Service Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Hierarchical Bayes Modeling of the Customer Satisfaction Index

Nobuhiko Terui, Shohei Hasegawa, Taemyung Chun, Kosuke Ogawa,

Hierarchical Bayes Modeling of the Customer Satisfaction Index

Nobuhiko Terui, Shohei Hasegawa, Taemyung Chun
Graduate School of Economics and Management, Tohoku University
Kawauchi, Aoba-ku, Sendai 980-8576, Japan
terui@econ.tohoku.ac.jp, shohei.hasegawa@gmail.com, sendai1000goon@yahoo.co.jp

Kosuke Ogawa
Graduate School of Business Administration, Hosei University
Fujimi 2-17-1, Chiyoda-ku, Tokyo 102-8160, Japan.
huko-ogawa@nifty.com

Customer Satisfaction Index has been developed in many countries including North America, Europe Asia last decades, which are based on Americal Customer Satisfaction Index (ACSI) by the University of Michigan, where the latent factor "Customer Satisfaction" related to the customer loyalty is estimated by a covariance structural model with six factors generated from 17 question items and PLS method. They apply the identical structural model to all companies in order to measure the national and industrial indexes that are used to compare the services in different companies as well as industries.

In this paper, by using the assumption that the same model must be applied to every company, we 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. Moreover, our approach has the additional advantages. (i)The volume of information (number of survey data) can be augmented, (ii)The index can be constructed without additional surveys for new company (forecasting) and not-surveyed company (missing observations), (iii)When aggregating individual index of each company up to the industrial index and national index, the communality assumption could increase the stability of the macro index.

Key words: customer satisfaction index; hierarchical Bayes; MCMC(Markov chain Monte Carlo)
History: Received Jul. 25, 2010; Received in revised form Oct. 3, 2010; Accepted Oct. 10, 2010; Online first publication Jan. 17, 2011

1. Introduction

The Customer Satisfaction Index (CSI) has been developed in many countries including North America, Europe and Asia in the last decade. Their methodology and modeling are based on the American Customer Satisfaction Index (ACSI) that was established in 1994 by the National Quality Research Center, Stephen M. Ross School of Business at the University of Michigan. They evaluate the products and services of specific private sector companies and the performances of federal agencies (public sector) and analyze the factors of economic growth. The ACSI is a uniform, national, cross-industry measure of satisfaction with the quality of goods and services available to household consumers in the United States. Investigation subjects include the service sectors of the economy, industries, and companies to build (i) a National index, (ii) a 10-sector index, (iii) a 40-industry index, and (iv) a 200-company index. They report that ACSI is predictive of corporate performance, growth in the gross domestic product (GDP), and changes in consumer spending.

Many research, including Homburg, Koschate, and Hoyer (2005), Keiningham, Perkins-Munn, and Evans (2003), Rust and Zahorik (1993) and Oliver (1997), has shown that customer satisfaction affects individual consumer's purchase behavior. The literatures have also discussed on the linkage of aggregate satisfaction scores to market share, stock price and other firm's profitability in e.g. Fornell et al. (2006), Anderson, Fornell, and Lehmann (1994), Anderson, Fornell, and Rust (1996), Mittal et al.(2005), Anderson, Fornell, and Mazvancheryl (2004) and Gruca and Rego (2005). More recently, Fornell, Rust and Dekimpe (2010) show how consumption is dependent on both customer satisfaction and service.

The CSI intends to be a benchmark for companies to compare themselves with others in their own or other industries. CSI models deal with "Customer Satisfaction (CS)" as the latent factor leading to customer loyalty, and it is estimated by structural equation models with six factors generated from 17 question items using the Partial Least Squares (PLS) method. They need to employ the identical structural model for all subjects—companies—in order to compare the customer satisfaction metrics for services in different companies, and their indexes are aggregated to industry and national levels to build national macro indexes.

However, each company is likely to be heterogeneous in structure on customer satisfaction measures. That is, the individual companies could exhibit not only parameter heterogeneity with the model but also structural heterogeneity. Thus, it is very likely that CSI models are not stable across the companies because some companies might not always have significant path coefficients.

On the other hand, the hierarchical Bayesian approach provides several advantages over classical methods for treating the heterogeneity of subjects in CSI modeling based on structural equation models. Bayesian methods allow the flexible incorporation of prior information and allow the estimation of individual-specific estimates, while accounting for uncertainty in such estimates. Specifically, in our modeling context, the hierarchical Bayes structural equation methodology provides individual-specific estimates of the factor scores, structural coefficients, and other model parameters.

In this study, we accommodate the necessity of using an identical structural model for every subject as prior information in the modeling. Specifically, we assume a common structure behind individual CSI models and deal with the path coefficients of each subject as random effects from common parameters by using hierarchical modeling to link the path coefficients of each subject in the form of a hierarchical regression model to estimate the customer satisfaction structure across companies.

In the literature, Ansari et. al (2000) discussed hierarchical Bayes (HB) modeling of structural equation models on the grounds that the absence of adequate accounting for unobserved heterogeneity results in misleading inferences. However, our motivation for employing this modeling is slightly different from theirs in that the HB approach is used to incorporate homogeneity inside an industry in order to accommodate prior restrictions of identical structures of customer satisfaction on heterogeneous individuals. Our approach allows for the appropriate pooling of information while taking into account heterogeneity, implying the shrinkage estimation of path coefficients.

We show that, representing "homogeneity or communality" inside an industry and "heterogeneity" outside that industry, the hierarchical Bayes modeling produces more significant path coefficients by shrinking the individual estimate toward common values than other methods of estimation. Our approach has three additional advantages, as well. (i) The volume of information (amount of survey data) can be augmented, which is important because there is no guarantee that questioning 200 customers for each company, as in the ACSI, is sufficient. (ii) The index can be constructed without additional surveys for new companies (forecasting) and not surveyed companies (missing observations). (iii) When aggregating the individual indexes of each company into the industrial index and national index, the communality assumption might increase the stability of the macro index.

2. CSI Modeling Using the Hierarchical Structural Equation Model

Our CSI model for the company h is defined by modifying the ACSI as in Figure 1.

