DSSR

Discussion Paper No.100

Threshold Measurement Model for Perceived Service Quality

Toshikuni Sato

August, 2019

Data Science and Service Research
Discussion Paper

Center for Data Science and Service Research
Graduate School of Economic and Management
Tohoku University
27-1 Kawauchi, Aobaku
Sendai 980-8576, JAPAN

Threshold Measurement Model for Perceived Service Quality

Toshikuni Sato*

Graduate School of Economics and Management, Tohoku University

Abstract

This study proposes a threshold measurement model based on the prospect theory and zone of tolerance for the SERVQUAL scale to measure the latent perceived service quality. The concept of zone of tolerance is where customers are willing to accept a service discrepancy within a standard they recognize. The discussion focuses on three stages of consumers' mental state and how they relate to observable perceived service quality. It then proposes a model that employs a threshold specification representing extent limit as a zone of tolerance. Because the value function in prospect theory describes human perception's dependence on the evaluation of differences, rather than absolute magnitudes, the proposed mode also integrates asymmetric and nonlinear properties. Empirical analysis was implemented using the data collected from several different service sectors, and the proposal model showed better performance as against other competitive models. The results provide an insight into the asymmetric and nonlinear latent structures of consumers' perceived service quality. Clustering was conducted by applying estimated thresholds and factor scores to obtain three different kinds of consumer segments.

Keywords: Nonlinear measurement model, Nonlinear factor analysis, Measuring perceived service quality, Zone of tolerance, Prospect theory

1. Introduction

The method of measuring service quality is an essential topic in management because the perceived service quality influences customer satisfaction in consumer behaviors (Cronin et al. 2000). If managers or marketers can obtain quality understanding of their consumers' perceived service quality, then the company can compare their position with its competitors. In the field of marketing, the SERVQUAL model (Parasuraman et al. 1985, 1988) is a primary method that utilizes measurement scale to measure the perceived service quality. Although there are many theoretical and statistical issues in the SERVQUAL scale, it fundamentally contributes to the existing service quality models (e.g., Cronin & Taylor 1992, Rust & Oliver 1994, Brady & Cronin 2001, Brady et al. 2002, Kang & James 2004).

The perceived service quality in SERVQUAL is defined as a discrepancy between expectations and perceived performances; therefore, the measurement scale of SERVQUAL is called the "difference score". Utilizing the difference score to measure perceived service quality is one of the issues in SERVQUAL, and the discussions have been ongoing for quite some time (e.g., Cronin & Taylor 1992; 1994, Parasuraman et al. 1994a; 1994b, Brady et al. 2002, Carrillat et al. 2007). Nearly all previous discussions regarding the issue of difference score have implicitly assumed linearity when observing the perceived service quality. Hence, the SERVQUAL model and other models are also defined within the linear measurement model based on the simple Classical Test Theory (CTT) (Traub 1997; Novick 1966; Lewis 2007).

In contrast, the prospect theory represents human judgments and perceptions as attuned to the evaluation of changes or differences, rather than absolute magnitudes (Kahneman & Tversky 1979, p.277), and defines value function with nonlinearity and asymmetry properties (Kahneman & Tversky 1979, p.279). Moreover, consumers may approve of the differences because they have a "zone of tolerance," defined as the extent to which customers recognize and are willing to accept the service discrepancies (Zeithaml et al. 1993, p.6). Because previous studies have not sufficiently discussed the relationship between these topics and the measurement models, it is necessary to address in the nonlinear mental process for perceived service quality evaluations.

In this paper, we discuss the functional form of the measurement model for observable perceived service quality, and reconsider the practical applications of the SERVQUAL model with difference score. Section 2 summarizes related literatures; section 3 introduces some extended SEVQUAL models. Section 4 presents the empirical results using the data from several service industries. Finally, the results from proposed models and future scope are discussed in section 5.

2. Related Literature

2.1. SERVQUAL Model

Service quality has different characteristics when compared with the quality of goods. The three basic characteristics of service quality are: intangibility, heterogeneity, and inseparability (Parasuraman et al. 1985, p.42; 1988, p.13). These characteristics make it difficult to measure service quality, thus inspiring many researchers to conceptualize a plethora of service quality models (e.g., Wolfinbarger & Gilly 2003; Parasuraman et al. 2005; Kang 2006; Lin & Hsieh 2011; Orel & Kara 2014; Blut 2016). The SERVQUAL method, which was developed in line with the expectation disconfirmation theory (Oliver 1980), is the first attempt to overcome these difficulties; Martínez & Martínez (2010) summarize the other representative service quality models (see also Grönroos 1984, Cronin & Taylor 1992, McDougall & Levesque 1995, Rust & Oliver 1994, Dabholkar et al. 1996, Brady & Cronin 2001; Kang & James 2004).

The SERVQUAL scale constitutes 22 questionnaires for each expectation and actual perception. Difference score is then calculated by subtracting the expectation score from the perception score. The SERVQUAL model identified as a factor analysis model with five dimensions (Figure 1). Although Parasuraman et al. (1993, 1994a, 1994b) confirm the validity of the SERVQUAL scale and model, the issues in this method have been widely discussed among many researchers (e.g., Babakus & Boller 1992; Cronin & Taylor 1992, 1994; Brown et al. 1993; Peter et al. 1993; Carman 1990; Prakash 1984). This paper briefly divides these issues into two parts, and suggests additional problems.

Figure 1: SERVQUAL model

2.2. Issues in SERVQUAL

The first issue is the measurement of service expectations. Based on the expectation disconfirmation theory and the assumption that consumers evaluate service quality depends on their subjectivity, Parasuraman et al. (1986) define perceived service quality as being result of a comparison between consumer expectation and the actual service performance. However, the difference score makes it difficult to specify the dissimilarities between service quality and satisfaction (Cronin & Taylor 1992), and results in a reduction of the reliability coefficient (Prakash 1984, Peter et al. 1993).

The second issue is the instability of dimensions. Although a factor analysis model requires original dimensions when the measurement scales being used, the SERVQUAL model with the difference score often provides different dimensions from the original five (Babakus & Boller 1992; Cronin & Taylor 1992,1994). This issue implies that the construct validity, such as the convergent and discriminant validity of SERVQUAL, is not sufficient. Therefore, Cronin and Taylor (1992; 1994) recommend a performance-only measurement, i.e., SERVQUAL scale without the expectation score, because the SERVPERF model is specified by a one-factor model with this measurement and reports better results when compared with the difference score (Cronin & Taylor 1992; 1994).

2.3. Nonlinearity and Zone of Tolerance

Few researchers discuss the nonlinear and asymmetric properties of perceived service quality. Based on the context of prospect theory, Mittal et al. (1998) examine the nonlinear effects of attribute-level performance on the overall satisfaction for services and products. They mention the possibility that the relationship between SERVQUAL dimensions and the overall quality is nonlinear and asymmetric (Mittal et al. 1998, p.34). Sivakumar et al. (2014) discuss the theoretical application of the prospect theory regarding the perceived service quality with expectations. They define service failure and delight as, service performances that fall below expectations and exceed expectations, respectively (Sivakumar et al. 2014, p.41). This is in line with expectation disconfirmation theory proposed by Parasuraman et al. (1985, p.48; 1988).

According to the prospect theory, the mental process of service failure and delight communicate the value function of the observable perceived service quality. The function is defined as concave for service delights and convex for service failures; the impact of service failures is more than that of service delights (Sivakumar et al. 2014; Kahaneman & Tversky 1979). Moreover, Zeithaml et al. (1993) and Parasuraman et al (1993) discuss the zone of tolerance. It is defined as a mental space between the adequate and desired service, which is the standard of services that the customer will accept and hopes to receive. The zone of tolerance indicates that a consumer’s mental space of perceived service quality has thresholds where they are wiling to accept the discrepancy. Although Teas (1993) proposed a modified SERVQUAL scale, comprising the measurement of ideal points corresponding to the thresholds, it would also be impactful to consider specifying the zone of tolerance as a model of measurement.

In spite of these two important mental properties, the original SERVQUAL model has been misspecified by linear measurement model; hence, the nonlinear measurement model should be investigated. The next section focuses on the nonlinearity and threshold for perceived service quality, and discusses the marketing applications of the SERVQUAL model with difference scores. A few nonlinear SERVQUAL models with threshold specifications based on the prospect theory and zone of tolerance are also proposed. Results of the empirical analysis provides an insight on the performance of proposed models and practical applications.

3. Model Development

3.1. Basic Concepts

According to the CTT, the linear measurement model with a construct is defined as the following equation:

$$z_{ji} = a_j t_i + \varepsilon_{ji}, \tag{1}$$

where i is the number of individuals, j is the number of items, z_{ji} is the observed score, t_i is the true score, ε_{ji} is the measurement error, and a_j is the item discrimination that indicates the effectiveness of the construct to the jth item. The true score replaces the latent variable, and a linear factor analysis is adapted to estimate this model. In contrast, this paper considers the latent nonlinear mental process between observed and latent variables as follows:

$$z_{ji} = f(t_i) + \varepsilon_{ji}. \tag{2}$$

To introduce the properties of prospect theory and zone of tolerance to the SERVQAL measurement model, the observed perceived service quality is through a nonlinear and asymmetric process when the latent discrepancy crosses the thresholds. Three types of difference scores are subsequently observed as perceived service qualities based on the value function with thresholds. The three types of difference score are as follows:

i. A positive difference score is observed when the latent positive discrepancies (service delights) cross over the positive threshold.
ii. A negative difference score is observed when the latent negative discrepancies (service failures) cross over the negative threshold.
iii. A difference score of 0 is observed when a consumer does not recognize the discrepancies or the latent discrepancies within the thresholds.

The proposed model uses the second-order SERVQUAL model (Figure 2) to express the aforementioned assumptions, and modifies this model based on a nonlinear factor analysis model (e.g., Zhu & Lee 1999).

