Multilevel Generalized Structured Component Analysis

Heungsun Hwang

McGill University

Yoshio Takane

McGill University

Naresh Malhotra

Georgia Institute of Technology

November 16, 2007

The work reported in this paper was supported by Grant 290439 and Grant A6394 from the Natural Sciences and Engineering Research Council of Canada to the first and second authors, respectively. We wish to thank Claes Fornell for generously providing us with the ACSI data. Requests for reprints should be sent to: Heungsun Hwang, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, QC, H3A 1B1, Canada. Email: heungsun.hwang@mcgill.ca.

Abstract

Generalized structured component analysis has been proposed as an alternative to partial least squares for path analysis with latent variables. In practice, observed and latent variables may often be hierarchically structured in that their individual-level scores are grouped within higher-level units. The observed and latent variable scores nested within the same higher-level group are likely to be more similar than those in different groups, thereby giving rise to the interdependence of the scores within the same group. Unless this interdependence is taken into account, obtained solutions are likely to be biased. In this paper, generalized structured component analysis is extended so as to account for the nested structures of both observed and latent variables. An alternating least-squares procedure is developed for parameter estimation. An empirical application concerning the measurements of customer-level customer satisfaction nested within different companies is presented to illustrate the usefulness of the proposed method.

Keywords: Generalized structured component analysis, multilevel analysis, alternating least squares, customer satisfaction.

1. Introduction

Generalized structured component analysis (GSCA) (Hwang & Takane, 2004) was proposed as an alternative method to partial least squares (PLS) (Lohmöller, 1989; Wold, 1966, 1973, 1975, 1982) for path analysis with latent variables defined as weighted composites of observed variables. GSCA so far assumes that all individual cases are independently selected from a population. In practice, however, data are often hierarchically structured in that their individual-level cases are grouped within higher-level units. For instance, an adolescent's level of substance use may be measured across different urban areas nested within different provinces. The standardized test scores of students may be observed across different schools. The functional magnetic resonance imaging data of each patient may be repeatedly taken over time (i.e., multiple time points are nested within each patient). The individual-level measures nested within the same group are likely to be more similar compared to those in different groups; thus, leading to the dependencies among the observations within the same group. Unless this interdependence of the individual-level observations is taken into account, obtained solutions are likely to be inaccurate (cf. Bryk & Raudenbush, 1992; Snijders & Bosker, 1999).

In addition to such a clustered nature of observed variables, latent variables may also be seen hierarchically structured in the context of path analysis with latent variables. For instance, in the National Longitudinal Survey of Youth data, antisocial behavior of children can be viewed as a latent variable, which is often calculated as a sum of mothers' responses to six items from the Behavior Problems Index (e.g., Curran, 1998). The antisocial behavior of each child can be repeatedly measured across multiple time points. The American Customer Satisfaction Index (ACSI) model (Fornell, Johnson, Anderson, Cha, & Bryant, 1996) represents a path analysis model, which estimates customer satisfaction as a latent variable of major interest, and also focuses on the inter-relationships among antecedent and consequent latent variables of customer satisfaction. In the ACSI model, customer-level satisfaction scores are measured for about 200 major companies nested within different sectors of the US economy (Fornell *et al*., 1996). Thus, it may be desirable to take into account the nested structure of individual latent variable scores as well as that of observed variable scores in path analysis with latent variables.

In this paper, GSCA is generalized so as to deal with the hierarchical structures of both observed and latent variables. Specifically, the proposed method permits both loadings and path coefficients to be assumed to vary across higher-level units. Moreover, it allows investigating cross-level or interaction effects of explanatory variables for loadings and path coefficients in different levels.

The paper is organized as follows. In Section 2, the proposed method is discussed in detail. This section focuses on the two-level GSCA model for simplicity. An alternating least squares algorithm is presented for parameter estimation. In Section 3, an empirical data set is used to illustrate the usefulness of the proposed method. The final section is devoted to discussing implications and further prospects of the proposed method.

2. The Proposed Method

2.1. The Two-Level GSCA Model

For simplicity, only the two-level GSCA model is discussed in this section. An extension to the three-level GSCA model is presented in the Appendix. Let $\mathbf{Z}_j$ denote an

N_j by P matrix of observed variables in the j-th group $(j = 1, \cdots, J)$. Let **V** and **W** denote weight matrices for endogenous and exogenous observed variables and latent variables, respectively. Let $\mathbf{A}_j$ denote a matrix of Level-1 loadings relating latent variables to observed variables, which are assumed to vary across groups. Let $\mathbf{B}_j$ denote a matrix of Level-1 path coefficients of latent variables, which are conceived to be different across groups. Let $\mathbf{E}_j$ denote a matrix consisting of all residuals. Then, the Level-1 GSCA model is given by

$$\begin{aligned} \mathbf{Z}_j\mathbf{V} &= \mathbf{Z}_j\mathbf{W}[\mathbf{A}_j, \mathbf{B}_j] + \mathbf{E}_j, \\ \mathbf{\Psi}_j &= \mathbf{\Gamma}_j\mathbf{T}_j + \mathbf{E}_j, \end{aligned} \tag{1}$$

where $\mathbf{\Psi}_j = \mathbf{Z}_j\mathbf{V}$, $\mathbf{\Gamma}_j = \mathbf{Z}_j\mathbf{W}$, and $\mathbf{T}_j = [\mathbf{A}_j,\ \mathbf{B}_j]$.

Next, let $\mathbf{Q}_j$ denote a matrix consisting of Level-2 exogenous observed variables, which explain the relationships between observed and latent variables in group j. Let **A** denote a matrix of Level-2 fixed loadings. Let $\mathbf{\Lambda}_j$ denote a matrix of Level-2 random loadings, which are assumed to vary across groups. Also, let $\mathbf{C}_j$ denote a matrix consisting of Level-2 exogenous observed variables, which account for the relationships between latent variables. Let **B** denote a matrix of Level-2 fixed path coefficients. Let $\mathbf{U}_j$ denote a matrix of Level-2 random effects of path coefficients, which may vary across groups. Then Level-2 models are given by

$$\begin{aligned} \mathbf{A}_j &= \mathbf{Q}_j\mathbf{A} + \mathbf{\Lambda}_j, \\ \mathbf{B}_j &= \mathbf{C}_j\mathbf{B} + \mathbf{U}_j. \end{aligned} \tag{2}$$

The Level-2 models indicate that Level-1 loadings and Level-1 path coefficients are conceived as varying over the population of a Level-2 unit, and also that such a variation is accounted for by the Level-2 characteristics, i.e., $\mathbf{Q}_j$ and $\mathbf{C}_j$. In (2), the variance of each random effect represents the inter-group variability of the corresponding loading or path coefficient.

