low reliability, and get strongly biased parameter estimates.

In chapter 2, the theoretical introduction of nonlinear relationship between satisfaction and loyalty will be demonstrated, and then the measure method of satisfaction and loyalty by one kind of comprehensive system (CSI model) will be shown.

2.1 Nonlinear Relationship between Customer Satisfaction and Loyalty

2.1.1 Nonlinear Relations of Satisfaction
There are many extant works on nonlinear relationships between satisfaction and outcome variables. They investigate customer satisfaction relationship with one of firm’s outcomes by using different metrics. Most of the studies show that satisfaction has a positive and nonlinear asymmetric impact on firm's outcomes. These studies also support different forms and different functional forms are supported by these studies. Dong et al (2011) find that the linear functional forms of the effect from customer satisfaction on repurchase intentions may not be suitable under some restricted circumstances. They consider how the functional form of the effect from satisfaction to repurchase intentions varies across segments, which is formed by the product categories and customer economic and demographic variables. Their data set is obtained from the Chinese Customer Satisfaction Index, and they also use quadratic and cubic regression models as the nonlinear functional forms. Their research shows that, the linear functional form performs best across 972 product-customer segments (51%), followed by the S-shaped and convex nonlinear forms, and finally, the inverse S-shape and concave forms. In addition, different product attributes and customer characteristics, including customer economic and demographic variables, and market characteristics will moderate the effect of satisfaction on repurchase intentions differently, in terms of all the linear, quadratic, and cubic functions.

Kumar et al (2014) elaborate several issues which should be carefully considered in analyzing the efficacy of customer satisfaction in explaining and predicting customer loyalty. They find that for decades companies all around the world have kept investing customer satisfaction, in order to increase customers’ loyalty, and hence, profitability as the final purpose and consequence. However, after detailed observation and analysis, they find this kind of relationship is not as strong as it is believed, and customer satisfaction is not strong enough to explain the change of loyalty. The major findings of their research show that, there is a positive relationship between customer satisfaction and loyalty, but the variance explained by
just satisfaction is rather small. Models containing other relevant variables as moderators, mediators, and antecedent variables, or all three are better predictors of loyalty than just customer satisfaction. Further, the satisfaction–loyalty relationship will be changed potentially over time. Their study also offers specific guidelines to measure the satisfaction and loyalty of customers.

Fornell (1992) empirically showed that the relationship between customer satisfaction and their repurchase intention of goods or services, an indicator of loyalty, is nonlinear, and the dissatisfaction has greater influence than satisfaction on customers’ repurchase intentions. He uses the data from Customer Satisfaction Barometer (CSB), the satisfaction data includes Swedish customers in more than 30 industries and 100 companies. He found that satisfaction level should be lower in industries where the supply is homogeneous and demand is heterogeneous. Satisfaction should be higher when the heterogeneity of demand is matched by the homogeneity of supply. If the industries are highly dependent on repeat business, these industries tend to have a higher level of customer satisfaction.

Levels of customer satisfaction and dissatisfaction affecting the purchase intention and consumption behavior are also examined by Mittal and Kamakura (2001). They find that it is underestimated that using the linear formulas to measure the influence of satisfaction, so they suggest the nonlinear relations moderated by consumer heterogeneity across the attributes. They present a nonlinear model relating satisfaction ratings and consumption loyalty, and the repurchase behavior also working as the measurement of loyalty. To deal with the problem that the ratings observed in typical customer satisfaction survey method are error-prone measures, they assume that the satisfaction levels vary systematically on the basis of consumer characteristics and demographics. The authors apply the model to 100,040 automotive customers. Results show the heterogeneity of the consumers with different characteristics and
demographics. For example, even if the customers have the same level of satisfaction, their repurchase are systematically different. It is also found in their research that the heterogeneity of response bias in satisfaction ratings vary by customer characteristics. Furthermore, the nonlinear functional form relating rated satisfaction to repurchase intent is different from the one relating it to repurchase behavior. Although the functional form shows the decreasing returns in the case of repurchase intent, the monotonically increasing returns is shown in the case of repurchase behavior.

Keiningham et al (2003) examine the relationship between satisfaction and actual share-of-wallet in a business-to-business environment. The authors find that there is a positive relationship with nonlinear property, and the greatest positive impact happens at the upper limit of satisfaction levels.

Bowman and Narayandas (2004) adapt to the service-profit chain (SPC) framework, so as to accommodate the characteristics of the commercial market, especially the complex decision making units, strategic supplier selection, and allocation of resources in the level of individual customers. They also expand the SPC framework to describe the complex relationship between suppliers account profitability, namely nonlinear differential response to customer contact and competition environment caused by such factors as specific. By controlling such factors to certain degree, they try to find out the reason why similar levels of customer management effort and performance lead to totally different customer profitability outcomes. The authors prove the property and importance of decreasing returns in customer management and satisfaction management, and they reinforce the notion of customer delight.

Using data from the Canadian banking industry, Cooil et al. (2007) provide the longitudinal examination of the impact of changes in customer satisfaction on the changes in share of wallet, and they also test the moderating effects of demographics, including customer age, income,
education, expertise, and length of relationship. Their data indicate a positive relationship between changes in satisfaction and share of wallet. In particular, the initial satisfaction level and the conditional percentile of change in satisfaction have significant relationship with the changes in share of wallet. Other variables of demographic and situational characteristics are also tested in their research, but no impact is found.

Homburg, Koschate and Hoyer (2005) suggest inverse-S shaped nonlinear function for willingness to pay in their experimental study. Their research reveals the existence of a strong and positive impact of customer satisfaction on willingness to pay, and they prove it by using a nonlinear functional form based on disappointment theory (i.e., an inverse S-shaped form). In addition, their study provides evidence for the stronger impact of cumulative satisfaction on willingness to pay.

On the other hand, Ngobo (1999) examine the relation between satisfaction and loyalty (purchase and word-of mouth intentions) and suggest S-shaped function for this relationship. Their study questions the positive effectiveness from the ‘customer delight’ or ‘100% satisfaction’ to customer loyalty. A two-threshold, utility-oriented model of the effects of satisfaction on loyalty is developed in their research. Their empirical results show that trying to delight customers or provide 100% satisfaction may not be worth the effort, because there are points where the effect of customer satisfaction on loyalty levels off.

It is worth stressing that, Homburg, Koschate and Hoyer (2005) also suggest S shaped nonlinear function with the subjects of German university students. After comparing between inverse-S shaped and S shaped function by using cubic model, they prove the superiority of inverse-S shaped. However, the inverse-S shaped function goes against the property of satiation and decreasing marginal return when customers are high satisfied. Similarly, Dong et al. (2011) prove the inverse-S shaped function in terms of 19 kinds of goods, and S shaped function in
terms of 128 kinds of goods. According to the characteristics of inverse-S shaped function, zero marginal return occurs around the middle region, and positive marginal return occurs at both sides of the inverse-S shaped function. Oppositely, as for S shaped function, positive marginal return appears in the middle region, and zero marginal return appears at both sides. So it is better to construct the model combining the characteristics of both models. Homburg, Koschate and Hoyer (2005) mention that it would be worthwhile to examine the threshold point between the zero marginal and positive marginal regions.

2.1.2 Non-linearity and Moderating Effects
Non-linearity has been investigated in the context of moderating effects on the relationship between satisfaction and outcome variables. In particular, Jones and Sasser (1995) clarify that the competitive environment of the market affects the nonlinear relationship between satisfaction and loyalty. Their empirical research involves more than 30 companies in five industries (telephone communication companies, airlines, hospitals, personal computers, and automobiles) and shows that the relationship between customer satisfaction and loyalty is greatly influenced by the competitive environment.

Mittal and Kamakura (2001) and Cooil et al. (2007) discuss Non-linearity in the context of the moderating effect of consumer characteristics, and explain how demographics and customer characteristics have the moderating effects on satisfaction and repurchase intention or share of wallet.

Recently, Eisenbeiss, Corneliben, Backhaus and Hoyer (2014) investigate the moderating effects of firm reputation and customer involvement, and the properties of Non-linearity and asymmetric returns on satisfaction to willingness-to-pay are also considered in their model. Their empirical research shows that a firm with a strong reputation benefits from a broader zone of tolerance, compared with a firm of minor reputation. The customers who are highly involved
will react more intensively to extreme changes in satisfaction than those customers who are less
involved.

Kumar et al (2013) mention the variable of Word-of-Mouth (WOM) as a moderator, and
WOM has both positive and negative effect. The real value of those customers most loyal to
an entity stems more from their impact on other customers in the marketplace than from their
individual purchase behavior. At the same time, Customer satisfaction is considered an
antecedent of WOM. Their research shows that positive WOM from satisfied customers
lowers the cost of attracting new customers and enhances the firm’s overall reputation, while
that from unsatisfied customers has the opposite effect. In addition, the relationship between
customer satisfaction and WOM is characterized by the presence of moderators and mediators.
While customer satisfaction has a positive effect on customer referral, other variables seem to
predict WOM better.

Similarly, Terui et al. (2011) use Recommendation Intention (RI) as moderator in CSI
model. By using the assumption that the same model must be applied to every company, they
link the path coefficients of each company as the hierarchical regression model to estimate the
structure for customer satisfaction across companies to show that, representing “communality”
inside industry and “heterogeneity” outside industry, the hierarchical Bayes modeling produces
more stable significant path coefficients. In all the industries, RI acts as positive mediator to
prove the influence of satisfaction on loyalty.

Deng et al. (2013) use customer complaints as moderator in ACSI model. Their study
integrates consumption emotions into the American Customer Satisfaction Index (ACSI) model
of hotel industry. Their database includes 212 customers of international tourist hotels. The
research results show that customer complaints, as the mediator between satisfaction and loyalty,
weaken the influence of satisfaction.
Bloemer and Ruyter (1998) demonstrate Non-linearity by incorporating involvement as the key parameter between customer satisfaction and loyalty. Based on customers’ expression of equal levels of satisfaction, their study shows that highly involved customers exhibit greater loyalty than customers with low involvement. The relationship between store image, store satisfaction and store loyalty is examined. A distinction is made between true store loyalty and spurious store loyalty and manifest and latent satisfaction with the store. The positive relationship between manifest store satisfaction and store loyalty is stronger than the positive relationship between latent store satisfaction and store loyalty. Furthermore, they hypothesize a direct as well as an indirect effect through satisfaction of store image on store loyalty. Second, the relationship between store image and store loyalty is mediated by store satisfaction. The evidence for a direct effect of store image on store loyalty is not found.

2.2 Customer Satisfaction Index (CSI) Model

2.2.1 Aim and Definition of CSI Model
To improve the reliability of satisfaction and loyalty measuring, we use the joint analysis of measurement and latent variable modeling (Fornell and Yi, 1992) to measure satisfaction and loyalty, with the impact of other antecedents on satisfaction is also considered.

The customer satisfaction index uses the only uniform measure of customer satisfaction that allows comparison between companies and benchmark across industries. It also illustrates how customer satisfaction is embedded in a system of cause–effect relationships. Furthermore, this index is significant as a leading indicator of the financial results of the company (Anderson et al., 2004; Fornell et al., 1996, 2010). They employ the adopted expectancy disconfirmation as a basic theory which was proposed by Oliver (1980), who propose a model which expresses consumer satisfaction as a function of expectation and expectancy dis-confirmation, and the
results from the two-stage field study support the scheme for consumers and non-consumers of a flu inoculation. It is a model in which the level of customer satisfaction is decided by the degree of dis-confirmation between perceived quality after a purchase and customer expectation before a purchase.

The CSI model describes that customer expectations drive perceived quality and perceived value, and these three latent variables generate customer satisfaction. Customer satisfaction in turn directly affects customers’ voice and loyalty. The model estimation employs 17 manifest variables, which are ordered categorical variables based on survey questions rated on a scale of 1–10 (low–high). The scores of customer satisfaction are factor scores for n sampled customers’ satisfaction, and they are reported as standardized metrics between 0 and 100 points. Full description of model is provided in Appendix A. Using identical structure model for companies allows us to compare satisfaction level between companies, and the changes in it over years.

Given these reasons, we use the system of structural equations that contain satisfaction and loyalty as latent variables, i.e., customer satisfaction index (CSI) model by Fornell, et al. (1996). The CSI is a type of market-based performance measure for firms, industries, economic sectors, and national economies. Fornell, et al. (1996) also illustrates the use of CSI in conducting benchmarking studies, both cross-sectional and over time. In their research, the customer satisfaction should be greater for goods than for services and, in turn, greater for services than for government agencies, as well as find cause for concern in the observation that customer satisfaction in the United States is declining, primarily because of decreasing satisfaction with services. The authors conclude with a discussion of the implications of CSI for public policymakers, managers, consumers, and marketing in general.

The frame of CSI modeling is shown below:
Figure 1: Linear CSI Model

The structural equation model assumes that the factor scores \( \omega_i = (\eta_{i1}, \eta_{i2}, \ldots, \eta_{is}, \xi_i) \) have the relation each other in terms of set of equations:

- **Perceived Quality (PQ):** \( \eta_i = \gamma_{i1}\xi + \delta_i \)

- **Perceived Value (PV):** \( \eta_2 = \gamma_{21}\eta_1 + \gamma_{2}\xi + \delta_2 \)

- **Customer Satisfaction (CS):** \( \eta_3 = \gamma_{31}\eta_1 + \gamma_{32}\eta_2 + \gamma_{3}\xi + \delta_3 \)

- **Recommendation Intention (RE):** \( \eta_4 = \gamma_{43}\eta_3 + \delta_4 \)

- **Customer Loyalty (CL):** \( \eta_5 = \gamma_{53}\eta_3 + \gamma_{54}\eta_4 + \delta_5 \)

### 2.2.2 Classical Estimation of Linear CSI Model

The linear CSI model can be estimated by several methods. One method is Maximum Likelihood (ML) estimation, which assumes that the sample data follow a multivariate normal
distribution, so the information of means and covariance matrix are used. As for how much
the latent variables contribute to the manifest variable, most SEM software, can compute the
modification indices to minimize the chi-square value. In addition, the large enough sample
size is also required in ML estimation (Hox and Bechger 1998). To make the likelihood based
on Gaussian distribution, the ML estimation is based on the relationships between covariance of
manifest variables $\text{Cov}(x)$ by its sample estimate, 
$$S = \frac{1}{T} \sum_{i=1}^{T} (x_i - \bar{x})(x_i - \bar{x})' .$$

Another widely used method is Partial Least Squares (PLS), which simultaneously models
the structural paths and measurement paths. The PLS algorithm allows each manifest variable
to vary in how much it contributes to the composite score of the latent variable. Unlike ML,
estimation, the PLS estimation does not require a strong assumption of normal likelihood for
ordered categorical data and sufficient samples, hence Fornell and Cha (1994) applied PLS
method in the ACSI model for categorical data.

On the other hand, both ML and PLS estimations are only available when CSI model
premises a linear relationship between latent variables. In our research, for the purpose of
efficient algorithm and stable estimation, we extend the model to propose a nonlinear structural
equation model which includes nonlinear function from satisfaction to loyalty, so another
method-Bayesian approach is used, and Bayesian algorithm is shown in the following chapter.
3. Nonlinear Hierarchical Bayesian CSI Model and Empirical Results

In view of the related literates, we will put forward four kinds of nonlinear functions, which contain the properties of zone of tolerance, decreasing return and loss aversion. Then an efficient hierarchical Bayesian algorithm will be used to combine the nonlinear functions with SEM modeling. To make comparison of the model effects with the four nonlinear models, we make use of goodness of fit-DIC and log of marginal likelihood, and the ratio of significant coefficients.

3.1. Nonlinear Model on the Satisfaction to Loyalty

In the original CSI, Customer loyalty (LOY) is a function of customer satisfaction (CS) and recommendation to others (RE) (“Voice” in original CSI) and the model is described as one equation in the set of equations of CSI model by

$$LOY_i = \beta CS_i + \alpha RE_i + \epsilon_i,$$

where $i$ means the index for respondent (customer), $\epsilon_i$ is the normally distributed error term, and $\beta, \alpha$ are path coefficients.

This linear specification is reasonable as local approximation to possibly more complicated relations, but it has limitations in failing to accommodate some characteristics discussed in the literature, i.e., (i) not constant return to scale, (ii) saturation effects and (iii) asymmetric response. These are well captured by nonlinear models.

We model the accommodation of these characteristics in a parsimonious way as follows:

$$LOY_i = \beta^{(+)I} \left\{ \frac{1}{1+\exp(-CS_i + r_0)} \right\} \frac{1}{2} + \beta^{(-I)} \left\{ 1-I \right\} \left\{ \frac{1}{1+\exp(-CS_i + r_0)} \right\} \frac{1}{2} + \alpha RE_i + \epsilon_i,$$

where $I$ is the indicator function taking 1 if $CS_i > r_0$ and zero otherwise, $r_0$ is identified as 0, the change point dividing highly satisfied and low satisfied regions. The shape of function
is depicted as “Asymmetric Logit” in Figure 2. This function is in line with Prospect Theory (Kahamena and Tversky, 1979) and empirically supported by Ngobo (1999) whose research objective and dependent variable of loyalty closely resemble ours in the literature.

Figure 2: Linear and Nonlinear CSI Model

(a) Asymmetric Linear
(b) Asymmetric Logit
(c) Threshold Linear
(d) Threshold Logit

We identify $r_0 = 0$ as an example in Figure 2.