Figure 2: Second order factor model for proposed model

3.2. Base Model for Proposed Model

The base model for second-order SERVQUAL (for $i = 1, \ldots, n$) is defined as

$$\mathbf{y}_i = \mathbf{\Lambda}\boldsymbol{\omega}_i + \boldsymbol{\varepsilon}_i, \tag{3}$$

$$\boldsymbol{\omega}_i = \mathbf{G}(\zeta_i) + \boldsymbol{\tau}_i, \tag{4}$$

where $\mathbf{y}_i = \{y_{i,1}, \cdots, y_{i,22}\}^T$ is a (22×1) random vector of observed variables for difference scores to measure "Tangibles $(j = 1, \cdots, 4)$," "Reliability $(j = 5, \cdots, 9)$," "Responsiveness $(j = 10, \cdots, 13)$," "Assurance $(j = 14, \cdots, 17)$," and "Empathy $(j = 18, \cdots, 22)$," $\mathbf{\Lambda}$ is a (22×5) factor loading matrix, $\boldsymbol{\omega}_i = \{\omega_{i,1}, \cdots, \omega_{i,5}\}^T$ is a (5×1) random vector of first-order latent variables corresponding to "Tangible $(k = 1)$," "Reliability $(k = 2)$," "Responsiveness $(k = 3)$," "Assurance $(k = 4)$," and "Empathy $(k = 5)$," $\boldsymbol{\varepsilon}_i$ is a random vector of error measurements assumed as $\boldsymbol{\varepsilon}_i \sim i.i.d.N(\mathbf{0}, \boldsymbol{\Psi}_\epsilon), \boldsymbol{\Psi}_\epsilon = diag\{\psi_{\epsilon,1}, \cdots, \psi_{\epsilon,22}\}$. For the second measurement equation, $\mathbf{G}(\ )$ is a function proposed in the next sections, and $\boldsymbol{\zeta}_i$ is a (1×1) second-order latent variable defined as baseline quality that indicates a latent common discrepancy and assumed as $\zeta_i \sim i.i.d.N(0, \sigma_\zeta^2)$, $\boldsymbol{\tau}_i$ is a random vector of error measurements assumed as $\boldsymbol{\tau}_i \sim i.i.d.N(\mathbf{0}, \boldsymbol{\Delta}_\tau), \boldsymbol{\Delta}_\tau = diag\{\delta_{\tau,1}, \cdots, \delta_{\tau,5}\}$, and $\boldsymbol{\epsilon} \perp \mathbf{Z} \perp \boldsymbol{\tau}$.

In this model, the first corresponding elements of $\boldsymbol{\Lambda}$ between observed variable and latent factor is fixed by 1, and the other corresponding and remaining elements of $\boldsymbol{\Lambda}$ are free parameters and the reaming elements of $\boldsymbol{\Lambda}$ are fixed by 0, respectively. The linear model for second-order equation is defined as $\mathbf{G}(\boldsymbol{\zeta}_i) = \boldsymbol{\Gamma}\boldsymbol{\zeta}_i$, where $\boldsymbol{\Gamma}$ is a (5×1) matrix of factor loadings. The above-stated model can be expressed as

$$\mathbf{y}_i = \boldsymbol{\Lambda}\{\mathbf{G}(\zeta_i) + \boldsymbol{\tau}_i\} + \boldsymbol{\varepsilon}_i = \boldsymbol{\Lambda}\mathbf{G}(\zeta_i) + \boldsymbol{\Lambda}\boldsymbol{\tau}_i + \boldsymbol{\varepsilon}_i. \tag{5}$$

Equation (5) explains that the proposed model is modified by adding a nonlinear term instead of assuming the factor correlations in the original linear SERVQUAL model. The next section proposes a few assumptions for $\mathbf{G}$.

3.3. Proposed Model

To express the zone of tolerance, let η^+ and η^- be a positive and negative threshold parameter, respectively. The threshold logistic model (TLGM) is defined as

$$\begin{aligned}\mathbf{G}(\zeta_i; \boldsymbol{\Gamma}^+, \boldsymbol{\Gamma}^-, \eta^+, \eta^-) = \boldsymbol{\Gamma}^+ I(\zeta_i - \eta^+ \geq 0)\left\{\frac{1}{1+\exp(-\zeta_i + \eta^+)} - \frac{1}{2}\right\} \\ + \boldsymbol{\Gamma}^- I(\zeta_i - \eta^- < 0)\left\{\frac{1}{1+\exp(-\zeta_i + \eta^-)} - \frac{1}{2}\right\},\end{aligned} \tag{6}$$

where $\boldsymbol{I}$ is the indicator function taking 1 if a condition in ( ) is satisfied. $\boldsymbol{\Gamma}^+ = \{\gamma_{1,k}{}^+, \cdots, \gamma_{1,5}{}^+\}^T$ and $\boldsymbol{\Gamma}^- = \{\gamma_{2,k}{}^-, \cdots, \gamma_{2,5}{}^-\}^T$ are assumed to be service delight and failure parameters, respectively. This model is specified by a logistic function because it uses one of the "S"-shaped curves as a value function, where $\boldsymbol{\Gamma}^-$ is expected to be larger than $\boldsymbol{\Gamma}^+$. The estimates for η^+ and η^-, which correspond to a lower and upper limits for zone of tolerance, represent a level for adequate and desired services, respectively.

In addition, the other two threshold models and three asymmetric models are considered to investigate a better functional form. The threshold linear model (TLM) and threshold quadratic model (TQM) are defined as

$$\begin{aligned}\mathbf{G}(\zeta_i; \boldsymbol{\Gamma}^+, \boldsymbol{\Gamma}^-, \eta^+, \eta^-) = \boldsymbol{\Gamma}^+ I(\zeta_i - \eta^+ \geq 0)\{\zeta_i - \eta^+\} \\ + \boldsymbol{\Gamma}^- I(\zeta_i - \eta^- < 0)\{\zeta_i - \eta^-\},\end{aligned} \tag{7}$$

$$\begin{aligned}\mathbf{G}(\zeta_i; \boldsymbol{\Gamma}^+, \boldsymbol{\Gamma}^-, \eta^+, \eta^-) = \boldsymbol{\Gamma}^+ I(\zeta_i - \eta^+ \geq 0)\left\{(\zeta_i - \eta^+)^2\right\} \\ + \boldsymbol{\Gamma}^- I(\zeta_i - \eta^- < 0)\left\{-(\zeta_i - \eta^-)^2\right\}.\end{aligned} \tag{8}$$

The asymmetric linear model (ALM), asymmetric quadratic model (AQM), and asymmetric logistic model (ALGM) are defined as

$$\mathbf{G}\left(\zeta_i ;\boldsymbol{\Gamma}^+ ,\boldsymbol{\Gamma}^- \right)=\boldsymbol{\Gamma}^+ I(\zeta_i \geq 0)\{\zeta_i\}+\boldsymbol{\Gamma}^- I(\zeta_i <0)\{\zeta_i\}, \tag{9}$$

$$\mathbf{G}\left(\zeta_i ;\boldsymbol{\Gamma}^+ ,\boldsymbol{\Gamma}^- \right)=\boldsymbol{\Gamma}^+ I(\zeta_i \geq 0)\{\zeta_i^2\}+\boldsymbol{\Gamma}^- I(\zeta_i <0)\{-\zeta_i^2\}, \tag{10}$$

$$\begin{aligned}\mathbf{G}\left(\zeta_i ;\boldsymbol{\Gamma}^+ ,\boldsymbol{\Gamma}^- \right)=&\boldsymbol{\Gamma}^+ I(\zeta_i \geq 0)\left\{\frac{1}{1+\exp(-\zeta_i)}-\frac{1}{2}\right\}\\&+\boldsymbol{\Gamma}^- I(\zeta_i <0)\left\{\frac{1}{1+\exp(-\zeta_i)}-\frac{1}{2}\right\}.\end{aligned} \tag{11}$$

Table 1 summarizes all of the prosed models for model comparison, and Figure 3 shows each function described by the threshold and asymmetric models.

Table 1: Summary of the proposed models
Figure 3: Proposed functions

4. Empirical Applications

4.1. Data Description

The data were gathered through a research company from two types of hotels, banks, and retail stores in Japan. The questionnaires were referred to Parasuraman et al. (1988; 1991; 1994b), and a total of 300 respondents were gathered in each service industry. Hotel B is a business hotel offering select services in low prices. Hotel A is a city hotel with some restaurants and shops located near a large station. Bank B is a local bank focusing on local customers and companies. Bank A is a megabank providing diverse services in domestic and overseas market. Retail B is a supermarket primarily selling commodities and food. Retail A is a department store with several specialty shops.

4.2. Model Estimation

The proposed models used the Bayesian estimation via Markov chain Monte Carlo (MCMC) algorithm with Gibbs sampling for estimation. In Gibbs sampling, conditional distributions for each random variable are considered; therefore, the algorithm of nonlinear factor analysis model is almost the same as that of the linear factor analysis model (Zhu & Lee 1999, Lee 2007), which is a major advantage of the Bayesian approach with the Gibbs sampler. However, simulating from $p(\boldsymbol{\zeta}_i| -)$, $p(\eta^+| -)$, and $p(\eta^-| -)$, which are nonstandard and complex, is not an easy task. Hence, to simulate from these distributions, the random walk Metropolis-Hastings (RW-MH) method is used within the MCMC algorithm.

4.3. Model Comparison

Tables 2 and 3 report that the model fits for each model were evaluated using the widely applicable information criterion (WAIC) (Watanabe 2010a, 2010b, Gelman et al. 2013) and the widely applicable Bayesian information criterion (WBIC) (Watanabe 2013). These indexes represent an information criterion for model selection in terms of prediction and the logarithm of Bayes marginal likelihood, respectively. The smaller WAIC indicates a more accurate model. The WBIC is interpreted as a minus logarithm of Bayes marginal likelihood (Watanabe 2013), so that the smaller WBIC also suggests a better model fitting.