The model describes that the customer expectation (CE) drive the perceived quality (PQ) as well as perceived value (PV) and these three latent variables cause the customer satisfaction (CS). Then CS affects directly positive voice (Po.V) as well as negative voice (Ne.V) and customer loyalty (CL) indirectly. The model is based on the Japanese customer satisfaction index, which was made and managed by SPRING (service productivity and innovations and growth) agency as the national project in Japan, whehre American customer satisfaction index is modified to be more adjustable to Japanese customer and industry after extensive experiments and preliminary research. We employ their model in the base model.

The model is expressed as the sets of equations, for $i = 1, \ldots, T_h$ (number of respondents), by

$$\begin{cases} Y_{hi} = \Lambda_h \omega_{hi} + \varepsilon_{hi}, & \varepsilon_{hi} \sim N(0, \Psi_{\varepsilon h}) \\ \eta_{hi} = \Pi_h \eta_{hi} + \Gamma_h \xi_{hi} + \delta_{hi}, & \delta_{hi} \sim N(0, \Psi_{\delta h}) \end{cases} \tag{1}$$

where variables are defined as follows.

$$\omega_{hi} = \left(\eta_{hi}', \xi_{hi}\right)';$$

ξ_{hi} : Customer Expectation(CE)

η_{1hi} : Perceived Quality(PQ)

η_{2hi} : Perceived Value(PV)

η_{3hi} : Customer Satisfaction (CS)

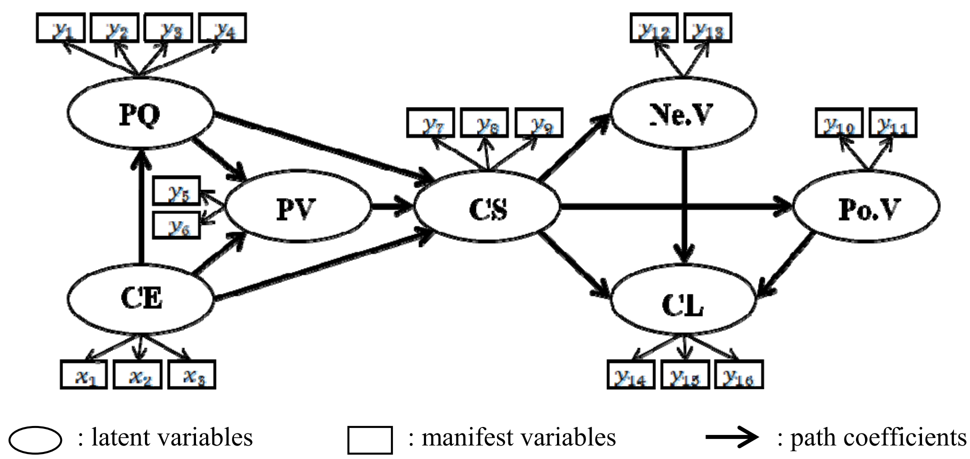
η_{4hi} : Positive Voice(Po.V)

η_{5hi} : Negative Voice(Ne.V)

η_{6hi} : Customer Loyalty(CL)

In the equation (1), we have $\Lambda_h : 19 \times 7$, $\Pi_h : 6 \times 6$ and $\Gamma_h : 6 \times 1$ matrices.

Figure 1 The Structure of Customer Satisfaction



We have 19 manifest variables, Y_{hi} , which are ordered categorical variables based on survey questions rated between 1(low) to 10(high) scale. The set of model parameters is $\{(\Lambda_h, \Pi_h, \Gamma_h), \Psi_{\varepsilon h}, \Psi_{\delta h}\}$, and we denote the coefficient parameter as $\theta_h = (\Lambda_h, \Pi_h, \Gamma_h)$. There are several ways of estimating model parameters, including the maximum likelihood and the generalized least squares methods, which estimate by using the relationships between covariances in the population

$$\mathrm{Cov}(Y_h) \equiv \Sigma(\theta_h) = f(\Lambda_h, \Pi_h, \Gamma_h), \tag{2}$$

by its sample estimate

$$S = \frac{1}{T}\sum_{i=1}^{T}\left(Y_i - \bar{Y}\right)\left(Y_i - \bar{Y}\right)', \tag{3}$$

and is implemented by the software "AMOS." In particular, the maximum likelihood estimate needs to have a strong assumption of Gaussian likelihood for ordered categorical variables. The ACSI uses the partial least squares (PLS) method initiated by Wold (1981) and applied to marketing research by Fornell and Cha (1994), without relying on the distributional assumption on the error terms.

On the other hand, the Bayes method, in Lee (2007) for example, evaluates the joint posterior density conditional on the data,

$$p(\Lambda_h, \Pi_h, \Gamma_h, \{\omega_{hi}\} \mid \{Y_{hi}\}). \tag{4}$$

The likelihood is very important for the Bayes method as it transforms categorical data Y_i to the continuous variable, X_i, which follows the specified Gaussian distribution in data augmentation. (See Tanner and Wang, 1987 and Lee, 2007.) For example, we introduce a set of thresholds over the domain of normal distribution to decompose into ten segments for 1 to 10 scale categorical data, so that the probability of each region corresponds to the probability mass of each ordered category. Then, when we have a categorical sample, we generate a normal deviate from the truncated normal distribution whose cut points are defined by the corresponding segment. The detailed algorithm is in the appendix.

We note that the latent variables $\{\omega_{hi}\}$ are dealt with using parameters and jointly estimated with other model parameters.

Our proposed method in the study is hierarchical Bayes (HB) modeling, which connects each model of the companies $h = 1,...,H$ such that

$$\boldsymbol{\beta}_h = \boldsymbol{\Theta}'\mathbf{z}_h + \boldsymbol{\eta}_h; \quad \boldsymbol{\eta}_h \sim N_k\left(\mathbf{0}, V_\beta\right), \tag{5}$$

where $\boldsymbol{\beta}_h = vec\left(\Pi_h, \Gamma_h\right)$ is eleven dimensional vector of path coefficients between latent variables for the company *h*. This implies a prior distribution on the path coefficients, and $\boldsymbol{\beta}_h$ is restricted by the common parameter ($\boldsymbol{\Theta}, V_\beta$) when the *k*-dimensional attribute data of the company *h,* i.e. $\mathbf{z}_h$, is given($\boldsymbol{\Theta}: 11\times k, \quad V_\beta : 11\times 11$). This prior specification is motivated in our problems as follows. CSI's methodology has a strong assumption that every company has the identical customer satisfaction structure in order to compare the services across different companies and industries, and thus aggregating to industrial and national levels. However, each company should have structural heterogeneity on customer satisfaction measures. The CSI model should not be stable across all companies in the sense that some of path coefficient estimates could not be significant for some companies.

