

The Estimate of Customer Satisfaction

La Stima della Customer Satisfaction

Pietro Giorgio Lovaglio

Dipartimento di Statistica,

Università Milano-Bicocca, P.zza Ateneo Nuovo 1, Edificio U6, 20123 Milano

piergiorgio.lovaglio@unimib.it

Riassunto: Nel presente lavoro si affronta il problema della stima di variabili latenti all'interno di un modello strutturale dal punto di vista applicativo per la stima della Customer Satisfaction (CS) del cliente in seguito all'erogazione di un servizio.

Lo scopo del presente lavoro è duplice: da un lato la proposta di un modello che pervenga alla stima dei punteggi latenti unici in un modello non parametrico basato su un criterio di ottimo globale, evitando l'approccio LISREL e PLS, dall'altra, partendo dall'indice di Customer Satisfaction americano (ACSI) la specificazione delle variabili da inserire nel modello distinguendo, come il modello prevede, le variabili manifeste tra indicatori delle cause ed indicatori in senso stretto della CS.

Keywords: Customer Satisfaction, Latent Variable, Structural Models

1. Models for the estimate of Customer Satisfaction

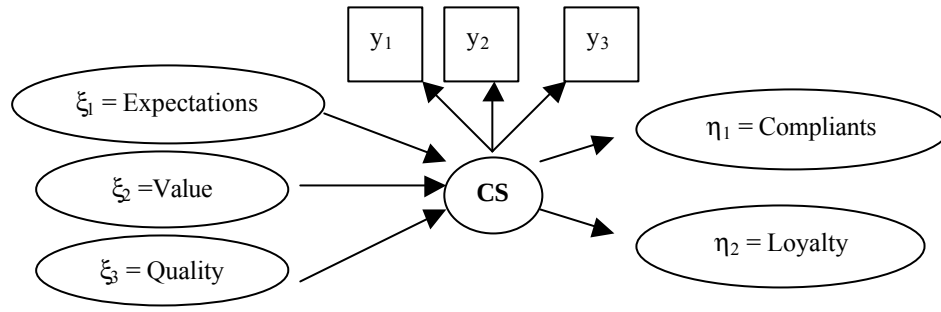
The Customer Satisfaction (CS) address the analysis of the satisfaction of the consumers towards products or their services offered; in the evaluation of services the conceptual models proposed for the indirect estimate of the CS go back to the theory of the gaps between service/ideal products and those perceived, separately for 5 fundamental dimensions *Servqual*, (Parasuman et al. 1991) or to perceived quality, *Serperf*, (Cronin et al., 1992, cited in Zanella 2000).

These approaches do not estimate the scores of CS simultaneously in a statistical model that describes the causal links between CS and his structural dimensions; other methodologies, based on linear, monotone or logistic regression, Rasch Analysis, (cited in Zanella 2000), use the overall perceived satisfaction as proxy of CS.

In the American Customer Satisfaction Index (ACSI, Anderson et al., 2000), an adaptation of the Swedish Customer Satisfaction Barometer (CSB, Fornell, 1992) the CS is estimated like latent variable (LV) by means of structural equations models: the choice of manifest variables (MV's) for each LV is based on well-established theories and approaches in customer behaviour.

In Fig.1 are depicted the structural equations of the ACSI model (LV's in the model) and the MV's of Customer Satisfaction: y_1 (disconfirmation of expectations), y_2 (distance from ideal service), y_3 (overall satisfaction).

Figure 1: Path diagram of the ACSI model



In the causal model CS is a latent construct generating the scores of y_i but in the estimation step, using PLS algorithm, (Tenenhaus, 1995) the scores of CS ($\mathbf{h}$) are performed like linear combination of $\bar{y}_i$ ($i=1,..,3$) averaged by clients where whose weights coincide with the reciprocal ones of the structural (w_i) coefficients of the factor model $\mathbf{y} = \mathbf{W}\mathbf{h} + \mathbf{e}$, which specifies the relationships between $\mathbf{h}$ and y_i .

In previous works (Lovaglio, 2001; Vittadini-Lovaglio, 2001) it has been shown the logic inconsistency of PLS (the LV's are linear transformations of manifest variables, but in the structural model they become underlying causes), the arbitrariness of Mode (A) and Mode (B) and finally the treatment of ordinal data as measures on interval scale.