Although the TLM is supported in Bank A and Retail B by the WAIC, the TLGM displays better WBIC for all service industries. This result also indicates that consumers' perceived service qualities are driven by latent nonlinear structure along with the thresholds.

Table 2: WAIC
Table 3: WBIC

4.4. Estimation Results

The Appendix provides the estimates of parameters in the TLGM, and Tables 4 and 5 show the estimated asymmetric and threshold parameters. Figure 4 describes the estimated function and distribution of unique factors for each SERVQUAL dimension.

Table 4: Estimated threshold parameters
Table 5: Estimated asymmetric parameters
Figure 4: Estimated function and uniqueness

Table 4 provides estimates for threshold parameters and the ranges reveal the latent zone of tolerance. This result indicates that consumers evaluate perceived service quality with an acceptable discrepancy between expectations and perceptions. The larger the positive threshold, the more difficult it is for consumers to experience service delight. In contrast, at a smaller negative threshold, consumers find it easier to tolerate service failure. When comparing the absolute value of each estimated threshold parameter in Table 4, the negative threshold in Hotel A, Bank B, Retail B, and Retail A are estimated to be larger than the positive threshold. They ,therefore, obtained better results, while both thresholds mostly displayed similar estimates in Hotel B and Bank A, respectively (see also Figure 4). The absolute value of estimates in Bank A is the largest; thus, indicating that customers might accept discrepancies more easily in Bank A. Hotel B is required to pay more attention to service failures because of the smallest absolute value of the negative threshold.

"Delight" and "Failure" in Table 5 indicate the estimates for delight and failure parameters, and the standardized coefficients are shown in std.D and std.F. According to the 95% highest probability density interval (HPDI), all estimates are not 0. P{D < F} shows that of the two parameters, failure is greater than delight. Although some failure parameters are smaller than delight parameters in Hotel B, Bank A, and Retail B, whereas all failure parameters are larger than delight parameters in Hotel A, Bank B, and Retail A, which is parallel to the assumptions of value function. These results indicate that delight and failure parameters primarily follow the prospect theory, and that service failure, which is a negative discrepancy, has significantly more influence on the observed perceived service quality than service delight. Therefore, consumers' evaluation process of perceived service quality has an asymmetric structure.

In Table 5, std.U indicates standardized estimated variances of each uniqueness factor for SEVQUAL dimensions that show the dependent efficacy of each SERVQUAL dimension. The precisions of the five dimensions' qualities are presumed to be unequal because corresponding distributions look different. The distribution becomes flat if the factor's uniqueness has a larger effect, whereas it becomes shaper with a smaller effect (see Figure 4). Smaller uniqueness indicates that the sub-dimension depends on the higher dimension (common factor) rather than on the uniqueness itself. On the contrary, larger uniqueness indicates that the sub-dimension has some unique features in comparison to other sub-dimensions. For example, estimates for Tangible in Hotel A (see Table 5) indicates that the Tangible factor is almost independent from the other factors, and has a larger effect than the baseline quality, so that it possesses larger uniqueness than the other factors.

4.5. Segmentation for the Customer by Threshold Parameters

Figure 5 illustrates the proportion of segments for the customers in each service industry divided by the threshold parameters. P.PSQ, or the top portion of the bar plots, describes the proportion of customers whose baseline quality (second-order factor) score exceeds the positive threshold, indicating that the customers perceived positive service quality. S.PSQ, in the middle of the bar plots, indicates the class of customers whose baseline quality score is inside both positive and negative threshold parameters, whereas N.PSQ, at the bottom of the bar plots, indicates the class of customers whose baseline quality score is less than the negative threshold parameter. These plots enable the comparison of potential perceived service quality for each service industry that does not meet customer expectations.

For example, over 30 % of customers in Hotel A perceived that services exceeded expectations. Bank A achieved better service perception than other service industries; however, almost all customers might evaluate that the service is neither good nor bad because of highly proportion of S.PSQ. In Hotel B, the each segment is divided as almost equally, and the proportion of customers who perceived negative service quality is the largest among these industries, which suggests that it may be useful to improve their services.

5. Implications and Conclusions

Three possible implications from the proposed model are investigated and future research is discussed in this study.

First, the common nonlinear effects and independent linear effects of each SERVQUAL dimension are estimated using the second-order factor analysis with nonlinear structure. In addition, a nonlinear structure for customers' perceived service quality is established by comparing several nonlinear measurement models. Second, a comparison of different magnitudes of effects between service delights and failures is possible by estimating the asymmetric parameters. Third, considering the threshold parameter in the measurement model, it is possible to estimate consumers' zone of tolerance. Moreover, the properties of the proposed model can be visualized by constructing a plot, as shown in Figure 4. The threshold parameters are also helpful in classifying the customers into three categories as in Figure 5.

In this study, the nonlinear and asymmetric measurement model with threshold is established to measure the perceived service quality. Finally, the threshold logistic model is specified, and demonstrates better results when compared with the original SERVQUAL model and the other candidate models. Moreover, using the difference score enables a proper interpretation of the threshold logistic model because both the prospect theory and zone of tolerance assume the evaluation with some reference point, such as expectation. Additional work is warranted to develop a nonparametric measurement model to explore and estimate the functional form directly. Finally, the construct validation must be extended to confirm the validity for the nonlinear measurement model. Future research can focus on investigating those issues.

Figures and Tables

Figure 1: SERVQUAL model

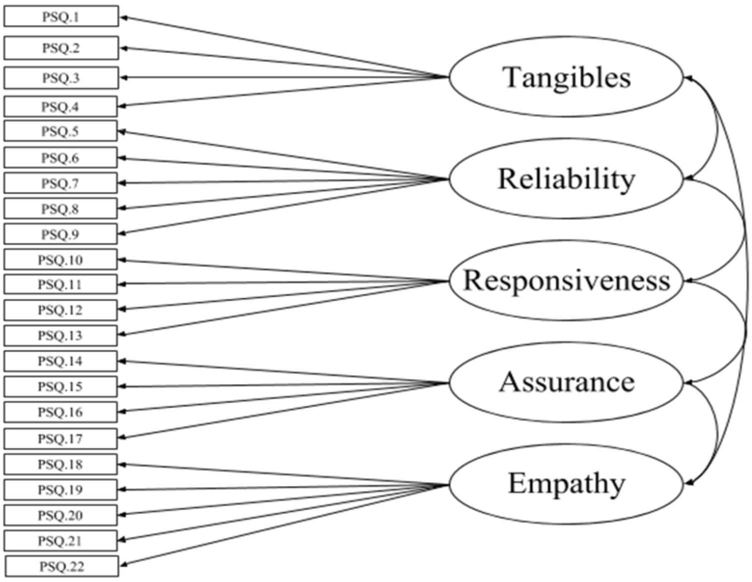
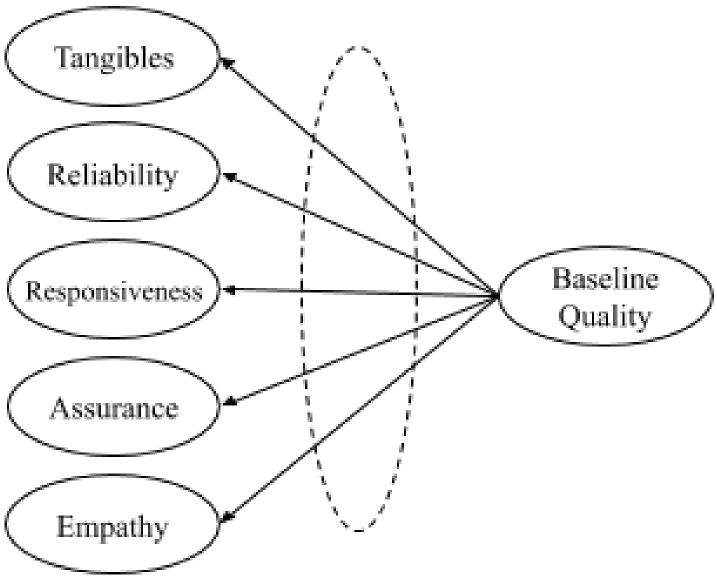


Figure 2: Second order factor model for proposed model



Note: the observed variables and error variables are abbreviated.