The prior specification (5), together with appropriate prior specifications for other parameters discussed in Lee (2007), are combined with the Gaussian likelihood based on the data augmentation above to constitute joint posterior density,

$$p\left(\{\Lambda_h\}, \{\Pi_h\}, \{\Gamma_h\}, \{\omega_{hi}\}, \boldsymbol{\Theta}, V_\beta \mid \{Y_{hi}\}, \{z_h\}\right). \tag{6}$$

The numerical evaluation of this density is conducted by Markov chain Monte Carlo (MCMC), and its algorithm is described in the appendix.

As a result, the index is composed as

$$\mathrm{CSI}_h^* = f\left(\boldsymbol{\beta}_h, \boldsymbol{\Theta} \mid \mathbf{z}_h, \{\mathbf{Y}_h\}_{h=1}^H\right) \tag{7}$$

3. Empirical Results

Our dataset is supplied by the Japanese CSI development working group managed by the Japanese agency of Service Productivity and Innovation Growth (SPRING). It includes three industries--cellular phone (four companies), convenience store (five) and hotel (12)--totaling 21 companies.

The CSI model was estimated using maximum likelihood (ML), generalized least squares (GLS), partial least squares (PLS), and Bayes across individual respondents for each of the three industries in our overall sample: (1) cellular phone (overall = 1991 respondents, C1 = 447, C2 = 447, C3 = 447, C4 = 300), (2) convenience store (overall = 2232 respondents, S1 = 456, S2 = 456, S3 = 360, S4 = 360, S5 = 300), and (3) Hotel (overall = 4027 respondents, H1~H12 = 300).

3.1 Industry level

As a preliminary examination, we report the estimation result for the industry level structure. That is, we aggregate individual companies data to each of the three industries and estimate the industry CS structure by four estimation methods, i.e. ML, GLS, PLS, and Bayes methods. As for the last method, we propose two kinds of hierarchical Bayes models. The first model assumes diffuse prior on the hyper parameters, e.g. as is discussed in Lee (2007), and the second model selects the hyper parameter so as to minimize the DIC (deviance information criteria), i.e. Bayesian model fit criteria.

Table 1 Number of Insignificant Path Coefficients Estimated at 5%

	ML	GLS	PLS	Bayes
Cellular Phone	1 (PQ->PV)	0	1 (Ne.V->CL)	1 (Ne.V->CL)
Convenience Store	1 (Ne.V->CL)	0	0	1 (Ne.V->CL)
Hotel	0	1 (Ne.V->CL)	0	1 (Ne.V->CL)

Table 2 Number of Path Coefficients That Are not Estimated as Significant at 5%

(a) Cellular Phone:

	ML	GLS	PLS	Bayes	HB1(c=1)	HB2(c=1e-5)
C1	6	2	1	1	1	0
C2	3	3	1	1	1	0
C3	2	2	1	1	2	0
C4	4	3	1	2	2	0

(b) Convenience Store

	ML	GLS	PLS	Bayes	HB1(c=1)	HB2(c=1e-5)
S1	3	5	1	1	1	0
S2	2	4	1	1	1	0
S3	1	4	1	1	1	0
S4	1	4	0	2	2	1
S5	2	2	1	1	1	0

(c) Hotel

	ML	GLS	PLS	Bayes	HB1(c=1)	HB2(c=1e-5)
H1	2	4	0	1	1	0
H2	1	2	1	1	2	0
H3	2	3	1	1	2	0
H4	1	1	1	1	2	0
H5	0	1	0	0	0	0
H6	3	2	1	1	0	0
H7	5	5	1	1	3	0
H8	3	3	1	2	3	0
H9	2	4	1	1	2	0
H10	2	2	2	2	2	0
H11	3	2	1	1	3	0
H12	4	4	3	3	3	0

Table 1 shows the number of path coefficient estimates without 5% significance in the model for each estimation method, where the insignificant path is indicated inside the bracket. For example, in the case of the cellular phone industry, the ML method calculated one insignificant coefficient of the path from PQ to PV, but the GLS produced significant estimates for all paths, and e.t.c.

The results show that all methods perform relatively well in that most of the path coefficients were estimated as significant.

3.2 Company level

Next, we apply our proposed HB modeling of the CSI to this dataset. We first explain the necessary variable and the hyper-parameter setting.

The demographic variable z_h comprises the constant term and the industrial dummy; 1: cell-phone, 2: convenience store, 3: hotel. This setting assumes two kinds of homogeneity for the individual company; the constant term reflects homogeneity across industries, and the industrial dummy expresses homogeneity inside the industry.

As for the hyper-parameter of the prior variance of β, implying the degree of heterogeneity, we establish two settings: non-informative weak prior (HB1) and prior minimizing DIC (deviance information criteria, Spiegelhalter et al, 2002) (HB2).

Table 2 indicates the number of path coefficients estimated as insignificant for each category. We observe the following: the ML and GLS perform worse because path coefficients were often estimated as not significant. On the other hand, the PLS, Bayes and HB are relatively stable, with HB2 being the best.

3.3 CSI Scores

After obtaining the estimate, we obtain the customer satisfaction score as the latent score value for each respondent, $\{\eta_{3h1}, \eta_{3h2}, \ldots, \eta_{3hn}\}$. The ACSI uses the mean of the latent score, $E[\eta_{3h}]$, as an index, leading to 100 scores by using the transformation:

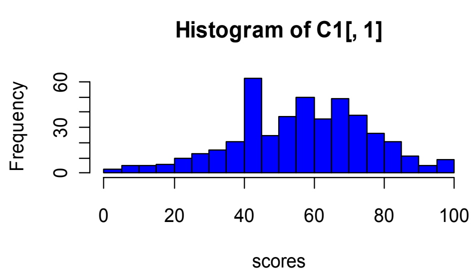
$$\mathrm{ACSI} = \frac{E[\eta_3] - Min[\eta_3]}{Max[\eta_3] - Min[\eta_3]} \times 100 \quad (8)$$

Tables 3-1 and 3-2 show summary statistics of the empirical distribution of respondents' scores $\{\eta_{31}{}^{[m]}, \eta_{32}{}^{[m]}, \ldots, \eta_{3n}{}^{[m]}\}$ when using different estimation methods: m = PLS, Bayes, HB1 and HB2. It is well known that the mean $E[\xi]$ is not an appropriate measure when the distribution is skewed, and the median could be more appropriate.

