Nonlinear Hierarchical Bayes Modeling of Customer Satisfaction Index

著者	XING AIJING
号	21
学位授与機関	Tohoku University
学位授与番号	経博(経営)第117号
URL	http://hdl.handle.net/10097/00121330

氏　名（本籍地）	ケイ　アイショウ XING AIJNG
学 位 の 種 類	博　士（経営学）
学 位 記 番 号	経博（経営）第 117 号
学位授与年月日	平成 28 年 9 月 26 日
学位授与の要件	学位規則第 4 条第 1 項該当
研 究 科 、専 攻	東北大学大学院経済学研究科（博士課程後期 3 年の課程） 経済経営学専攻
学 位 論 文 題 目	Nonlinear Hierarchical Bayes Modeling of Customer Satisfaction Index （顧客満足度指数の非線形階層ベイズモデリング）
博士論文審査委員	（主査） 教　授　照井　伸彦　　准教授　石　垣　　司 教　授　松田　安昌

論 文 内 容 の 要 旨

1. Research Purpose

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

Two research methods are employed in the analysis. The first one uses a hierarchical Bayes nonlinear model to improve the goodness of fit. The second uses finite mixture model to accommodate structural heterogeneity across companies.

2. Literature Review

Over the past several decades, company managers have realized the importance of customer loyalty, which leads to stronger competitive position resulting in 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 is applied in some empirical literatures (V. Kuma et al 2013).

The relationship between customer satisfaction and loyalty has been one of the widely studied relationships in marketing and services literature (Dong et al 2011; Kumar et al 2013). The premise that customer satisfaction, a construct that underlies customers' perceptions regarding their overall consumption experiences (Anderson and Salisbury 2003), significantly impacts customer loyalty, a construct that drives customer retention and repurchase behavior, is key to firms' customer orientation. It is this relationship that forms the basis for measuring marketing effectiveness (Fornell 1992, Bolton and Lemon 1999, Anderson et al. 2004), and for firms' market and 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 creating of firm value through marketing action.

Besides estimating the relationship between satisfaction and loyalty, how to measure the two variables is an important research subject. In most prior research, satisfaction and loyalty variables are directly measured by using a survey directed to respondents. The impact of other antecedents on satisfaction is also not taken into 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 contain some degrees of errors (Andrews, 1984), and satisfaction 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 retrospectively to estimate the relationship, we tend to have the results with low reliability and strongly biased parameter estimates.

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 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. The CSI is a type of market-based performance measure for firms, industries, economic sectors, and national economies. Fornell, et al. (1996) also illustrate the use of CSI in conducting benchmarking studies, both cross-sectional and over time.

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 tries to contribute to the modeling literature by the modeling nonlinear structural equations that include nonlinear terms of endogenous latent variable. In addition, we employ a hierarchical Bayes model and 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 using the 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).

According to these above literatures, we examine the relationship between customer satisfaction and customer loyalty by using the survey data used for developing 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.

As for functional forms of the nonlinear relationship, we consider piecewise 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 common with ours.

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. In addition, the enough large sample size is also required in ML estimation (Hox and Bechger 1998). 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. The ML estimation requires a strong assumption of normal likelihood and sufficient samples, hence Fornell and Cha (1994) applied the partial least squares (PLS) method in the ACSI model for ordered 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, 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.

3. Nonlinear Hierarchical Bayesian CSI Model and Empirical Results

3.1 Nonlinear Hierarchical Bayesian CSI Model

CSI model premises a linear relationship between latent variables, and we extend the model to propose a nonlinear structural equation model which includes nonlinear function from satisfaction to loyalty. In the original CSI, Customer loyalty (LOY) is a responsive variable of a linear function, and customer satisfaction (CS) acts as explanatory variable. 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. So we make the model with the accommodation of these characteristics in a parsimonious way. We put forward four candidates of nonlinear forms: "Asymmetric Linear", "Threshold Linear","Asymmetric Logit" and "Threshold Logit".

As for the hierarchical Bayesian algorithm, the CSI model assumes that every company has the identical path coefficient of customer satisfaction to loyalty 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. This heterogeneity appears to produce the result that some path coefficient estimates are not significant for some companies, which leads to reduce 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 on the goodness of fit, but also by the number of significant estimates of path coefficient. That is, HB model produces 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.