Figure 3: Proposed functions

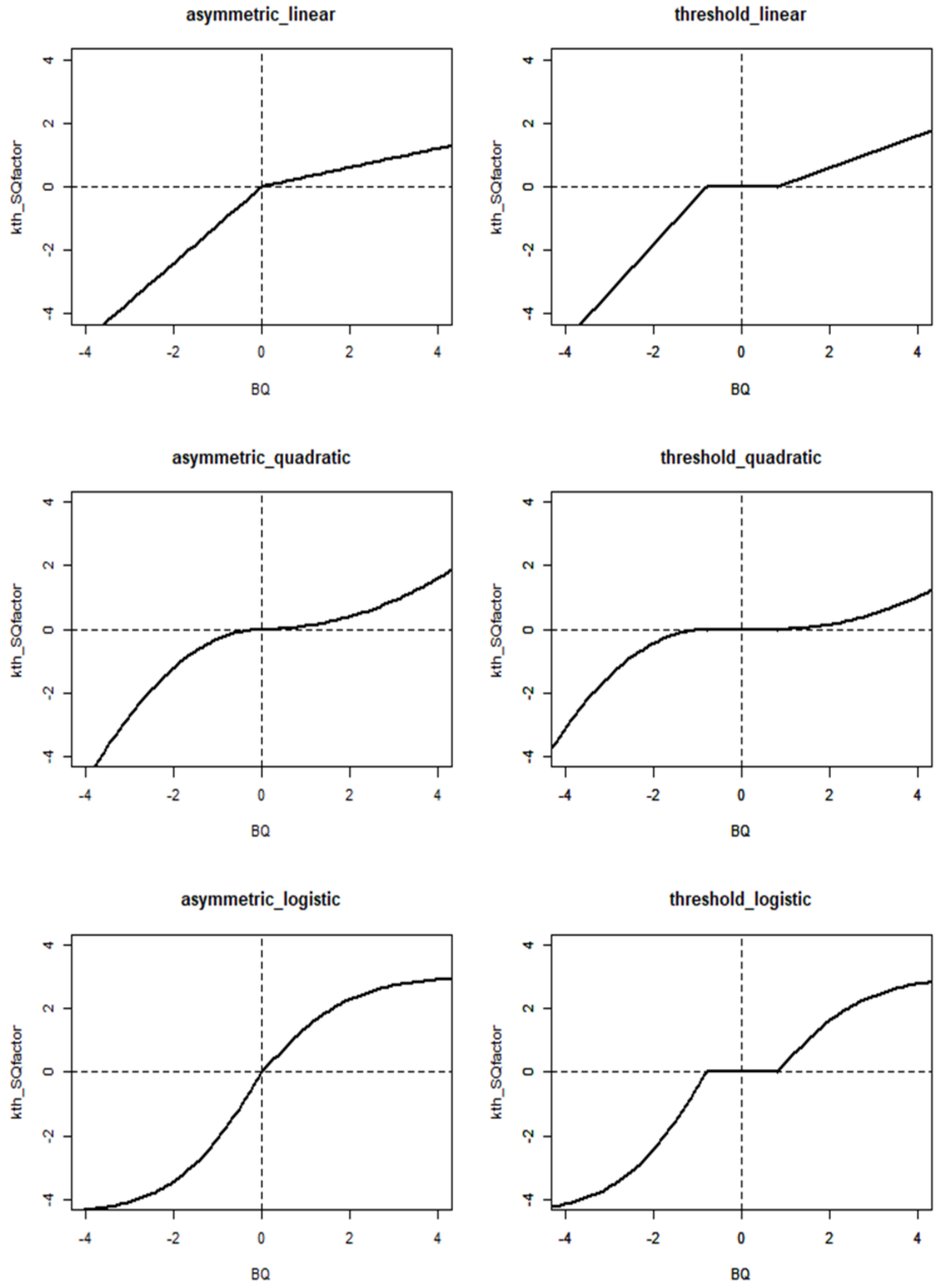
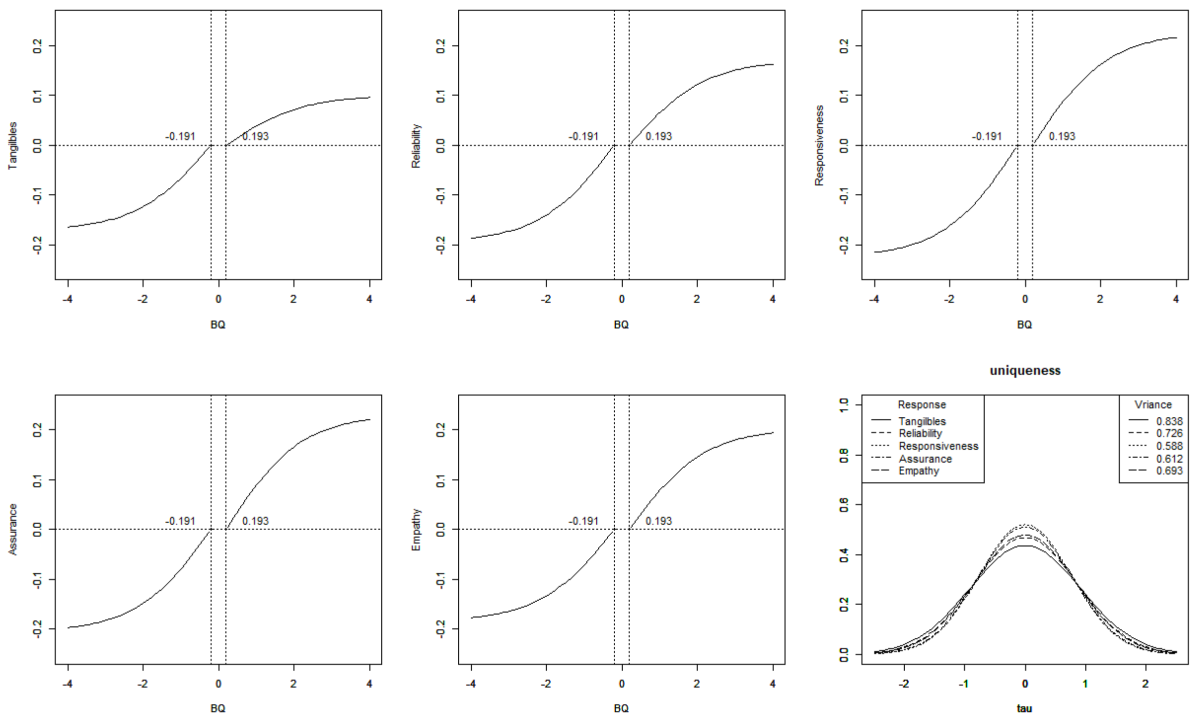
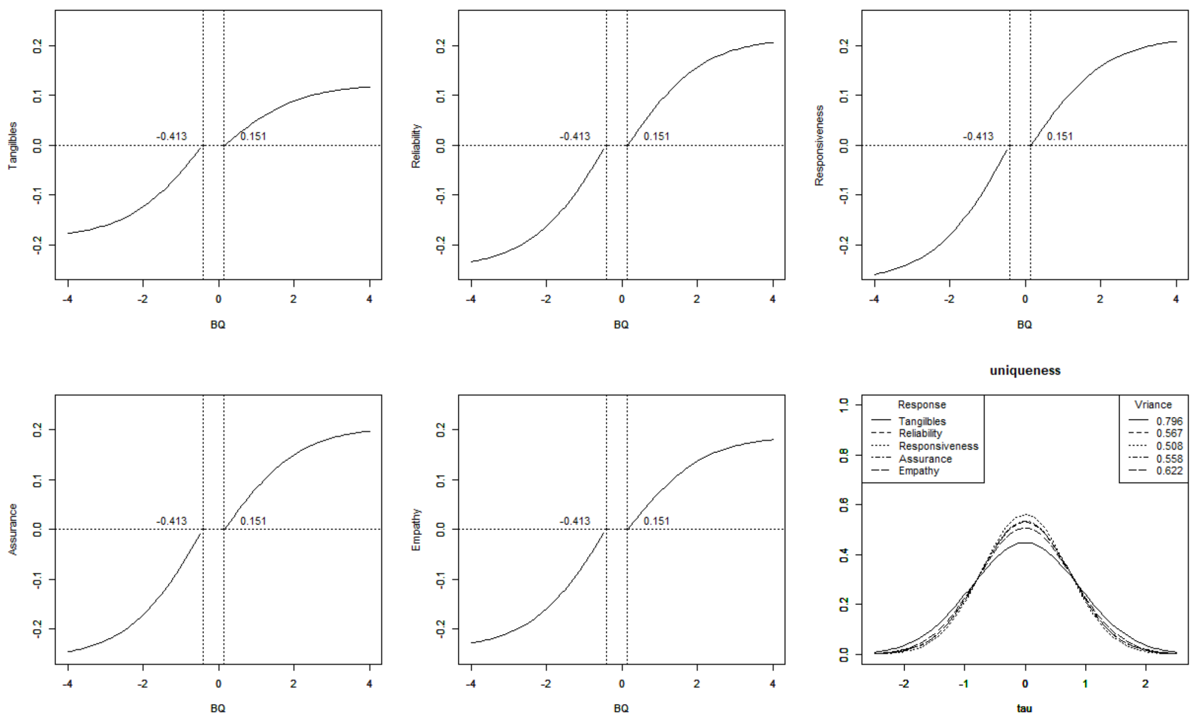


Figure 4: Estimated function and uniqueness

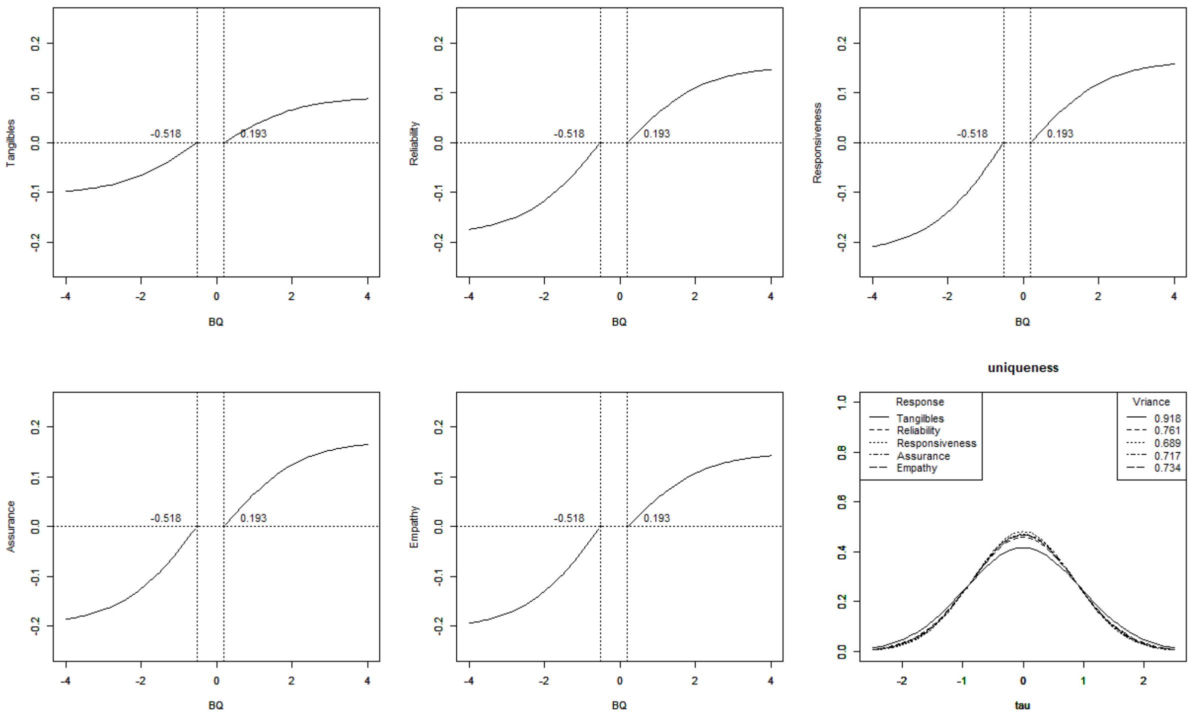
Hotel B (TLGM)