Finally, the model in combined form is as follows:

$$\begin{aligned} \mathbf{Z}_j\mathbf{V} &= \mathbf{Z}_j\mathbf{W}[\mathbf{Q}_j\mathbf{A} + \mathbf{\Lambda}_j, \mathbf{C}_j\mathbf{B} + \mathbf{U}_j] + \mathbf{E}_j \\ &= [\mathbf{\Gamma}_j\mathbf{Q}_j\mathbf{A} + \mathbf{\Gamma}_j\mathbf{\Lambda}_j, \mathbf{\Gamma}_j\mathbf{C}_j\mathbf{B} + \mathbf{\Gamma}_j\mathbf{U}_j] + \mathbf{E}_j. \end{aligned} \tag{3}$$

In this combined model, it is seen that **A** and **B** represent the cross-level or interaction effects between Level-1 latent/observed variables ($\mathbf{\Gamma}_j$) and Level-2 characteristics ($\mathbf{Q}_j$ and $\mathbf{C}_j$). It is noteworthy that the proposed model is essentially the same form as the model of GSCA for a single group. Indeed, (3) reduces to the GSCA model when $J = 1$, $\mathbf{Q}_j = \mathbf{I}$, $\mathbf{\Lambda}_j = \mathbf{0}$, $\mathbf{C}_j = \mathbf{I}$, and $\mathbf{U}_j = \mathbf{0}$. Therefore, the proposed model is an extension of the single-level GSCA model, which takes into account the multilevel structure of observed and latent variables. We shall call this proposed method Multilevel GSCA (MGSCA) hereafter.

To illustrate MGSCA, we shall use the ACSI model which is a well-known generic model for measuring customer satisfaction in the United States (Fornell *et al*, 1996). The ACSI model for the *j*-th group is depicted in Figure 1.

Insert Figure 1 about here

In Figure 1, $\mathbf{z}_j$'s represent vectors of the following fourteen observed variables in the *j*-th group: $\mathbf{z}_{1j}$ = customer expectations about overall quality, $\mathbf{z}_{2j}$ = customer expectations about reliability, $\mathbf{z}_{3j}$ = customer expectations about customization, $\mathbf{z}_{4j}$ = overall quality, $\mathbf{z}_{5j}$ = reliability, $\mathbf{z}_{6j}$ = customization, $\mathbf{z}_{7j}$ = price given quality, $\mathbf{z}_{8j}$ = quality given price, $\mathbf{z}_{9j}$ =

overall customer satisfaction, $\mathbf{z}_{10j}$ = confirmation of expectations, $\mathbf{z}_{11j}$ = distance to ideal product or service, $\mathbf{z}_{12j}$ = formal or informal complaint behavior, $\mathbf{z}_{13j}$ = repurchase intention, and $\mathbf{z}_{14j}$ = price tolerance. The measures and scales of these observed variables are presented in Fornell *et al*. (1996).

The $\boldsymbol{\eta}_i$'s are vectors of latent variables and denoted as follows: $\boldsymbol{\eta}_{1j}$ = customer expectations (CE), $\boldsymbol{\eta}_{2j}$ = perceived quality (PQ), $\boldsymbol{\eta}_{3j}$ = perceived value (PV), $\boldsymbol{\eta}_{4j}$ = customer satisfaction (CS), $\boldsymbol{\eta}_{5j}$ = customer complaints (CC), and $\boldsymbol{\eta}_{6j}$ = customer loyalty (CL). In the figure, straight arrows are used to signify that the variable at the base of an arrow affects the variable at the head of the arrow whereas straight lines are used to represent the weighted relations between observed and latent variables. Moreover, the symbol '+' in parentheses positioned over a path stands for a positive relationship between two latent variables, while the symbol '-' represents a negative relationship.

As shown in Figure 1, the *i*-th latent variable in group *j* is defined as a weighted composite of several observed variables, that is, $\boldsymbol{\eta}_{ij} = \sum_p \mathbf{z}_{pj} w_p$, where w_p is the weight for the *p*-th observed variable in group *j* $(i = 1,\cdots,6;\ p = 1,\cdots,14)$. Note that the weight for $\mathbf{z}_{12j}$ is fixed to one because this is the only observed variable for $\boldsymbol{\eta}_{5j}$ (CC). In the ACSI model, the Level-1 measurement model that specifies the relationships between observed and latent variables is given as follows:

$$\begin{aligned}
[\mathbf{z}_{1j},\mathbf{z}_{2j},\mathbf{z}_{3j}] &= \boldsymbol{\eta}_{1j}[a_{1j},a_{2j},a_{3j}] + [\mathbf{e}_{1j},\mathbf{e}_{2j},\mathbf{e}_{3j}],\\
[\mathbf{z}_{4j},\mathbf{z}_{5j},\mathbf{z}_{6j}] &= \boldsymbol{\eta}_{2j}[a_{4j},a_{5j},a_{6j}] + [\mathbf{e}_{4j},\mathbf{e}_{5j},\mathbf{e}_{6j}],\\
[\mathbf{z}_{7j},\mathbf{z}_{8j}] &= \boldsymbol{\eta}_{3j}[a_{7j},a_{8j}] + [\mathbf{e}_{7j},\mathbf{e}_{8j}],\\
[\mathbf{z}_{9j},\mathbf{z}_{10j},\mathbf{z}_{11j}] &= \boldsymbol{\eta}_{4j}[a_{9j},a_{10j},a_{11j}] + [\mathbf{e}_{9j},\mathbf{e}_{10j},\mathbf{e}_{11j}],\\
\mathbf{z}_{12j} &= \boldsymbol{\eta}_{5j} + \mathbf{e}_{12j},\\
[\mathbf{z}_{13j},\mathbf{z}_{14j}] &= \boldsymbol{\eta}_{6j}[a_{13j},a_{14j}] + [\mathbf{e}_{13j},\mathbf{e}_{14j}],
\end{aligned} \tag{4}$$