2. Proposal for the Estimate of Customer Satisfaction

The aim of this paper is to propose a general model to estimate a latent variable Customer satisfaction with manifest categorical variables. The approach joins the all available information about CS in the structural models specified in CSB and ACSI models with gap theory of SERVQUAL.

The current proposal leaves approaches of structural equations (Bentler's) definition of LV (cited in Lovaglio, 2000) since it involves the dual problem of indeterminacy of the latent scores and normality hypothesis on the MV (in the parameters estimation step) and on the LV underlying the ordinal variables (in the polychoric correlations estimation step).

For the simultaneous estimate of latent scores $\mathbf{h}$ and parameters of the measurement model in a recent work (Lovaglio, 2000), the author proposes the Reduced Rank Regression which uses two blocks of (means zero) MV: $\mathbf{X}$ ($n \times p$) non stochastic matrix and $\mathbf{Y}$ ($n \times q$) stochastic matrix, of p and q continuous MV, respectively causes and indicators (effects) of $\mathbf{h}$, following the well-known econometric terms:

$$\mathbf{Y} = \mathbf{X}\mathbf{A}\mathbf{B} + \mathbf{E} = \mathbf{h}\mathbf{B} + \mathbf{E} \quad \text{Vec}(\mathbf{E}) \sim (\mathbf{0}_{nq}, \mathbf{S}_q \otimes \mathbf{I}_n) \\ \text{with constraints: } \mathbf{A}'\mathbf{X}'\mathbf{X}\mathbf{A} = \mathbf{I} \quad \text{rank}(\mathbf{A}) = r \quad (1)$$

where $\mathbf{E}$ is the errors matrix, r ($< \min(p, q)$) the rank of $\mathbf{h}$ ($n \times r$), Vec the usual operator which stacks columns, $\mathbf{A}$ a matrix ($p \times r$) of coefficients that defines uniquely (but an orthogonal rotation) $\mathbf{h} = \mathbf{X}\mathbf{A}$ of rank r , $\mathbf{B}$ ($r \times q$) regression matrix of $\mathbf{Y}$ on $\mathbf{h}$: the estimate of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{S}$ in (1) can be reached avoiding distributional hypotheses on $\mathbf{Y}$, specifying the rank of $\mathbf{h}$, and the error structure of $\mathbf{E}$, maximizing a well-defined fit function.

In a previous work (Vittadini-Lovaglio, 2001), it was shown that manipulating the

structural equations of a structural model in a PLS framework (only for exogenous LV), we can estimate the scores of an endogenous LV $\mathbf{h}$ (CS in Fig.1) as the linear combination of manifest indicators of his exogenous LV ($\mathbf{x}=[\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3]$ in Fig.1) that simultaneously best fits indicators of $\mathbf{h}$ ($y_i \ i=1,..,3$), maximising a linear combination of multiple $R_i^2(i=1,..,q)$ weighted with the variances of $\mathbf{y}_i$; by the same criteria we can estimate the scores of other endogenous LV ($\mathbf{h}_1, \mathbf{h}_2$), specifying opportunely the MV in the two blocks, (Table.1): this consideration justifies the indicators of $\mathbf{x}$ as $\mathbf{X}$ in (1).

The scores of LV's are performed by making rational use of prior information in the structural model: especially the paths between $\mathbf{h}$ and $\mathbf{x}$, and those linking the indicators of different latent variables, avoiding the logic inconsistency in the use of MV'S, maximizing a global optimum criteria.

Table 1 specifies causes and indicators for the estimation of CS, $\mathbf{h}_1, \mathbf{h}_2$ via equation (1).

Table 1: Causes and indicators for the estimate of endogenous LV in the ACSI model.

LV	Indicators of CS (Y_{CS})	Causes of CS (X_{CS})
Customer Satisfaction	-Disconfirmation of expectations -Distance from ideal service -Overall satisfaction	
ξ_1 = Expectations		-SERVQUAL dimensions
ξ_2 = Perceived value		-Price for fixed quality -Quality for fixed Price
ξ_3 = Perceived quality		-SERVQUAL dimensions
	Indicators of $\mathbf{h}_1$ and $\mathbf{h}_2$ (Y_i)	Causes of $\mathbf{h}_1$ and $\mathbf{h}_2$ (X_i)
η_1 = Compliants	-Number of complaint handled? -Care of a staff about complaint?	-Indicators of ξ_j ($j=1, 2, 3$) -Estimated CS
η_2 = Loyalty	-Buy the same product again? -Level of difference (%) for another service -Recommendation of the product	-Indicators of ξ_j ($j=1, 2, 3$) -Estimated CS