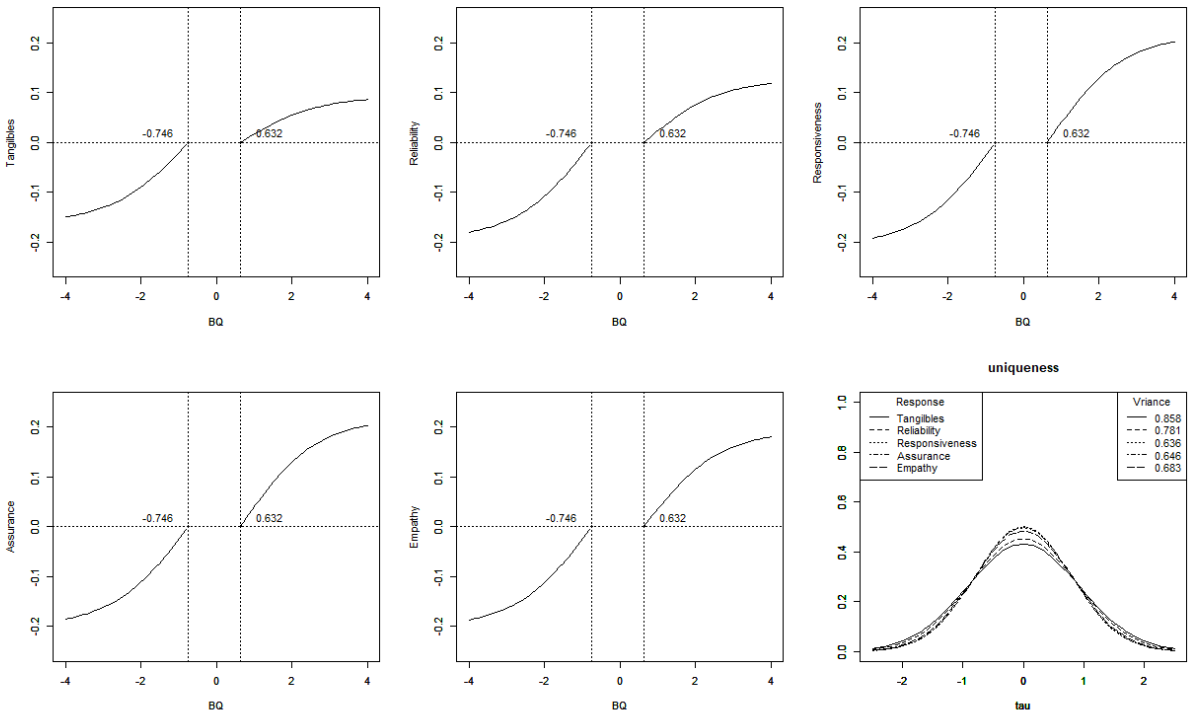
Hotel A (TLGM)



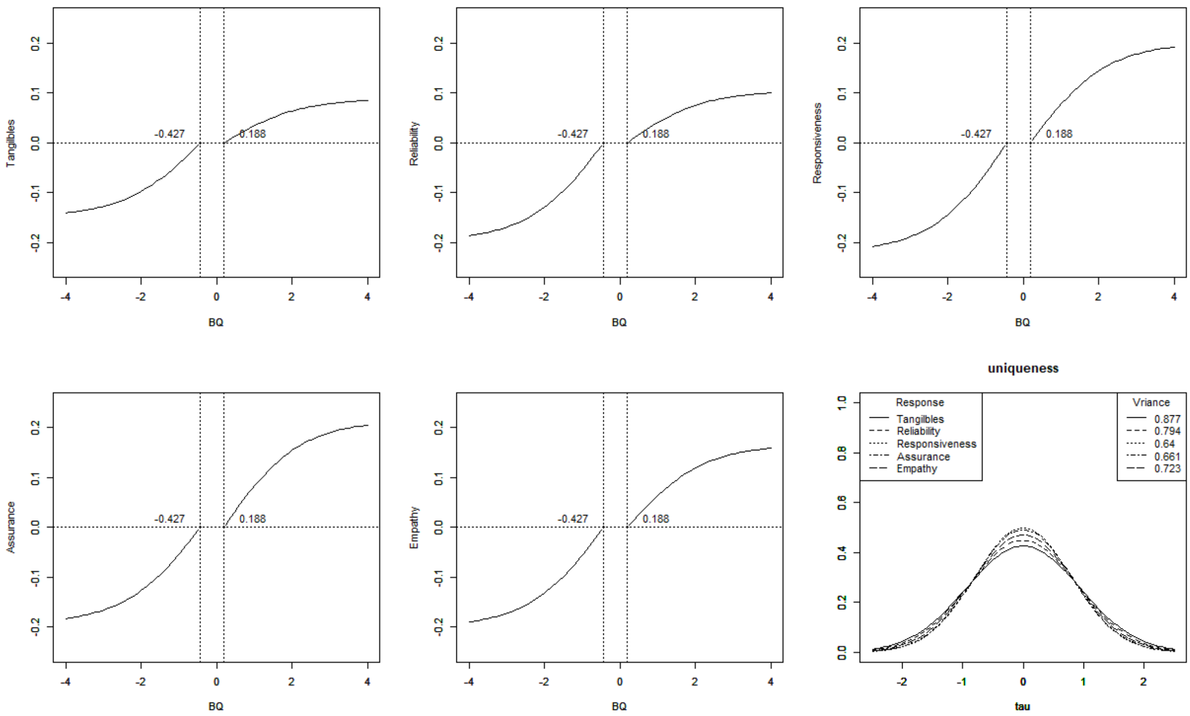
Bank B (TLGM)



Bank A (TLGM)



Retail B (TLGM)



Retail A (TLGM)