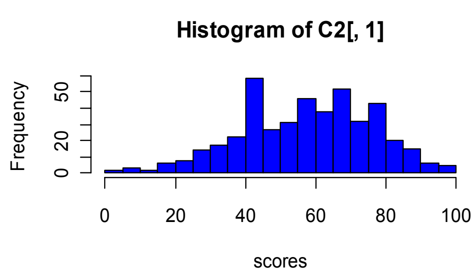
Figure 2 Histogram of PLS estimates for Customer Satisfaction Index

Cellular phone

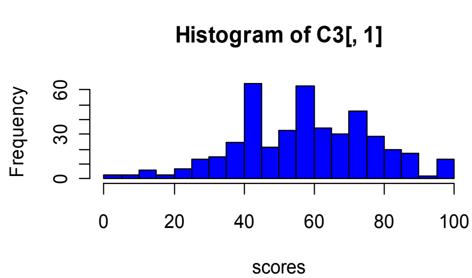
*C1



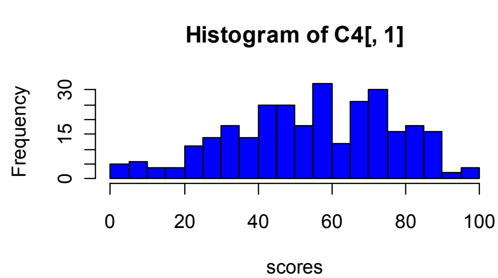
*C2



*C3

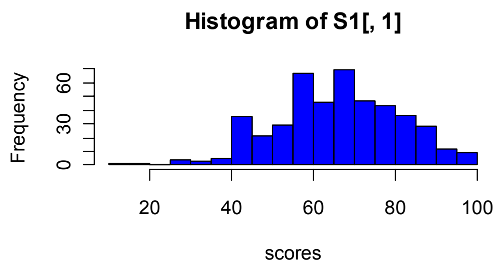


*C4

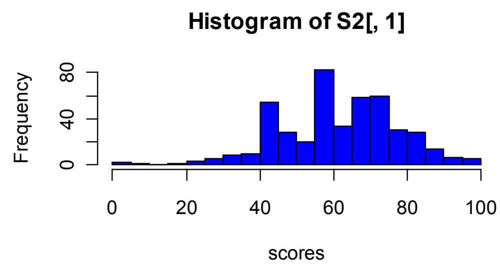


Convenience store

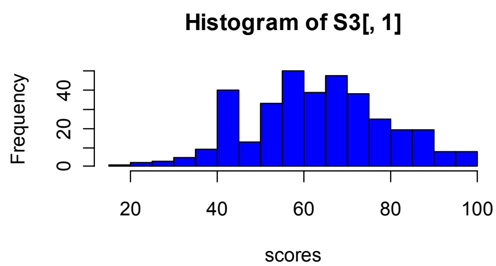
*S1



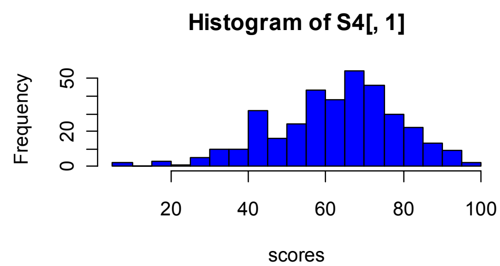
*S2



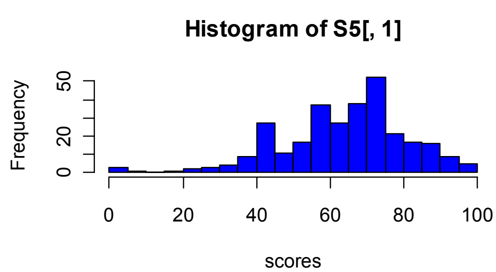
*S2



*S4

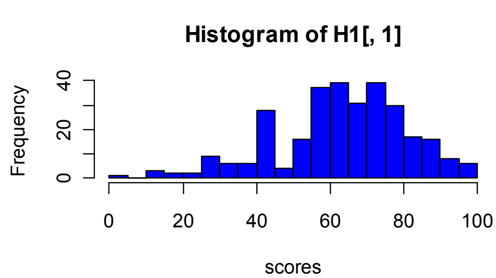


*S5

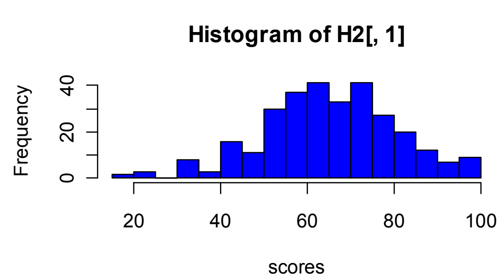


Hotel

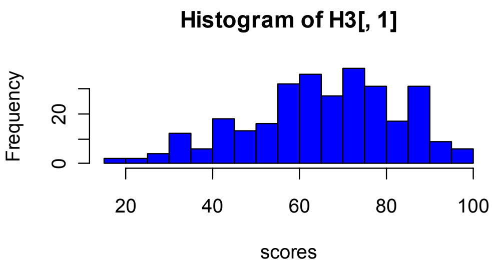
*H1



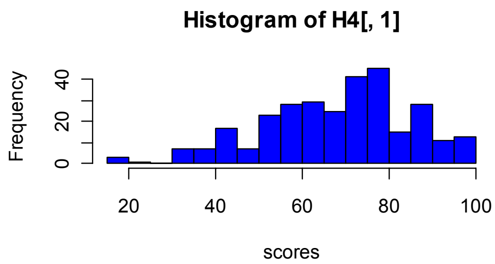
*H2



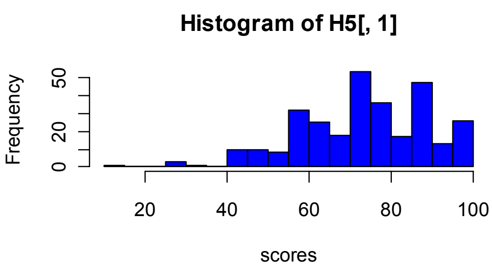
*H3



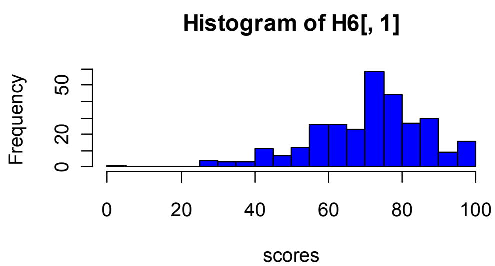
*H4



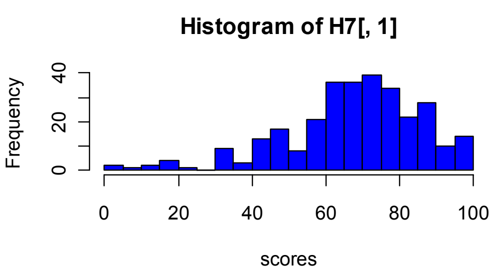
*H5



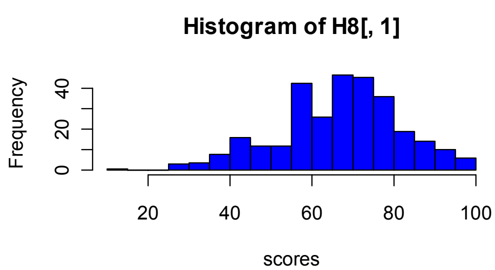
*H6



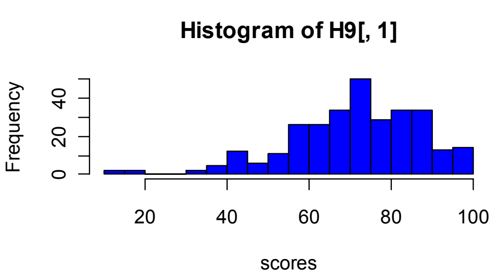
*H7



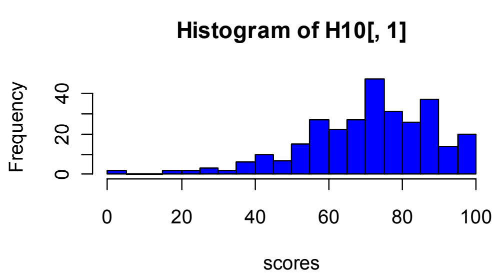
*H8



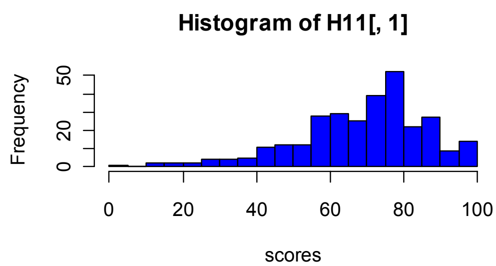
*H9



*H10



*H11



*H12

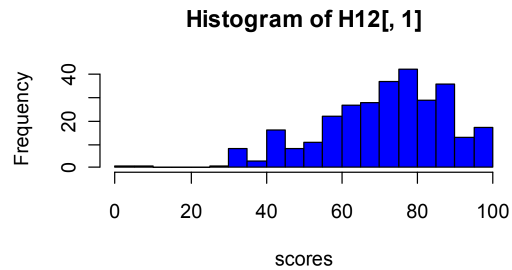


Table 3-1 Summary Statistics of CSI Score Distribution

	PLS			Bayes			HB 1			HB 2		
	Mean	Median	STD	Mean	Median	STD	Mean	Median	STD	Mean	Median	STD
CellularPhone－All	56.56	58.35	19.92	52.11	51.76	17.86	N.A	N.A	N.A	N.A	N.A	N.A
C1	56.10	57.80	19.57	51.48	51.43	18.43	51.93	51.80	18.21	51.51	51.25	18.36
C2	57.35	58.75	18.87	52.70	53.32	18.18	52.81	53.18	18.27	52.39	52.90	18.17
C3	56.93	56.99	19.64	52.88	52.66	18.27	52.79	52.68	18.16	52.85	52.88	18.11