We set the assumption as that: all companies obey an identical nonlinear structure which is chosen from the four candidate models, and these companies have heterogeneous values of coefficients. So the next step is to identify the most suitable nonlinear models. We use Deviance Information Criterion (DIC) and Log of Marginal Likelihood (LML) as goodness of fit.

3.2. Data

The dataset 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). We also transform the 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.

3.3. Empirical Results of HB Nonlinear Modeling

We estimated the parameter using Markov chain Monte Carlo (MCMC). We compare (i) linear model, (ii) logit model, (iii) asymmetric linear model, and (iv) asymmetric logit model in their HB estimations. We also set the original CSI model– the independent linear model as a benchmark 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 note 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 as the former, and we employ DIC which discounts the number of parameter more appropriately than LML.

In addition, the effect of HB modeling that borrows other company's data on the estimates is evident. The number of insignificant estimates in the 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.

4. Finite Mixture Model and Empirical Results

4.1 Structural Heterogeneity and Finite Mixture Model

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

In many situations, it is practical and meaningful to identify the suitable nonlinear model for each company. We propose an approach of modeling the customers' satisfaction affect the loyalty in different nonlinear structural, 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 Linear Model" and "Threshold Logit Model", and company managers should not invest the resources to those customers who are fallen in the zone of tolerance.

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. 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.

4.2. Empirical Results of Heterogeneous Model

We estimated the parameters of finite mixture model. We compare (i) linear model, (ii) HB linear model, (iii) HB logit model, (iv) finite mixture model. After 10,000 iterations of MCMC, and the 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. In each MCMC iteration, we simulate the state indicator Si of customer i of company h, and confirm the state indicator Sh of company h according to the most preferred nonlinear form among the customers. 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, 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

5. Concluding Remarks

In this study, we use both HB nonlinear 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 on 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 for estimation to improve the measurement, with identical model structure applied. In all, this HB nonlinear model's contributions to the modeling literatures 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 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. As managerial implications, 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 company. We also set the change point as parameter, to divide the high satisfied and low satisfied customers, and this method can help company managers to segment the customer groups more effectively. And this kind of structural heterogeneous model has better the goodness of fit compared with homogeneous modeling, so the assumption the 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 structural 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 homogeneous 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 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 modelingof 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.

論文審査結果の要旨

ケイ アイショウ氏は、サービス経済における企業の重要業績指標であり、同時にサービス科学およびマーケティング分野での重要テーマである顧客満足に関するマーケティング・サイエンスの研究を行っている。

第1章の序論に続いて第2章では、顧客満足度とロイヤルティの関係性についての先行研究を丹念に調べ、研究の独自性および重要性を論証している。第3章では、顧客満足とロイヤルティの関係性を統計モデリングの観点から研究し、従来の当該変数のみの局所的分析から複数の潜在変数間の因果関係を構造方程式により表現する顧客満足度指数モデルを用いた包括的アプローチにより捉え、従来の線形構造方程式モデルを非線形モデルへの拡張した新しいモデルを提案している。提案する非線形モデルは、顧客満足がロイヤルティへ与える効果の限界(飽和)を表す上限値および下限値の評価、不満足顧客への影響がより大きい損失回避行動の検証、効果が無い顧客を識別する許容範囲の推定、などを可能とする特徴を有している。モデル化では、各企業のモデルを結合させて安定した推定を可能とする階層ベイズモデリングも取り入れた分析を行っている。さらに第4章では各企業の非線形関数の違いを捉えるために有限混合モデルを適用し、モデル係数のみならず非線形構造の異質性を同時に取れた拡張モデルを提案している。第5章では総括と今後の課題について論述している。

当該分野の学術研究に貢献する新しい分析モデルを複数提案していることに加え、日本の顧客満足度指数データを用いた実証分析については、非線形階層ベイズモデルが最も支持されることを確認した上で、企業がロイヤルティ獲得戦略を考える際に、モデルから推定される期待限界効果曲線と顧客満足度指数分布の両者を重ね合わせることにより、どの顧客にターゲットを絞るのがロイヤルティを向上させるうえで最も効果的かについての示唆を与えるなど、企業の戦略策定にも有用な結果を導出する分析を行っている。

その研究成果は、国際共同研究としてディスカッションペーパーとして刊行し、さらに審査付き国際学会や国内研究集会において研究報告を行っている。

以上により、本論文は博士（経営学）の学位を授与するに値する論文であると認定する。