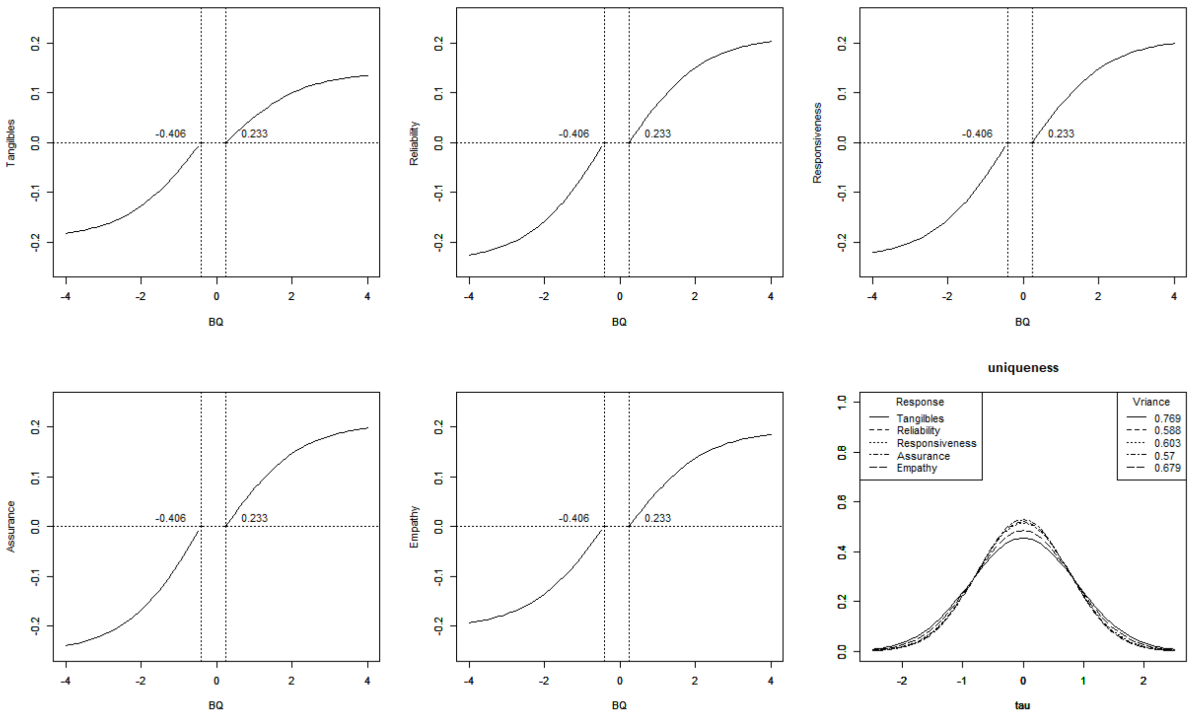


Figure 5: Proportion of segments

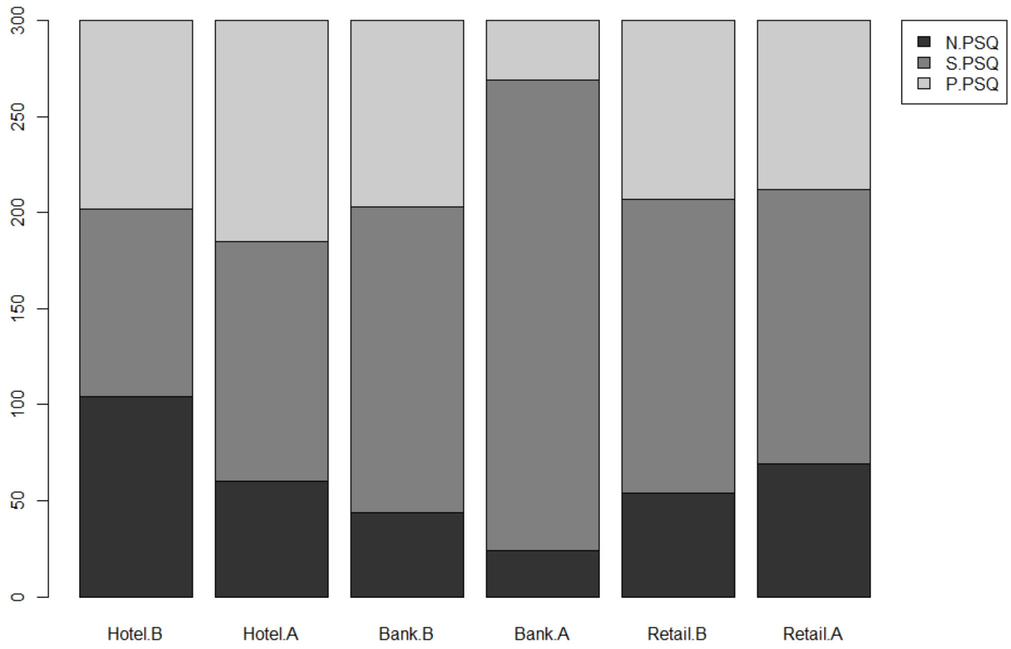


Table 1: Summary of the proposed models

Model	Nonlinearity	Asymmetry	Threshold	Function
1_factor	No	No	No	Linear
Original	No	No	No	Linear
2nd_order	No	No	No	Linear
ALM	No	Yes	No	Linear
AQM	Yes	Yes	No	Quadratic
ALGM	Yes	Yes	No	Logistic
TLM	No	Yes	Yes	Linear
TQM	Yes	Yes	Yes	Quadratic
TLGM	Yes	Yes	Yes	Logistic

Note: 1_factor is the first-order factor analysis model with only one latent variable. Original is the SERVQUAL model proposed by Parasuraman et al. (1988), and 2nd_order is the linear second-order factor analysis model.

Table 2: WAIC

	original	1_factor	2nd_factor	ALM	AQM	ALGM	TLM	TQM	TLGM	result
Hotel B	15545.63	16687.35	15528.51	15515.58	15512.13	15513.99	15487.12	15528.36	15456.88	TLGM
Hotel A	14341.73	15511.40	14334.90	14317.28	14322.45	14324.11	14288.40	14321.04	14281.28	TLGM
Bank B	16057.84	17047.89	16041.60	16007.22	15997.80	16012.55	15971.49	16003.18	15968.51	TLGM
Bank A	14261.95	15428.69	14252.64	14254.08	14227.90	14273.41	14183.03	14227.22	14197.54	TLM
Retail B	15759.58	16983.10	15751.12	15741.73	15717.25	15793.99	15694.44	15721.36	15751.09	TLM
Retail A	14854.60	15787.66	14848.88	14846.96	14825.10	14850.17	14801.11	14834.09	14787.83	TLGM

Table 3: WBIC

	original	1_factor	2nd_factor	ALM	AQM	ALGM	TLM	TQM	TLGM	result
Hotel B	7402.83	8219.37	7420.27	7389.84	7380.46	7348.89	7361.30	7387.29	7331.83	TLGM
Hotel A	6818.23	7666.57	6812.96	6782.97	6799.91	6752.82	6789.55	6811.24	6746.56	TLGM
Bank B	7649.33	8415.56	7656.09	7618.58	7619.35	7605.23	7601.53	7615.12	7568.46	TLGM
Bank A	6756.16	7587.34	6736.31	6745.13	6733.92	6701.56	6717.43	6740.27	6682.58	TLGM
Retail B	7490.73	8380.04	7482.02	7475.15	7485.33	7476.88	7463.93	7478.27	7456.48	TLGM
Retail A	7090.21	7796.33	7069.99	7078.22	7047.49	7016.30	7043.05	7049.40	7007.85	TLGM

Table 4: Estimated threshold parameters

		negative		positive	
Hotel B	[	-0.191	,	0.193	]
Hotel A	[	-0.413	,	0.151	]
Bank B	[	-0.518	,	0.193	]
Bank A	[	-0.746	,	0.632	]
Retail B	[	-0.427	,	0.188	]
Retail A	[	-0.406	,	0.233	]

Table 5: Estimated asymmetric parameters

	Delight	95%HPDI	Failure	95%HPDI	P{D < F}	std.D	std.F	std.U
Hotel B								
Tangilbles←BQ	1.000		1.737	[1.019 , 2.438]	0.984	0.200	0.342	0.838
Reliability←BQ	1.299	[0.790 , 1.878]	1.496	[0.975 , 2.055]	0.709	0.340	0.389	0.726
Responsiveness←BQ	2.146	[1.539 , 2.784]	2.143	[1.568 , 2.781]	0.497	0.453	0.449	0.588
Assurance←BQ	2.122	[1.538 , 2.781]	1.905	[1.325 , 2.469]	0.293	0.462	0.411	0.612
Empathy←BQ	2.003	[1.350 , 2.648]	1.839	[1.260 , 2.553]	0.352	0.405	0.369	0.693
Hotel A								
Tangilbles←BQ	1.000		1.745	[1.112 , 2.424]	0.993	0.245	0.374	0.796
Reliability←BQ	1.692	[1.215 , 2.206]	2.183	[1.631 , 2.726]	0.932	0.431	0.493	0.567
Responsiveness←BQ	1.987	[1.480 , 2.522]	2.818	[2.217 , 3.409]	0.993	0.435	0.547	0.508
Assurance←BQ	2.399	[1.740 , 3.023]	3.397	[2.704 , 4.133]	0.986	0.412	0.517	0.558
Empathy←BQ	1.835	[1.269 , 2.408]	2.648	[2.035 , 3.355]	0.981	0.376	0.481	0.622
Bank B								
Tangilbles←BQ	1.000		1.351	[0.467 , 2.212]	0.779	0.184	0.207	0.918
Reliability←BQ	1.129	[0.542 , 1.685]	1.615	[0.988 , 2.280]	0.888	0.307	0.369	0.761
Responsiveness←BQ	1.302	[0.724 , 1.927]	2.068	[1.401 , 2.798]	0.965	0.330	0.442	0.689
Assurance←BQ	1.680	[1.006 , 2.409]	2.277	[1.557 , 3.072]	0.875	0.346	0.395	0.717
Empathy←BQ	1.366	[0.734 , 2.052]	2.239	[1.470 , 2.979]	0.966	0.298	0.412	0.734
Bank A								
Tangilbles←BQ	1.000		1.882	[1.141 , 2.685]	0.990	0.185	0.322	0.858
Reliability←BQ	1.000		1.656	[1.114 , 2.267]	0.993	0.254	0.388	0.781
Responsiveness←BQ	2.608	[1.889 , 3.328]	2.659	[1.924 , 3.445]	0.544	0.434	0.413	0.636
Assurance←BQ	2.979	[2.114 , 3.774]	2.915	[2.065 , 3.733]	0.447	0.436	0.399	0.646
Empathy←BQ	2.800	[1.890 , 3.604]	3.117	[2.264 , 4.061]	0.699	0.387	0.403	0.683
Retail B								
Tangilbles←BQ	1.000		1.923	[1.058 , 2.733]	0.988	0.175	0.178	0.877
Reliability←BQ	0.953	[0.392 , 1.553]	2.052	[1.292 , 2.876]	0.991	0.201	0.210	0.794
Responsiveness←BQ	1.359	[0.854 , 1.869]	1.694	[1.159 , 2.273]	0.843	0.409	0.402	0.640
Assurance←BQ	1.956	[1.276 , 2.637]	2.010	[1.371 , 2.677]	0.544	0.442	0.429	0.661
Empathy←BQ	1.490	[0.890 , 2.127]	2.057	[1.318 , 2.762]	0.898	0.328	0.332	0.723
Retail A								
Tangilbles←BQ	1.000		1.499	[1.000 , 2.107]	0.970	0.282	0.383	0.769
Reliability←BQ	2.173	[1.530 , 2.779]	2.650	[1.951 , 3.278]	0.858	0.425	0.476	0.588
Responsiveness←BQ	1.892	[1.284 , 2.469]	2.302	[1.628 , 2.920]	0.852	0.417	0.466	0.603
Assurance←BQ	2.428	[1.788 , 3.123]	3.223	[2.483 , 3.987]	0.952	0.414	0.504	0.570
Empathy←BQ	1.866	[1.225 , 2.460]	2.138	[1.527 , 2.803]	0.746	0.387	0.407	0.679

Note 1: std.D, std.F, and std.U are standard coefficients for Delight, Failure, and Uniqueness, respectively.
Note 2: Bank A fixed two parameters to avoid improper solutions, whereas the other industries fixed only one parameter in factor loadings.