where a_{pj} is a Level-1 loading for the p-th observed variable in group j, and $\mathbf{e}_{pj}$ is a vector of the residual of $\mathbf{z}_{pj}$. Note that the Level-1 loading for $\mathbf{z}_{12j}$ is set to one across groups, i.e., $a_{12j} = 1$. Moreover, the Level-1 structural model that specifies the relationships among latent variables in group j is given as follows:

$$\begin{aligned}
\boldsymbol{\eta}_{2j} &= \boldsymbol{\eta}_{1j} b_{1j} + \mathbf{d}_{2j}, \\
\boldsymbol{\eta}_{3j} &= \boldsymbol{\eta}_{1j} b_{2j} + \boldsymbol{\eta}_{2j} b_{4j} + \mathbf{d}_{3j}, \\
\boldsymbol{\eta}_{4j} &= \boldsymbol{\eta}_{1j} b_{3j} + \boldsymbol{\eta}_{2j} b_{5j} + \boldsymbol{\eta}_{3j} b_{6j} + \mathbf{d}_{4j}, \\
\boldsymbol{\eta}_{5j} &= \boldsymbol{\eta}_{4j} b_{7j} + \mathbf{d}_{5j}, \\
\boldsymbol{\eta}_{6j} &= \boldsymbol{\eta}_{4j} b_{8j} + \boldsymbol{\eta}_{5j} b_{9j} + \mathbf{d}_{6j},
\end{aligned} \tag{5}$$

where b_{kj} is a Level-1 path coefficient ($k = 1, \cdots, 9$), and $\mathbf{d}_{ij}$ is a vector of the residual of $\boldsymbol{\eta}_{ij}$.

Next, Level-2 models may be specified for the ACSI model as follows: $a_{pj} = a_p + \lambda_{pj}$ and $b_{kj} = b_k + u_{kj}$, where a_p is a Level-2 fixed loading, λ_{pj} is a Level-2 random loading, b_k is a Level-2 fixed path coefficient, and u_{kj} is a Level-2 random path coefficient. These models are contemplated because no Level-2 exogenous variables are available to predict across-level variations in Level-1 loadings and path coefficients. Then, the combined measurement model for the Level-1 loadings is specified as follows:

$$\begin{aligned}
[\mathbf{z}_{1j}, \mathbf{z}_{2j}, \mathbf{z}_{3j}] &= \boldsymbol{\eta}_{1j}[a_1, a_2, a_3] + \boldsymbol{\eta}_{1j}[\lambda_{1j}, \lambda_{2j}, \lambda_{3j}] + [\mathbf{e}_{1j}, \mathbf{e}_{2j}, \mathbf{e}_{3j}], \\
[\mathbf{z}_{4j}, \mathbf{z}_{5j}, \mathbf{z}_{6j}] &= \boldsymbol{\eta}_{2j}[a_4, a_5, a_6] + \boldsymbol{\eta}_{2j}[\lambda_{4j}, \lambda_{5j}, \lambda_{6j}] + [\mathbf{e}_{4j}, \mathbf{e}_{5j}, \mathbf{e}_{6j}], \\
[\mathbf{z}_{7j}, \mathbf{z}_{8j}] &= \boldsymbol{\eta}_{3j}[a_7, a_8] + \boldsymbol{\eta}_{3j}[\lambda_{7j}, \lambda_{8j}] + [\mathbf{e}_{7j}, \mathbf{e}_{8j}], \\
[\mathbf{z}_{9j}, \mathbf{z}_{10j}, \mathbf{z}_{11j}] &= \boldsymbol{\eta}_{4j}[a_9, a_{10}, a_{11}] + \boldsymbol{\eta}_{4j}[\lambda_{9j}, \lambda_{10j}, \lambda_{11j}] + [\mathbf{e}_{9j}, \mathbf{e}_{10j}, \mathbf{e}_{11j}], \\
\mathbf{z}_{12j} &= \boldsymbol{\eta}_{5j} + \mathbf{e}_{12j}, \\
[\mathbf{z}_{13j}, \mathbf{z}_{14j}] &= \boldsymbol{\eta}_{6j}[a_{13}, a_{14}] + \boldsymbol{\eta}_{6j}[\lambda_{13j}, \lambda_{14j}] + [\mathbf{e}_{13j}, \mathbf{e}_{14j}].
\end{aligned} \tag{6}$$

The combined structural model for the Level-1 path coefficients consists of:

$$
\begin{aligned}
\boldsymbol{\eta}_{2j} &= \boldsymbol{\eta}_{1j} b_1 + \boldsymbol{\eta}_{1j} u_{1j} + \mathbf{d}_{2j}, \\
\boldsymbol{\eta}_{3j} &= \boldsymbol{\eta}_{1j} b_2 + \boldsymbol{\eta}_{1j} u_{2j} + \boldsymbol{\eta}_{2j} b_4 + \boldsymbol{\eta}_{2j} u_{4j} + \mathbf{d}_{3j}, \\
\boldsymbol{\eta}_{4j} &= \boldsymbol{\eta}_{1j} b_3 + \boldsymbol{\eta}_{1j} u_{3j} + \boldsymbol{\eta}_{2j} b_5 + \boldsymbol{\eta}_{2j} u_{5j} + \boldsymbol{\eta}_{3j} b_6 + \boldsymbol{\eta}_{3j} u_{6j} + \mathbf{d}_{4j}, \\
\boldsymbol{\eta}_{5j} &= \boldsymbol{\eta}_{4j} b_7 + \boldsymbol{\eta}_{4j} u_{7j} + \mathbf{d}_{5j}, \\
\boldsymbol{\eta}_{6j} &= \boldsymbol{\eta}_{4j} b_8 + \boldsymbol{\eta}_{4j} u_{8j} + \boldsymbol{\eta}_{5j} b_9 + \boldsymbol{\eta}_{5j} u_{9j} + \mathbf{d}_{6j}.
\end{aligned} \tag{7}
$$

Suppose $\mathbf{Z}_j = [\,\mathbf{z}_{1j}, \mathbf{z}_{2j}, \cdots, \mathbf{z}_{14j}\,]$ and $\mathbf{E}_j = [\,\mathbf{e}_{1j}, \mathbf{e}_{2j}, \cdots, \mathbf{e}_{14j}, \mathbf{d}_{1j}, \mathbf{d}_{2j}, \mathbf{d}_{3j}, \mathbf{d}_{4j}, \mathbf{d}_{5j}\,]$.