The nonlinear function captures nonlinear effects in terms of logistic function. That is, it
has asymmetric response around the inflection point, i.e., $CS = r^*_0$, which is fixed for identification of model, and also the marginal return to LOY is changing at every level of CS. The upper limit represents the satiation level and the lower limit indicates the baseline independent of CS. They are respectively provided by $\frac{1}{2} \beta^{(+)}$ and $-\frac{1}{2} \beta^{(-)}$,

$$LOY_{CS \to +\infty} = \beta^{(+)} \left( \frac{1}{1 + 0} - \frac{1}{2} \right) + \alpha RE_i = \frac{1}{2} \beta^{(+)} + \alpha RE_i,$$

$$LOY_{CS \to -\infty} = \beta^{(-)} \left( \frac{1}{1 + \infty} - \frac{1}{2} \right) + \alpha RE_i = -\frac{1}{2} \beta^{(-)} + \alpha RE_i.$$ 

Thus the estimates of coefficients $\beta^{(+)}$ and $\beta^{(-)}$ determine the maximum and minimum levels of loyalty caused by satisfaction. They also define the speed to reach these limits in respective regimes. They provide several interesting implications for service management.

On the other hand, the company with larger $\beta^{(+)}$ value will attain the attainable limit of loyalty quickly by additional effort on satisfaction as long as it stays in the positive regime.

On the other hand, the smaller value of $\beta^{(-)}$ shows that the increasing rate of return is slower although the lowest limit might not be all that detrimental when it is in negative regime. That is, the model suggests that, given a level of CS score, if the marginal return of CS increases, the attainable limit of loyalty will also increase. We call model (2) as “asymmetric logit model.”

When we replace the logistic function by simply $CS_i$ as a piece-wise linear term, we obtain the “asymmetric linear model” which we use for comparison purposes in the empirical application,

$$LOY_i = \beta^{(+)} I\{CS_i\} + \beta^{(-)} (1 - I)\{CS_i\} + \alpha RE_i + \varepsilon_i. \quad (3)$$

This model, which is named as “Asymmetric Linear” in Figure 2, defines different slopes of
linear response in respective regimes. It approximates nonlinear relation by piece-wise linear functions, and it does not have attainable limits since it models the relationship locally, even though it has more useful information than a simple linear model.

In addition, we consider the inverse S-shaped models, as is discussed in Homburg et al. (2005), by introducing thresholds, \( r_1 \) and \( r_2 \), in the satisfaction domain, which define the zone of tolerance.

\[
LOY_i = \beta^{(+,r_1)} \{CS_i\} + \beta^{(r_2,-)} (1-I) \{CS_i\} + \alpha RE_i + \epsilon_i, \tag{4}
\]

where \( \beta^{(+,r_1)} \) takes some value when \( \max \{CS_i\} > r_1 > 0 \) and zero otherwise, \( \beta^{(r_2,-)} \) takes some value when \( 0 > r_2 > \min \{CS_i\} \) and zero otherwise. We call this “threshold linear model”. Similarly, “threshold logit model” is specified as

\[
LOY_i = \beta^{(+,r_1)} \left[ \frac{1}{1+\exp(-CS_i + r_1)} - \frac{1}{2} \right] + \beta^{(r_2,-)} (1-I) \left[ \frac{1}{1+\exp(-CS_i + r_2)} - \frac{1}{2} \right] + \alpha RE_i + \epsilon_i, \tag{5}
\]

We compare these alternative models in the empirical application. The functions of these models are depicted in Figure 2 for comparison.

We employ these nonlinear models just for the customer satisfaction to loyalty relationship and keep linear structural equations for other relations in CSI model. Full description of our model is given at Appendix A.

3.2. Structural Equation Modeling with Bayesian Approach

In behavioral, educational, medical and social sciences, substantive theory usually involves two kinds’ variables, namely manifest (observed) and latent (potential) variables. Manifest variables are those can be measured directly, such as the questionnaire scores in the linear CSI model. As for the latent variables, such as the satisfaction and loyalty, they cannot be measures directly, but can be partially measured by a linear combination of some manifest
variables.

As Lee (2007) mentioned in his theory, in formulating various SEMs, and in developing the Bayesian methods, the emphasis is placed on the raw individual random observations rather than on the sample covariance matrix. Compared with classical SEM estimation method, such as ML and PLS, Bayesian approach leads to a direct estimation of latent variables, and gives better results in small samples, and most importantly, it can be applied in complicated situations, like inducing nonlinear path between latent variables.

The standard linear SEM includes two parts:

Measurement model \[ x_i = \Lambda \omega_i + \epsilon_i, \quad i = 1, \ldots, n, \quad (6) \]

Structural model \[ \eta_i = \Pi \eta_i + \Gamma \xi_i + \delta_i, \quad (7) \]

where \( x_i \) represents the manifest variables, \( \Lambda \) represents the unknown factor loading matrix, \( \omega_i = (\xi_i, \eta_{1i}, \eta_{2i}, \ldots, \eta_{ki}) \) is latent variable vector, \( \Pi \) and \( \Gamma \) are unknown parameter matrix of regression coefficients, the \( \epsilon_i \) and \( \delta_i \) are error terms.

The structural model provides the prior distribution density of latent variables \( \omega_i \), and measurement model provides the likelihood. The parameters of measurement model \( \theta_x \) and structural model \( \theta_\omega \) can be estimated by using standard Bayesian linear regression. The Gibbs sampler is used to estimate the latent variables and parameters, at the \( j \)th iteration:

(i) Generate \( \omega_i^{(j)} \) from \[ p(\omega_i^{(j)} \mid x_i, \omega_i^{(j-1)}, \theta_x) \propto p(x_i \mid \omega_i^{(j)}, \theta_x^{(j-1)}) p(\omega_i^{(j)} \mid \theta_\omega^{(j-1)}) \],

(ii) Generate \( \theta_x^{(j)} \) from \[ p(\theta_x^{(j)} \mid x_i, \omega_i^{(j)}) \propto p(x_i \mid \omega_i^{(j)}, \theta_x^{(j)}) p(\theta_x^{(j)}) \],

(iii) Generate \( \theta_\omega^{(j)} \) from \[ p(\theta_\omega^{(j)} \mid \omega_i^{(j)}) \propto p(\omega_i^{(j)} \mid \theta_\omega^{(j)} p(\theta_\omega) \).

Based on this Gibbs sampler loop, we induce the nonlinear function form in step (iii), and the details of the above conditional posterior density are shown in Appendix A.
3.3. Hierarchical Bayes Modeling for Stable Estimation

The CSI model assumes that every company has the identical structure on customer satisfaction for the purpose of comparing the services across different companies, and thus is aggregated to industry groups and national levels. However, each company should have heterogeneity on customer satisfaction measures. The heterogeneity appears to produce the result that some path coefficient estimates are not significant for some companies, which reduces the credibility of scores. To overcome this problem, Terui et al. (2011) proposed the hierarchical Bayes modeling of customer satisfaction index to increase reliability of model estimates not only for the goodness of fit, but also the number of significant estimates of path coefficient. That is, HB model produces a larger number of significant estimates in the model and better goodness of fit than independently estimated model. This result comes from the property that HB modeling borrows information of neighbors by pooling data to get the stable estimate of parameters on the assumption that they share homogeneity in some aspects regardless of independent information.

In this study, we employ HB modeling which relates the model of each company $h = 1, ..., H$ such that

$$
\beta_h = \Theta^T z_h + \eta_h; \quad \eta_h \sim N_k \left(0, V_{\beta}\right),
$$

where $\beta_h$ is the vector of parameters $\theta$, and $\theta_{\omega}$ in measurement and structure model, and $z_h$ is attribute data for the company $h$, and $\eta_h$ is error term. This is prior distribution on the path coefficients $\beta_h$, and this means that the path coefficients are not independent and restricted by the common parameter $(\Theta, V_{\beta})$. The prior specification (6), together with appropriate prior specifications for other parameters including latent variables of $\omega_{hi}$, are combined with the Gaussian likelihood based to constitute joint posterior density,
The numerical evaluation of this density is conducted by Markov chain Monte Carlo (MCMC), and its algorithm is described together with prior specifications in Appendix B.

3.4. Model Selection

Since we have put forward four candidates of nonlinear models in 3.1 and HB modeling method in 3.3, the next step is to set goodness of fit to identify the best candidate. Let’s set the structural homogeneous assumption as that: all companies obey an identical nonlinear structure which is chosen from the four candidate models, while the value of path coefficient of each company is different. We use Deviance Information Criterion (DIC) and Log of Marginal Likelihood (LML) as goodness of fit.

The deviance information criterion (DIC) is a hierarchical modeling generalization of the AIC (Akaike information criterion) and BIC (Bayesian information criterion, also known as the Schwarz criterion). It is particularly useful in Bayesian model selection problems where the posterior distributions of the models have been obtained by Markov chain Monte Carlo (MCMC) simulation. Like AIC and BIC, it is an asymptotic approximation as the sample size becomes large. Because the numbers of estimators in the four kind of nonlinear models are different, DIC is efficient for the problem of overfitting, which means increasing the number of parameters in the model almost always improves the goodness of the fit. Bradley and Thomas (2009) show the measurement of DIC.

\[
\text{DIC} = \bar{D} + p_D = 2\bar{D} - D(\bar{\theta}),
\]

where \( p_D = E_{\theta|Y}[D] - D(E_{\theta|Y}[\theta]) = \bar{D} - D(\bar{\theta}) \), it means the complexity of a model by the effective number of parameters.

\[
D(\theta) = -2 \log f(Y | \theta) + 2 \log h(Y),
\]
where \( f(Y \mid \theta) \) is the likelihood function for observed data \( Y \) given the parameter vector \( \theta \), and \( h(Y) \) is some standardizing function of the data alone (which thus has no impact on model selection). In this approach, the fit of a model is summarized by the posterior expectation of deviance \( \bar{D} = E_{\theta \mid Y} [D] \), while the complexity of a model is captured by the effective number of parameters \( p_D \). In model selection, DIC deals with the trade-off between the goodness of fit of the model and the complexity (number of estimators) of the model. The model with smaller DIC value is selected. The specific form of DIC in our research is

\[
DIC = 2 \sum_{j=1}^{J} \left( -2 \log f(Y \mid \theta_j^0, M_k) + 2 \log h(Y) \right) - \left( -2 \log f(Y \mid \overline{\theta}_k, M_k) + 2 \log h(Y) \right)
\]

\[
\propto \frac{-4}{J} \sum_{j=1}^{J} \left( \log f(Y \mid \theta_j^0, M_k) \right) + 2 \log f(Y \mid \overline{\theta}_k, M_k).
\]

Marginal likelihood is also one kind of criterion, it is often used in Bayesian model selection. Newton and Raftery (1994) put forward the harmonic mean of the marginal likelihood \( \hat{p}(Y \mid M_k) \) to check the model fit and convergence, and it is called as Newton-Raftery (NR) estimate. In our research, we use log of marginal likelihood

\[
\log(\hat{p}(Y \mid M_k)).
\]

\[
LML_k = \log(\hat{p}(Y \mid M_k)) = \left( \frac{1}{J} \sum_{j=1}^{J} \frac{1}{\log(p(Y \mid \theta_j^0, M_k))} \right)^{-1}.
\] (12)

Based upon Importance-Sampling Approaches, when \( J \to \infty \), the harmonic mean of the marginal likelihood value converges to be the correct value \( \log(p(Y \mid M_k)) \) by using posterior draws.

The model with larger LML is selected. We will estimate the nonlinear CSI models and choose the candidate model according to DIC and LML.
3.5. Data and Data Processing

3.5.1 Data Description

The data set is available from the Japanese CSI development working group managed by the Japanese Agency of Service Productivity and Innovation Growth. We use the data for survey conducted in 2008 year, and it includes 21 companies in three industries—mobile telecommunications (4), convenience stores (5), hotels (12). The sample sizes used in analysis are: mobile telecommunications (company1 = 456, company2 = 456, company3 = 360), convenience stores (company1 =456, company2 = 456, company3 = 360), hotels (company1 = 300, company2 = 300, company3 = 300).

3.5.2 Data Transformation

We employ Bayesian inference on the model estimation as is explained in Lee (2007) and Terui et al. (2011) on the grounds of distributional property of observations as well as derived distribution of estimated satisfaction scores. The data are measured by 10 point Likert scale. Thus the ordered categorical data are not consistent with normality assumption. On the other hand, the structural equation model is developed on the assumption of normality on variables. The American CSI model employs PLS method for model estimation since PLS does not assume any distribution on the error terms to estimate the model parameters. However, there is no free lunch. In fact, Terui et al. (2011) compares the estimates by Bayesian MCMC method with those by PLS, and it demonstrates that the distribution of estimated satisfaction score by PLS method is mostly skewed, also, the score distribution evaluated by Bayesian MCMC algorithm is stable and symmetric. The satisfaction score is calculated by taking sample means of estimated respondent’s scores, and being standardized to be inside 0 and 100 points, and the satisfaction score must be reasonable only when the distribution is symmetrical.
In order to be consistent with our inference below, we first transform ordered categorical data into continuous variable, which follows the specified normal distribution by way of data augmentation (Lee, 2007; Terui et al., 2011). We introduce a set of cut points across the normal distribution to decompose it into 10 segments that may be categorized on a scale of 1 to 10. Thus, the probability of each region corresponds to the probability mass of each ordered category. When we have a categorical sample, we generate the samples from the truncated normal distribution whose cut points are defined by the corresponding segment.

3.5.3 Full Conditional Posterior Density and Multi-Move Gibbs Sampler

The algorithm for Bayesian inference of linear structural equation model is given by Lee (2007). By using the special properties of CSI model that the latent variables $\omega = (\xi, \eta_1, \eta_2, \eta_3, \eta_4)'$ are determined sequentially by the initial driving force of “expectation $\xi$”, in the way that $\eta_1 \rightarrow \eta_2 \rightarrow \eta_3 \rightarrow \eta_4 \rightarrow \eta_5$ and also the nonlinear equation of $\eta_5$ (LOY) by $\eta_3$ (CS) is positioned in the last. The efficient algorithm is available for generating posterior distribution of latent variables.

We first decompose the set of latent variables into linear and nonlinear parts

$$\omega_1 = (\xi, \eta_1, \eta_2, \eta_3, \eta_4)'$$

and $\eta_5$. Multi-move Gibbs sampler is induced for $\omega_1$.

In particular, for the conditional prior, according to the structural model, the covariance of latent variables is 0. The joint prior density of latent variables is

$$p(\xi, \eta_1, \eta_2, \eta_3, \eta_4, \eta_5) = p(\xi) p(\eta_1 | \xi) p(\eta_2 | \eta_1, \xi) p(\eta_3 | \eta_1, \eta_2, \xi) p(\eta_4 | \eta_2, \xi) p(\eta_5 | \eta_3, \eta_4),$$

and the prior of $\omega_1 = (\xi, \eta_1, \eta_2, \eta_3, \eta_4)'$ is

$$p(\xi, \eta_1, \eta_2, \eta_3, \eta_4) = p(\xi) p(\eta_1 | \xi) p(\eta_2 | \eta_1, \xi) p(\eta_3 | \eta_1, \eta_2, \xi) p(\eta_4 | \eta_3),$$
and the conditional prior of \( \eta_5 \) is
\[
p(\eta_5 | \xi, \eta_1, \eta_2, \eta_3, \eta_4) = p(\eta_5 | \eta_3, \eta_4),
\]
(15)

As for the conditional likelihood, the observed data \( x_i = [x_{i1}, x_{i2}, x_{i3}, \ldots, x_{i17}] \) of customer \( i \) obey the independent multivariate normal distribution, according to the measurement model.

Particular, we set the manifest variables of \( \omega_i = (\xi, \eta_1, \eta_2, \eta_3, \eta_4)' \)
as \( x_i^{[1]} = [x_{i1}, x_{i2}, x_{i3}, \ldots, x_{i11}, x_{i15}, x_{i16}, x_{i17}] \), and set the manifest variables of \( \eta_5 \) as \( x_i^{[2]} = [x_{i12}, x_{i13}, x_{i14}] \), and then the conditional likelihood is
\[
p(x_i^{[1]} | \eta_1, \eta_2, \eta_3, \eta_4, \xi, \theta) = p(x_i^{[1]} | \eta_1, \eta_2, \eta_3, \eta_4, \xi, \theta)p(x_i^{[2]} | \eta_5, \theta),
\]
(16)

Then we briefly express the joint prior density is defined by \( p(\omega) = p(\omega_i) p(\eta_5 | \omega_i) \) to derive marginal posterior density of linear latent variables \( \omega_i \) and conditional density of nonlinear latent variable of \( \eta_5 \) on \( \omega_i \)

(i) \( p(\omega_i | x, \theta) \propto p(\omega_i) p^{[\eta_5]}(x | \omega_i, \theta), \)
(17)

(ii) \( p(\eta_5 | \omega_i, x, \theta) \propto p(\eta_5 | \omega_i) p^{[\eta_5]}(x | \omega_i, \theta), \)
(18)

where \( \theta \) means the set of model parameters including factor loadings, path coefficients, and variances, and \( x \) is data. The algorithm for linear part (i) is given by Lee (2007). The multi-move sampler is available for the vector of latent variables. The path coefficient parameters are defined as linear in our model, and the algorithm for linear structural equation model is available, together with other parameters of factor loadings and variance, in Lee (2007).