Appendix: MCMC Algorithm for Threshold Logistic Model

A.1. Details of Second-order Measurement Equation for Threshold Logistic Model

The second-order measurement equation in base model (4) can be expressed as

$$\begin{aligned}\boldsymbol{\omega}_{\mathbf{i}} &= \mathbf{G}(\boldsymbol{\zeta}_{\mathbf{i}}) + \boldsymbol{\tau}_{\mathbf{i}} \\ &= \boldsymbol{\Gamma}\mathbf{F}(\boldsymbol{\zeta}_{\mathbf{i}}) + \boldsymbol{\tau}_i\end{aligned}, \tag{12}$$

where,

$$\boldsymbol{\Gamma}^T = \begin{bmatrix} \gamma_{1,1}^{+} & \gamma_{2,1}^{+} & \gamma_{3,1}^{+} & \gamma_{4,1}^{+} & \gamma_{5,1}^{+} \\ \gamma_{1,2}^{-} & \gamma_{2,2}^{-} & \gamma_{3,2}^{-} & \gamma_{4,2}^{-} & \gamma_{5,2}^{-} \end{bmatrix}, \tag{13}$$

$$\mathbf{F}(\boldsymbol{\zeta}_{\mathbf{i}}) = \left\{ f^{+}\left(\zeta_i - \eta^{+}\right), f^{-}\left(\zeta_i - \eta^{-}\right)\right\}^T, \tag{14}$$

$$f^{+}(x) = I\{x \geq 0\}\left\{\frac{1}{1+\exp(-x)} - \frac{1}{2}\right\}, \tag{15}$$

and

$$f^{-}(x) = I\{x < 0\}\left\{\frac{1}{1+\exp(-x)} - \frac{1}{2}\right\}. \tag{16}$$

A.2. Prior Distribution

Parameter	Settings
$\Lambda_j \mid \psi_{\varepsilon.j} \sim N\left(\Lambda_{j,0}, \psi_{\varepsilon.j} H_{\lambda.j,0}\right)$	$\Lambda_{j,0} = \mathbf{0}$, $H_{\lambda.j,0} = I$
$\psi_{\varepsilon,j} \sim IG\left(\alpha_{\varepsilon,j,0}, \beta_{\varepsilon,j,0}\right)$	$\alpha_{\varepsilon,j,0} = 0.01$, $\beta_{\varepsilon,j,0} = 0.01$
$\Gamma_k \mid \delta_{\tau.k} \sim N\left(\Gamma_{k,0}, \delta_{\tau.k} H_{\gamma,k,0}\right)$	$\Gamma_{k,0} = \mathbf{0}$, $H_{\gamma,k,0} = I$
$\delta_{\tau,k} \sim IG\left(\alpha_{\tau,k,0}, \beta_{\tau,k,0}\right)$	$\alpha_{\tau,k,0} = 0.01$, $\beta_{\tau,k,0} = 0.01$
$\eta^{+} \mid \mathbf{Z} \sim trN_{[0, \max(\mathbf{Z})]}(\eta_0, \nu_0)$	$\eta_0 = 0$, $\nu_0 = 1$
$\eta^{-} \mid \mathbf{Z} \sim trN_{[\min(\mathbf{Z}), 0]}(\eta_0, \nu_0)$	

A.3. Full Conditional Distribution

[1]
$$\begin{aligned}&\left[\boldsymbol{\omega}_i \mid \boldsymbol{\Lambda}, \boldsymbol{\Psi}_{\boldsymbol{\varepsilon}}, \boldsymbol{\Delta}_{\boldsymbol{\tau}}, \boldsymbol{\zeta}_i, \boldsymbol{\Gamma}, \boldsymbol{\eta}^{+}, \boldsymbol{\eta}^{-}, \mathbf{y}_i\right] \\ &\sim N\left[\left(\boldsymbol{\Delta}_{\boldsymbol{\tau}}^{-1} + \boldsymbol{\Lambda}^{\mathbf{T}}\boldsymbol{\Psi}_{\boldsymbol{\varepsilon}}^{-1}\boldsymbol{\Lambda}\right)^{-1}\left(\boldsymbol{\Lambda}^{\mathbf{T}}\boldsymbol{\Psi}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{y}_i + \boldsymbol{\Delta}_{\boldsymbol{\tau}}^{-1}\mathbf{G}_i\right), \left(\boldsymbol{\Delta}_{\boldsymbol{\tau}}^{-1} + \boldsymbol{\Lambda}^{\mathbf{T}}\boldsymbol{\Psi}_{\boldsymbol{\varepsilon}}^{-1}\boldsymbol{\Lambda}\right)^{-1}\right].\end{aligned} \tag{17}$$

[2]
$$\left[\boldsymbol{\Lambda}_j \mid \psi_{\varepsilon,j}, \boldsymbol{\Omega}, \mathbf{Y}\right] \sim N\left(\mathrm{a}_{\lambda,j}, \psi_{\varepsilon,j}\mathrm{A}_{\lambda,j}\right) \tag{18}$$

where, $\mathrm{A}_{\lambda,j} = \left(H_{\lambda,j,0}^{-1} + \boldsymbol{\Omega}\boldsymbol{\Omega}^T\right)^{-1}$ and $\mathrm{a}_{\lambda,j} = \mathrm{A}_{\lambda,j}\left(H_{\lambda,j,0}^{-1}\boldsymbol{\Lambda}_{j.0} + \boldsymbol{\Omega}\mathbf{Y}_j\right)$.

[3]
$$\left[\psi_{\varepsilon,j} \mid \boldsymbol{\Omega}, \mathbf{Y}\right] \sim IG\left(\alpha_{\varepsilon,j,0} + \frac{n}{2}, \beta_{\varepsilon,j}\right). \tag{19}$$

where, $\beta_{\varepsilon,j} = \beta_{\varepsilon,j,0} + \frac{1}{2}\left(\mathbf{Y}_j^T\mathbf{Y}_j - \mathrm{a}_{\lambda,j}^T\mathrm{A}_{\lambda,j}^{-1}\mathrm{a}_{\lambda,j} + \boldsymbol{\Lambda}_{j.0}^T H_{\lambda j,0}^T \boldsymbol{\Lambda}_{j,0}\right)$.

Set $t(=1,\cdots,T)$ as a number of MCMC iterations and the RW-MH algorithm for $p(\zeta_i|\boldsymbol{\Gamma},\boldsymbol{\Delta}_\tau,\sigma_\zeta^2,\boldsymbol{\omega}_i,\eta^+,\eta^-)$, $p(\eta^+|\mathbf{Z},\boldsymbol{\Gamma},\boldsymbol{\Delta}_\tau,\sigma_\zeta^2,\boldsymbol{\Omega},\eta^-)$, and $p(\eta^-|\mathbf{Z},\boldsymbol{\Gamma},\boldsymbol{\Delta}_\tau,\sigma_\zeta^2,\boldsymbol{\Omega},\eta^+)$ are following

[4] $$\zeta_i^t \sim N\left(\zeta_i^{t-1},\sigma_\zeta\right). \tag{20}$$

The probability of accepting is

$$\min\left[\frac{p\left(\zeta_i^t \mid \boldsymbol{\Gamma},\boldsymbol{\Delta}_\tau,\sigma_\zeta^2,\boldsymbol{\omega}_i,\eta^+,\eta^-\right)}{p\left(\zeta_i^{t-1} \mid \boldsymbol{\Gamma},\boldsymbol{\Delta}_\tau,\sigma_\zeta^2,\boldsymbol{\omega}_i,\eta^+,\eta^-\right)},1\right]. \tag{21}$$

[5] $$\eta^{+,t} \sim trN_{[0,\max(\mathbf{Z})]}\left(\eta_k^{+,t-1},\nu_{\eta^+}\right). \tag{22}$$

[6] $$\eta^{-,t} \sim trN_{[\min(\mathbf{Z}),0]}\left(\eta^{-,t-1},\nu_{\eta^-}\right). \tag{23}$$

The probability of accepting is

$$\min\left[\frac{p\left(\eta^{+,t},\eta^{-,t} \mid \boldsymbol{\Gamma},\boldsymbol{\Delta}_\tau,\sigma_\zeta^2,\boldsymbol{\Omega},\mathbf{Z}\right)p\left(\eta^{+,t-1}\middle|\eta^{+,t}\right)p\left(\eta^{-,t-1}\middle|\eta^{-,t}\right)}{p\left(\eta^{+,t-1},\eta^{-,t-1} \mid \boldsymbol{\Gamma},\boldsymbol{\Delta}_\tau,\sigma_\zeta^2,\boldsymbol{\Omega},\mathbf{Z}\right)p\left(\eta^{+,t}\middle|\eta^{+,t-1}\right)p\left(\eta^{-,t}\middle|\eta^{-,t-1}\right)},1\right]. \tag{24}$$

σ_ζ, ν_{η^+}, and ν_{η^-} are step-size parameters which are given so that each acceptance rate becomes approximately 0.25 (Gelman et al. 1995; Zhu & Lee 1999).

[7] $$\left[\boldsymbol{\Gamma}_k \mid \delta_{\tau,k},\mathbf{Z},\boldsymbol{\Omega},\eta^+\right] \sim N\left(\mathrm{a}_{\gamma,k},\delta_{\tau,k}\mathrm{A}_{\gamma,k}\right). \tag{25}$$

where, $\mathrm{A}_{\gamma,k}=\left(H_{\gamma,k,0}^{-1}+\mathbf{F}_{(k,k+1)}\mathbf{F}_{(k,k+1)}^T\right)^{-1}$,

and $\mathrm{a}_{\gamma,k}=\mathrm{A}_{\gamma,k}\left(H_{\gamma,k,0}^{-1}\boldsymbol{\Gamma}_{k.0}+\mathbf{F}_{(k,k+1)}\boldsymbol{\Omega}_k\right)$.

[8] $$\left[\delta_{\tau,\mathrm{k}} \mid \mathbf{Z},\boldsymbol{\Omega},\eta^+\right] \sim IG\left(\alpha_{\tau,\mathrm{k},0}+\frac{n}{2},\beta_{\tau,\mathrm{k}}\right), \tag{26}$$

where, $\beta_{\tau,k}=\beta_{\tau,k,0}+\frac{1}{2}\left(\boldsymbol{\Omega}_k^T\boldsymbol{\Omega}_k-\mathrm{a}_{\tau,k}^T\mathrm{A}_{\gamma,k}^{-1}\mathrm{a}_{\gamma,k}+\boldsymbol{\Gamma}_{k.0}^T H_{\gamma,k,0}^T\boldsymbol{\Gamma}_{k,0}\right)$.