The two-level GSCA model for the ACSI data is then expressed as

$$
\mathbf{Z}_j
\begin{bmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&w_4&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&w_5&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&w_6&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&w_7&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&w_8&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&w_9&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&w_{10}&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&w_{11}&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&w_{13}\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&w_{14}
\end{bmatrix}
= \mathbf{Z}_j
\begin{bmatrix}
w_1&0&0&0&0&0\\
w_2&0&0&0&0&0\\
w_3&0&0&0&0&0\\
0&w_4&0&0&0&0\\
0&w_5&0&0&0&0\\
0&w_6&0&0&0&0\\
0&0&w_7&0&0&0\\
0&0&w_8&0&0&0\\
0&0&0&w_9&0&0\\
0&0&0&w_{10}&0&0\\
0&0&0&w_{11}&0&0\\
0&0&0&0&1&0\\
0&0&0&0&0&w_{13}\\
0&0&0&0&0&w_{14}
\end{bmatrix}
$$

$$
\left(
\begin{bmatrix}
a_1+\lambda_{1j}&a_2+\lambda_{2j}&a_3+\lambda_{3j}&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&a_4+\lambda_{4j}&a_5+\lambda_{5j}&a_6+\lambda_{6j}&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&a_7+\lambda_{7j}&a_8+\lambda_{8j}&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&a_9+\lambda_{9j}&a_{10}+\lambda_{10j}&a_{11}+\lambda_{11j}&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&a_{13}+\lambda_{13j}&a_{14}+\lambda_{14j}
\end{bmatrix},
\right.
$$

$$
\left.
\begin{bmatrix}
b_1+u_{1j}&b_2+u_{2j}&b_3+u_{3j}&0&0\\
0&b_4+u_{4j}&b_5+u_{5j}&0&0\\
0&0&b_6+u_{6j}&0&0\\
0&0&0&b_7+u_{7j}&b_8+u_{8j}\\
0&0&0&0&b_9+u_{9j}\\
0&0&0&0&0
\end{bmatrix}
\right) + \mathbf{E}_j. \tag{8}
$$

This can be re-expressed as

$$\mathbf{Z}_j\mathbf{V} = \mathbf{Z}_j\mathbf{W}[\mathbf{Q}_j\mathbf{A} + \boldsymbol{\Lambda}_j, \mathbf{C}_j\mathbf{B} + \mathbf{U}_j] + \mathbf{E}_j,$$
$$\boldsymbol{\Psi}_j = \boldsymbol{\Gamma}_j\mathbf{T}_j + \mathbf{E}_j, \tag{9}$$

where $\mathbf{V} = \begin{bmatrix} 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&0&0&0&0&w_4&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&0&0&0&w_5&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0&0&0&0&0&w_6&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&w_7&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&w_8&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&w_9&0&0 \\ 0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&w_{10}&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&w_{11}&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&w_{13} \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&w_{14} \end{bmatrix}$,

$\mathbf{W} = \begin{bmatrix} w_1&0&0&0&0&0 \\ w_2&0&0&0&0&0 \\ w_3&0&0&0&0&0 \\ 0&w_4&0&0&0&0 \\ 0&w_5&0&0&0&0 \\ 0&w_6&0&0&0&0 \\ 0&0&w_7&0&0&0 \\ 0&0&w_8&0&0&0 \\ 0&0&0&w_9&0&0 \\ 0&0&0&w_{10}&0&0 \\ 0&0&0&w_{11}&0&0 \\ 0&0&0&0&1&0 \\ 0&0&0&0&0&w_{13} \\ 0&0&0&0&0&w_{14} \end{bmatrix}$, $\mathbf{Q}_j = \mathbf{I}$,

$\mathbf{A} = \begin{bmatrix} a_1&a_2&a_3&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&a_4&a_5&a_6&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&a_7&a_8&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&a_9&a_{10}&a_{11}&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&a_{13}&a_{14} \end{bmatrix}$,

$$\mathbf{\Lambda}_j = \begin{bmatrix} \lambda_{1j} & \lambda_{2j} & \lambda_{3j} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{4j} & \lambda_{5j} & \lambda_{6j} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{7j} & \lambda_{8j} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{9j} & \lambda_{10j} & \lambda_{11j} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{13j} & \lambda_{14j} \end{bmatrix}, \mathbf{C}_j = \mathbf{I},$$

$$\mathbf{B} = \begin{bmatrix} b_1 & b_2 & b_3 & 0 & 0 \\ 0 & b_4 & b_5 & 0 & 0 \\ 0 & 0 & b_6 & 0 & 0 \\ 0 & 0 & 0 & b_7 & b_8 \\ 0 & 0 & 0 & 0 & b_9 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{U}_j = \begin{bmatrix} u_{1j} & u_{2j} & u_{4j} & 0 & 0 \\ 0 & u_{3j} & u_{5j} & 0 & 0 \\ 0 & 0 & u_{6j} & 0 & 0 \\ 0 & 0 & 0 & u_{7j} & u_{8j} \\ 0 & 0 & 0 & 0 & u_{9j} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In (9), $\mathbf{\Psi}_j = [\mathbf{Z}_j, \boldsymbol{\eta}_{2j}, \boldsymbol{\eta}_{3j}, \boldsymbol{\eta}_{4j}, \boldsymbol{\eta}_{5j}, \boldsymbol{\eta}_{6j}]$ and $\mathbf{\Gamma}_j = [\boldsymbol{\eta}_{1j}, \boldsymbol{\eta}_{2j}, \boldsymbol{\eta}_{3j}, \boldsymbol{\eta}_{4j}, \boldsymbol{\eta}_{5j}, \boldsymbol{\eta}_{6j}]$. Note that $\lambda_{12j} = 0$ because the Level-1 loading for $\mathbf{z}_{12j}$ is constrained to be equal to one across groups.