C4	54.53	55.50	22.03	50.26	49.16	18.63	49.89	49.04	18.52	50.09	49.20	18.52
Convenience Store-All	62.97	63.41	16.06	57.91	58.17	15.54	N.A	N.A	N.A	N.A	N.A	N.A
S1	60.26	60.99	17.51	54.40	54.03	18.74	54.52	54.52	18.68	54.59	54.44	18.62
S2	61.40	62.58	16.10	56.65	57.06	15.82	56.76	57.14	15.75	56.64	57.15	15.75
S3	54.54	54.72	19.42	49.36	49.69	18.87	48.85	49.36	18.76	48.91	49.55	18.80
S4	59.55	60.57	17.51	56.70	57.24	16.96	57.19	57.90	16.95	56.91	57.62	17.03
S5	63.51	66.61	17.03	56.98	58.26	17.92	56.56	57.31	17.68	57.37	58.43	17.70
Hotel-All	68.41	70.49	17.01	63.13	64.01	16.99	N.A	N.A	N.A	N.A	N.A	N.A
H1	63.27	63.95	17.73	60.27	59.84	17.38	60.29	60.33	17.35	60.18	60.13	17.28
H2	56.79	55.71	19.16	53.67	53.18	19.62	54.16	53.59	19.39	54.10	53.45	19.53
H3	57.39	59.74	21.28	52.85	53.43	19.94	52.25	52.77	20.35	51.88	52.41	20.27
H4	61.88	64.86	20.12	55.43	55.94	21.09	55.16	55.96	20.77	55.38	56.34	20.93
H5	68.72	69.86	18.64	59.19	58.60	19.71	59.35	58.93	19.85	59.50	59.18	19.73
H6	70.94	74.07	15.74	61.02	61.56	19.01	60.97	58.93	19.85	60.97	59.18	19.73
H7	67.57	69.87	18.76	58.23	59.06	19.04	58.35	59.27	18.99	58.07	59.06	19.00
H8	62.60	62.84	16.91	59.46	59.77	17.77	59.14	59.66	17.67	58.97	59.34	17.52
H9	67.58	68.45	18.05	58.49	59.81	19.93	58.27	59.56	20.06	57.99	59.33	19.80
H10	69.48	73.01	18.51	60.48	61.14	18.71	60.77	61.57	18.68	60.75	61.49	18.51
H11	67.28	69.90	18.36	57.90	58.85	19.03	57.71	58.98	19.14	57.42	58.51	19.13
H12	70.92	74.00	17.17	62.61	63.45	18.14	62.39	63.20	18.06	62.45	63.25	17.93