3. The Specification of Manifest Variables for $\mathbf{x}_1$ and $\mathbf{x}_3$.

To analyse the real incidence of the gaps (perceived Vs. ideal service) in terms of satisfaction and to enable the distributor to know what expectations to guarantee, it seems more rational to insert, for every specified dimensions j , the gaps directly as covariate $\mathbf{z}_j = \mathbf{g}_j(\mathbf{x}_{3j}-\mathbf{x}_{1j})$, instead of the perceived quality and expectation indicators of each dimension j , by means of the following answer scale:

Very beneath my expectations		In line with my expectations			Very beyond my expectations	
1	2	3	4	5	6	7

to the question “How much the quality of the j -dimension is in line with your expectations?”. Because most of variables in the analysis are ordinal and they are not measures on interval scale (Zanella, 2000), the equation (1) has to be generalized to categorical and mixed manifest variables.

The model (1) has been extended to categorical and/or mixed manifest variables (Lovaglio, 2001) in an optimal scaling framework. In the simple hypothesis $\text{Var}(\mathbf{E})=\mathbf{I}_q \otimes \mathbf{I}_n$; the estimate of CS ($\mathbf{h}^*$) is performed by the algorithm as a weighted

combination of its manifest variables optimally scaled (y_i^{os})

$$h_j^* = S_i y_i^{os} c_{ji} \quad (2)$$

where the weights maximize the same optimum in (1).

The algorithm simultaneously estimates the scores of CS, the structural parameter $\mathbf{B}$ in the equation (1) and the coefficients defining every new-scaled categorical variable.

4. Conclusion

A criticism of PLS (Garthwaite, 1994) is that there seems to be no well-defined modeling problem for which it provides the optimal solution, other than the arbitrary criteria to optimize; the estimate of (2) is the same of estimate of CS with PLS (see Sez.1) because the weights of $\mathbf{Y}=(y_1, y_2, y_3)$ are performed from the causal paths between $\mathbf{Y}$ and the indicators of his exogenous LV, but (2) maximises a global optimum criteria. Model (1), with continuous MV's, for LV strictly endogenous, provides consistent and asymptotically normal estimators (like Instrumental Variables estimators): to estimate $\mathbf{h}_1$ and $\mathbf{h}_2$, model (1) has to be applied twice: this is analogous to two simultaneous equations system with r_i equations (indicators of $\mathbf{h}_i$) where the correlation between errors of $\mathbf{h}_i$ and CS may causes inconsistent estimator; nevertheless the estimator of CS by model (1) make CS an instrumental variable for every indicators of $\mathbf{h}_i$ vanishing the correlation between errors and CS and having asymptotic properties, as Two Stages Least Squares Estimator or Instrumental Variable Estimator.

References

- Anderson E.W., Fornell C. (2000) Foundations of the American Customer Satisfaction Index, *Total Quality Management*, 11, 869-882.
- Fornell C. (1992) A National Customer Satisfaction Barometer: The Swedish Experience, *Journal of Marketing*, Vol. 56, 6-21.
- Garthwaite P.H. (1994) An interpretation of Partial Least Squares, *Journal of American Statistical Association*, 89, 122-127.
- Lovaglio P.G. (2001) La stima di outcome latenti, *Atti della Riunione intermedia SIS, Processi e metodi statistici di valutazione*, Roma 2001.
- Lovaglio P.G. (2000) La Stima di Variabili Latenti da Variabili Osservate Miste, *Statistica*, in press.
- Parasuraman A., Berry L.L., Zeithaml V.A. (1991), Refinement and Reassessment of the SERVQUAL Scale, *Journal of Retailing*, vol.67, n.4, pp.420-450.
- Tenenhaus M. (1995) *La Régression PLS: Théorie et Pratique*. Editions Technip, Paris.
- Vittadini G., Lovaglio P.G. (2001) *The estimate of latent variables in a structural model: an alternative approach to PLS*, 2nd International Symposium on PLS and Related Methods (PLS'01), October 1-3, 2001 Capri (Italy), CISIA-CERESTA, Montreuil, France.
- Zanella A. (2000) *Valutazione della qualità e customer satisfaction: il ruolo della statistica*. Atti giornate di studio sul tema, Bologna, 24-9-1999, Vita e Pensiero, 1-24.