2.2. Parameter Estimation

To estimate model parameters, we seek to minimize the following least squares (LS) criterion:

$$\begin{aligned} f &= \sum_{j=1}^{J} \mathrm{SS}(\mathbf{Z}_j\mathbf{V} - \mathbf{Z}_j\mathbf{W}[\mathbf{Q}_j\mathbf{A} + \mathbf{\Lambda}_j, \mathbf{C}_j\mathbf{B} + \mathbf{U}_j]) \\ &= \sum_{j=1}^{J} \mathrm{SS}\left(\mathbf{Z}_j\mathbf{V} - \mathbf{Z}_j\mathbf{W}\left[\mathbf{Q}_j, \mathbf{I}, \mathbf{C}_j, \mathbf{I}\right] \begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{\Lambda}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{U}_j \end{bmatrix}\right) \\ &= \sum_{j=1}^{J} \mathrm{SS}(\mathbf{Z}_j\mathbf{V} - \mathbf{Z}_j\mathbf{W}\mathbf{G}\mathbf{R}\mathbf{H}_j), \end{aligned} \quad (10)$$

with respect to **V**, **W**, **A**, **B**, $\mathbf{\Lambda}_j$ and $\mathbf{U}_j$, subject to $\boldsymbol{\eta}_i'\boldsymbol{\eta}_i = 1$, $\sum_j \lambda_{pj} = 0$, and $\sum_j u_{kj} = 0$ for identification, where SS(**X**) = trace($\mathbf{X}'\mathbf{X}$), **G** = [$\mathbf{Q}_j$, **I**, $\mathbf{C}_j$, **I**], $\mathbf{R} = \begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}$, and $\mathbf{H}_j = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{\Lambda}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{U}_j \end{bmatrix}$.

An alternating least squares (ALS) algorithm (de Leeuw, Young, & Takane, 1976) is developed to minimize (10). The proposed ALS algorithm can be viewed as a simple extension of that for GSCA. Specifically, this algorithm repeats the following three main steps until convergence is obtained:

<u>Step 1</u>. **V** and **W** are updated for fixed **A**, **B**, $\mathbf{\Lambda}_j$ and $\mathbf{U}_j$. This step is equivalent to minimizing

$$f = \sum_{j=1}^{J} \mathrm{SS}(\mathbf{Z}_j\mathbf{V} - \mathbf{Z}_j\mathbf{W}\mathbf{T}_j), \quad (11)$$

with respect to **V** and **W**, where $\mathbf{T}_j = \mathbf{G}\mathbf{R}\mathbf{H}_j$. This criterion is essentially equivalent to that for GSCA. Thus, the same ALS algorithm as that for GSCA is used to update **V** and **W** (see Hwang & Takane, 2004).

<u>Step 2</u>. **A** and **B** (or equivalently **R**) are updated for fixed **V**, **W**, $\mathbf{\Lambda}_j$ and $\mathbf{U}_j$. Criterion (10) can be re-written as

$$f = \sum_{j=1}^{J} \mathrm{SS}(\mathrm{vec}(\mathbf{Z}_j\mathbf{V}) - (\mathbf{H}_j' \otimes \mathbf{Z}_j\mathbf{W}\mathbf{G})\mathrm{vec}(\mathbf{R})), \quad (12)$$

where vec(**X**) denotes a supervector consisting of all columns of **X** one below another, and $\otimes$ denotes a Kronecker product. Let **r** denote the vector formed by eliminating any fixed (e.g., zero or one) elements from vec(**R**). Let $\mathbf{\Omega}_j$ denote the matrix formed by eliminating the columns of $\mathbf{H}_j'\otimes\mathbf{Z}_j\mathbf{WG}$ corresponding to the fixed elements in vec(**R**). Let $\boldsymbol{\omega}_{vj}$ denote the column vector of $\mathbf{H}_j'\otimes\mathbf{Z}_j\mathbf{WG}$ corresponding to the v-th fixed element q_v (e.g., q_v = 0 or 1) in vec(**R**). Then, the LS estimate of **r** is obtained by

$$\hat{\mathbf{r}} = \left\{\sum_{j=1}^{J}\mathbf{\Omega}_j'\mathbf{\Omega}_j\right\}^{-1}\left\{\sum_{j=1}^{J}\mathbf{\Omega}_j'\left(\mathrm{vec}(\mathbf{Z}_j\mathbf{V}) - \sum_v \boldsymbol{\omega}_{vj}q_v\right)\right\}. \tag{13}$$

The updated **R** is reconstructed from $\hat{\mathbf{r}}$. The regular inverse may be replaced by the Moore-Penrose inverse if $\sum_{j=1}^{J}\mathbf{\Omega}_j'\mathbf{\Omega}_j$ is singular.

Step 3. $\mathbf{\Lambda}_j$ and $\mathbf{U}_j$ (or equivalently $\mathbf{H}_j$) is updated for fixed **V**, **W**, **A**, and **B**. Criterion (10) can also be re-written as

$$f = \sum_{j=1}^{J}\mathrm{SS}(\mathrm{vec}(\mathbf{Z}_j\mathbf{V}) - (\mathbf{I}\otimes\mathbf{Z}_j\mathbf{WGR})\mathrm{vec}(\mathbf{H}_j)). \tag{14}$$

Let $\mathbf{h}_j$ denote the vector formed by eliminating any fixed elements from vec($\mathbf{H}_j$). Let $\mathbf{\Xi}_j$ denote the matrix formed by eliminating the columns of $\mathbf{I}\otimes\mathbf{Z}_j\mathbf{WGR}$ corresponding to the fixed elements in vec($\mathbf{H}_j$). Let $\boldsymbol{\xi}_{lj}$ denote the column vector of $\mathbf{I}\otimes\mathbf{Z}_j\mathbf{WGR}$ corresponding to the l-th fixed element m_{lj} (e.g., m_{lj} = 0 or 1) in vec($\mathbf{H}_j$). Then, the LS estimate of $\mathbf{h}_j$ is obtained by

$$\hat{\mathbf{h}}_j = \left(\mathbf{\Xi}_j'\mathbf{\Xi}_j\right)^{-1}\mathbf{\Xi}_j'\left(\mathrm{vec}(\mathbf{Z}_j\mathbf{V}) - \sum_l \boldsymbol{\xi}_{lj}m_{lj}\right). \tag{15}$$

The updated $\mathbf{H}_j$ is reconstructed from $\hat{\mathbf{h}}_j$. Again the regular inverse may be replaced by the Moore-Penrose inverse if $\mathbf{\Xi}_j'\mathbf{\Xi}_j$ is singular.