The nonlinear part (ii) is the product of normal prior and normal likelihood, and thus the
posterior density is analytically derived by using conjugate property. The details are explained in the algorithm section in Appendix B.

Finally, the conditional posterior density of model parameter, \( p(\theta | \omega, x) \), is available in Lee (2007). Since this is the same structure conditional on latent variables \( \omega \) with linear structural equation model, the details of full conditional posterior density are provided in Appendix B.

3.6. Empirical Results of HB Nonlinear CSI Model

3.6.1. Model Comparison

We estimated the parameter using Markov chain Monte Carlo (MCMC) by the use of Gibbs sampling. This section reports results of the comparison between models by comparing the values of Deviance Information Criterion (DIC), an information criterion of Bayesian analysis as well as log of marginal likelihood (LML).

We compare (i) linear model, (ii) logit model, (iii) asymmetric linear model, and (iv) asymmetric logit model in their HB estimations. As a benchmark model, we also set the original CSI model, denoted by (0) independent linear model.

Table 1: Model Comparison: DIC and log of Marginal Likelihood

<table>
<thead>
<tr>
<th></th>
<th>Independent</th>
<th>HB</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Linear</td>
<td>Logit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Symmetric</td>
<td>Asymmetric</td>
</tr>
<tr>
<td>DIC</td>
<td>238539.1</td>
<td>234960.5</td>
<td>234912.3</td>
</tr>
<tr>
<td>LML</td>
<td>-98516.1</td>
<td>-96844.0</td>
<td>-96577.7</td>
</tr>
</tbody>
</table>

Table 1 shows the calculated values of DIC and LML for the different models, and we also add the independent linear, symmetric linear and logit model. First of all, both measures support the HB models than independent linear model, and the advantage of HB modeling is
more evident for the measure of LML. The comparison between linear and nonlinear models supports nonlinear models by both criteria, and within groups, asymmetric response models are supported more than symmetric models: HB asymmetric linear model is better than HB linear model, and HB asymmetric logit model performs better than HB logit model in case of DIC.

We notice that LML of HB logit model slightly shows better fit than HB asymmetric logit model. However, the latter model contains double number of response parameter of the former, and we employ DIC which discounts the number of parameter more appropriately than LML.

Table 2 tabulates the number of path coefficients that were not significant in the sense of 95% highest probability density (HPD) region for respective models. The total number of insignificant estimates in each model is shown at the bottom of table. The effect of HB modeling that borrows other company’s data on the estimates is evident. The number of insignificant estimates in (0) independent model is drastically reduced from 27 to 16 for (i) HB linear; 17 for (ii) HB asymmetric linear; 13 for (iii), (iv) HB (asymmetric) logit models. The heterogeneity in industry is evident to see that hotels have relatively more insignificant estimates.

Table 2: The Effect of HB modeling on Estimate of Path Coefficients

<table>
<thead>
<tr>
<th></th>
<th>Independent</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Linear</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Symmetric</td>
<td>Asymmetric</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>14.29%(27/189)</td>
<td>8.47%(16/189)</td>
</tr>
<tr>
<td>Mobile Telecommunication</td>
<td></td>
<td>19.44%(7/36)</td>
<td>5.56%(2/36)</td>
</tr>
<tr>
<td>Convenience stores</td>
<td></td>
<td>4.44%(2/45)</td>
<td>2.22%(1/45)</td>
</tr>
</tbody>
</table>

The number means the percentage of significant estimates in the model. The ratio is given in parenthesis.

The examination of results on model fit criteria in Table 1 and the number of significant

30
parameters estimate in Table 2 shows the order of better models in terms of fitness is HB asymmetric logit, HB logit, HB asymmetric linear, and HB linear models. The asymmetric response is better supported and furthermore the model gets more advantageous if the saturation effects are incorporated in the model.

3.6.2. Parameter Estimates of Nonlinear Term from CS to LOY

Table 3.1 shows the estimates (posterior mean) of coefficient of $\beta^{(+) and \beta^{(-)}$ of nonlinear term from satisfaction to loyalty for individual companies. 95% HPD region is also given next to the estimate. The industry level estimates given in Table 3.2 are derived from posterior means of industrial dummy in the hierarchical model.
Table 3.1: Parameter Estimates (Company Level)

<table>
<thead>
<tr>
<th>Company</th>
<th>( \beta(-) )</th>
<th>HPD</th>
<th>( \beta(+) )</th>
<th>HPD</th>
<th>( \Pr(\beta(-) &gt; \beta(+)) )</th>
<th>Band Width</th>
<th>( \alpha )</th>
<th>HPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>1.63</td>
<td>[0.98, 2.28]</td>
<td>0.91</td>
<td>[0.29, 1.56]</td>
<td>91.12%</td>
<td>1.27</td>
<td>0.22</td>
<td>[0.12, 0.33]</td>
</tr>
<tr>
<td>M2</td>
<td>1.62</td>
<td>[1.06, 2.18]</td>
<td>0.82</td>
<td>[0.25, 1.41]</td>
<td>95.54%</td>
<td>1.22</td>
<td>0.24</td>
<td>[0.15, 0.34]</td>
</tr>
<tr>
<td>M3</td>
<td>1.46</td>
<td>[0.82, 2.04]</td>
<td>0.82</td>
<td>[0.19, 1.44]</td>
<td>89.28%</td>
<td>1.14</td>
<td>0.24</td>
<td>[0.15, 0.34]</td>
</tr>
<tr>
<td>M4</td>
<td>1.68</td>
<td>[1.09, 2.28]</td>
<td>0.99</td>
<td>[0.42, 1.56]</td>
<td>92.34%</td>
<td>1.33</td>
<td>0.23</td>
<td>[0.13, 0.33]</td>
</tr>
<tr>
<td>Industry mean</td>
<td>1.60</td>
<td>0.88</td>
<td>92.07%</td>
<td>1.24</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>1.79</td>
<td>[1.18, 2.37]</td>
<td>1.43</td>
<td>[0.85, 2.01]</td>
<td>77.56%</td>
<td>1.61</td>
<td>0.30</td>
<td>[0.21, 0.39]</td>
</tr>
<tr>
<td>C2</td>
<td>1.75</td>
<td>[1.25, 2.26]</td>
<td>1.57</td>
<td>[1.06, 2.12]</td>
<td>66.70%</td>
<td>1.66</td>
<td>0.15</td>
<td>[0.07, 0.23]</td>
</tr>
<tr>
<td>C3</td>
<td>1.70</td>
<td>[0.99, 2.34]</td>
<td>1.40</td>
<td>[0.72, 2.08]</td>
<td>70.92%</td>
<td>1.55</td>
<td>0.22</td>
<td>[0.11, 0.35]</td>
</tr>
<tr>
<td>C4</td>
<td>2.00</td>
<td>[1.46, 2.54]</td>
<td>1.77</td>
<td>[1.23, 2.37]</td>
<td>69.46%</td>
<td>1.89</td>
<td>0.26</td>
<td>[0.17, 0.35]</td>
</tr>
<tr>
<td>C5</td>
<td>1.84</td>
<td>[1.30, 2.37]</td>
<td>1.71</td>
<td>[1.19, 2.25]</td>
<td>61.88%</td>
<td>1.78</td>
<td>0.25</td>
<td>[0.16, 0.35]</td>
</tr>
<tr>
<td>Industry mean</td>
<td>1.81</td>
<td>1.58</td>
<td>69.30%</td>
<td>1.69</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>1.42</td>
<td>[0.97, 1.88]</td>
<td>1.19</td>
<td>[0.74, 1.66]</td>
<td>75.38%</td>
<td>1.31</td>
<td>0.26</td>
<td>[0.16, 0.37]</td>
</tr>
<tr>
<td>H2</td>
<td>1.31</td>
<td>[0.83, 1.80]</td>
<td>0.93</td>
<td>[0.42, 1.44]</td>
<td>85.44%</td>
<td>1.12</td>
<td>0.31</td>
<td>[0.21, 0.44]</td>
</tr>
<tr>
<td>H3</td>
<td>1.44</td>
<td>[1.06, 1.83]</td>
<td>1.06</td>
<td>[0.63, 1.50]</td>
<td>89.36%</td>
<td>1.25</td>
<td>0.28</td>
<td>[0.20, 0.44]</td>
</tr>
<tr>
<td>H4</td>
<td>1.34</td>
<td>[0.92, 1.75]</td>
<td>0.95</td>
<td>[0.54, 1.38]</td>
<td>89.58%</td>
<td>1.15</td>
<td>0.29</td>
<td>[0.20, 0.39]</td>
</tr>
<tr>
<td>H5</td>
<td>1.28</td>
<td>[0.77, 1.76]</td>
<td>1.10</td>
<td>[0.57, 1.63]</td>
<td>68.60%</td>
<td>1.19</td>
<td>0.35</td>
<td>[0.23, 0.46]</td>
</tr>
<tr>
<td>H6</td>
<td>1.48</td>
<td>[1.08, 1.91]</td>
<td>0.87</td>
<td>[0.45, 1.28]</td>
<td>97.66%</td>
<td>1.18</td>
<td>0.21</td>
<td>[0.12, 0.31]</td>
</tr>
<tr>
<td>H7</td>
<td>1.54</td>
<td>[1.09, 2.02]</td>
<td>1.06</td>
<td>[0.58, 1.54]</td>
<td>91.78%</td>
<td>1.30</td>
<td>0.33</td>
<td>[0.23, 0.45]</td>
</tr>
<tr>
<td>H8</td>
<td>1.26</td>
<td>[0.75, 1.76]</td>
<td>0.79</td>
<td>[0.26, 1.30]</td>
<td>89.30%</td>
<td>1.03</td>
<td>0.31</td>
<td>[0.20, 0.42]</td>
</tr>
<tr>
<td>H9</td>
<td>1.36</td>
<td>[0.91, 1.83]</td>
<td>1.11</td>
<td>[0.63, 1.63]</td>
<td>77.10%</td>
<td>1.24</td>
<td>0.22</td>
<td>[0.13, 0.34]</td>
</tr>
<tr>
<td>H10</td>
<td>1.54</td>
<td>[1.06, 2.05]</td>
<td>0.92</td>
<td>[0.41, 1.42]</td>
<td>95.36%</td>
<td>1.23</td>
<td>0.36</td>
<td>[0.25, 0.48]</td>
</tr>
<tr>
<td>H11</td>
<td>1.47</td>
<td>[1.04, 1.92]</td>
<td>1.10</td>
<td>[0.62, 1.57]</td>
<td>87.08%</td>
<td>1.28</td>
<td>0.27</td>
<td>[0.17, 0.38]</td>
</tr>
<tr>
<td>H12</td>
<td>1.44</td>
<td>[1.03, 1.89]</td>
<td>0.99</td>
<td>[0.54, 1.46]</td>
<td>91.80%</td>
<td>1.22</td>
<td>0.24</td>
<td>[0.15, 0.34]</td>
</tr>
<tr>
<td>Industry mean</td>
<td>1.41</td>
<td>1.01</td>
<td>86.54%</td>
<td>1.21</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total mean</td>
<td>1.54</td>
<td>1.12</td>
<td>83.49%</td>
<td>1.33</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Parameter Estimates (Industry Level)

<table>
<thead>
<tr>
<th>Industry</th>
<th>( \beta(-) )</th>
<th>HPD</th>
<th>( \beta(+) )</th>
<th>HPD</th>
<th>( \Pr(\beta(-) &gt; \beta(+)) )</th>
<th>Band Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile Telecommunication</td>
<td>1.60</td>
<td>[0.97, 2.22]</td>
<td>0.88</td>
<td>[0.26, 1.52]</td>
<td>92.02%</td>
<td>0.72</td>
</tr>
<tr>
<td>Convenience Stores</td>
<td>1.81</td>
<td>[1.25, 2.35]</td>
<td>1.58</td>
<td>[1.02, 2.14]</td>
<td>71.22%</td>
<td>0.24</td>
</tr>
<tr>
<td>Hotels</td>
<td>1.41</td>
<td>[1.03, 1.78]</td>
<td>1.01</td>
<td>[0.62, 1.41]</td>
<td>91.94%</td>
<td>0.40</td>
</tr>
</tbody>
</table>
The posterior mean of parameter estimate and 95% HPD region are given for respective parameters. The table contains the column for band width \( \left| \frac{1}{2} \beta^{(+)} + \frac{1}{2} \beta^{(-)} \right| \) for attainable limits.

First of all, path coefficients are significant for all companies as HDP region does not include zero with the level of 95% probability. Second, we observe the estimate of \( \beta^{(-)} \) is greater than that of \( \beta^{(+)} \) for all cases. More precisely, the posterior probability \( \Pr \{ \beta^{(-)} > \beta^{(+)} \} \) is given in the fifth column of the Table 3.1 to show that it holds with high probability for most companies, and Table 3.2 shows that probability of at industry level is the highest 92.02\% for mobile telecommunication and the lowest 71.22\% for convenient store industry. These coefficients respectively determine the lower limit \( -\frac{1}{2} \beta^{(+)} \) and the upper limit \( \frac{1}{2} \beta^{(+)} \) of loyalty over satisfaction dimension, and it turns out that the speed of reaching the limit is slower in positive regime than negative regime. This means that the loss aversion is observed for every company across industries. The customers recognizing dissatisfaction induce great depreciation of loyalty compared with the same amount of increase of satisfaction in positive regime.

Finally, we observe that these estimates of convenience store industry are relatively larger than those in other industry. The mean value of the estimated difference \( \beta^{(-)} - \beta^{(+)} \) is 0.72 for mobile telecommunication, 0.24 for convenience store; and 0.40 for hotels. This implies that the loss aversion is most pronounced in telecommunication industry, and next is hotels although the situation is rather heterogeneous within hotels. Next we consider the band width \( \left| \frac{1}{2} \beta^{(+)} + \frac{1}{2} \beta^{(-)} \right| \) between the upper and lower limits. This is a measure of importance of satisfaction on the variation of loyalty. The mean value of band width is estimated as 1.24 for
mobile telecommunication, 1.69 for convenience store and 1.21 for hotels. These suggest that the satisfaction in convenience stores industry is most likely to produce significant impact on loyalty. On the other hand, mobile telecommunication companies are not relatively well placed to gain the loyalty by means of satisfaction.

3.6.3. Estimated Functional Form

Figure 3 depicts the figure of estimated functional form of CS → LOY. We observe asymmetric responses in respective regimes and upper limits are smaller than negative of lower limits, implying loss aversion in relationship between satisfaction and loyalty. This is the most evident for M2 of mobile telecommunications as the difference is -0.40 (= 0.41 - 0.81). The opposite situation happens for C5 in convenience stores, i.e., -0.07 (=0.92-0.85).

3.6.4. Distribution of Satisfaction Score

We express the levels of satisfaction and loyalty on 100-point scale as is usually reported in customer satisfaction index (CSI). The CSI score is calculated as the standardized factor
scores for customer satisfaction by \( \frac{CS_i - \min[CS_i]}{\max[CS_i] - \min[CS_i]} \times 100 \). Figure 3 depicts the empirical distribution of respondent’s scores for each company. In the figure, the statistics of mean, median, and standard deviations as well as number of respondents are shown as legends. The score distributions are heterogeneous among companies. However, the distributions are relatively stable and symmetric since the difference of mean and median is small for every company. This is consistent with the study by Terui et al. (2011) on the ground of estimation after transformation of original categorical data to normal distributed data by data augmentation for Bayes modeling. Thus the sample mean would be reasonable estimate of CSI score even for nonlinear model.

Figure 4: Distribution of Satisfaction Score

The specific figures of 21 companies are shown in Appendix C.

3.6.5. Marginal Effects and Indirect Effect of Satisfaction

Loyalty is determined not only by satisfaction, but also by the intention to recommend to others. According to the model (2), the marginal effects of satisfaction and recommendation intention are respectively measured by
The marginal effect of satisfaction is not constant, changing with the level of satisfaction. In contrast, the marginal effect of recommendation intention is constant $\hat{\alpha}$. Figure 5 depicts these effects. The marginal effect of satisfaction among unsatisfied customers (negative regime) is increasing up to $\frac{1}{4}\hat{\beta}^{-}$ from the left, and then it is decreasing from $\frac{1}{4}\hat{\beta}^{+}$ toward zero.