The above results are valid for situations where all elements of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ are free parameters. As an example, consider that $\boldsymbol{\Lambda}_j^T$ and the jth row of $\boldsymbol{\Lambda}$ contain fixed parameters. Let $\boldsymbol{c}_j$ be the corresponding $1\times q$ row vector such that $c_{jk}=0$ if λ_{jk} is a fixed parameter, and $c_{jk}=1$ if λ_{jk} is an unknown parameter. As for $j=1,\cdots,p$ and $k=1,\cdots,q$, let $r_j=c_{j1}+\cdots+c_{jq}$ be the number of unknown parameters in $\boldsymbol{\Lambda}_j^T$, $\boldsymbol{\Lambda}_j^{*T}$ be a $1\times r_j$ row vector that contains the only unknown parameters in $\boldsymbol{\Lambda}_j^T$, and $\boldsymbol{\Omega}_j^*$ be a $r_j\times n$ submatrix of $\boldsymbol{\Omega}$ such that all the rows corresponding to $c_{jk}=0$ are deleted. Let, $\mathbf{Y}_j^{*T}=\left(y_{1,j}^*,\cdots,y_{n,j}^*\right)$ with

$$y_{i,j}^*=y_{i,j}-\sum_{k=1}^{q}\lambda_{j,k}\omega_{k,i}\left(1-c_{jk}\right). \tag{27}$$

Equation (27) subtracts the constant terms from $\mathbf{Y}_j$. Hence, the conditional distributions with $\mathbf{\Lambda}_j$, $\mathbf{Y}_j$, and $\mathbf{\Omega}$ in part of $(\mathbf{\Lambda}, \mathbf{\Psi}_\epsilon)$ must be replaced by $\mathbf{\Lambda}_j^*$, $\mathbf{Y}_j^*$, and $\mathbf{\Omega}_j^*$, respectively. This procedure is also adapted in full conditional distribution for $(\mathbf{\Gamma}, \Delta_\tau)$ because $\gamma_{1,1}^+$ and $\gamma_{1,2}^+$ are fixed by 1 in Bank A, and $\gamma_{1,1}^+$ is fixed by 1 in the other industries to avoid improper solutions. Moreover, σ_ζ^2, the variance of ζ_i, is fixed by 1 to identify this model. The basic and related algorithm is explained in Xing et al. (2016), Song & Lee (2010), Lee (2007), and Zhu & Lee (1999).

References

Babakus, Emin and Gregory W. Boller (1992), "An Empirical Assessment of the SERVQUAL Scale," *Journal of Business Research*, 24(3), 253-268.

Blut, Markus (2016), "E-Service Quality: Development of a Hierarchical Model," *Journal of Retailing*, 92(4), 500-517.

Brady, Michael K. and J. Joseph Cronin, Jr. (2001), "Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach," *Journal of Marketing*, 65(3), 34-49.

Brady, Michael K., J. Joseph Cronin, Jr., and Richard R. Brand, "Performance-only measurement of service quality: a replication and extension," Journal of Business Research, 55(1), 17-31.

Brown, Tom J., Gilbert A. Churchill Jr., and J. Paul Peter (1993), "Research Note: Improving the Measurement of Service Quality," *Journal of Retailing*, 69(1), 127-147.

Carman, James M. (1990), "Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions," *Journal of Retailing*, 66(1), 33-55.

Carrillat, François A., Fernando Jaramillo, and Jay P. Mulki (2007), "The Validity of the SERVQUAL and SERVPERF Scales," *International Journal of Service Industry Management*, 18(5), 472-490.

Cronin, J. Joseph, Jr., Michael K. Brady, and G. Tomas M. Hult (2000), "Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer Behavioral Intentions in Service Environments," *Journal of Retailing*, 76(2), 193-218.

Cronin, J. Joseph, Jr. and Steven A. Taylor (1994), "SERVPERF Versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality," *Journal of Marketing*, 58(1), 125-131.

Cronin, J. Joseph, Jr. and Steven A. Taylor (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56(3), 55-68.

Dabholkar, Pratibha A., Dayle I. Thorpe, and Joseph O. Rentz (1996), "A Measure of Service Quality for Retail Stores: Scale Development and Validation," *Journal of the Academy of Marketing Science*, 24(1), 3-16.

Gelman, Andrew, G. O. Roberts, and W. R. Gilks (1995), "Efficient Metropolis Jumping Rules," In J. M. Bernardo, J. O. Berger, A. P. Dawid & A. F. M. Smith (Eds), *Bayesian Statistics 5*, Oxford: Oxford University Press, 599-607.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin (2013), "*Bayesian Data Analysis*," 3rd ed., Chapman & Hall/CRC Texts in Statistical Science, 165-195.

Grönroos, Christian (1984), "A Service Quality Model and its Marketing Implications," *European Journal of Marketing*, 18(4), 36-44.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263-292.

Kang, Gi-Du and Jeffrey James (2004), "Service Quality Dimensions: An Examination of Grönroos's Service Quality Model," *Managing Service Quality*, 14(4), 266-277.

Kang, Gi-Du (2006), "The Hierarchical Structure of Service Quality: Integration of Technical and Functional Quality," *Managing Service Quality*, 16(1), 37-50.

Lee, Sik-Yum (2007), *Structural Equation Modeling: A Bayesian Approach*, Chichester: John Wiley & Sons, Ltd.

Lewis, C. (2006), "Selected Topics in Classical Test Theory," *Handbook of Statistics*, 26: *Psychometrics*, 29-43.

Lin, Jiun-Sheng Chris and Pei-Ling Hsieh (2011), "Assessing the Self-service Technology Encounters: Development and Validation of SSTQUAL Scale," *Journal of Retailing*, 87(2), 194-206.

Martínez, Jose A. and Laura Martínez (2010), "Some Insights on Conceptualizing and Measuring Service Quality," *Journal of Retailing and Consumer Services*, 17(1), 29-42.

McDougall, G. H. G. and Terrence J. Levesque (1995), "A Revised View of Service Quality Dimensions: An Empirical Investigation," *Journal of Professional Services Marketing*, 11(1), 189–210.

Mittal, Vikas, William T. Ross Jr., and Patrick M. Baldasare (1998), "The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions," *Journal of Marketing*, 62(1), 33-47.

Novick, Melvin. R. (1966), "The Axioms and Principal Results of Classical Test Theory," *Journal of Mathematical Psychology*, 3(1), 1-18.

Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17(4), 460-469.

Orel, Fatma Demirci and Ali Kara (2014), "Supermarket Self-Checkout Service Quality, Customer Satisfaction, and Loyalty: Empirical Evidence from an Emerging Market," *Journal of Retail and Consumer Services*, 21(2), 118-129.

Parasuraman, A., Leonard L. Berry, and Valarie A. Zeithaml (1993), "More on Improving Service Quality Measurement," *Journal of Retailing*, 69(1), 140-147.

Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49(4), 41-50.

Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry (1988), "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, 64(1), 12-40.

Parasuraman, A., Leonard L. Berry, and Valarie A. Zeithaml (1991), "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing*, 67(4), 420-450.

Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry (1994a), "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research," *Journal of Marketing*, 58(1), 111-124.

Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry (1994b), "Alternative Scale for Measuring Service Quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria," *Journal of Retailing*, 70(3), 201-230.

Parasuraman, A., Valarie A. Zeithaml and Arvind Malhotra (2005), "E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality," *Journal of Service Research*, 7(3), 213-233.

Peter, J. Paul, Gilbert A. Churchill Jr., and Tom J. Brown (1993), "Caution in the Use of Difference Scores in Consumer Research," *Journal of Consumer Research*, 19(4), 655-662.

Prakash, Ved (1984), "Validity and Reliability of the Confirmation of Expectations Paradigm as a Determinant of Consumer Satisfaction," *Journal of the Academy of Marketing Science*, 12(4), 63-76.

Rust, Ronald T. and Richard L. Oliver (1994), "Service quality: insights and managerial implications from the frontier," In R.T. Rust & R. L. Oliver (Eds), *Service Quality: New Directions in Theory and Practice*, London: Sage Publications, 1-19.

Sivakumar, K., Mei Li, and Beibei Dong (2014), "Service Quality: The Impact of Frequency, Timing, Proximity, and Sequence of Failures and Delights," *Journal of Marketing*, 78(1), 41-58.

Song, Xin-Yuan and Sik-Yum Lee (2012), *Basic and Advanced Bayesian Structural Equation Modeling: With Applications in the Medical and Behavioral Sciences*, Chichester: John Wiley & Sons, Ltd.

Teas, R. Kenneth (1993), “Expectations, Performance Evaluation, and Consumers’ Perceptions of Quality,” *Journal of Marketing*, 57(4), 18-34.

Traub, R. E. (1997), “Classical Test Theory in Historical Perspective,” *Educational Measurement: Issues and Practice*, 16(4), 8-14.

Watanabe, Sumio (2010a), “Equations of States in Singular Statistical Estimation,” *Neural Networks*, 23(1), 20-34.

Watanabe, Sumio (2010b), “Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory,” *Journal of Machine Learning Research*, 11(Dec), 3571-3594.

Watanabe, Sumio (2013), “A Widely Applicable Bayesian Information Criterion,” *Journal of Machine Learning Research*, 14(Mar), 867-897.

Wolfinbarger, Mary and Mary C. Gilly (2003), “eTailQ: Dimensionalizing, Measuring and Predicting Etail Quality,” *Journal of Retailing*, 79(3), 183-198.

Xing, Aijing, Nobuhiko Terui, and P. K. Kannan (2016), “How customer satisfaction affects loyalty: Insights from nonlinear hierarchical Bayes modeling of customer satisfaction index,” *Data Science and Service Research Discussion Paper (Tohoku University)*, No. 56, 1-37.

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1993), “The Nature and Determinants of Customer Expectations of Service,” *Journal of the Academy of Marketing Science*, 21(1), 1-12.

Zhu, Hong-Tu and Sik-Yum Lee (1999), “Statistical Analysis of Nonlinear Factor Analysis Models,” *British Journal of Mathematical and Statistical Psychology*, 52(2), 225-242.