A few remarks concerning the proposed algorithm are in order. First, the ALS algorithm monotonically decreases the value of criterion (10) which, in turn, is also bounded from below. This algorithm is therefore convergent. However, the algorithm does not guarantee that the convergence point is the global minimum. To safeguard against this so-called convergence to non-global minimum problem, one may repeat the ALS procedure with a large number of random initial estimates of parameters. One then compares the obtained function values after convergence and subsequently chooses the solution associated with the smallest one. Second, when N_j is large relative to P, the proposed algorithm can be made more efficient by a procedure similar to that for GSCA. Let $\mathbf{Z}_j = \mathbf{\Theta}_j\mathbf{\Delta}'_j$ be portions of the QR decomposition of $\mathbf{Z}_j$ pertaining to the column space of $\mathbf{Z}_j$, where $\mathbf{\Theta}_j$ is an N_j by P semi-orthonormal matrix, so that $\mathbf{\Theta}_j'\mathbf{\Theta}_j = \mathbf{I}$, and $\mathbf{\Delta}'_j$ is a P by N upper-triangular matrix. Then, minimizing (10) reduces to minimizing

$$f = \sum_{j=1}^{J} \mathrm{SS}(\mathbf{\Delta}_j'\mathbf{V} - \mathbf{\Delta}_j'\mathbf{W}\mathbf{T}_j)\,. \tag{16}$$

It is computationally more convenient to minimize (16) than (10) because the size of $\mathbf{\Delta}'_j$ is usually much smaller than that of $\mathbf{Z}_j$. Finally, the proposed algorithm can be readily extended so as to estimate parameters of a higher-level GSCA model. For example, as shown in the Appendix, the three-level GSCA model entails an additional matrix of parameters compared to the two-level model. This matrix of new parameters can be easily updated by adding a sub-optimization step similar to Step 2 or Step 3.

Similarly to GSCA, MGSCA measures the overall goodness of fit of a hypothesized model by the portion of the total variance of all endogenous variables explained by model specifications. This is given by

$$\text{FIT} = 1 - \frac{\sum_{j=1}^{J} \text{SS}(\mathbf{Z}_j \mathbf{V} - \mathbf{Z}_j \mathbf{W} \mathbf{T}_j)}{\sum_{j=1}^{J} \text{SS}(\mathbf{Z}_j \mathbf{V})}. \tag{17}$$

This index ranges from 0 to 1. The closer to 1 its value is, the more variance of endogenous variables is accounted for.

Also like GSCA, MGSCA employs the bootstrap method (Efron, 1982) in order to estimate the standard errors of parameter estimates. In this method, random samples (bootstrap samples) of $\mathbf{Z}_j$ are repeatedly sampled from the original data matrix with replacement. MGSCA is applied to each bootstrap sample to obtain the estimates of parameters. Then, the bootstrapped standard errors of the estimates are calculated across entire bootstrap samples. The bootstrapped standard errors are used to assess the reliability of the estimates. The critical ratios (i.e., the parameter estimates divided by their standard errors) can be used to examine the significance of the parameter estimates (e.g., a parameter estimate having a critical ratio greater than two in absolute value is considered significant at .05 level).

3. Empirical Application: The ACSI Data

The present example consists of customer-level measures of the fourteen observed ACSI variables for thirteen American financial-services companies including banks and insurance companies. In other words, the observations of customers on the ACSI variables were nested within the companies. The number of customers who have used the services offered by one of the thirteen companies was 3096. The number of customers within each company ranged from 100 to 400 (the average number of customers per company = 238).

We considered a two-level GSCA model for this data, in which the loadings for customer-level (Level-1) observed variables and the path coefficients of customer-level latent variables were assumed to vary freely across companies (Level-2). More specifically, we applied the same two-level model as (8) to the data because no Level-2 or company-level characteristics were available from this data.

The specified two-level model provided an overall goodness of fit of .80, indicating that it accounted for about 80% of the total variance of all endogenous variables. Table 1 provides weight (w's) estimates for the observed variables in the ACSI model. It also presents the standard errors of the parameter estimates, calculated based on 100 bootstrapped samples.

Insert Table 1 about here

It was shown that the weight estimates for each latent variable were similar to each other and all turned out to be significant. This indicates that all observed variables contributed equally well to determining their latent variables in the model.

Table 2 presents the fixed (a_p) loadings for the observed variables. It was shown that the estimated loadings appeared high and were significant. This suggests that the latent variables seemed to be well constructed in that they accounted for a large portion of the variances of the observed variables. Table 2 also provides the fixed (b_k) path coefficient estimates and their standard errors. The interpretations of the fixed effects appear consistent with the relationships among the latent variables hypothesized in the ACSI model. That is, CE had significant and positive influences on PQ (b_1 = .49; s.e. = .02), PV (b_2 = .15; s.e. = .01), and CS (b_3 = .06; s.e. = .01). In turn, PQ showed significant and positive effects on PV (b_4 = .65; s.e. = .01) and CS (b_5 = .53; s.e. = .02). Also, PV exhibited a significant and positive effect on CS (b_6 = .40; s.e. = .02). In turn, CS had a significant and positive impact on CL (b_8 = .69; s.e. = .01), while a significant and negative impact on CC (b_7 = -.38; s.e. = .02) was also apparent. Finally, CC showed a significant and negative effect on CL (b_9 = -.08; s.e. = .02).

Insert Table 2 about here

Table 3 shows the variance estimate of each Level-2 random loading (λ_{pj}) across the companies. This variance estimate represents inter-group variability in the loading. It was found that the variances of the random loadings for all observed variables turned out to be significant. This suggests that there existed substantial company-wise differences in each of the loadings. Table 3 also exhibits the variance estimate of the Level-2 random path coefficient of each latent variable (u_{kj}). The variance estimates of all random path coefficients turned out to be significant, suggesting substantial differences in each of the path coefficients across the companies.