Table 3-2 Skewness and Kurtosis

	PLS		Bayes		HB1		HB2	
	skewness	**kurtosis**	**skewness**	**kurtosis**	**skewness**	**kurtosis**	**skewness**	**kurtosis**
C1	−0.27	−0.01	−0.04	−0.05	−0.04	−0.05	−0.04	−0.05
C2	−0.29	−0.21	−0.08	−0.30	−0.08	−0.29	−0.09	−0.30
C3	−0.20	−0.02	−0.01	−0.05	−0.01	−0.06	−0.01	−0.06
C4	−0.29	−0.49	−0.09	−0.21	−0.09	−0.21	−0.09	−0.20
S1	−0.15	−0.17	−0.02	−0.31	−0.01	−0.31	−0.02	−0.31
S2	−0.35	0.41	−0.10	−0.07	−0.10	−0.05	−0.10	−0.06
S3	−0.01	−0.26	0.06	−0.05	0.07	−0.07	0.08	−0.08
S4	−0.47	0.21	−0.22	0.02	−0.22	0.02	−0.21	0.01
S5	−0.69	−0.69	−0.20	0.04	−0.21	0.04	−0.21	0.04
H1	−0.56	0.41	−0.22	−0.04	−0.21	−0.05	−0.22	−0.05
H2	−0.24	0.21	−0.04	−0.09	−0.04	−0.10	−0.04	−0.11
H3	−0.36	−0.35	−0.13	−0.47	−0.12	−0.47	−0.12	−0.47
H4	−0.43	0.04	−0.14	−0.37	−0.14	−0.36	−0.14	−0.36
H5	−0.47	0.21	−0.10	−0.25	−0.11	−0.23	−0.11	−0.24
H6	−0.78	1.28	−0.25	−0.16	−0.26	−0.18	−0.27	−0.16
H7	−0.87	1.20	−0.26	0.01	−0.27	0.00	−0.26	−0.01
H8	−0.36	0.25	−0.21	−0.12	−0.22	−0.12	−0.22	−0.11
H9	−0.72	1.00	−0.22	−0.15	−0.22	−0.15	−0.22	−0.16
H10	−0.80	1.00	−0.24	0.01	−0.25	0.04	−0.25	0.04
H11	−0.80	0.87	−0.23	0.15	−0.23	0.13	−0.23	0.12
H12	−0.70	0.74	−0.23	−0.15	−0.23	−0.16	−0.24	−0.15

Figure 2 shows the histogram of the empirical distribution of PLS estimates of CSI. Table 3-1 and 3-2 show the difference between the sample mean and the median, and the skewness and kurtosis measures. The skewness and kurtosis are the deviation measures from normal distribution. According to these tables and figures, we observe that the score distribution is highly skewed in the case of the PLS. This might happen because the PLS does not assume the normal distribution of manifest variables (ordered categorical data). The CSI score (8) is defined by the sample average of generated values of the respondent's latent variable of customer satisfaction, which is rescaled to fall between zero and 100. As is well known in statistics, the sample mean is not always reasonable estimate unless we could assume the distribution is neither symmetric nor normal.

On the other hand, in the case of the Bayes method, the ordered categorical data is first transformed into the normal distribution variable so that the update mechanism can work. As a result, the score distribution becomes compatible with normal distribution.

4. Conclusion and Remarks

In this article, we discussed the problems in the modeling of the customer satisfaction index, in particular the most popular one, the ACSI. Then we suggested that the assumption of an identical structural model for every company must be strong in that some of the path coefficients are not always estimated as significant, and the index constructed by the standardized sample mean of factor score estimates might not be reasonable because of the skewness.

In order to accommodate these problems, we proposed hierarchical Bayes modeling, which accommodates the strong restriction that each company should have the identical structure of the customer satisfaction index in order to compare the indices with other companies or to aggregate their indices to industrial and national levels. We examined the performance of the modeling methods by comparison, and we have summarized the results as follows. The ML is not good because it uses Gaussian likelihood for categorical data. The GLS shares nearly the same properties with the ML. The PLS performs better because it does not depend the distributional property of data. The Bayes method performs better because it transforms data to continuous Gaussian variables. The HB produces the most stable significant path coefficients, showing "communality" inside an industry and "heterogeneity" outside that industry.

In addition to these statistical properties, our proposed method has other advantages: (i) The volume of information (amount of investigation data) can be augmented, which is important because there is no guarantee that questioning 200 customers for each company, as in the ACSI, is sufficient. (ii) The index can be constructed without additional surveys for new companies (forecasting) or/and not-surveyed companies (missing observations)

$$\mathrm{CSI}_{h+1}{}^{*} = f\left(\tilde{\boldsymbol{\beta}}_{h+1}, \boldsymbol{\Theta} \mid \tilde{\mathbf{z}}_{h+1}, \left\{\mathbf{Y}_h\right\}_{h=1}^{H}\right),$$

if the attribute $\tilde{\mathbf{z}}_{h+1}$ of company "h+1" is available. (iii) When aggregating the individual index of each company into an industrial index and national index, the homogeneity assumption might increase the stability of the macro index.

$$\mathrm{CSI}_{\Omega}{}^{*}=\sum_{h\in\Omega} \varpi_h \mathrm{CSI}_h{}^{*}$$

We leave these problems for future research.