Satisfaction and recommendation have positive impact to loyalty since their marginal effects are positive over the domain of satisfaction for every company because of positive parameter estimates reported in Table 4. However, which is more influential depends on the level of satisfaction. According to the relation depicted in Figure 4, the satisfaction is more influential on loyalty than recommendation for the customers with the central level of satisfaction. The recommendation intention, in turn, is more effective for extreme customers. The interval of satisfaction where satisfaction is more important is reported as $[d_1, d_2]$, and the transformed score by 100 points scale as $[b_1, b_2]$ in Table 4. The percentage of customers who belong to this interval is also given as “Customer Ratio” and it shows that the satisfaction is more effective for 89% customers for convenient stores; 54% customers for telecommunication industry and 37% customers for hotel industry which has strong heterogeneity inside.
Figure 5: Marginal and Indirect Effects
<table>
<thead>
<tr>
<th>Company</th>
<th>4β(-)</th>
<th>4β(+)</th>
<th>α</th>
<th><a href="CS">d1,d2</a></th>
<th>[d1,d2](CSI score)</th>
<th>Customer Ratio</th>
<th>4αγ</th>
<th><a href="CS">b1,b2</a></th>
<th>[b1,b2](CSI score)</th>
<th>Customer Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>0.41</td>
<td>0.23</td>
<td>0.22</td>
<td>[-1.64, 0.32]</td>
<td>[25.3, 54.4]</td>
<td>61.07%</td>
<td>0.07</td>
<td>[-3.04, 2.37]</td>
<td>[4.6, 84.8]</td>
<td>98.21%</td>
</tr>
<tr>
<td>M2</td>
<td>0.41</td>
<td>0.21</td>
<td>0.24</td>
<td>[-1.51, 0.00]</td>
<td>[22.7, 52.3]</td>
<td>45.41%</td>
<td>0.12</td>
<td>[-2.40, 1.48]</td>
<td>[5.3, 81.2]</td>
<td>96.20%</td>
</tr>
<tr>
<td>M3</td>
<td>0.37</td>
<td>0.20</td>
<td>0.24</td>
<td>[-1.32, 0.00]</td>
<td>[27.0, 45.9]</td>
<td>44.07%</td>
<td>0.15</td>
<td>[-2.00, 1.09]</td>
<td>[17.2, 61.6]</td>
<td>87.25%</td>
</tr>
<tr>
<td>M4</td>
<td>0.42</td>
<td>0.25</td>
<td>0.23</td>
<td>[-1.66, 0.59]</td>
<td>[21.5, 52.5]</td>
<td>65.00%</td>
<td>0.13</td>
<td>[-2.38, 1.68]</td>
<td>[11.5, 67.6]</td>
<td>96.00%</td>
</tr>
<tr>
<td>Industry mean</td>
<td>0.40</td>
<td>0.22</td>
<td>0.23</td>
<td></td>
<td></td>
<td>53.89%</td>
<td>0.12</td>
<td></td>
<td></td>
<td>94.41%</td>
</tr>
<tr>
<td>C1</td>
<td>0.45</td>
<td>0.36</td>
<td>0.30</td>
<td>[-1.33, 0.87]</td>
<td>[22.4, 66.7]</td>
<td>77.85%</td>
<td>0.13</td>
<td>[-2.50, 2.23]</td>
<td>[0.00, 94.0]</td>
<td>99.12%</td>
</tr>
<tr>
<td>C2</td>
<td>0.44</td>
<td>0.39</td>
<td>0.15</td>
<td>[-2.27, 2.14]</td>
<td>[27.4, 80.8]</td>
<td>98.46%</td>
<td>0.05</td>
<td>[-3.46, 3.35]</td>
<td>[13.0, 95.5]</td>
<td>99.34%</td>
</tr>
<tr>
<td>C3</td>
<td>0.42</td>
<td>0.35</td>
<td>0.22</td>
<td>[-1.69, 1.39]</td>
<td>[8.8, 60.2]</td>
<td>90.28%</td>
<td>0.06</td>
<td>[-3.36, 3.15]</td>
<td>[0.00, 89.6]</td>
<td>99.17%</td>
</tr>
<tr>
<td>C4</td>
<td>0.50</td>
<td>0.44</td>
<td>0.26</td>
<td>[-1.70, 1.52]</td>
<td>[18.5, 81.3]</td>
<td>92.50%</td>
<td>0.11</td>
<td>[-2.79, 2.65]</td>
<td>[0.00, 100]</td>
<td>100.00%</td>
</tr>
<tr>
<td>C5</td>
<td>0.46</td>
<td>0.43</td>
<td>0.25</td>
<td>[-1.62, 1.51]</td>
<td>[31.2, 77.8]</td>
<td>90.00%</td>
<td>0.08</td>
<td>[-3.05, 2.97]</td>
<td>[9.9, 99.6]</td>
<td>99.33%</td>
</tr>
<tr>
<td>Industry mean</td>
<td>0.45</td>
<td>0.39</td>
<td>0.24</td>
<td></td>
<td></td>
<td>89.82%</td>
<td>0.08</td>
<td></td>
<td></td>
<td>99.39%</td>
</tr>
<tr>
<td>H1</td>
<td>0.36</td>
<td>0.30</td>
<td>0.26</td>
<td>[-1.16, 0.75]</td>
<td>[42.9, 69.7]</td>
<td>65.33%</td>
<td>0.13</td>
<td>[-2.22, 2.00]</td>
<td>[27.8, 87.2]</td>
<td>97.00%</td>
</tr>
<tr>
<td>H2</td>
<td>0.33</td>
<td>0.23</td>
<td>0.31</td>
<td>[-0.43, 0.00]</td>
<td>[35.9, 45.5]</td>
<td>24.00%</td>
<td>0.22</td>
<td>[-1.34, 0.56]</td>
<td>[15.8, 57.8]</td>
<td>74.00%</td>
</tr>
<tr>
<td>H3</td>
<td>0.36</td>
<td>0.26</td>
<td>0.28</td>
<td>[-1.00, 0.00]</td>
<td>[21.6, 40.2]</td>
<td>34.33%</td>
<td>0.21</td>
<td>[-1.51, 0.93]</td>
<td>[12.1, 57.5]</td>
<td>79.67%</td>
</tr>
<tr>
<td>H4</td>
<td>0.34</td>
<td>0.24</td>
<td>0.29</td>
<td>[-0.76, 0.00]</td>
<td>[30.4, 43.2]</td>
<td>27.67%</td>
<td>0.17</td>
<td>[-1.76, 1.22]</td>
<td>[13.6, 63.6]</td>
<td>87.67%</td>
</tr>
<tr>
<td>H5</td>
<td>0.32</td>
<td>0.27</td>
<td>0.35</td>
<td>[ 0.00, 0.00]</td>
<td>[48.6, 48.6]</td>
<td>0.00%</td>
<td>0.20</td>
<td>[-1.45, 1.19]</td>
<td>[22.2, 70.1]</td>
<td>86.33%</td>
</tr>
<tr>
<td>H6</td>
<td>0.37</td>
<td>0.22</td>
<td>0.21</td>
<td>[-1.58, 0.40]</td>
<td>[24.9, 57.6]</td>
<td>64.00%</td>
<td>0.14</td>
<td>[-2.11, 1.35]</td>
<td>[16.1, 73.3]</td>
<td>91.67%</td>
</tr>
<tr>
<td>H7</td>
<td>0.38</td>
<td>0.27</td>
<td>0.33</td>
<td>[-0.78, 0.00]</td>
<td>[39.4, 51.3]</td>
<td>25.67%</td>
<td>0.21</td>
<td>[-1.64, 1.00]</td>
<td>[26.4, 66.4]</td>
<td>80.00%</td>
</tr>
<tr>
<td>H8</td>
<td>0.32</td>
<td>0.20</td>
<td>0.31</td>
<td>[-0.35, 0.00]</td>
<td>[46.6, 53.6]</td>
<td>18.00%</td>
<td>0.19</td>
<td>[-1.49, 0.42]</td>
<td>[23.8, 61.9]</td>
<td>64.00%</td>
</tr>
<tr>
<td>H9</td>
<td>0.34</td>
<td>0.28</td>
<td>0.22</td>
<td>[-1.35, 0.96]</td>
<td>[24.3, 65.0]</td>
<td>79.00%</td>
<td>0.16</td>
<td>[-1.83, 1.54]</td>
<td>[15.9, 75.1]</td>
<td>91.67%</td>
</tr>
<tr>
<td>H10</td>
<td>0.38</td>
<td>0.23</td>
<td>0.36</td>
<td>[-0.49, 0.00]</td>
<td>[43.7, 52.0]</td>
<td>17.67%</td>
<td>0.26</td>
<td>[-1.32, 0.00]</td>
<td>[29.6, 52.0]</td>
<td>41.33%</td>
</tr>
<tr>
<td>H11</td>
<td>0.37</td>
<td>0.27</td>
<td>0.27</td>
<td>[-1.11, 0.00]</td>
<td>[29.8, 49.5]</td>
<td>32.00%</td>
<td>0.20</td>
<td>[-1.67, 1.19]</td>
<td>[19.7, 70.7]</td>
<td>87.33%</td>
</tr>
<tr>
<td>H12</td>
<td>0.36</td>
<td>0.25</td>
<td>0.24</td>
<td>[-1.34, 0.44]</td>
<td>[32.4, 59.3]</td>
<td>56.33%</td>
<td>0.17</td>
<td>[-1.86, 1.29]</td>
<td>[24.6, 72.1]</td>
<td>88.00%</td>
</tr>
<tr>
<td>Industry mean</td>
<td>0.35</td>
<td>0.25</td>
<td>0.29</td>
<td></td>
<td></td>
<td>37.00%</td>
<td>0.19</td>
<td></td>
<td></td>
<td>80.72%</td>
</tr>
<tr>
<td>Total mean</td>
<td>0.39</td>
<td>0.28</td>
<td>0.26</td>
<td></td>
<td></td>
<td>52.79%</td>
<td>0.15</td>
<td></td>
<td></td>
<td>87.78%</td>
</tr>
</tbody>
</table>
Following the customer satisfaction model, RE is determined by CS as a structural equation

\[ RE_i = \gamma CS_i + e_i. \]  

(20)

The indirect effect of CS to LOY is defined as marginal effect of estimated RE by CS, \( \hat{RE}_i = \hat{\gamma} CS_i, \) and it is given by \( \hat{\alpha} \hat{\gamma}. \) It means the effect of CS by way of RE to LOY.

In (20), CS is assumed to have a positive effect on RE, and in fact \( \hat{\gamma} \) is estimated positive for every company. Then the indirect effect of CS by way of RE to LOY is interpreted as marginal effect of RE discounted by \( \hat{\gamma}. \) We also define the direct effect of CS as the marginal effect in (9), and we compare these effects over the domain of satisfaction. It is evident that direct effect is much more influential for loyalty by the interpretation of estimated indirect effects above. We define, similarly before, the interval where direct effect of CS is greater than indirect effect, and their “Customer Ratio”. These are reported in Table 4. It shows that the direct effect is most pervasive with over 99% customers in wider range of interval in convenience stores; in particular, it is dominant (100%) for the company C4. The hotels have rather significant impacts of indirect effect, 88% in average. The indirect effect is more important for the hotel H10 (41.33% customer ratio) and comparable for H8 (64%). The CS campaign leading to recommendation would be necessary for these companies.

3.6.6. Managerial Implications

We consider two kinds of measure for managerial implications derived from our models. The first measure is the expected incremental loyalty that is defined by the expected values of incremental loyalty on unit change of satisfaction with respective to customer distribution on CSI scores. This measure is useful for overall evaluation of loyalty program when we assume that the firm approaches every customer. Under the limited budget for loyalty program, the
second measure finds a customer segment optimizing the incremental loyalty.

(i) Expected Incremental Loyalty

According to satisfaction and loyalty scores obtained from the empirical study of our model, company managers can review their situations and formulate their strategies. First, the expected incremental loyalty (EIL) can be used to forecast future profitability of loyalty program by combining estimated response function and customers distribution over the same domain. Based on the empirical distribution of CSI scores in 4.6, we first set the cut-off point vector (0, 5, 10, 15, 20, ..., 100) for CSI score dimension, and calculate the frequency of each cell to get the empirical distribution of CSI scores \( p(cs_i | data), i=1,2,...,20 \). Then we define a middle point vector (cs1, cs2, cs3, ..., cs19, cs20) by (2.5, 7.5, 12.5, ..., 97.5) for calculating estimate of marginal incremental loyalty, \( f'(cs_i) \) defined in (19). Then the EIL is formally defined by

\[
EIL = \sum_{i=1}^{20} f'(cs_i)p(cs_i | data).
\]  

EIL shows the future profitability when loyalty program has a full access to their customers.

Table 5 at the first column shows the measure of EIL for individual companies. The industry of convenience store has the highest EIL, and it will get the largest loyalty increment when the customers’ satisfaction level is improved. Another point is the difference between EIL and the band width given in Table 3.1. The band width of company M1 is a little lower than H11, having 1.27 and 1.28 respectively. However, M1 has higher EIL (0.2741) than H11 (0.2693). It means the extensible space of loyalty in M1 is not as wide as those in H11, however customers in M1 have more concentration around neutral point where it has the most sensitive change of loyalty.

(ii) Targeted Customer Interval for Efficient Loyalty Program

Next, we consider a situation that company might just offer loyalty program to a limited
proportion of customers due to their budget constraint. Our model provides the framework to consider how they should target some customers effectively subject to their budget for loyalty program. Assume that a company manager prospects that she/he is allowed to provide loyalty program for only 30% customers. Then the problem is to specify the set of customers under constrained optimization,

\[
\max \sum_{i = 1}^{CS} f_i (CS_i),
\]

\[
s.t. P\{CS_a \leq CS \leq CS_b\} = 0.3.
\]

Figure 6: Frequency and Increment Loyalty
We call the interval \([CS_a, CS_b]\) 30% targeted customer interval (TCI), implying that the customer segment maximizes incremental loyalty induced by loyalty program. Figure 6 shows the smoothed frequency distribution of customer’s CSI scores on the left, and the marginal loyalty curve over CSI score dimension on the right. The customers whose CSI scores are located in the interval of \([CS_a, CS_b]\), are most suitable to be the targets of loyalty program. 

TCI is constructed in the same way as the highest probability density (HPD) region for Bayesian confidence interval. That is, we incorporate customers into TCI in order with higher incremental loyalty until the interval contains 30% customers.

### Table 5: EIL and TCI(30%)

<table>
<thead>
<tr>
<th>Company</th>
<th>EIL</th>
<th>TCI(30%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>0.2741</td>
<td>[39.7, 49.6]</td>
</tr>
<tr>
<td>M2</td>
<td>0.2515</td>
<td>[38.1, 52.2]</td>
</tr>
<tr>
<td>M3</td>
<td>0.2369</td>
<td>[35.2, 45.7]</td>
</tr>
<tr>
<td>M4</td>
<td>0.2777</td>
<td>[32.7, 44.3]</td>
</tr>
<tr>
<td>mean(variance)</td>
<td>0.2604(0.000372)</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>0.3467</td>
<td>[34.4, 49.1]</td>
</tr>
<tr>
<td>C2</td>
<td>0.3491</td>
<td>[46.9, 55.8]</td>
</tr>
<tr>
<td>C3</td>
<td>0.3229</td>
<td>[22.8, 37.0]</td>
</tr>
<tr>
<td>C4</td>
<td>0.4001</td>
<td>[38.1, 53.5]</td>
</tr>
<tr>
<td>C5</td>
<td>0.3673</td>
<td>[47.0, 58.1]</td>
</tr>
<tr>
<td>mean(variance)</td>
<td>0.3572(0.000825)</td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>0.2765</td>
<td>[49.1, 59.1]</td>
</tr>
<tr>
<td>H2</td>
<td>0.2453</td>
<td>[31.2, 45.5]</td>
</tr>
<tr>
<td>H3</td>
<td>0.2614</td>
<td>[25.1, 40.1]</td>
</tr>
<tr>
<td>H4</td>
<td>0.2442</td>
<td>[28.6, 42.0]</td>
</tr>
<tr>
<td>H5</td>
<td>0.2530</td>
<td>[35.5, 48.4]</td>
</tr>
<tr>
<td>H6</td>
<td>0.2432</td>
<td>[38.7, 50.9]</td>
</tr>
<tr>
<td>H7</td>
<td>0.2625</td>
<td>[37.1, 51.1]</td>
</tr>
<tr>
<td>H8</td>
<td>0.2260</td>
<td>[40.3, 53.4]</td>
</tr>
<tr>
<td>H9</td>
<td>0.2627</td>
<td>[34.0, 48.0]</td>
</tr>
<tr>
<td>H10</td>
<td>0.2489</td>
<td>[38.3, 51.9]</td>
</tr>
<tr>
<td>H11</td>
<td>0.2693</td>
<td>[33.0, 49.4]</td>
</tr>
<tr>
<td>H12</td>
<td>0.2601</td>
<td>[39.4, 52.6]</td>
</tr>
<tr>
<td>mean(variance)</td>
<td>0.2544(0.000189)</td>
<td></td>
</tr>
</tbody>
</table>

The second column of Table 5 shows the TCI for individual companies. Under the assumptions of limited access and identification of CSI scores of their customers, every company can find the customers to be targeted for their loyalty program.
4. Finite Mixture Model of Structural Heterogeneity and Empirical Results

On the foundation of the MCMC algorithm in Chapter 3, we will induce the finite mixture modeling to solve the problem of structural heterogeneity, which assumes that all companies tend to have various types of nonlinear functions, rather than identical function. Then we will show some managerial implications with the case of some companies.

4.1. Structural Heterogeneity

Standard methods of understanding customer behavior in marketing, such as HB modeling, allow for differences in customer sensitivity across companies, but often assume that the structural sensitivity of these companies is fixed. As we put forward four kinds of nonlinear model in chapter 3, we assume this kind of structural homogeneity for all companies.

In many situations, this structural homogeneous assumption may not be valid. Both the importance of variables, and the manner that they are combined to form an overall measure of value for an offer, can change. It is practical and meaningful to identify the suitable nonlinear model for each company. In this chapter we propose an approach of modeling the customers’ satisfaction affect the loyalty in different nonlinear structures, which allows identification of subsections of customers. This information is useful in customer relationship management when particular customers will be most likely to respond in the form of “Threshold Logit Model”, and company managers should not invest the resources to those customers who are fallen in the consumption tolerance interval. Finite mixture modeling is used to implement this structural heterogeneous assumption of dealing with individual differences, and borrow the information from other companies and industries more effectively.