Insert Table 3 about here

To summarize, the proposed method was applied to two-level ACSI data, where customer-level measurements were clustered within different companies. Explicitly taking the nested structure into account, the proposed method provided fixed or average loadings between observed and latent variables in addition to fixed path coefficients between latent variables, specified in the ACSI model. The fixed effect estimates were shown to validate the relationships hypothesized in the ACSI model. Moreover, the proposed method allowed investigating company-wise invariance of the loadings and path coefficients by looking at the variance estimates of the corresponding random effects. Substantial inter-company variations of all loadings and path coefficients were revealed from this data. The single-level GSCA may also conduct the test of invariance through multi-group comparison analyses. However, this involves applying a separate ACSI model to each company. Thus, when the number of groups is large or the number of cases within each group is small, the single-level approach may become less attractive than the proposed method. Finally, the present example was not able to empirically illustrate the interaction effects of exogenous variables across different unit levels because no company-level explanatory variables were available from this data set.

4. Concluding Remarks

An extension of GSCA was proposed that takes into account the interdependence of hierarchical data. The proposed method allows for the modeling of fixed and random effects of both loadings and path coefficients in the path analytic model with latent variables. A straightforward least-squares algorithm was proposed for parameter estimation in the proposed method. Although this algorithm is more complicated than the original ALS algorithm for GSCA, it still appears computationally efficient as it has hardly suffered from slow or no convergence according to our experience. The usefulness of the proposed method was empirically demonstrated with customer-level customer satisfaction data clustered within different companies.

In spite of its important technical implications, MGSCA is not free from limitations. Most critical perhaps is the fact that it treats the random-effects (i.e., λ_{pj} and u_{kj}) as if they were fixed parameters. This is the usual approach in ANOVA (cf. Searle, 1971). In theory, this approach seems somewhat restrictive because it regards higher-level units as unique entities and focuses on the effects of the higher-level units on endogenous variables (Snijders & Bosker, 1999). Despite this narrow point of view regarding random effects, this approach is well suited to the distributional-free optimization procedure of the proposed method because it does not require any distributional assumptions on the random effects, e.g., *iid* normal. If this distributional assumption is unlikely to hold, the proposed approach can be a suitable choice. In this regard, it seemed appropriate to apply MGSCA to the ACSI data in the previous section because the assumption of normality is known to be almost always violated in customer satisfaction measures (Fornell, 1995; Anderson & Fornell, 2000). MGSCA can also emerge as a reasonable method when the number of a higher-level unit is small (say, $J \leq 10$) or the number of cases within the higher-level unit is relatively large (say, ≥ 100) because it is justifiable to treat random effects as fixed parameters in these cases (Snijders & Bosker, 1999). This condition was also satisfied in the ACSI data, in which the number of customers within each company was large.

A number of topics may be considered to enhance the capability of MGSCA. For instance, MGSCA is currently geared for the analysis of continuous variables. It may be effectively extended so as to deal with discrete variables through data transformations. In particular, the optimal scaling approach (Gifi, 1990; Young, 1981) is deemed promising because it can be readily coupled with the estimation algorithm of MGSCA. Moreover, MGSCA may be viewed as an *a priori* classification approach in the sense that it incorporates known information on group-level heterogeneity of respondents into the modeling of path-analytic model parameters. In many instances, the information on group-level heterogeneity can also be obtained by identifying clusters of respondents through the analysis of the data as in finite mixture models (McLachlan & Peel, 2000; Wedel & Kamakura, 1998). This *post-hoc* approach may be integrated into MGSCA for more sophisticated analyses. This extension may involve combining MGSCA with cluster analysis. Finally, MGSCA may serve to evolve new multilevel breeds of various extant multivariate techniques, for example, multilevel principal components analysis, multilevel canonical correlation analysis, etc, because their single-level counterparts can be viewed as special cases of GSCA (Hwang & Takane, 2004). All of these possibilities warrant further investigation and provide the fodder for future theoretical and empirical work.

Appendix: The Three-Level GSCA Model

Let $\mathbf{Z}_{js}$ denote a matrix of Level-1 observed variables in the j-th Level-2 unit nested within the s-th Level-3 unit $(s = 1, \cdots, S)$. Let $\mathbf{A}_{js}$ and $\mathbf{B}_{js}$ denote matrices of Level-1 loadings and path coefficients, respectively. The Level-1 model is given by

$$\mathbf{Z}_{js}\mathbf{V} = \mathbf{Z}_{js}\mathbf{W}[\mathbf{A}_{js}, \mathbf{B}_{js}] + \mathbf{E}_{js} . \quad \text{(A1)}$$

Next, let $\mathbf{Q}_{js}$ and $\mathbf{C}_{js}$ denote matrices of Level-2 exogenous observed variables for loadings and path coefficients, respectively. Let $\mathbf{A}_s$ and $\mathbf{B}_s$ denote matrices of Level-2 loadings and path coefficients, respectively. Let $\mathbf{\Lambda}_{js}$ and $\mathbf{U}_{js}$ denote matrices of Level-2 random loadings and path coefficients, respectively. The Level-2 model is given by

$$\begin{aligned} \mathbf{A}_{js} &= \mathbf{Q}_{js}\mathbf{A}_s + \mathbf{\Lambda}_{js} \\ \mathbf{B}_{js} &= \mathbf{C}_{js}\mathbf{B}_s + \mathbf{U}_{js} \end{aligned} . \quad \text{(A2)}$$

Also, let $\mathbf{Y}_s$ and $\mathbf{D}_s$ denote matrices of Level-3 exogenous variables for loadings and path coefficients, respectively. Let $\mathbf{A}$ and $\mathbf{B}$ denote matrices of Level-3 fixed loadings and path coefficients, respectively. Let $\mathbf{\Lambda}_s$ and $\mathbf{U}_s$ denote matrices of Level-3 random loadings and path coefficients, respectively. The Level-3 model is given by

$$\begin{aligned} \mathbf{A}_s &= \mathbf{Y}_s\mathbf{A} + \mathbf{\Lambda}_s \\ \mathbf{B}_s &= \mathbf{D}_s\mathbf{B} + \mathbf{U}_s \end{aligned} . \quad \text{(A3)}$$