Acknowledgement

The authors are grateful to two anonymous referees for their useful comments. The financial support of by the Japanese Ministry of Education Scientific Research Grants (A) 21243030 is acknowledged by Terui and (B) 21330105 by Ogawa and Terui.

Appendix MCMC Algorithms

(1) Prior density

Parameter of the Measurement equation

$$\left[\Lambda_{hk} \mid \psi_{\varepsilon hk}\right] \sim N\left(\Lambda_{0k}, \psi_{\varepsilon hk} H_{0hk}\right)$$

$$\psi_{\varepsilon hk}^{-1} \sim Gamma\left(\alpha_{0\varepsilon k}, \beta_{0\varepsilon k}\right)$$

where Λ_{hk} is k th row of Λ_h

$\psi_{\varepsilon hk}$ is k th diagonal element of $\Psi_{\varepsilon h}$

Parameter of the System equation

$$\left[\Lambda_{\omega hk} \middle| \psi_{\delta hk}, \Theta_k, z_h\right] \sim N\left(\Theta_k' z_h, \psi_{\delta hk} V_{\beta k}\right)$$

$$\psi_{\delta hk}^{-1} \sim Gamma\left(\alpha_{0\delta k}, \beta_{0\delta k}\right)$$

$$\Phi_h^{-1} \sim Wishart\left(R_0, \rho_0\right)$$

where $\Lambda_{\omega hk}$ is k th row of $\Lambda_{\omega h} = \left(\Pi_h \quad \Gamma_h\right)$

$\psi_{\delta hk}$ is k th diagonal element of $\Psi_{\delta h}$

Θ_k is rows of Θ, which corresponding to $\Lambda_{\omega hk}$

$V_{\beta k}$ is block matrix of V_β, which corresponding to $\Lambda_{\omega hk}$

Parameter of the Hierarchical equation

$$\left[\Theta \mid V_\beta\right] = \left[vec\left(\Theta\right) \mid V_\beta\right] \sim N\left(vec\left(\overline{\Theta}\right), V_\beta \otimes D_0^{-1}\right)$$

$$V_\beta^{-1} \sim Wishart\left(\nu_0, V_0\right)$$

(2) Conditional posterior density for MCMC:

(a) Transform categorical data Y_{hi} to continuous data X_{hi} (Data Augmentation)

Set the company *h*'s cut-off parameter vector $\boldsymbol{\alpha}_{hk} = \left(\alpha_{hk0}, \alpha_{hk1}, \cdots, \alpha_{hk9}, \alpha_{hk10}\right)'$ for the rating distribution of question *k* such that $-\infty = \alpha_{hk0} < \alpha_{hk1} < \cdots < \alpha_{hk9} < \alpha_{hk10} = \infty$, by

$$\alpha_{hkj} = \Phi^{-1}\left(\sum_{l=1}^{j} n_{hkl} \big/ T_h\right),$$

where $k = 1,...,19$: questions, $j = 1,...,10$: ten scale rating between 1(low) and 10(high), Φ^{-1} is the inverse of cumulative distribution function of standard normal distribution, n_{hkl} is number of respondents who answered *l* in question k and T_h is the number of respondent for the subject *h*.

Then, if the respondent *i* answers question *k* by *j* rating, i.e., $y_{hik} = j$, generate x_{hik} from the truncated standard normal distribution over the region $\alpha_{hkj-1} < x_{hik} < \alpha_{hkj}$, i.e.

$$x_{hik} \sim N_{\left[\alpha_{hkj-1}, \alpha_{hkj}\right]}\left(0,1\right).$$

(b) Measurement equation

$$\left[\omega_{hi} \mid y_{hi}, \Lambda_h, \Phi_h, \Pi_h, \Gamma_h, \Psi_{\varepsilon h}, \Psi_{\delta h}\right] \sim N\left[\left(\Sigma_h^{-1} + \Lambda_h^T \Psi_{\varepsilon h}^{-1} \Lambda_h\right)^{-1} \Lambda_h^T \Psi_{\varepsilon h}^{-1} y_{hi}, \left(\Sigma_h^{-1} + \Lambda_h^T \Psi_{\varepsilon h}^{-1} \Lambda_h\right)^{-1}\right]$$

where

$$\Sigma_h = \begin{bmatrix} \left(I - \Pi_h\right)^{-1}\left(\Gamma_h \Phi_h \Gamma_h^T + \Psi_{\delta h}\right)\left(I - \Pi_h\right)^{-T} & \left(I - \Pi_h\right)^{-1} \Gamma_h \Phi_h \\ \Phi_h \Gamma_h^T \left(I - \Pi_h\right)^{-T} & \Phi_h \end{bmatrix}$$

set $\Omega_h = \left(\omega_1, \cdots, \omega_{n_h}\right)$, $\Omega = \left(\Omega_1, \cdots, \Omega_H\right)$

$$\left[\psi_{\varepsilon hk}^{-1} \mid X_h, \Omega_h, \Lambda_h\right] \sim Gamma\left[\frac{T_h}{2} + \alpha_{0\varepsilon}, \beta_{\varepsilon k}\right]$$

$$\left[\Lambda_{hk} \mid X_h, \Omega_h, \psi_{\varepsilon hk}\right] \sim N\left[a_{yk}^*, \psi_{\varepsilon hk} A_{yk}^*\right]$$

where

$$A_{yk}^* = \left(H_{0yk}^{-1} + \Omega_{hk}^T \Omega_{hk}\right)^{-1}$$

$$a_{yk}^* = A_{yk}^* \left(H_{0k}^{-1} \Lambda_{0k} + \Omega_{hk}^T X_{hk}\right)$$

$$\beta_{\varepsilon k} = \beta_{0\varepsilon k} + \left(X_{hk} - \Omega_h^T \Lambda_{hk}^T\right)^T \left(X_{hk} - \Omega_h^T \Lambda_{hk}^T\right) / 2$$