4.2. Finite Mixture Model

Modeling based on finite mixture distributions is a rapidly developing area with the range of
applications exploding. Finite mixture models are nowadays applied in such diverse areas as biometrics, genetics, medicine, and marketing whereas Markov switching models are applied especially in economics and finance. Finite mixture distributions arise in a natural way as marginal distribution for statistical models involving discrete latent variables such as clustering or latent class models. This extension to Markov mixture models is able to deal with many features of practical time series, for example, spurious long-range dependence and conditional heteroscedasticity. Finite mixture models provide a straightforward, but very flexible extension of classical statistical models.

Among various mixture models, mixtures with all components assumed to be normally distributed are the most commonly used by practitioners. This can be explained by the existence of well-developed statistical theory for Gaussian distributions as well as the fact that normal distributions arise naturally in many applications. Unlike original Finite Mixture Model in clustering, Sylvia Fruehwirth-Schnatter (2006) implements Finite Mixture Model in regression model. The given components are scalar or vector, rather than normal distribution, and these components share a common variance.

According to the introductions of Sylvia Fruehwirth-Schnatter (2006), we replace the model (2) by following equation:

\[
\begin{align*}
\text{LOY}_i &= G(CS_i, S) + \alpha R_{E_i} + \varepsilon_i, \\
G(CS_i, S) &= \begin{cases} 
\beta^{(1)} I(CS_i - r_0) + \beta^{(2)}(1-I)(CS_i - r_0) 
\quad \text{Pr}(S = 1) = \pi_1, \\
\beta^{(3)} I(CS_i - r_1) + \beta^{(3)-2}(1-I)(CS_i - r_2) 
\quad \text{Pr}(S = 2) = \pi_2, \\
\beta^{(4)} I(1 + \exp(-CS_i + r_1)) - \frac{1}{2} + \beta^{(5)}(1-I) \left\{ \frac{1}{1 + \exp(-CS_i + r_0)} - \frac{1}{2} \right\} 
\quad \text{Pr}(S = 3) = \pi_3, \\
\beta^{(6)} I(1 + \exp(-CS_i + r_1)) - \frac{1}{2} + \beta^{(7)}(1-I) \left\{ \frac{1}{1 + \exp(-CS_i + r_2)} - \frac{1}{2} \right\} 
\quad \text{Pr}(S = 4) = \pi_4,
\end{cases}
\end{align*}
\]
where \( \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \), \( r_1 > 0 \), \( r_2 < 0 \). \( S \) is the state indicator, and \([ r_1 \quad r_2 ]\) is zone of tolerance. The posterior densities of parameters \( \theta_m = \{ r, \beta^{(+)}, \beta^{(-)} \} \) are

\[
p(\theta_m | LOY, CS, RE_i) \propto \left( \prod_{i=1}^{j} f(LOY_i | CS_i, RE_i, \theta_m) \right) \times p(\theta_m), \quad (24)
\]

The conjugate prior distribution of \( \pi \) is \( \pi(\pi_1, \pi_2, \pi_3, \pi_4) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \), and the hyper parameters are set as \( \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1 \). So the posterior distribution of \( \pi \) is:

\[
(\pi_1, \pi_2, \pi_3, \pi_4) \sim \text{Dirichlet}(\alpha_1 + n_1, \alpha_2 + n_2, \alpha_3 + n_3, \alpha_4 + n_4), \quad (25)
\]

\( n_1 \) means the number of customers whose state indicator \( S_i = 1 \), which is given in last draw of MCMC. The posterior density of \( S_i \) of customer i is:

\[
P(S_i = m | \theta_m, \pi_m, LOY_i, CS_i, RE_i) = \frac{\pi_m f(LOY_i | CS_i, RE_i, \theta_m)}{\sum_{h=1}^{M} \pi_h f(LOY_i | CS_i, RE_i, \theta_h)}, \quad (26)
\]

where \( M = 4 \), it is the number of candidate states (nonlinear models). \( f(LOY_i | CS_i, RE_i, \theta_m) \) is the probability density, we consider the heterogeneity of \( S \) in customer level. The details of the algorithm are in Appendix B.

As for the goodness of fit, DIC and LML are also available. We will make comparison between the structural heterogeneous and HB models.

4.3. Empirical Results of Structural Heterogeneous Model

4.3.1. Model comparison

We estimated the parameters of finite mixture model. This section reports results of the comparison between models by comparing the values of DIC and LML. We compare (i) linear model, (ii) HB linear model, (iii) HB logit model, (iv) finite mixture model.
Table 6: Goodness of Fit

<table>
<thead>
<tr>
<th></th>
<th>Independent Symmetric</th>
<th>Linear Asymmetric</th>
<th>Linear Symmetric</th>
<th>Linear Asymmetric</th>
<th>Logit Symmetric</th>
<th>Logit Asymmetric</th>
<th>HB Threshold</th>
<th>HB Threshold</th>
<th>Finite Mixture</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIC</td>
<td>238539.1</td>
<td>234960.5</td>
<td>234912.3</td>
<td>234963.2</td>
<td>234908.7</td>
<td>234908.5</td>
<td>234937.7</td>
<td>238888.412</td>
<td></td>
</tr>
<tr>
<td>LML</td>
<td>-98516.1</td>
<td>-96844.0</td>
<td>-96577.7</td>
<td>-96578.3</td>
<td>-96568.3</td>
<td>-96575.5</td>
<td>-96570.8</td>
<td>-96564.218</td>
<td></td>
</tr>
</tbody>
</table>

After 10,000 iterations of MCMC, according to Table 6, Finite Mixture performs best in both DIC and LML, which means the heterogeneous assumption is preferred. So we can identify the nonlinear form of each company.

4.3.2. Parameter Estimates of Nonlinear Term from CS to LOY

In each MCMC iteration, we simulate the state indicator $S_i$ of customer $i$ of company $h$, and confirm the state indicator $S_h$ of company $h$ according to the most preferred nonlinear form among the customers. For example, in Figure 7, $\#\{S_h = 3\}$ gets the highest frequency in Company M1, and $\#\{S_h = 1\}$ gets the highest frequency in Company M2. Figure 8 shows the histograms of all 21 companies.

Figure 7: Histogram of State Indicator of Company M1 and M2
Table 7 shows the parameter estimates of each company. First of all, according to the model type, the number of companies for each nonlinear form is: “Asymmetric Linear”(9), “Threshold Linear”(6), “Asymmetric Logit”(4), “Threshold Logit”(2). Second, most estimates of nonlinear path coefficients are significant for 95% HPD region test. Third, except for company C5 and H5, the property of loss aversion is proved in most companies. Fourth, we estimate the change point of “Asymmetric Linear” and “Asymmetric Logit”, and the boundary of tolerance interval of “Threshold Linear” and “Threshold Logit”. Based on the location the change point and distribution of satisfaction, we calculate the Segment ratio of each interval.
### Table 7: Parameter Estimates

<table>
<thead>
<tr>
<th>Industry</th>
<th>Model Type</th>
<th>( \beta(-) )</th>
<th>HPD(95%)</th>
<th>( \beta(+) )</th>
<th>HPD(95%)</th>
<th>Loss aversion probability</th>
<th>Change point</th>
<th>Change point(100 scale)</th>
<th>Segment Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Communication operators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>3</td>
<td>1.75</td>
<td>[1.07, 2.52]</td>
<td>0.81</td>
<td>[0.20, 1.42]</td>
<td>96.04%</td>
<td>-0.15</td>
<td>47.36</td>
<td>0.45:0.55</td>
</tr>
<tr>
<td>M2</td>
<td>1</td>
<td>0.35</td>
<td>[0.20, 0.53]</td>
<td>0.15</td>
<td>[0.01, 0.30]</td>
<td>94.52%</td>
<td>0.00</td>
<td>52.24</td>
<td>0.49:0.51</td>
</tr>
<tr>
<td>M3</td>
<td>1</td>
<td>0.27</td>
<td>[0.12, 0.47]</td>
<td>0.14</td>
<td>[0.01, 0.28]</td>
<td>87.12%</td>
<td>-0.03</td>
<td>45.07</td>
<td>0.49:0.51</td>
</tr>
<tr>
<td>M4</td>
<td>1</td>
<td>0.40</td>
<td>[0.24, 0.58]</td>
<td>0.21</td>
<td>[0.08, 0.34]</td>
<td>95.20%</td>
<td>-0.11</td>
<td>42.58</td>
<td>0.46:0.54</td>
</tr>
<tr>
<td>C1</td>
<td>3</td>
<td>1.87</td>
<td>[1.25, 2.56]</td>
<td>1.35</td>
<td>[0.80, 1.89]</td>
<td>85.94%</td>
<td>-0.13</td>
<td>46.11</td>
<td>0.45:0.55</td>
</tr>
<tr>
<td>C2</td>
<td>2</td>
<td>0.45</td>
<td>[0.29, 0.62]</td>
<td>0.42</td>
<td>[0.26, 0.60]</td>
<td>61.80%</td>
<td>[-0.27, 0.26]</td>
<td>50.58, 57.02</td>
<td>0.39:0.22:0.39</td>
</tr>
<tr>
<td><strong>Convenience stores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>4</td>
<td>2.39</td>
<td>[1.59, 3.28]</td>
<td>2.03</td>
<td>[1.25, 2.81]</td>
<td>73.30%</td>
<td>[-0.51, 0.53]</td>
<td>28.88, 46.25</td>
<td>0.30:0.45:0.25</td>
</tr>
<tr>
<td>C4</td>
<td>2</td>
<td>0.56</td>
<td>[0.37, 0.79]</td>
<td>0.57</td>
<td>[0.37, 0.79]</td>
<td>50.38%</td>
<td>[-0.25, 0.20]</td>
<td>46.77, 55.58</td>
<td>0.37:0.19:0.44</td>
</tr>
<tr>
<td>C5</td>
<td>1</td>
<td>0.36</td>
<td>[0.24, 0.49]</td>
<td>0.42</td>
<td>[0.29, 0.58]</td>
<td><strong>26.56%</strong></td>
<td>0.10</td>
<td>56.96</td>
<td>0.53:0.47</td>
</tr>
<tr>
<td><strong>Hotels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>4</td>
<td>2.03</td>
<td>[1.44, 2.71]</td>
<td>1.58</td>
<td>[1.06, 2.14]</td>
<td>87.22%</td>
<td>[-0.53, 0.22]</td>
<td>51.81, 62.28</td>
<td>0.28:0.31:0.41</td>
</tr>
<tr>
<td>H2</td>
<td>1</td>
<td>0.23</td>
<td>[0.05, 0.45]</td>
<td>0.14</td>
<td>[-0.01, 0.29]</td>
<td>75.60%</td>
<td>-0.09</td>
<td>43.49</td>
<td>0.48:0.52</td>
</tr>
<tr>
<td>H3</td>
<td>1</td>
<td>0.34</td>
<td>[0.22, 0.49]</td>
<td>0.22</td>
<td>[0.10, 0.35]</td>
<td>89.85%</td>
<td>-0.02</td>
<td>39.44</td>
<td>0.48:0.52</td>
</tr>
<tr>
<td>H4</td>
<td>3</td>
<td>1.43</td>
<td>[1.00, 1.96]</td>
<td>0.81</td>
<td>[0.42, 1.21]</td>
<td>96.88%</td>
<td>-0.22</td>
<td>40.50</td>
<td>0.44:0.56</td>
</tr>
<tr>
<td>H5</td>
<td>1</td>
<td>0.21</td>
<td>[0.05, 0.40]</td>
<td>0.24</td>
<td>[0.09, 0.42]</td>
<td><strong>42.30%</strong></td>
<td>-0.02</td>
<td>48.02</td>
<td>0.53:0.47</td>
</tr>
<tr>
<td>H6</td>
<td>2</td>
<td>0.48</td>
<td>[0.31, 0.70]</td>
<td>0.13</td>
<td>[-0.05, 0.31]</td>
<td>99.40%</td>
<td>[-0.41, 0.51]</td>
<td>43.62, 58.80</td>
<td>0.29:0.44:0.27</td>
</tr>
<tr>
<td>H7</td>
<td>2</td>
<td>0.48</td>
<td>[0.30, 0.71]</td>
<td>0.32</td>
<td>[0.15, 0.53]</td>
<td>88.86%</td>
<td>[-0.46, 0.33]</td>
<td>45.29, 57.41</td>
<td>0.29:0.36:0.35</td>
</tr>
<tr>
<td>H8</td>
<td>1</td>
<td>0.24</td>
<td>[0.04, 0.52]</td>
<td>0.08</td>
<td>[-0.07, 0.25]</td>
<td>85.60%</td>
<td>-0.22</td>
<td>49.19</td>
<td>0.41:0.59</td>
</tr>
<tr>
<td>H9</td>
<td>1</td>
<td>0.26</td>
<td>[0.11, 0.43]</td>
<td>0.25</td>
<td>[0.12, 0.41]</td>
<td>51.22%</td>
<td>-0.01</td>
<td>45.95</td>
<td>0.48:0.52</td>
</tr>
<tr>
<td>H10</td>
<td>2</td>
<td>0.48</td>
<td>[0.30, 0.69]</td>
<td>0.23</td>
<td>[0.03, 0.46]</td>
<td>95.96%</td>
<td>[-0.31, 0.58]</td>
<td>45.92, 60.89</td>
<td>0.37:0.37:0.26</td>
</tr>
<tr>
<td>H11</td>
<td>2</td>
<td>0.47</td>
<td>[0.30, 0.70]</td>
<td>0.35</td>
<td>[0.19, 0.53]</td>
<td>85.22%</td>
<td>[-0.42, 0.32]</td>
<td>42.20, 55.50</td>
<td>0.30:0.30:0.40</td>
</tr>
<tr>
<td>H12</td>
<td>3</td>
<td>1.57</td>
<td>[1.12, 2.10]</td>
<td>0.85</td>
<td>[0.41, 1.29]</td>
<td>98.06%</td>
<td>-0.14</td>
<td>51.63</td>
<td>0.43:0.57</td>
</tr>
</tbody>
</table>

Model Type: 1 Asymmetric Linear, 2 Threshold Linear, 3 Asymmetric Logit, 4 Threshold Logit.
In particular, the graphs of each type of nonlinear model are shown below.

**Asymmetric Linear:**

![Figure 9: Asymmetric Linear](image)

**Table 8: Companies of Asymmetric Linear**

<table>
<thead>
<tr>
<th>company</th>
<th>$\beta_-$</th>
<th>$\beta_+$</th>
<th>Change point ($r_0$)</th>
<th>Change point(100 scale)</th>
<th>Segment Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>M2</td>
<td>0.35</td>
<td>0.15</td>
<td>0.00</td>
<td>52.24</td>
<td>0.49:0.51</td>
</tr>
<tr>
<td>M3</td>
<td>0.27</td>
<td>0.14</td>
<td>-0.03</td>
<td>45.07</td>
<td>0.49:0.51</td>
</tr>
<tr>
<td>M4</td>
<td>0.40</td>
<td>0.21</td>
<td>-0.11</td>
<td>42.58</td>
<td>0.46:0.54</td>
</tr>
<tr>
<td>C5</td>
<td>0.36</td>
<td>0.42</td>
<td>0.10</td>
<td>56.96</td>
<td>0.53:0.47</td>
</tr>
<tr>
<td>H2</td>
<td>0.23</td>
<td>0.14</td>
<td>-0.09</td>
<td>43.49</td>
<td>0.48:0.52</td>
</tr>
<tr>
<td>H3</td>
<td>0.34</td>
<td>0.22</td>
<td>-0.02</td>
<td>39.44</td>
<td>0.48:0.52</td>
</tr>
<tr>
<td>H5</td>
<td>0.21</td>
<td>0.24</td>
<td>-0.02</td>
<td>48.02</td>
<td>0.53:0.47</td>
</tr>
<tr>
<td>H8</td>
<td>0.24</td>
<td>0.08</td>
<td>-0.22</td>
<td>49.19</td>
<td>0.41:0.59</td>
</tr>
<tr>
<td>H9</td>
<td>0.26</td>
<td>0.25</td>
<td>-0.01</td>
<td>45.95</td>
<td>0.48:0.52</td>
</tr>
</tbody>
</table>
Threshold Linear:

Figure 10: Threshold Linear

![Threshold Linear Graph]

Table 9 Companies of Threshold Linear

<table>
<thead>
<tr>
<th>company</th>
<th>β(-)</th>
<th>β(+)</th>
<th>Change point [r_2, r_1]</th>
<th>Change point (100 scale)</th>
<th>Segment Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>0.45</td>
<td>0.42</td>
<td>[-0.27, 0.26]</td>
<td>[50.58, 57.02]</td>
<td>0.39:0.22:0.39</td>
</tr>
<tr>
<td>C4</td>
<td>0.56</td>
<td>0.57</td>
<td>[-0.25, 0.20]</td>
<td>[46.77, 55.58]</td>
<td>0.37:0.19:0.44</td>
</tr>
<tr>
<td>H6</td>
<td>0.48</td>
<td>0.13</td>
<td>[-0.41, 0.51]</td>
<td>[43.62, 58.80]</td>
<td>0.29:0.44:0.27</td>
</tr>
<tr>
<td>H7</td>
<td>0.48</td>
<td>0.32</td>
<td>[-0.46, 0.33]</td>
<td>[45.29, 57.41]</td>
<td>0.29:0.36:0.35</td>
</tr>
<tr>
<td>H10</td>
<td>0.48</td>
<td>0.23</td>
<td>[-0.31, 0.58]</td>
<td>[45.92, 60.89]</td>
<td>0.37:0.37:0.26</td>
</tr>
<tr>
<td>H11</td>
<td>0.47</td>
<td>0.35</td>
<td>[-0.42, 0.32]</td>
<td>[42.20, 55.50]</td>
<td>0.30:0.30:0.40</td>
</tr>
</tbody>
</table>

Asymmetric Logit:

Figure 11: Asymmetric Logit

![Asymmetric Logit Graph]
Table 10: Companies of Asymmetric Logit

<table>
<thead>
<tr>
<th>company</th>
<th>β(-)</th>
<th>β(+)</th>
<th>Change point (r₂)</th>
<th>Change point(100 scale)</th>
<th>Segment Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>1.75</td>
<td>0.81</td>
<td>-0.15</td>
<td>47.36</td>
<td>0.45:0.55</td>
</tr>
<tr>
<td>C1</td>
<td>1.87</td>
<td>1.35</td>
<td>-0.13</td>
<td>46.11</td>
<td>0.45:0.55</td>
</tr>
<tr>
<td>H4</td>
<td>1.43</td>
<td>0.81</td>
<td>-0.22</td>
<td>40.50</td>
<td>0.44:0.56</td>
</tr>
<tr>
<td>H12</td>
<td>1.57</td>
<td>0.85</td>
<td>-0.14</td>
<td>51.63</td>
<td>0.43:0.57</td>
</tr>
</tbody>
</table>

Threshold Logit:

Figure 12: Threshold Logit

![Threshold Logit](image)

Table 11: Companies of Threshold Logit

<table>
<thead>
<tr>
<th>company</th>
<th>β(-)</th>
<th>β(+)</th>
<th>Change point [r₂,r₁]</th>
<th>Change point(100 scale)</th>
<th>Segment Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3</td>
<td>2.39</td>
<td>2.03</td>
<td>[-0.51,0.53]</td>
<td>[28.88, 46.25]</td>
<td>0.30:0.45:0.25</td>
</tr>
<tr>
<td>H1</td>
<td>2.03</td>
<td>1.58</td>
<td>[-0.53,0.22]</td>
<td>[31.81, 62.28]</td>
<td>0.28:0.31:0.41</td>
</tr>
</tbody>
</table>

The functional shapes of each company are shown in Appendix D.