The model in combined form is then as follows:

$$\begin{aligned} \mathbf{Z}_{js}\mathbf{V} &= \mathbf{Z}_{js}\mathbf{W}[\mathbf{Q}_{js}\mathbf{Y}_s\mathbf{A} + \mathbf{Q}_{js}\mathbf{\Lambda}_s + \mathbf{\Lambda}_{js}, \mathbf{C}_{js}\mathbf{D}_s\mathbf{B} + \mathbf{C}_{js}\mathbf{U}_s + \mathbf{U}_{js}] + \mathbf{E}_{js} \\ &= \mathbf{Z}_{js}\mathbf{W}[\mathbf{Q}_{js}\mathbf{Y}_s, \mathbf{Q}_{js}, \mathbf{I}, \mathbf{C}_{js}\mathbf{D}_s, \mathbf{C}_{js}, \mathbf{I}] \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_s & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \\ \mathbf{\Lambda}_{js} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{U}_{js} \end{bmatrix} + \mathbf{E}_{js}. \end{aligned} \quad \text{(A4)}$$

References

Anderson, E. W., & Fornell, C. (2000). Foundations of the American customer satisfaction index. *Total Quality Management*, 11, 869-882.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: applications and data analysis methods*. Newbury Park: Sage Publications.

Curran, P. J. (1998). Introduction to hierarchical linear models of individual growth: An applied example using the SAS data system. Paper presented at the first international institute on developmental science, University of North Carolina, Chapel Hill.

de Leeuw, J., Young, F. W., & Takane, Y. (1976). Additive structure in qualitative data: An alternating least squares method with optimal scaling features. *Psychometrika*, 41, 471-503.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Philadelphia: SIAM.

Fornell, C. (1995). The quality of economic output: Empirical generalizations about its distribution and association to market share. *Marketing Science*, 14, 203-211.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996), The American customer satisfaction index: Nature, purpose, and findings. *Journal of Marketing*, 60, 7-18.

Gifi, A. (1990). *Nonlinear multivariate analysis*. Chichester: Wiley.

Hwang, H., & Takane, Y. (2004), Generalized structured component analysis. *Psychometrika, 69, 81-99*.

Lohmöller, J. B. (1989). *Latent variable path modeling with partial least squares*. New York: Springer-Verlag.

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: John Wiley & Sons.

Searle, S. R. (1971). *Linear models*. New York: John Wiley & Sons, Inc.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling.* London: Sage Publications.

Wedel, M., & Kamakura, W. A. (1998). *Market segmentation: conceptual and methodological foundations*. Boston: Kluwer Academic Publishers.

Wold, H. (1966). Estimation of principal components and related methods by iterative least squares. In P. R. Krishnaiah (Ed.), *Multivariate analysis* (pp. 391-420). New York: Academic Press.

Wold, H. (1973). Nonlinear iterative partial least squares (NIPALS) modeling: Some current developments. In P. R. Krishnaiah (Ed.), *Multivariate analysis* (pp. 383-487). New York: Academic Press.

Wold, H. (1975). Path models with latent variables: the NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, and V. Cappecchi (Eds.), *Quantitative sociology: International perspectives on mathematical statistical model building* (pp. 307-357). New York: Academic Press.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog and H. Wold (Eds.), *Systems under indirect observations: Causality, structure, prediction, Part 2* (pp. 1-54). Amsterdam: North-Holland.

Young, F. W. (1981). Quantitative analysis of qualitative data. *Psychometrika*, 46, 357-388.

Table 1. The weight estimates and their standard errors (in the parentheses) in the ASCI model.

w_1	.45	(.01)
w_2	.47	(.01)
w_3	.35	(.01)
w_4	.40	(.00)
w_5	.40	(.00)
w_6	.35	(.00)
w_7	.52	(.00)
w_8	.52	(.00)
w_9	.38	(.00)
w_{10}	.37	(.00)
w_{11}	.36	(.00)
w_{12}	1.00	
w_{13}	.53	(.00)
w_{14}	.53	(.00)

Table 2. The Level-2 fixed loading and path coefficient estimates and their standard errors (in the parentheses) in the ASCI model.

Level-2 fixed loading	a_1	.83	(.01)
	a_2	.86	(.01)
	a_3	.66	(.02)
	a_4	.90	(.01)
	a_5	.90	(.00)
	a_6	.80	(.01)
	a_7	.96	(.00)
	a_8	.96	(.00)
	a_9	.93	(.00)
	a_{10}	.91	(.00)
	a_{11}	.89	(.01)
	a_{12}	1.00	
	a_{13}	.94	(.00)
	a_{14}	.95	(.00)
Level-2 fixed path coefficient	b_1	.49	(.02)
	b_2	.15	(.01)
	b_3	.06	(.01)
	b_4	.65	(.01)
	b_5	.53	(.02)
	b_6	.40	(.02)
	b_7	-.38	(.02)
	b_8	.69	(.01)
	b_9	-.08	(.02)

Table 3. The variance estimates of the Level-2 random loadings and random path coefficients and their standard errors (in the parentheses) in the ASCI model.

Variance of level-2 random loading	λ_{1j}	7.99 (3.09)
	λ_{2j}	13.37 (4.37)
	λ_{3j}	16.82 (5.90)
	λ_{4j}	5.57 (2.02)
	λ_{5j}	6.26 (2.02)
	λ_{6j}	8.45 (3.36)
	λ_{7j}	1.75 (.78)
	λ_{8j}	1.78 (.81)
	λ_{9j}	4.62 (1.75)
	λ_{10j}	4.12 (1.91)
	λ_{11j}	4.33 (1.74)
	λ_{12j}	-
	λ_{13j}	4.52 (1.42)
	λ_{14j}	4.39 (1.38)
Variance of level-2 random path coefficient	u_{1j}	.011 (.004)
	u_{2j}	.008 (.003)
	u_{3j}	.005 (.001)
	u_{4j}	.007 (.003)
	u_{5j}	.011 (.004)
	u_{6j}	.010 (.003)
	u_{7j}	.008 (.003)
	u_{8j}	.008 (.003)
	u_{9j}	.015 (.006)

Figure 1: A Level-1 ACSI model