X_{hk}^T is k th row of X_h

Ω_{hk}^T is k th row of Ω_h

(c) Structural equation

$$\left[\psi_{\delta hk}^{-1} \mid \Omega_h, \Lambda_{\omega h}\right] \sim Gamma\left[\frac{T_h}{2} + \alpha_{0\delta}, \beta_{\delta k}\right]$$

$$\left[\Lambda_{\omega hk} \mid \Omega_h, \psi_{\delta hk}, \Theta, z_h, V_\beta\right] \sim N\left[a_{\omega k}^*, \psi_{\delta hk} A_{\omega k}^*\right]$$

$$\left[\Phi_h^{-1} \mid \Omega_h\right] \sim Wishart\left[\xi_h^T \xi_h + R_0^{-1}, T_h + \rho_0\right]$$

where

$$A_{\omega k}^* = \left(V_{\beta k}^{-1} + \Omega_{hk}^{*\,T} \Omega_{hk}^*\right)^{-1}$$

$$a_{\omega k}^* = A_{\omega k}^* \left(V_{\beta k}^{-1} \Theta_k' z_h + \Omega_{hk}^{*\,T} \eta_{hk}\right)$$

$$\beta_{\delta k} = \beta_{0\delta k} + \left(\eta_{hk} - \Omega_h^T \Lambda_{\omega hk}^T\right)^T \left(\eta_{hk} - \Omega_h^T \Lambda_{\omega hk}^T\right) / 2$$

$\Omega_{hk}^{*\,T}$ is rows of Ω_h, which corresponding to $\Lambda_{\omega hk}$

(c) HB regression

$$\left[\Theta \mid V_\beta, \Lambda_\omega, z\right] = \left[vec(\Theta) \mid V_\beta, \Lambda_\omega, z\right] \sim N\left(\tilde{d}, V_\beta \otimes W\right)$$

$$\left[V_\beta^{-1} \mid \Lambda_\omega, \Theta, z\right] \sim Wishart(v_0 + H, V_0 + S)$$

where

$$W = \left(z'z + D_0\right)^{-1}$$

$$\tilde{d} = vec\left(\tilde{D}\right) \qquad \tilde{D} = W^{-1}\left(z'B + D_0 vec\left(\bar{\Theta}\right)\right) \qquad B = \left(\beta_1, \cdots, \beta_{\mathrm{H}}\right)'$$

$$S = \sum_{h=1}^{\mathrm{H}} \left(\beta_h - \Theta' z_h\right)\left(\beta_h - \Theta' z_h\right)'$$

References

Anderson, Eugene W., Claes Fornell, and Donald R. Lehmann. 1994. “Customer Satisfaction, Market Share and Profitability: Findings from Sweden,” *Journal of Marketing*, 58, 53–66.

Anderson, Eugene W., Claes Fornell, and Roland T. Rust (1996), “Customer Satisfaction, Productivity and Profitability: Differences Between Goods and Services,” *Marketing Science*, 16, 129–45.

Anderson, Eugene W., Claes Fornell,and Sanal Mazvancheryl (2004), “Customer Satisfaction and Shareholder Value,” *Journal of Marketing*, 68, 172–85.

Ansari, A., K. Jedidi and S. Jagpal (2000), “A Hierarchical Bayesian Methodology for Treating Heterogeneity in Structural Equation Models,” Marketing Science, 19:328–347

Fornell, C. and J. Cha (1994), “Partial Least Squares,” in Richard P. Bagozzi, ed., *Advanced Methods of Marketing Research,* Blackwell Publishers, Oxford: England.

Fornell, Claes, Sunil Mithas, Forrest Morgeson, and M.S. Krishnan (2006), “Customer Satisfaction and Stock Prices: High Returns, Low Risk,” *Journal of Marketing*, 70, 3–14.

Fornell, Claes, and Donald T. Rust and Marnik G. Dekimpe (2010), “Customer Satisfaction and Consumer Spending Growth,” Journal of Marketing Research, XLVII, 28-35.

Homburg, Christian, Nicole Koschate, and Wayne D. Hoyer (2005), “Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay,” *Journal of Marketing*, 69, 84–96.

Keiningham, Timothy L., Tiffany Perkins-Munn, and Heather Evans (2003), “The Impact of Customer Satisfaction on Share-of-Wallet in a Business-to-Business Environment,” Journal of Service Research, 6, 37–50.

Lee, S.Y. (2007), *Structural Equation Modeling: A Bayesian Approach*, Wiley, New York

Mittal, Vikas, Eugene W. Anderson, Akin Sayrak, and Pandu Tadikamalla (2005), “Dual Emphasis and the Long-Term Financial Impact of Customer Satisfaction,” *Marketing Science,* 24, 544–58.

National Quality Research Center (NQRC) (2005), “American customer satisfaction index (ACSI): methodology report,” Stephen M. Ross School of Business at the University of Michigan.

Oliver, Richard L. (1997), Satisfaction: A Behavioral Perspective on the Consumer. Boston: Richard D. Irwin/McGraw-Hill.

Rust, Ronald T. and Anthony J. Zahorik (1993), “Customer Satisfaction, Customer Retention, and Market Share,” *Journal of Retailing*, 69, 193–215.

Spiegelhalter, D. J., N.G. Best, B.P. Carlin, and A. van der Linde (2002), “Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society. Series B*, 64:583–639.

Tanner MA, WH. Wong (1987), “The calculation of posterior distributions by data augmentation,” *Journal of the American Statistical Society*, 82:528–50.

Wold, H. (1981), *The fix-point approach to interdependent systems.* Amsterdam: North Holland.

Nobuhiko Terui is a Professor at Graduate School of Economics and Management, Tohoku University, Japan. He has published in journals such as *Marketing Science, Quantitative Marketing and Economics, Journal of Interactive Marketing, Journal of Time Series Analysis*, *International Journal of Forecasting, Econometric Theory*, *Geographical Analysis* and others. His current research interests are in the statistical modeling of consumer behavior, marketing dynamics and related decision problems.

Shohei Hasegawa is a Ph.D. student at Graduate School of Economics and Management, Tohoku University, Japan. He received his master's degree in Management from Tohoku University. His research focuses on dynamic modeling of consumer behavior.

Taemyung Chun is a Ph.D. student at graduate school of economics and management, Tohoku University, Japan. He completed his master degree in 2009 at Tohoku University. His research interests are customer satisfaction and customer behavior.

Kosuke Ogawa is Dean and Professor of Marketing at Hosei Business School of Innovation Management, Hosei University, Japan. He has published the article in *Marketing Science* and more than one-hundred academic papers mainly regarding marketing research, marketing data analysis, and brand marketing.