4.3.3. Measure of Tolerance Interval

Both “Threshold Linear” (Model Type 2) and “Threshold Logit” (Model Type 4) have consumption tolerance intervals, so we pick up these companies and measure the degree of tolerance interval. If customers fall into the consumption tolerance interval, they are still unwilling to purchase goods even their satisfaction are improved.
Table 12: Companies with Tolerance Interval

<table>
<thead>
<tr>
<th>Company</th>
<th>Model Type</th>
<th>$\beta(-)$</th>
<th>$\beta(+))</th>
<th>Change point</th>
<th>Spacing</th>
<th>Change point (100 scale)</th>
<th>Spacing (100 cscale)</th>
<th>Segment Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>2</td>
<td>0.45</td>
<td>0.42</td>
<td>[-0.27,0.26]</td>
<td>0.53</td>
<td>[50.58, 57.02]</td>
<td>6.44</td>
<td>0.39:0.22:0.39</td>
</tr>
<tr>
<td>C4</td>
<td>2</td>
<td>0.56</td>
<td>0.57</td>
<td>[-0.25,0.20]</td>
<td>0.45</td>
<td>[46.77, 55.58]</td>
<td>8.81</td>
<td>0.37:0.19:0.44</td>
</tr>
<tr>
<td>H6</td>
<td>2</td>
<td>0.48</td>
<td>0.13</td>
<td>[-0.41,0.51]</td>
<td>0.92</td>
<td>[43.62, 58.80]</td>
<td>15.18</td>
<td>0.29:0.44:0.27</td>
</tr>
<tr>
<td>H7</td>
<td>2</td>
<td>0.48</td>
<td>0.32</td>
<td>[-0.46,0.33]</td>
<td>0.79</td>
<td>[45.29, 57.41]</td>
<td>12.12</td>
<td>0.29:0.36:0.35</td>
</tr>
<tr>
<td>H10</td>
<td>2</td>
<td>0.48</td>
<td>0.23</td>
<td>[-0.31,0.58]</td>
<td>0.89</td>
<td>[45.92, 60.89]</td>
<td>14.97</td>
<td>0.37:0.37:0.26</td>
</tr>
<tr>
<td>H11</td>
<td>2</td>
<td>0.47</td>
<td>0.35</td>
<td>[-0.42,0.32]</td>
<td>0.74</td>
<td>[42.20, 55.50]</td>
<td>13.30</td>
<td>0.36:0.30:0.40</td>
</tr>
<tr>
<td>C3</td>
<td>4</td>
<td>2.39</td>
<td>2.03</td>
<td>[-0.51,0.53]</td>
<td>1.04</td>
<td>[28.88, 46.25]</td>
<td>17.37</td>
<td>0.36:0.45:0.25</td>
</tr>
<tr>
<td>H1</td>
<td>4</td>
<td>2.03</td>
<td>1.58</td>
<td>[-0.53,0.22]</td>
<td>0.75</td>
<td>[51.81, 62.28]</td>
<td>10.47</td>
<td>0.28:0.31:0.41</td>
</tr>
</tbody>
</table>

In Table 12, “Change Point” means the boundary of tolerance interval, “Spacing” means the distance between the two change points, “Segment Ratio” means the ratio of customers who fall into the intervals of low satisfied, tolerance interval, and highly satisfied respectively. For example, company C4’s segment ratio of tolerance interval is 0.19, which is the lowest ratio compared with other companies, and the spacing of C4 is 0.45, which means the narrow spacing contributes to the reducing of tolerance interval ratio. On the other hand, the company C3 has biggest spacing of 1.04 and tolerance interval ratio of 0.45, which means the broad spacing contributes to the increasing of tolerance interval ratio. In addition, C4 also has higher segment ratio (0.44) in highly satisfied customer, and C3 has just 0.25 of highly satisfied customer ratio. So we can say that C4 has better structure of segment ratio.

It is obvious that company managers prefer low tolerance interval ratio, which will improve the efficiency of loyalty improvement derived from satisfaction promotion. If we make promotion of satisfaction for those customers of company C3 and C4, more customers in C4 are willing to repurchase goods.

The high tolerance interval ratio has double-sides effect. When the gross satisfaction is improved, the tolerance interval ratio will slow down the increase of gross loyalty, while if the
gross satisfaction is decreased, the tolerance interval ratio will also slow down the decrease of gross loyalty.

4.3.4. Managerial Implications and Suggestions

(i) Expected Incremental Loyalty and Targeted Customer Interval for Efficient Loyalty Program

The expected increment of loyalty we discussed in chapter 3.6.6 is still available to predict the increment of loyalty for this heterogeneous structural model. In Table 13, company C4 get the highest EIL, it means that even if company C4 has the property of tolerance interval, the narrow tolerance interval and the high marginal return can also contribute to the gross increment of loyalty level. As for the industry level of EIL, the convenient store industry has the highest EIL, which is identical with the Table 5.

Table 13: EIL of Heterogeneous Model

<table>
<thead>
<tr>
<th>Company</th>
<th>Model Type</th>
<th>EIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>3</td>
<td>0.2462</td>
</tr>
<tr>
<td>M2</td>
<td>1</td>
<td>0.2412</td>
</tr>
<tr>
<td>M3</td>
<td>1</td>
<td>0.2053</td>
</tr>
<tr>
<td>M4</td>
<td>1</td>
<td>0.3109</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>0.2509</td>
</tr>
<tr>
<td>C1</td>
<td>3</td>
<td>0.3348</td>
</tr>
<tr>
<td>C2</td>
<td>2</td>
<td>0.3527</td>
</tr>
<tr>
<td>C3</td>
<td>4</td>
<td>0.2621</td>
</tr>
<tr>
<td>C4</td>
<td>2</td>
<td>0.4494</td>
</tr>
<tr>
<td>C5</td>
<td>1</td>
<td>0.3939</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>0.3586</td>
</tr>
<tr>
<td>H1</td>
<td>4</td>
<td>0.2331</td>
</tr>
<tr>
<td>H2</td>
<td>1</td>
<td>0.1853</td>
</tr>
<tr>
<td>H3</td>
<td>1</td>
<td>0.2788</td>
</tr>
<tr>
<td>H4</td>
<td>3</td>
<td>0.2202</td>
</tr>
<tr>
<td>H5</td>
<td>1</td>
<td>0.2211</td>
</tr>
<tr>
<td>H6</td>
<td>2</td>
<td>0.1820</td>
</tr>
<tr>
<td>H7</td>
<td>2</td>
<td>0.2839</td>
</tr>
<tr>
<td>H8</td>
<td>1</td>
<td>0.1487</td>
</tr>
<tr>
<td>H9</td>
<td>1</td>
<td>0.2571</td>
</tr>
<tr>
<td>H10</td>
<td>2</td>
<td>0.2302</td>
</tr>
<tr>
<td>H11</td>
<td>2</td>
<td>0.2603</td>
</tr>
<tr>
<td>H12</td>
<td>3</td>
<td>0.2301</td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td>0.2276</td>
</tr>
</tbody>
</table>

As for the “Targeted Customer Interval for Efficient Loyalty Program”, the discrimination depends on the model type in Table 13. If companies obey the type 1 (asymmetric linear) and
type 2 (threshold linear), because of the constant marginal return (loyalty increment), company managers should aim at the customers whose satisfaction fall into the interval with higher marginal return. On the other hand, if companies obey type 3 (asymmetric logit) and type 4 (threshold logit), because of the decreasing marginal return, company managers are supposed to calculate the frequency of satisfaction and marginal return, and then identify the targeted customers, just like the Figure 6 shows. The example will be shown in the (ii) Managerial Suggestions.

(ii) Managerial Suggestions
Given the estimates of coefficients and zone of tolerance, company managers can make the marketing strategy to take advantage and avoid disadvantage. We choose company C3 as illustration. In convenient store industry, C3 obey the threshold logit model and get the lowest EIL, while C4 obey the threshold linear model and get the highest EIL.

Figure 13: Nonlinear Type of C3 and C4

Under the same level of coordinate value, people might have the illusion that the steep slope and large variation of C3 will lead to more loyalty increment. However, if we take the segment ratio in Table 7 into consideration, C3 has 45% customers in the zone of tolerance, and C4 just has 19% customers in it. So we can say the narrow distributed CS and extensive zone of tolerance make the company C3 get low increase of loyalty, even if its marginal return is
more than company C4.

The solution to make C3 get higher increment of LOY is using “Targeted Customer Interval for Efficient Loyalty Program” mentioned in chapter 3.6.6, giving up the customers whose CS fall into the zone of tolerance and low marginal LOY interval, and focus on the customers with high marginal return.

\[
\frac{\partial \text{LOY}}{\partial \text{CS}_i} = \hat{\beta}^{(+)} I \frac{\exp(-\text{CS}_i + r_i)}{(1 + \exp(-\text{CS}_i + r_i))^3} \text{ if } CS_i > r_i > 0,
\]

\[
\frac{\partial \text{LOY}}{\partial \text{CS}_i} = \hat{\beta}^{(-)} \frac{\exp(-\text{CS}_i + r_2)}{(1 + \exp(-\text{CS}_i + r_2))^3} \text{ if } CS_i < r_2 < 0,
\]

\[
\frac{\partial \text{LOY}}{\partial \text{CS}_i} = 0 \text{ if } r_2 < CS_i < r_i.
\]

Figure 14 Increment and Frequency of LOY of C3

The Figure 14 shows the increment (marginal return) of loyalty, and managers can fix the “Target Marginal LOY” to identify the interval [CS_a1 , CS_a2] and [CS_b1, CS_b2], within which the customers have more sensitive increment of LOY. The right Figure-histogram of LOY, indicates the number of customers whose CS fall into the targeted interval [CS_a1 , CS_a2] and [CS_b1, CS_b2].

As for other companies with different shapes of nonlinear function, this method of
“Targeted Customer Interval for Efficient Loyalty Program” is also available for them.
5. Concluding Remarks

In this study, we use both HB and structural heterogeneous models to investigate the effects of customer satisfaction on loyalty by focusing on nonlinear characteristics represented as attainable limit of loyalty induced by satisfaction, asymmetric response between satisfied and dissatisfied customers, and not-constant marginal returns over the domain of satisfactions. There are a few extant works investigating the relation between satisfaction and loyalty, in particular, and this is the first model to measure nonlinear relation based on a uniform measure of customer satisfaction index in terms of system equation by using structure that the loyalty is determined by customer satisfaction in the connections of related other constructs. Many scholars used post-estimate method or just pick up the manifest data from CSI model. However, as is discussed in Fornell (1992), the investigation by using the system approach leads to higher reliability than the results obtained under the perspective being limited to two variables.

In the first research, we firstly introduced hierarchical Bayes modeling into estimation to improve the measurement, and the identical model structure is applied. In all, this HB model’s contributions to the modeling literature are that (i) nonlinear term is embedded in the structural model of customer satisfaction index, and (ii) hierarchical Bayes modeling of nonlinear structural equation model for measuring customer satisfaction index to accommodate the homogeneity of surveyed companies. To our knowledge, this is the first study on nonlinear structural equation model which includes nonlinear term of endogenous latent variable. We propose an efficient algorithm of MCMC, i.e., multi-move sampler for latent variables by using Gibbs sampling. In the empirical application, we compared comprehensive sets of specifications and the asymmetric nonlinear function with attainable limits is best supported by two kinds of criteria, goodness of fit measures and the number of significant parameter
estimates. We obtained managerial implications for loyalty management such as attainable limits; customer’s loss aversion response; asymmetric marginal returns between satisfied and dissatisfied customers, i.e., increasing for unsatisfied customers and decreasing for satisfied customers, direct effect of customer satisfaction is more significant than recommendation in general. We derived the measures for efficient loyalty program by combining information of estimated response curve of satisfaction to loyalty and empirical distribution of customers on the dimension of CSI scores under assumptions of fully and limited access to customers.

Furthermore, in the second research, we induced Finite Mixture modeling to accommodate the structural heterogeneity of companies. We also set the change point as parameter, to divide the highly satisfied and low satisfied customers, and this method can help company managers to segment the customer groups more effectively. This kind of structural heterogeneous model has better goodness of fit compared with HB modeling, so the assumption that companies with various structural types of nonlinear forms is preferred. The methodological contribution of this structural heterogeneous modeling is the efficient combination of Finite Mixture and structure equation model. We set the heterogeneity states among customers in each company, and choose the most widely nonlinear form as the state of that company. As for the theoretical contribution, we prove the four nonlinear forms are suitable for different companies. The suitable nonlinear forms are in the order of “Asymmetric Linear”, “Threshold Linear”, “Asymmetric Logit” and “Threshold Logit”. In addition, there eight companies have the property of consumption tolerance interval, so we measure the spacing of tolerance interval. We also measure the expected increment of loyalty of each company, and the EIL is similar with the EIL measured in HB modeling.

Other studies by using not loyalty but other outcome, for example, willingness to pay (Homburg et al., 2005), suggested the inverse S-shaped function which means having negligible
change for customers with medium level satisfaction in consistent with the concept of zone of tolerance. The inverse S-shaped function represents unrealistic situation since unlimited effect can be expected for highly satisfied (delighted) customers, this is the reason that we did not set the inverse S-shaped function as candidate model in our research. Then the nonlinear function with neutral zone as well as attainable limits can be devised by modifying S-shaped function so that it has three regimes by two additional parameters which split the domain of satisfaction to plug zone of tolerance at the mid regime, and loss and gain regimes with attainable limits at the extremes, that is identical with our threshold logit model.

In addition, the term label switching has been introduced into the finite mixture modeling of Bayesian estimation, and it has to be addressed explicitly because in the course of sampling from the mixture posterior distribution, the labeling of the unobserved categories changes (Sylvia 2006). So the label switching will be our future research object.
References


Bradley, P. C. and Thomas, A. L (2009), Bayesian Methods for Data Analysis, CRA, Boca Raton.


Research, 6, 37–50.

Appendix A: Full Description of Model and Inference Procedure

The heterogeneous modeling algorithm consists of HB modeling algorithm and Finite Mixture algorithm, so in Appendix A and Appendix B we will just show the details of heterogeneous modeling algorithm.

The CSI model assumes six latent variables \( \omega = (\eta_1, \eta_2, \ldots, \eta_5) \) and these are extracted by 17 questions of survey. For the vector of question items \( y_i = (y_{i1}, \ldots, y_{i17}) \), which are ordered categorical variables, we first transform them into continuous data \( x_i = (x_{i1}, \ldots, x_{i17}) \) following normal distribution. This transformation is conducted by data-augmentation when \( y_i = (y_{i1}, \ldots, y_{i17}) \) is given at the conditional posterior density in Appendix B.

The structural equation model has measurement model to extract the latent variables from data, and structural model which describes the relation between latent variables. Then we set the measurement model for \( x_i \) by factor model and we define the structure on the factors as structural model.

(i) Measurement model

The observable vector of \( x_i = (x_{i1}, \ldots, x_{i17}) \) has a factor analytic representation with six factors

\[
    x_i = \Lambda \omega + \epsilon_i, \quad i = 1, \ldots, n, \tag{A1}
\]

where \( \Lambda \) represents the factor loading matrix, \( \omega = (\xi, \eta_1, \eta_2, \ldots, \eta_5) \) is factor score vector for \( i \) represents. The error term vector \( \epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{i17}) \) is assumed to follow

\[
    \epsilon_i \sim N_{17}(0, \Psi_{\epsilon}) \quad \text{where} \quad \Psi_{\epsilon} = \text{diag}\{\sigma_1, \ldots, \sigma_{17}\}.
\]

(ii) Structural model
The structural equation model assumes that the factor scores \( \omega_i = (\eta_{1i}, \eta_{2i}, \ldots, \eta_{si}, \xi_i) \) have the relation each other in terms of set of equations:

Perceived Quality: \( \eta_i = \gamma_{1i} \xi + \delta_i \)  \hspace{1cm} (A2)

Perceived Value: \( \eta_2 = \gamma_{21} \eta_1 + \gamma_{20} \xi + \delta_2 \)  \hspace{1cm} (A3)

Customer Satisfaction: \( \eta_3 = \gamma_{31} \eta_1 + \gamma_{32} \eta_2 + \gamma_{36} \xi + \delta_3 \)  \hspace{1cm} (A4)

Recommendation Intention: \( \eta_4 = \gamma_{43} \eta_3 + \delta_4 \)  \hspace{1cm} (A5)

Customer Loyalty: \( \eta_5 = G(\eta_3, S) + \gamma_{54} \eta_4 + \delta_5 \)  \hspace{1cm} (A6)

\[
G(\eta_3, S) = \begin{cases} 
\beta^{(1)} \left( \frac{1}{1 + \exp(-\eta_3 + r_0)} - \frac{1}{2} \right) + \beta^{(3)} (1 - I) \left( \frac{1}{1 + \exp(-\eta_3 + r_0)} - \frac{1}{2} \right) & \text{Pr}(S = 1) = \pi_1, \\
\beta^{(4)} I (\eta_3 - r_1) + \beta^{(5)} (1 - I) (\eta_3 - r_2) & \text{Pr}(S = 2) = \pi_2, \\
\beta^{(6)} I \left( \frac{1}{1 + \exp(-\eta_3 + r_1)} - \frac{1}{2} \right) + \beta^{(7)} (1 - I) \left( \frac{1}{1 + \exp(-\eta_3 + r_1)} - \frac{1}{2} \right) & \text{Pr}(S = 3) = \pi_3, \\
\beta^{(8)} I \left( \frac{1}{1 + \exp(-\eta_3 + r_2)} - \frac{1}{2} \right) + \beta^{(9)} (1 - I) \left( \frac{1}{1 + \exp(-\eta_3 + r_2)} - \frac{1}{2} \right) & \text{Pr}(S = 4) = \pi_4,
\end{cases}
\]

where \( \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1, r_1 > 0, r_2 < 0, \) and \( \delta = (\delta_1, \ldots, \delta_5)' \sim N_5(0, \Psi), \ \Psi = \text{diag}\{\psi_1, \ldots, \psi_5\}. \)

More specifically, the model describes that customer expectation \( \xi \) drives perceived quality \( \eta_1 \) and perceived value \( \eta_2 \). These three latent variables next generate customer satisfaction \( \eta_3 \), which directly affects recommendation intention \( \eta_4 \) and customer loyalty \( \eta_5 \).

Equation (A6) indicates the extension of this study.

**Bayesian Inference of Nonlinear SEM**

The structural models (A2)-(A6) play a role of prior for the likelihood defined by the measurement model (A1) for Bayesian inference. The joint prior density of
\[ \omega = (\eta_1, \eta_2, \ldots, \eta_s, \xi) \] is decomposed by using their recursive relation between endogenous latent variable \((\eta_1, \eta_2, \ldots, \eta_s)\) by

\[ p(\omega) = p(\xi) p(\eta_1 | \xi) p(\eta_2 | \eta_1, \xi) p(\eta_3 | \eta_1, \eta_2, \xi) p(\eta_4 | \eta_3) p(\eta_5 | \eta_3, S, \eta_4) \]  \hspace{1cm} (A7)

On the other hand, we denote the likelihood function for \(\omega\) conditional on parameters \(\theta = (\Lambda, \Psi)\) and data \(x\) as \(p(x | \omega, \theta) = \prod_{i=1}^{n} p(x_i | \omega, \theta)\), then full conditional posterior density is as follows:

(i) \(p(\xi | x, \theta) \propto p(\xi) p^{[z]}(x | \omega, \theta)\)

(ii) \(p(\eta_1 | \xi, x, \theta) \propto p(\eta_1 | \xi) p^{[\eta_1]}(x | \omega, \theta)\)

(iii) \(p(\eta_2 | \eta_1, \xi, x, \theta) \propto p(\eta_2 | \eta_1, \xi) p^{[\eta_2]}(x | \omega, \theta)\)

(iv) \(p(\eta_3 | \eta_1, \eta_2, \xi, x, \theta) \propto p(\eta_3 | \eta_1, \eta_2, \xi) p^{[\eta_3]}(x | \omega, \theta)\)

(v) \(p(\eta_4 | \eta_3, x, \theta) \propto p(\eta_4 | \eta_3) p^{[\eta_4]}(x | \omega, \theta)\)

(vi) \(p(\eta_5 | \eta_3, \eta_4, x, \theta) \propto p(\eta_5 | \eta_3, \eta_4) p^{[\eta_5]}(x | \omega, \theta)\)

where \(p^{[z]}(x | \omega, \theta)\) means the part of joint likelihood regarding the latent variable \(z\).

These conditional posteriors are analytically evaluated to be normal distribution since both prior and likelihood functions are normal density.

Then, starting from initial value \(\omega^{(0)} = (\eta_1^{(0)}, \eta_2^{(0)}, \ldots, \eta_s^{(0)}, \xi^{(0)})\), we iterate the Gibbs sampling from the conditional posterior density to obtain the joint posterior density \(p(\omega | x, \theta)\). This is a single move-sampler for MCMC.

The multi-move sampler is available for our model by the use of recursive system of CSI model to derive more efficient algorithm by using linearity of subsystem on \(\omega_i = (\eta_1, \eta_2, \eta_3, \eta_4, \xi)^t\). We set the joint prior density by
\[ p(\omega) = p(\omega_1)p(\eta_3 | \eta_3, \eta_4), \]  

(A8)

and the conditional posterior density is obtained by multi-move sampler for \( \omega_1 \) following by Lee (2007) since full conditional posterior density is as follows:

(i) \( p(\omega_1 | x, \theta) \propto p(\omega_1)p^{(\omega_1)}(x | \omega, \theta) \),

(ii) \( p(\eta_3 | \eta_3, \eta_4, x, \theta) \propto p(\eta_3 | \eta_3, \eta_4)p^{(\eta_3)}(x | \omega, \theta) \).

The details of algorithm are described in Appendix B.
Appendix B: MCMC Algorithms

The prior setting and conditional posterior density are described in this appendix for our model.

The measurement model connecting observed data and latent variables in the form of factor model (A1), and structural model relating latent variables (A2)-(A6) are compactly written by

\[
\begin{align*}
    x_{hi} &= \Lambda_h \alpha_{hi} + \epsilon_{hi}; \\
    \eta_{hi} &= \Pi_d \eta_{hi} + \Gamma_h \xi_{hi} + \delta_{hi} = \Lambda_{sd} \alpha_{hi} + \delta_{hi}; \\
    \delta_{hi} &\sim N(0, \Psi_h), \xi_{hi} \sim N(0, \Phi_h) \\
\end{align*}
\]  
(B1)

\[
\eta_{hi} = G(\eta_{hi}, S_h) + \gamma_{ic} \eta_{ic} + \delta_{ic} 
\]  
(B2)

(1) Prior Density

The diffuse priors are set on the model parameters, and these are shown in next table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{hk}$</td>
<td>$\sim N(\mu_0, V_0)$</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>$= 0_{Kx1}, V_0 = I_{KxK} \times 100$</td>
</tr>
<tr>
<td>$\Lambda'_{hk}$</td>
<td>$\sim N(\Lambda_0, \Psi_{chh} H_0)$</td>
</tr>
<tr>
<td>$\Lambda_0$</td>
<td>$= 100, H_0 = 100$</td>
</tr>
<tr>
<td>$\psi_{chh}$</td>
<td>$\sim IG(\alpha_{c0}, \beta_{c0})$</td>
</tr>
<tr>
<td>$\alpha_{c0}$</td>
<td>$= 2, \beta_{c0} = 2$</td>
</tr>
<tr>
<td>$\Lambda_{\phi h j}$</td>
<td>$\sim N(\Theta'<em>{j}, \psi</em>{\phi hj} H_{\phi 0})$</td>
</tr>
<tr>
<td>$H_{\phi 0}$</td>
<td>$= I_{J \times J} \times 100$</td>
</tr>
<tr>
<td>$\psi_{\phi hj}$</td>
<td>$\sim IG(\alpha_{\phi 0}, \beta_{\phi 0})$</td>
</tr>
<tr>
<td>$\alpha_{\phi 0}$</td>
<td>$= 2, \beta_{\phi 0} = 2$</td>
</tr>
<tr>
<td>$\Phi_h$</td>
<td>$\sim IW(R_0^{-1}, \rho_0)$</td>
</tr>
<tr>
<td>$R_0^{-1}$</td>
<td>$= 3, \rho_0 = 1$</td>
</tr>
<tr>
<td>$[\Theta</td>
<td>V_y] = [\text{vec}(\Theta)</td>
</tr>
<tr>
<td>$\Theta$</td>
<td>$= 0_{H \times Z}, D = I_{H \times H} \times 100$</td>
</tr>
<tr>
<td>$V_{\phi h}$</td>
<td>$\sim IG(v_{\phi 0}, \psi_{\phi hj})$</td>
</tr>
<tr>
<td>$v_{\phi 0}$</td>
<td>$= 2, V_{\phi 0} = 2$</td>
</tr>
<tr>
<td>$\pi_h$</td>
<td>$\sim Dirichlet(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$</td>
</tr>
<tr>
<td>$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$</td>
<td></td>
</tr>
</tbody>
</table>
where \( \Lambda_k \) is the \( k \)th row of \( \Lambda \), \( \psi_{ek} \) is \( k \)th element of \( \Psi_e \) \( (k = 1, \ldots, 17) \), \( \Lambda_{ej} \) is the \( j \)th row of \( \Lambda_e \), and \( \psi_{\delta j} \) is \( j \)th element of \( \Psi_\delta \) \( (j = 1, \ldots, 5) \). \( H \) means the number of companies, and \( V_{\gamma h} \) is covariance matrix of path coefficient in structure model of company \( h \). \( L \) is the number of path coefficients. \( Z \) is the number of variables of attributes for company, and we use industrial dummy variables for this. \( \pi_h \) is mixing proportion of finite mixture.

(2) Conditional Posterior Density

In the below, the subscripts of heterogeneous parameters are not used for readability where not confused.

(i) Measurement model.

(a) \( x_i | y_i, \mu, \Lambda, \omega, \Psi_e (i = 1, \ldots, n) \) (Data Augmentation)

We transform categorical data \( y_i \) to continuous data \( x_i \) by

\[
x_i \sim N(\alpha_{y_i \cdot}, \alpha_y)(\mu + \Lambda \omega, \Psi_e),
\]

where the cut-off points vector \( \alpha_k = (\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{k8}, \alpha_{k9})' \) for the rating distribution of question \( k \) are determined by

\[
\alpha_p = \phi^{-1}\left( \sum_{i=1}^{n}{I(y_i \leq p) / n} \right), \quad p = 1, \ldots, 9,
\]

where \( \phi^{-1} \) is the inverse of cumulative distribution function of standard normal distribution, and \( I(y_i \leq p) \) is the indicator function, if \( y_i \leq p \), \( I(y_i \leq p) = 1 \). However, we set

\[
\alpha_0 = -\infty, \quad \alpha_{10} = \infty.
\]

(b) \( \omega_i | x_i, \mu, \Lambda, \Psi_e, \Lambda_j, \Psi_\delta, \Phi, (i = 1, \ldots, n) \)
\[ \omega_i = (\eta_{i1}, \eta_{i2}, \eta_{i3}, \eta_{i4}, \xi_{ki})' \]

(b.1) \[ \omega_i^{[-5]} = (\eta_{i1}, \eta_{i2}, \eta_{i3}, \eta_{i4}, \xi_{ki})' \]

\[
\omega_i^{[-5]} \sim \mathcal{N}\left( \Lambda^{[-5]} \right) \]

where \[ \Sigma = \left( I - \Pi^{[-5]} \right)^{-1} (I - \Pi^{[-5]} \Phi(I - \Pi^{[-5]})^{-1} \Phi(I - \Pi^{[-5]})) \]

\[ \Theta^{[-5]} \] means the parameter and data matrix respect to \[ \omega_i^{[-5]} \].

(b.2) \[ \omega_i^{[5]} = \eta_{si} \]

\[
\begin{align*}
\eta_{si} & \sim \left( \begin{pmatrix} l/\psi_{\delta^5} + \lambda_i^{[5]} \psi_{\delta}^{[-1]} \psi_i \end{pmatrix} \right) \left( \begin{pmatrix} \psi_{\delta12} & 0 & 0 \\ 0 & \psi_{\delta13} & 0 \\ 0 & 0 & \psi_{\delta14} \end{pmatrix} \right) \left( \begin{pmatrix} l/\psi_{\delta^5} + \lambda_i^{[5]} \psi_{\delta}^{[-1]} \lambda_i \end{pmatrix} \right) \\
& \sim \mathcal{N} \left( \Lambda^{[5]} \psi_{\delta}^{[-1]} \lambda_i \right) \\
\end{align*}
\]

where \[ \bar{\eta}_{si} = G(\eta_{3i}, S) + \gamma_{si} \eta_{4i} \] ( \[ G(\eta_{3i}, S) \] is mentioned in equation A6),

\[ x_i^{[5]} = [x_{i12}, x_{i13}, x_{i14}]', \lambda_i^{[5]} = \begin{bmatrix} l & \lambda_{i13} & \lambda_{i4} \end{bmatrix} \]

(c) \[ \mu_k | x_i, \Lambda, \omega, \Psi \omega (k = 1, \ldots, 17) \]

\[ \mu_k \sim \mathcal{N} \left( (V_0^{-1} + m\psi_{\epsilon k}^{-1})^{-1} (V_0^{-1} \mu_0 + \psi_{\epsilon k}^{-1} \sum_{i=1}^{n} (x_i - \Lambda_i \omega_i)), (V_0^{-1} + m\psi_{\epsilon k}^{-1})^{-1} \right). \]

(d) \[ \Lambda_k | x_i, \omega, \Psi \omega (k = 1, \ldots, 17) \]

\[ \Lambda_k' \sim \mathcal{N} \left( (H_0^{-1} + \psi_{\epsilon k}^{-1} \Omega_k' \Omega_k')^{-1} (H_0^{-1} \Lambda_0 + \Omega_k' x_k), (H_0^{-1} + \psi_{\epsilon k}^{-1} \Omega_k' \Omega_k')^{-1} \right). \]

\[ \Omega_k' = (\omega_{k1}', \cdots, \omega_{knk}')' \]

where \[ \Omega_k' \] is the parameter and data matrix respect to \[ \Lambda_k' \].
(e) \( \psi_{e_k} \mid x_i, \Lambda, \omega \sim IG(n/2 + \alpha_{e0}, \beta_{e_k}) \) \hspace{1cm} (k = 1, \ldots, 17) \hspace{1cm} (B9)

where \( \beta_{e_k} = \beta_{e0} + \sum_{i=1}^{n} (x_{ik} - \mu_k - \Lambda_k \omega)^2 / 2 \).

(ii) Structural model

(f) \( \Lambda_{ej} \mid \omega_j, \Psi_\delta \hspace{1cm} (j = 1, \ldots, 5) \)

Let \( \Lambda_{ej} = \begin{bmatrix} \Lambda_{ej}^{[1]} \\ \Lambda_{ej}^{[2]} \end{bmatrix} \), where \( \Lambda_{ej}^{[1]} \) is the linear part, \( j=1,2,3,4, \) and \( \Lambda_{ej}^{[2]} \) is the nonlinear part, then we estimate \( \Lambda_{ej}^{[1]} \) and separately \( \Lambda_{ej}^{[2]} \).

(f.1) \( \begin{bmatrix} \Lambda_{ej}^{[1]} \\ \Omega^{[1]} \\ \psi_{\delta k}^{[1]} \\ \Theta^{[1]} \\ z_h \\ H_{\omega 0}^{[1]} \end{bmatrix} \sim N(a^{*-}_{ej}, \psi_{\delta k}^{[1]} A^{*}_{ej}) \), \hspace{1cm} (B10)

where \( a^{*-}_{ej} = (H_{\omega 0k}^{-1} + \Omega_k^{[1]} \Theta_k^{[1]} \omega_{ik} - \Omega_k^{[1]} \eta_k \omega_{ik} \), and \( \Omega_k^{[1]} \) is the k th row of \( \Omega^{[1]} \), which corresponding to \( \Lambda_{ej}^{[1]} \cdot j=1,2,3,4. \)

(f.2) \( \Lambda_{ej}^{[2]} = (0, 0, \gamma_{53,S}, \gamma_{54}, 0) \)

where \( \gamma_{53,S} = \begin{cases} (\gamma_{53,S=1}, \gamma_{53,S=2}, \gamma_{53,S=3}, r_{0,S=1}) & \text{if } S = 1 \\ (\gamma_{53,S=1}, \gamma_{53,S=2}, \gamma_{53,S=3}, r_{0,S=2}) & \text{if } S = 2 \\ (\gamma_{53,S=1}, \gamma_{53,S=3}, r_{0,S=4}) & \text{if } S = 3 \\ (\gamma_{53,S=4}, r_{2,S=4}) & \text{if } S = 4 \end{cases} \)

(f.2.1) The prior distribution of \( \gamma_{54} \) is \( N(\Theta_{\gamma 54} z_h, H_{\gamma 54}^{[2]} \) \) and then we have

\( \gamma_{54} \mid \eta_{3i}, \eta_{4i}, \eta_{5i}, S, \gamma_{53}, \psi_{\delta S}, \Theta_{\gamma 54}, z_h, H_{\gamma 54}^{[2]} \)

\( \gamma_{54} \sim N\left(\left(H_{\gamma 54}^{[2]} \Theta_{\gamma 54} z_h + \psi_{\delta 5} \eta_4 \right)^{-1} \left(H_{\gamma 54}^{[2]} \Theta_{\gamma 54} z_h + \psi_{\delta 5} \eta_4 \right)^{-1} \right) \). \hspace{1cm} (B11)

(f.2.2) \( \gamma_{53} \) of Model 1, Model 2, Model 3 and Model 4.

(f.2.1) Model 1(Asymmetric linear. Given \( S=1 \)):
For the change point of $r_{0,S=1}$, based on M-H algorithm, the prior distribution is normal distribution $N(\Theta, \gamma_{\gamma_{0,S=4}} z_h, H_{\gamma_{0,S=4}}^{[2]}).$ Posterior density is

$$p\left(r_{0,S=1} \mid - \right) \propto l(\hat{\eta}_{3} \mid \eta_{3}, \gamma_{53,S=2}, \gamma_{53,S=2}^{-1}, r_{0,S=1}, \eta_{4}, \gamma_{54}, \psi_{\delta 5}) p\left(r_{0,S=1} \mid \Theta, \gamma_{0,S=4} z_h, H_{\gamma_{0,S=4}}^{[2]} \right).$$

(B12)

For $\gamma_{53,S=1}^{(\pm)}$ and $\gamma_{33,S=1}^{(\pm)}$, the prior distribution is $N(\Theta, \gamma_{53,\pm} z_h, H_{\gamma_{53,\pm}}^{[2]}$) and $N(\Theta, \gamma_{53,\pm} z_h, H_{\gamma_{53,\pm}}^{[2]}),$ they share a common prior variance. Then we have

$$\gamma_{53,S=1}^{(\pm)} \mid \eta_{3}, \eta_{4}, \eta_{5}, r_{0,S=1}, \gamma_{54}, \psi_{\delta 5}, \Theta, \gamma_{53,\pm} z_h, H_{\gamma_{53,\pm}}^{[2]}$$

$$\gamma_{53,S=1}^{(\pm)} \sim N \left( \left[ H_{\gamma_{53,\pm}}^{[2]} \right]^{-1} \psi_{\delta 5}^{(\pm)} \eta_{3}^{(\pm)} \right) \left( \left[ H_{\gamma_{53,\pm}}^{[2]} \right]^{-1} \psi_{\delta 5}^{(\pm)} \eta_{3}^{(\pm)} \right)$$

(B13)

$$\gamma_{53,S=1}^{(\pm)} \mid \eta_{3}, \eta_{4}, \eta_{5}, r_{0,S=1}, \gamma_{54}, \psi_{\delta 5}, \Theta, \gamma_{53,\pm} z_h, H_{\gamma_{53,\pm}}^{[2]}$$

$$\gamma_{53,S=1}^{(\pm)} \sim N \left( \left[ H_{\gamma_{53,\pm}}^{[2]} \right]^{-1} \psi_{\delta 5}^{(\pm)} \eta_{3}^{(\pm)} \right) \left( \left[ H_{\gamma_{53,\pm}}^{[2]} \right]^{-1} \psi_{\delta 5}^{(\pm)} \eta_{3}^{(\pm)} \right)$$

(B14)

where $\eta_{3}^{(\pm)} = \eta_{3} - r_{0,S=1}$, when $\eta_{3} \geq r_{0,S=1}$ and $\eta_{3}^{(\pm)} = \eta_{3} - r_{0,S=1}$, when $\eta_{3} < r_{0,S=1}$.

(f 2.2.2) Model 2(Threshold linear). Given $S=2$:

For boundary point of tolerance interval $r_{1,S=2}$ and $r_{2,S=2}$, based on M-H algorithm, the prior distribution is truncated normal $N[0, \max \eta_{3}^{(\pm)}) ||(\Theta, \gamma_{53,\pm} z_h, H_{\gamma_{53,\pm}}^{[2]}$) and $N[\min \eta_{3}^{(\pm)} - 0) ||(\Theta, \gamma_{53,\pm} z_h, H_{\gamma_{53,\pm}}^{[2]}),$ likelihood is given in equation A6. Posterior density is

$$p\left(r_{1,S=2} \mid - \right) \propto l(\hat{\eta}_{5} \mid \eta_{3}, \gamma_{53,S=2}, \gamma_{53,S=2}^{-1}, r_{1,S=2}, \eta_{4}, \gamma_{54}, \psi_{\delta 5}) p\left(r_{1,S=2} \mid \Theta, \gamma_{53,\pm} z_h, H_{\gamma_{53,\pm}}^{[2]} \right),$$

(B15)
$$p(r_{2,S=2} \mid -) \propto l(\eta_5 \mid \eta_3, y_{S=2}^{(+)}, y_{S=2}^{(-)}, r_{1,S=2}, \eta_4, \gamma_{S=4}, \psi_{S=5}) p(r_{2,S=2} \mid \Theta', y_{S=2}^{(+)}, z_h, H_{y_{S=2}}^{[2]}).$$

(B16)

For $y_{S=2}^{(+)}$ and $y_{S=2}^{(-)}$, the prior distribution is $N(\Theta', y_{S=2}^{(+)}, z_h, H_{y_{S=2}}^{[2]})$ and $N(\Theta', y_{S=2}^{(-)}, z_h, H_{y_{S=2}}^{[2]})$, they share a common prior variance. Then we have

$$y_{S=2}^{(+)} \sim N \left( H_{y_{S=2}}^{[2]} - \psi_{S=2}^{(+)}, \psi_{S=2}^{(+)}, \Theta', y_{S=2}^{(+)} z_h, H_{y_{S=2}}^{[2]} \right),$$

$$y_{S=2}^{(-)} \sim N \left( H_{y_{S=2}}^{[2]} - \psi_{S=2}^{(-)}, \psi_{S=2}^{(-)}, \Theta', y_{S=2}^{(-)} z_h, H_{y_{S=2}}^{[2]} \right),$$

where $\eta_{S=2}^{(\ast)} = \eta_{S=2}^{(+)} - r_{S=2}$.

(B17)

$$y_{S=2}^{(-)} \sim N \left( H_{y_{S=2}}^{[2]} - \psi_{S=2}^{(-)}, \psi_{S=2}^{(-)}, \Theta', y_{S=2}^{(-)} z_h, H_{y_{S=2}}^{[2]} \right),$$

$$y_{S=2}^{(\ast)} \sim N \left( H_{y_{S=2}}^{[2]} - \psi_{S=2}^{(\ast)}, \psi_{S=2}^{(\ast)}, \Theta', y_{S=2}^{(\ast)} z_h, H_{y_{S=2}}^{[2]} \right),$$

where $\eta_{S=2}^{(\ast)} = \eta_{S=2}^{(-)} - r_{S=2}$.

(f 2.2.3) Model 3 (Asymmetric logit). Given $S=3$:

For the change point of $r_{0,S=3}$, based on M-H algorithm, the prior distribution is normal distribution $N(\Theta', y_{0,S=3} z_h, H_{y_{0,S=3}}^{[2]})$. Posterior density is

$$p \left( r_{0,S=3} \mid - \right) \propto l(\eta_5 \mid \eta_3, y_{S=3}^{(+)} y_{S=3}^{(-)}, r_{0,S=3}, \eta_4, \gamma_{S=4}, \psi_{S=5}) p \left( r_{0,S=3} \mid \Theta', y_{0,S=3} z_h, H_{y_{0,S=3}}^{[2]} \right).$$

(B19)

For $y_{S=3}^{(+)}$ and $y_{S=3}^{(-)}$, the prior distribution is $N(\Theta', y_{S=3}^{(+)}, z_h, H_{y_{S=3}}^{[2]})$ and $N(\Theta', y_{S=3}^{(-)}, z_h, H_{y_{S=3}}^{[2]})$, they share a common prior variance. Then we have

$$y_{S=3}^{(+)} \mid \eta_5^{(+)} \eta_5^{(+)}, \eta_4, \gamma_{S=3}^{(+)} \gamma_{S=3}^{(-)}, \Theta', y_{S=3}^{(+)}, z_h, H_{y_{S=3}}^{[2]}.$$
\[ \gamma^{(+)}_{3,5,S=3} \sim N \left( \left( H_{3,5,S=3}^{-1} + \psi \delta^{(-)} \eta \right)^{-1} \left( H_{3,5,S=3}^{-1} \Theta_{3,5,S=3}^{-1} \gamma_{3,5,S=3}^{-1} \left( \eta - \gamma_{3,5,S=3} \right) \right) \right) \]

where \( \eta = \frac{1}{1 + \exp(-h_{k}^{(-)} + r_{0,S=3})} \frac{1}{2} \).

\[ \gamma^{(-)}_{3,5,S=3} | \eta \sim N \left( \left( H_{3,5,S=3}^{-1} + \psi \delta^{(-)} \eta \right)^{-1} \left( H_{3,5,S=3}^{-1} \Theta_{3,5,S=3}^{-1} \gamma_{3,5,S=3}^{-1} \left( \eta - \gamma_{3,5,S=3} \right) \right) \right) \]

for boundary point of tolerance interval \( r_{1,S=4} \) and \( r_{2,S=4} \), the prior distribution is

\[ N[0, \text{max}({\eta^{(+)}}_{3,S=4})] \left( \Theta_{3,5(S=4),S=4} z_{h}, H_{3,5,S=4}^{(2)} \right) \text{ and } N[\text{min}({\eta^{(-)}}_{3,S=4}), 0] \left( \Theta_{3,5(-),S=4} z_{h}, H_{3,5,S=4}^{(2)} \right) \]

for \( \gamma^{(+)l}_{3,5,S=4} \) and \( \gamma^{(-r,2)}_{3,5,S=4} \), the prior distribution is \( N(\Theta^{l}_{r_{1,S=4},S=4} z_{h}, H^{(2)}_{r_{1},S=4}) \) and \( N(\Theta^{r}_{r_{2,S=4},S=4} z_{h}, H^{(2)}_{r_{2},S=4}) \).

Then we have

\[ \gamma^{(+,1)}_{3,5,S=4} \sim N \left( \left( H_{3,5,S=4}^{-1} + \psi \delta^{(-)} \eta \right)^{-1} \left( H_{3,5,S=4}^{-1} \Theta_{3,5(1),S=4}^{-1} \gamma_{3,5,S=4}^{-1} \left( \eta - \gamma_{3,5,S=4} \right) \right) \right) \]

(22)

(23)

(24)
where \( \eta_{i3}^* = \frac{1}{1 + \exp(-\eta_{i3}^{(r_{1,1})} + r_{1,S=4}^1)} - \frac{1}{2} \).

\[
\gamma_{53,S=4}^{(-r_{2,2})} | \eta_{5}^{(-r_{1,1})}, \eta_{4}, \eta_{5}, \gamma_{54}, \gamma_{S}, \psi_{S}, \Theta_{\gamma_{53(-,S=4)}, z_{B}, H_{53,S=4}^{(2)}} = N(H_{53,S=4}^{(2)} + \psi_{S}^{-1} \eta_{3}^{*} \eta_{B}^{*})
\]

(B25)

where \( \eta_{3}^* = \frac{1}{1 + \exp(-\eta_{3}^{(r_{1,1})} + r_{2,S=4}^1)} - \frac{1}{2} \).

(f.3) \( (\pi_{1}, \pi_{2}, \pi_{3}, \pi_{4}) \sim \text{Dirichlet}(\alpha_{1} + n_{1}, \alpha_{2} + n_{2}, \alpha_{3} + n_{3}, \alpha_{4} + n_{4}) \),

\( n_{m} = \sum_{i=1}^{I} \# \{ S_{i} = m \} \) given in (e.2) of last draw. \( I \) is the number of customers of the observed company.

(f.4) \( \Pr(S_{i} = m | \pi_{3}, \eta_{3}, \eta_{4}, \eta_{5}, \gamma_{54}, \psi_{S}, \psi_{S}, \Theta) = \frac{\pi_{m}f((\eta_{5i} - \gamma_{54} \eta_{4i}) | G(\eta_{3i}, S_{i} = m), \psi_{S})}{\sum_{m=1}^{M} \pi_{m}f((\eta_{5i} - \gamma_{54} \eta_{4i}) | G(\eta_{3i}, S_{i} = m), \psi_{S})} \),

where \( M=4, \) and \( f((\eta_{5i} - \gamma_{54} \eta_{4i}) | G(\eta_{3i}, S_{i} = m), \psi_{S}) \) is the pdf of \( \eta_{5i} - \gamma_{54} \eta_{4i} \), based on equation B2, \( (\eta_{5i} - \gamma_{54} \eta_{4i}) \sim N(G(\eta_{3i}, S_{i} = m), \psi_{S}) \). Choosing the value of \( S_{i} \) of each customer according to the posterior probabilities of four candidates.

(g) \( \psi_{j, \omega_{j}} | \Lambda_{\omega_{j}}, \omega_{j} \) \( (j = 1, \ldots, 5) \)

\[
\psi_{j, \omega_{j}} \sim IG \left( \alpha_{\omega_{j}} + n/2, \beta_{\omega_{j}} + \sum_{i=1}^{n} (\eta_{j} - \Lambda_{\omega_{j}} \omega_{j})^{2}/2 \right) \text{ for } j = 1, \ldots, 4, \tag{B26}
\]

\[
\psi_{S, \omega_{S}} \sim IG \left( \alpha_{\omega_{S}} + n/2, \beta_{\omega_{S}} + \sum_{i=1}^{n} (\eta_{S} - G(\gamma_{S3}, \gamma_{3}, \eta_{S}) - \gamma_{S4} \eta_{4i})^{2}/2 \right), \tag{B27}
\]

where \( n_{i} \) is the number of rows of \( \omega_{j} \).

(h) \( \Phi | \omega_{j} \)

\[
\Phi \sim IW(\exp(\hat{\xi}_{\omega_{j}}^{+} + R_{0}^{-1}, n + \rho_{0}), \tag{B28}
\]

75
where $\xi = (\xi_1, \cdots, \xi_n)'$.

(iii) Hierarchical Bayes Regression

(i) $\left[ \Theta | V_j, \Lambda_{\omega}, z_h \right] = \left[ \text{vec}(\Theta) | V_j, \Lambda_{\omega}, z_h \right] \sim N(\tilde{d}, V_j \otimes W) \ (h = 1, 2, \ldots, 21)$ \hspace{1cm} (B29)

where $W = (z_h' z_h + D)^{-1}$, \hspace{0.5cm} $\tilde{d} = \text{vec}(\tilde{D})$, \hspace{0.5cm} $\tilde{D} = W^{-1} (z_h' B + D \text{ vec}(\Theta))$, \hspace{0.5cm} $B = (\tilde{\gamma}_v' \tilde{\gamma}_2' \cdots \tilde{\gamma}_L')$.

\[ \tilde{\gamma}_h = (\gamma_1, \gamma_2, \ldots, \gamma_j, \gamma_{j+1}, \ldots, \gamma_{L-1}, \gamma_L) \]

(j) $\left[ v_j^* | \Theta, z_h, \tilde{\gamma}_h, V_j, V_j^0 \right] \sim \text{Gamma} \left( v_0 + \frac{L}{2}, V_j^0 + \frac{(\tilde{\gamma}_h - z_h' \Theta)'(\tilde{\gamma}_h - z_h' \Theta)}{2} \right)$. \hspace{1cm} (B30)
Appendix C: Figures of CSI Scores

Mobile Telecommunication:

**M1**
- Mean = 49.60
- Std. Dev. = 13.966
- N = 447
- Reference point = 49.65
- Median = 49.08

**M2**
- Mean = 52.38
- Std. Dev. = 16.94
- N = 447
- Reference point = 52.25
- Median = 52.57

**M3**
- Mean = 45.60
- Std. Dev. = 13.344
- N = 447
- Reference point = 45.90
- Median = 45.47

**M4**
- Mean = 44.40
- Std. Dev. = 14.406
- N = 300
- Reference point = 44.35
- Median = 44.32

Convenient Store:

**C1**
- Mean = 43.26
- Std. Dev. = 17.437
- N = 455
- Reference point = 43.10
- Median = 43.43

**C2**
- Mean = 54.76
- Std. Dev. = 11.555
- N = 456
- Reference point = 54.91
- Median = 55.10
Hotel

C 3

C 4

C 5

C 6

H 1

H 2
Appendix D: Shapes of Nonlinear Functions

Mobile Telecommunication:

Convenient Store: