Title

Bias in Factor Score Regression and a Simple Solution

Authors
Hoshino, Takahiro
Bentler, Peter M.

Publication Date
2011-02-12

Peer reviewed

Bias in Factor Score Regression and a Simple Solution

Takahiro Hoshino
Department of Economics and Business Administration,
Nagoya University

Peter M. Bentler
Departments of Psychology and Statistics,
University of California, at Los Angeles

TAKAHIRO HOSHINO

Department of Economics and Business Administration,

Nagoya University,

Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan

E-mail: bayesian@jasmine.ocn.ne.jp

Abstract

Not only in social sciences but also in chemometrics, the paths (e.g., regression coefficients, correlations) between latent variables are often estimated by regarding the estimated latent variable scores as observed variables. Such methods are often called “factor score regression”. Recently partial least square (PLS) path modeling is used for the same purpose. Similarly, the latent variable scores estimated for discrete observed variables are used to infer relationships between the latent variables and other external variables. It is widely known that such estimators are generally biased.

In this paper, we investigate theoretically why factor score regression estimators are generally biased, and we propose a new method for estimation of paths between latent variables using the estimated latent variables. We prove that the proposed estimators are consistent for continuous indicators. We also show in the simulation studies that our method greatly reduces bias when using these estimated latent variable scores for discrete indicators.

Keywords: Structural Equation Modeling, Item Response Theory, Factor Analysis, Partial Least Squares, Item Parceling

1 Introduction

In social sciences, the main interest of research often lies in quantitive expression of relationships between latent variables that represent (psychological or sociological) constructs. Structural equation modeling (SEM) is suitable for inference and description of these relationships. As an important extension of Spearman's early ideas on factor analysis (Bartholomew, 2007), and built especially upon the many early contributions of Jöreskog (e.g., 1977), today SEM includes a variety of statistical models that are highly relevant to psychological research (e.g., Bauer & Curran, 2004; Bollen, 2002; Bollen & Curran, 2006; Lee, 2007; Yuan & Bentler, 2007). It is well known that SEM is frequently applied in the psychological (e.g., MacCallum & Austin, 2000) and related sciences (e.g., Hays, Revicki, & Coyne, 2005).

In contrast to simultaneous ML estimation of measurement and structural relations in SEM, many researcher prefer to estimate the latent variable scores and then use traditional regression methods to determine their relations.

In this paper we investigate the reason why the estimators of parameters regarding the relationships between latent variables (i.e., correlations between latent variables) are biased, when the estimated latent variable scores are used as if they are observed. We also propose a simple, valid estimation method using the estimated latent variable scores.

We consider a situation where two or more blocks of variables are observed on the same subjects. For example, consider a very simple SEM model depicted in Figure 1 (two blockes of variables and four variables for each block). There are several variants of estimation methods using estimated factor scores, but generally a stepwise estimation method can be divided into three steps:

(1) Estimate the parameters in the measurement part of the model (submodel 1 and 2 in Figure 1).

(2) Estimate the factor scores in exploratory (or confirmatory) factor analysis using the estimated parameters obtained in the step (1).

(3) Perform regression analysis, path analysis or factor analysis (submodel 3) where the variables of the model are the estimated factor scores.

The above procedure to estimate regression coefficients is usually called "factor score regression" because the estimated factor scores are used in regression analysis, or a series of regressions, as if they were observed variables. For examples, see e.g., Gass (1996) and Zammuner (1998). In the latter section we will consider the case where the observed variables are discrete and the latent variable scores are estimated using Item Response Theory (IRT).

Controversies About Structural Equation Modeling

Although simultaneous estimation in SEM can avoid the need for factor score regression, this approach may not be feasible. Many researchers have pointed out that there remain some problems in the practical application of SEM to data, especially when mixed continuous and ordinal variables are used, models are large, and simultaneous estimation methods such as maximum likelihood or least squares methods are used. These issues will be summed up with the following four points:

(1) The intractability of simultaneous estimation methods when continuous and discrete variables are mixed.

If all observed variables are continuous, then simultaneous estimation methods only require estimates of moments, i.e., the sample mean vector and sample covariance matrix. However, including binary or ordinal variables also requires considering response patterns as is typically done in item response theory (IRT). A wedding of SEM and IRT methods is being developed. The generalized linear and nonlinear methodologies described in De Boeck and Wilson (2004) provide a partial approach because they allow the prediction of IRT item parameters and person parameters by external variables. The generalized linear latent and mixed (GLLAMM) modeling framework of Skrondal and Rabe-Hesketh (2004; Rabe-Hesketh, Skrondal, & Pickles, 2004) also is promising since it allows IRT/SEM combinations theoretically by a unification and extension of multilevel and latent variable models to allow latent variable structural equation models in the context of measurement models that permit a wide range of link functions and variable types. However, their approach requires numerical integration and calculation of the likelihood, which is difficult to impossible when the model is complex or the number of variables is large. Thus “estimation can be quite slow, especially if there are several random effects” (Rabe-Hesketh & Skrondal, 2005, p. 128). Since computational time is proportional to number of cases and the square of number of parameters, this methodology is not yet useful for larger models. The models considered by Lee and Song (2003) are similarly general in that they allow nonlinear relations among normal latent variables and relation of latent to observed variables, along with a variety of indicator types (e.g., binary, ordinal). Their approach to model estimation and evaluation is based on Bayesian methodology including the Gibbs sampler and MCMC. An advantage of their approach is that greater precision can be expected in smaller samples, but estimation and evaluation requires accepting prior distributions on parameters as well as a computational burden that is as large as that of GLLAMM. Thus although there has been an important recent effort to joining IRT and SEM, the summary of Moustaki, Jöreskog, and Mavridis still holds: “On the other hand, IRT models have been developed recently and there is no flexible software available for fitting those models. If one wants to fit a model with many factors, one will probably have to use LISREL, Mplus or EQS” (2004, p. 507).

Unfortunately, while methods such as Muthén’s (1984) and Lee, Poon and Bentler’s (1995) approaches based on polychoric and polyserial correlations are applicable to the case where there are continuous and ordered categorical observed variables, and they are easy to implement in Mplus and EQS, these methods are not applicable when nominal observed variables are included in the model. In theory, simultaneous estimation such as that based on maximum likelihood estimation is applicable to SEM with mixed continuous and nominal observed variables (Moustaki & Knott, 2000; Sammel, Ryan & Legler, 1997). But again, simultaneous estimation methods are generally very difficult for researchers to use in applied areas because the evaluation of the likelihood requires numerical integration which becomes impractical in reasonably sized models.

(2) The effect of structural relation parameter estimates on the estimation of the measurement model.

It has been known for a long time (e.g., Kumar & Dillon, 1987) that partial misspecification in a model causes large biases in the estimates of several free model parameters in SEM. In addition, Burt (1973; 1976) and Hunter & Gerbing (1982) noted that simultaneous estimation of parameters in the measurement and structural portions of a model could lead to bias. They recommended estimating these two sets of parameters separately. As an example, consider a typical SEM with multiple indicators. In most cases, the meaning of the latent variables or factors in the model is determined from the measurement model, that is, the estimated relationships between observed indicators and factors. However, when a simultaneous estimation method, such as maximum likeihood or generalized least squares, is used, the parameterization of the structural relations may greatly affect the estimates in the measurement equations. To avoid such confounding, Anderson & Gerbing (1988) proposed a two-step model checking procedure that first confirms the measurement model with a saturated structural model so that the structural relations have no impact on the measurement model; then, with an appropriate measurement model, the researcher's substantive structural relations model of interest is added. Of course, this procedure does not solve the problem in that the estimates in the measurement model, and hence the meaning of the factors, can change greatly when other factors (and their indicators) are added to the model.

Simultaneous estimation in SEM has more arguable properties than might be found in comparable econometric simultaneous equation estimation, because in models with observed variables only there can be no confounding with latent variables such as are used in SEM. In econometrics, some stepwise estimation methods, such as limited maximum likelihood (ML) estimation or two stage least squares estimation, have been proposed for estimating each measurement equation independently. Bollen (1996) has been an advocate for using a general version of this idea in latent variable models, proposing that this would allow isolating specification errors in particular parts of the model. Nonetheless, he did not develop a special procedure for isolating structural versus measurement model misspecifications. The most complete theoretical development of two-stage estimation in SEM was given by Yuan and Chan (2002). They proposed a model segregation approach where a model is completed by a second set of parameters contingent on the estistence of a first set of estimated parameters, and provided a detailed statistical development. Their approach could be applied so that the measurement model is completely estimated at the first stage, and the latent structural relations, estimated at the second stage, could not influence the meaning of the factors. However, these developments in two-stage estimation have not considered the mixture of continuous and ordinal variables with which we are concerned. In particular, they have not been developed to allow the use of item response theory type measurement models along with latent variable linear relations models. A statistically sound stepwise estimation method to accomplish this purpose would be useful in social sciences.

(3) The need for data reduction.

The number of observed variables can be very large, e.g., Ashton, Lee, and Goldberg (2004) studied 1,710 English personality-descriptive adjectives, and one can envision circumstances where the number of latent factors extracted also can be quite large. While the number of latent variables will be far less than the number of observed variables, especially when psychological scales or test items are used as indicators, in such situations it would be desirable to first estimate the factors and then to do all subsequent analyses on the factor score estimates alone. These estimated latent variables would substantially reduce data size and facilitate data handling as well as modeling. However, modeling with estimated factor scores can lead to biased conclusions about the latent structural relations unless the methodology within which such factor score estimation is embedded can eliminate any sources of bias.

(4) The existence of improper solutions.

It is well known that unless a sufficient number of good indicators of each factor is available, improper solutions (negative variance estimates, also known as Heywood cases) can occur frequently. Jöreskog (1967) reported that 9 out of 11 classical data sets possessed improper solutions, and Anderson and Gerbing (1984) reported that with correct models, their simulation study found that 24.9% of replications had improper solutions. An important consequence is that test statistics no longer have their assumed distributions and model evaluation becomes difficult (see, e.g., Stoel, Garre, Dolan, & vanden Wittenboer, 2006). There is an extensive literature on the reasons for improper solutions (e.g., Chen, Bollen, Paxton, Curran, & Kirby, 2001), but it is also known that “factor score regression” can reduce the occurence of improper solutions through its division and separate treatment of the two parts of a complete model: the measurement model and the structural relations model.

Can Factor Score Regression/PLS/Item Parceling resolve the problems?

In order to avoid the problems stemming from simultaneous estimation in SEM, a frequently used method is a kind of stepwise estimation using estimated factor scores. A methodology such as this is encouraged by statistical packages such as SPSS FACTOR that make it easy to save “factor scores” or component scores for use in subsequent analyses. However, since such “factor scores” are linear combinations of variables, they contain error and are biased. Hence factor score regression procedures using such scores produce biased estimates, usually underestimating the relationship between factors (see section 3).

In an important paper, Skrondal & Laake (2001) pointed out the bias problem and proposed a modified version of the above estimation procedure especially for regression analysis between latent variables. Their proposed method (i) estimates the factor scores for dependent factors by the Bartlett method, (ii) estimates the factor scores for independent factors by the regression method (we review these factor score estimates below; see also Yanai & Ichikawa, 2007, pp. 287-289), and (iii) uses these estimated factor scores as if they were observed variables. They demonstrated the consistency of their proposed method, but the method has some disadvantages:(1) The method is not available for models that have more than three groups of factors (and indicators), (2) Independent factors and dependent factors must be pre-specified before the analysis, (3) Their method underestimates the correlations between factors, (4) Their method is not available when there are discrete variables in the model, (5) Their method neglects the effect of estimation of parameters of the factor models, and (6) they have not extended the theory to deal with higher-order factors. Their paper provided a solution for a certain (but restricted) case, but they did not clarify the cause of bias arising from using estimated factor scores.

A second class of methods that estimate factor scores is that based on partial least squares (PLS). Originally developed as an alternative to maximum likelihood emphasizing prediction and "soft modeling" rather than statistical efficiency and confirmatory modeling (Wold, 1973, 1982), it proposes to use estimates of latent factors based on blocks of variables using simple weighting schemes along with optimal regression prediction of these latent variables across blocks of variables. "PLS methods actually enrich a causal scheme with data analysis features and can be directly used in situations where the classical ordinary least squares (OLS) criterion for regression and the maximum likelihood estimation for structural equation models are not feasible because of critical situations in the data, such as too many variables or too few observations or a too strong correlation between variables or the presence of missing data" (Vinzi & Lauro, 2005, p. 1). As noted by McDonald (1996), however, "The PLS methods are difficult to describe and extremely difficult to evaluate partly because PLS constitutes a set of ad hoc algorithms that have generally not been formally analyzed, or shown to possess any clear global optimizing properties (except in the well understood case of just two composites), and partly because these devices are represented as a form of path analysis with latent variables, and it can be difficult to determine what properties of latent variable models they possess, if any" (p. 240). This critique has become somewhat less cogent as the methodology is vibrant and continues to be developed in algebraic and algorithmic ways (e.g., Hwang & Takane, 2004; a dozen articles in the January 2005 issue of *Computational Statistics & Data Analysis*). Certainly, PLS and ML-based estimates of latent factors can be quite close, as can be coefficients (e.g., Tenenhaus et al., 2005). However, as far as we can tell, these methods have the same drawback as factor score regression as described above. That is, PLS estimates will be correct only under the joint conditions of consistency (sample size becomes large) and consistency at large (the number of indicators per latent variable becomes large; Jöreskog and Wold, 1982). In practice, the correlations between the latent variables will tend to be underestimated (Dijkstra, 1983).

Actually, McDonald (1996) developed six alternative factor score regression methods as variants of PLS that optimized specific criteria. However, their statistical properties also were not developed.

A drawback for our purposes is that, generally speaking, these variants of factor score regression also do not extend to models with categorical indicators.

A third and popular method related to factor score regression is the "Item Parceling" methodology. Item Parceling involves summing or averaging item scores from two or more items and using these parcel scores (or "scale score" in personality psychology) as observed "latent variable scores" to estimate the relationships between latent variables (Bandalos, 2002). In contrast to PLS, these composite scores are fixed and not interatively updated. The rationale for the use of item parcels in SEM is as follows (Bandalos & Finney, 2001): (1) The reliability of item parcels will be greater than the raw scales (Cattell & Burdsal, 1975; Kishton & Widaman, 1994), (2) Even when the data contain raw items that are nonnormally distributed or/and coarsely categorized, item parcels based on a large number of items often can be regarded as normally distributed, and normal theory maximum likelihood and generalized least squares estimation techniques are applicable to such data, (3) Item parceling can reduce the number of variables in the analysis, thus also reducing the ratio of variables to subjects, which will lead to more stable estimates, and (4) Item parceling will typically lead to better model fit than estimation using the raw items (Thompson & Melancon, 1996). In spite of these advantages, item parceling has been criticized. There are at least two problems: (1) The resulting parameter estimates are sometimes biased, and then typically they are underestimated (Bandalos, 2002), and (2) The item parceling method does not always produce stable estimates (MacCallum, Widaman, Zhang & Hong, 1999; Marsh, Hau, Balla & Grayson, 1998). Although there are some theoretical results (e.g., Yuan, Bentler, & Kano, 1997), most conclusions on this methodology are mainly due to simulation studies. Further theoretical analyses on item parceling and factor score regression are still needed.

Content of the paper

In section 2, the model assumptions are made. These assumptions appear to be slightly restrictive: however, they are the same for factor score regression or item parceling. In section 3, we discuss the theoretical investigation of the sources of biases in factor score regression (also in PLS and the item parceling method). For notational convenience, we restrict our attention to a relatively simple model: however, the results apply to general cases. In section 4, we propose an alternative estimation method based on the estimated latent variables. The consistency of the proposed method is also shown. In section 5, we provide simulation studies in various model setup to justify the validity of the proposed method when the number of subjects and observed variables are finite. The concluding remarks and discussions are provided in the last section.

2 Model

We assume the model setup which is usually made in factor score regression and the item parceling method: Consider $J > 1$ measurement equations in which each equation measures different latent variables.

In the j-th measurement model, each ($P_j \times 1$) observed variable vector $\boldsymbol{x}_j$ is independently defined in terms of the Q_j-component factor vector f_j, and the P_j-component error vector $\boldsymbol{e}_j$ by the j-th measurement model,

$$\boldsymbol{x}_j = g_j(\boldsymbol{f}_j) + \boldsymbol{e}_j \quad (j = 1, \cdots, J), \tag{1}$$

where $g_j(\cdot)$ is a linear or nonlinear function.

The distribution of $\boldsymbol{x}_j$ in the j-th measurement model is also defined independently of the other measurement models and the structural model.

Measurement models vary with the level of measurement. If each element of $\boldsymbol{x}_j$ is a continuous variable, the j-th measurement equation (Eqn.(1)) is usually expressed as a linear factor analysis model, as below:

$$\boldsymbol{x}_j = \boldsymbol{\alpha}_j + \boldsymbol{\Lambda}_j f_j + \boldsymbol{e}_j, \quad \mathcal{E}(\boldsymbol{e}_j) = 0 \ \ and \ \ Var(\boldsymbol{e}_j) = \boldsymbol{\Psi}_j, \tag{2}$$

where $\boldsymbol{e}_j$ follows the multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Psi}_j$: and $\boldsymbol{\Lambda}_j$ is the factor loading matrix.

If $\boldsymbol{x}_j$ is dichotomous, the three-parameter logistic item response model (Lord & Novick, 1968;

Embretson & Reise, 2000) is employed:

$$Pr(\boldsymbol{x}_{jk} = 1) = c_{jk} + \frac{c_{jk}}{1 + \exp(-Da_{jk}(f_j - b_{jk}))} \tag{3}$$

where a_{jk}, b_{jk}, and c_{jk} are the k-th item parameters of $\boldsymbol{x}_j$. Sometimes, the probit item response model is also employed.

If $\boldsymbol{x}_j$ is nominal or polytomous, the nominal response model proposed by Bock (1972) or the graded response model proposed by Samejima (1969) is also available.

We further assume that the joint distribution of factor vectors, $p(\boldsymbol{f}_1, \cdots, \boldsymbol{f}_J | \boldsymbol{\xi}_F)$ follows multivariate normal distribution, where $\boldsymbol{\xi}_F$ denotes the parameter vector.

Therefore, the joint distribution of $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_J$ is

$$p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_J \mid \boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_J, \boldsymbol{\xi}_F) = \int \cdots \int \Big\{ \prod_{J=1}^{J} p_j(\boldsymbol{x}_j \mid \boldsymbol{f}_j, \boldsymbol{\xi}_j) \Big\} p(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_J \mid \boldsymbol{\xi}_F) \prod_{J=1}^{J} d\boldsymbol{f}_j \tag{4}$$

where $p_j(\boldsymbol{x}_j \mid \boldsymbol{f}_j, \boldsymbol{\xi}_j)$ denotes the conditional distribution of $\boldsymbol{x}_j$ with the given value of $\boldsymbol{f}_j$, and $\boldsymbol{\xi}_j$ is the parameter vector in the conditional distribution.

If $J = 2$, the entire model (Eqn.(4)) is equivalent to the LISREL model proposed by Jöreskog (1970). If $J > 2$, the entire model can be considered as a "multiple indicator model," a submodel of SEM.

$\boldsymbol{\xi}_F$ contains the mean vector and the covariance matrix of factors,

$$\boldsymbol{\nu} = \begin{pmatrix} \boldsymbol{\nu}_1 \\ \boldsymbol{\nu}_2 \\ \vdots \\ \boldsymbol{\nu}_J \end{pmatrix}, \qquad \boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\Phi}_{11} & \boldsymbol{\Phi}_{12} & \cdots & \boldsymbol{\Phi}_{1J} \\ \boldsymbol{\Phi}_{21} & \boldsymbol{\Phi}_{22} & \cdots & \boldsymbol{\Phi}_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Phi}_{J1} & \boldsymbol{\Phi}_{J2} & \cdots & \boldsymbol{\Phi}_{JJ} \end{pmatrix}, \tag{5}$$

where $\boldsymbol{\nu}_j$ is the mean vector and $\boldsymbol{\Phi}_{jk}$ is the covariance matrix of $\boldsymbol{f}_j$ and $\boldsymbol{f}_k$.

Usually, the concern is not the covariance matrix of factors, but the parameters of the structural equation. In this study, the objective of inference is to estimate parameters in the structural part of the model, τ. The mean vector and the covariance matrix of factors, $\boldsymbol{\nu}$ and $\boldsymbol{\Phi}$, are structured by $\boldsymbol{\tau}$, as $\boldsymbol{\nu}(\boldsymbol{\tau})$ and $\boldsymbol{\Phi}(\boldsymbol{\tau})$.

Random Effect Model and Fixed Effect Model

For clarification, we define "random effect model" and "fixed effect model" as follows. The j-th measurement model is called a "random effect model" if the factor scores follow multivariate normal distribution. The distribution of $\boldsymbol{x}_j$ can be expressed as follows:

$$p(\boldsymbol{x}_j \mid \boldsymbol{\xi}_j, \boldsymbol{\xi}_{F_j}) = \int p(\boldsymbol{x}_j \mid \boldsymbol{f}_j, \boldsymbol{\xi}_j) p(\boldsymbol{f}_j \mid \boldsymbol{\xi}_{F_j}) d\boldsymbol{f}_j. \tag{6}$$

where $\boldsymbol{\xi}_{F_j}$ is the parameter vector of the marginal distribution of $\boldsymbol{f}_j$ (i.e., the factor mean vector $\boldsymbol{\nu}_j$ and factor covariance matrix $\boldsymbol{\Phi}_{jj}$).

The j-th measurement model is called a "fixed effect model" if the factor scores are not random variants but incidental parameters (Neyman & Scott, 1948). The distribution of $\boldsymbol{x}_j$ can be expressed as:

$$p(\boldsymbol{x}_j \mid \boldsymbol{f}_j, \boldsymbol{\xi}_j). \tag{7}$$

$\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_J, \boldsymbol{\xi}_F$ are usually called the "structural parameters," as distinguished from incidental parameters, $\boldsymbol{f}$. If we employ the random effect model, the objects of the inference are not incidental parameters but structural parameters. In the fixed effect model, both the structural and incidental parameters are usually estimated simultaneously.

It should be noted that if there are large number of subjects and the incidental parameters follow some distribution, these two models are virtually the same (see Kiefer & Wolfowitz, 1956; Lindsay, Clogg & Grego, 1991).

Further, we employed the fixed effect model and assumed that there are a large number of subjects and that the incidental parameter vector $\boldsymbol{f}$ follows multivariate normal distribution. Henceforth we refer to the employed model as the "fixed effect Kiefer & Wolfowitz type model."

In this paper, it is also assumed that each random effect measurement model is identifiable, and the structural parameters can be estimated.

3 Bias of Estimators due to the Estimated Factor Scores: Factor Analysis model

This section discusses why the estimates obtained using factor score regression are generally biased.

To be more concrete, we restate the factor score regression procedure as follows:

(1) Employ the random effect model (Eqn.(6)) and obtain ML estimate structural parameters in each measurement model $\boldsymbol{\xi}_j, \boldsymbol{\xi}_{F_j}$. (For example, in the model depicted in Fiugre 1, regard submodel 1 and 2 as the random effect model and estimate the parameters in these models.)

(2) Fix the parameters at the estimates, and then estimate the factor scores for each subject in each measurement model. (Estimate factor score of f_X and that of f_Y independently.)

(3) Regard the estimated factor scores as the observed variables, then estimate the parameters regarding the relationships between the factors (i.e., regression coefficients, correlations or factor loadings for higher factors). (Estimate parameters in submodel 3 by regarding factor scores of f_X abd f_Y as if they were observed.)

This procedure includes two sources of bias: (i) neglect of the uncertainty of the estimators of the structural parameters in the first step, and (ii) overestimation of the variances of factors in the third step. In this section, we focus our attention on the latter source of bias. Let the parameters of the measurement part be known. The effects of the estimated structural parameters are not considered in this section; please see the appendix for this (Proof of Proposition 3).

For the purpose of our demonstration, we assume two measurement models in which each factor vector is measured by some observed continuous indicators. Each measurement model is expressed as a factor analysis model (Eqn.(2)) :

$$\boldsymbol{x}_j = \boldsymbol{\Lambda}_j \boldsymbol{f}_j + \boldsymbol{e}_j, \quad \boldsymbol{f}_j \sim N(\boldsymbol{\nu}_j, \boldsymbol{\Phi}_{jj}) \ \ and \ \ \boldsymbol{e}_j \sim N(0, \boldsymbol{\Psi}_j), \quad j = 1, 2. \tag{8}$$

For notational simplicity, $\boldsymbol{\alpha}_j$ was fixed at zero for each measurement model. Let $\boldsymbol{\Phi}_{12}$ be the covariance between $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$.

There are several estimation methods for factor scores in the factor analysis model. For illustrative purposes, two representative estimators are considered here: (1) Bartlett's method and (2) the regression method. The former can be considered the ML estimator of the factor score vector in the fixed effect factor analysis, while the latter can be regarded as the Bayes posterior mean estimator. These two estimators as well as other estimators are expressed as the product of a matrix and the observed variable vector.

Let $\hat{\boldsymbol{f}}_1 = \boldsymbol{A}_1\boldsymbol{x}_1$ be the estimate of $\boldsymbol{f}_1$ and $\hat{\boldsymbol{f}}_2 = \boldsymbol{A}_2\boldsymbol{x}_2$ be the estimate of $\boldsymbol{f}_2$. Therefore, $Cov(\boldsymbol{x}_1, \boldsymbol{x}_2) = \boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_2^t$ and the joint distribution of the estimated factor follow multivariate normal distribution:

$$\begin{pmatrix} \hat{\boldsymbol{f}}_1 \\ \hat{\boldsymbol{f}}_2 \end{pmatrix} \sim N(\begin{pmatrix} \boldsymbol{A}_1\boldsymbol{\mu}_1 \\ \boldsymbol{A}_2\boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{A}_1(\boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_1^t + \boldsymbol{\Psi}_1)\boldsymbol{A}_1^t & \boldsymbol{A}_1(\boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_2^t)\boldsymbol{A}_2^t \\ \boldsymbol{A}_2(\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{21}\boldsymbol{\Lambda}_1^t)\boldsymbol{A}_1^t & \boldsymbol{A}_2(\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{22}\boldsymbol{\Lambda}_2^t + \boldsymbol{\Psi}_2)\boldsymbol{A}_2^t \end{pmatrix}). \tag{9}$$

Using the well known relationship between the regression model and the multivariate normal distribution, the expectation of the regression coefficitent is as follow:

$$Cov(\hat{\boldsymbol{f}}_1, \hat{\boldsymbol{f}}_2) \times V(\hat{\boldsymbol{f}}_2)^{-1} = \boldsymbol{A}_2(\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{21}\boldsymbol{\Lambda}_1^t)\boldsymbol{A}_1^t\big(\boldsymbol{A}_2(\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{22}\boldsymbol{\Lambda}_2^t + \boldsymbol{\Psi}_2)\boldsymbol{A}_2^t\big)^{-1} \tag{10}$$

The expectation of the correlation matrix between $\hat{\boldsymbol{f}}_1$ and $\hat{\boldsymbol{f}}_2$ is

$$\big(\boldsymbol{A}_1(\boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_1^t + \boldsymbol{\Psi}_1)\boldsymbol{A}_1^t\big)^{-1/2}\boldsymbol{A}_2(\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{21}\boldsymbol{\Lambda}_1^t)\boldsymbol{A}_1^t\big(\boldsymbol{A}_2(\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{22}\boldsymbol{\Lambda}_2^t + \boldsymbol{\Psi}_2)\boldsymbol{A}_2^t\big)^{-1/2} \tag{11}$$

If Bartlett's estimators are used to calculate each factor score,

$$\boldsymbol{A}_j = (\boldsymbol{\Lambda}_j^t\boldsymbol{\Psi}_j^{-1}\boldsymbol{\Lambda}_j)^{-1}\boldsymbol{\Lambda}_j^t\boldsymbol{\Psi}_j^{-1} \quad (j = 1, 2). \tag{12}$$

If regression estimators are used to calculate each factor score,

$$\boldsymbol{A}_j = \boldsymbol{\Phi}_{jj}\boldsymbol{\Lambda}_j^t\boldsymbol{\Sigma}_j^{-1} = (\boldsymbol{\Phi}_{jj}^{-1} + \boldsymbol{\Lambda}_j^t\boldsymbol{\Psi}_j^{-1}\boldsymbol{\Lambda}_j)^{-1}\boldsymbol{\Lambda}_j^t\boldsymbol{\Psi}_j^{-1} \quad (j = 1, 2). \tag{13}$$

The expectation of the estimator of the mean, the regression coefficient, the correlaton matrix, and the covariance matrix is expressed in Table 1.

Note that in this section we neglect the fact that the true values of the structural parameters such as factor loadings are unknown and estimated. Therefore, the equivalence of the parameter and its

Table 1: Estimates by various estimation methods using factor scores

Parameter	True	Bartlett
mean of $\boldsymbol{f}_1$	$\boldsymbol{\nu}_1$	$\boldsymbol{\nu}_1$
mean of $\boldsymbol{f}_2$	$\boldsymbol{\nu}_2$	$\boldsymbol{\nu}_2$
Regression Coefficient	$\boldsymbol{\Phi}_{12}\boldsymbol{\Phi}_{22}^{-1}$	$\boldsymbol{\Phi}_{12}(\boldsymbol{\Phi}_{22} + (\boldsymbol{\Lambda}_2^t\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{\Lambda}_2)^{-1})^{-1}$
Covariance matrix	$\boldsymbol{\Phi}_{12}$	$\boldsymbol{\Phi}_{12}$
Parameter	**Regression**	**Skrondal & Laake**
mean of $\boldsymbol{f}_1$	$\boldsymbol{\Phi}_{11}\boldsymbol{\Lambda}_1^t\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\nu}_1$	$\boldsymbol{\Phi}_{11}\boldsymbol{\Lambda}_1^t\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\nu}_1$
mean of $\boldsymbol{f}_2$	$\boldsymbol{\Phi}_{22}\boldsymbol{\Lambda}_2^t\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\nu}_2$	$\boldsymbol{\nu}_2$
Regression Coefficient	$\boldsymbol{\Phi}_{11}(\boldsymbol{\Lambda}_1^t\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Lambda}_1)\boldsymbol{\Phi}_{12}\boldsymbol{\Phi}_{22}^{-1}$	$\boldsymbol{\Phi}_{12}\boldsymbol{\Phi}_{22}^{-1}$
Covariance matrix	$\boldsymbol{\Phi}_{11}\boldsymbol{\Lambda}_1^t\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_2^t\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{22}$	$\boldsymbol{\Phi}_{11}\boldsymbol{\Lambda}_1^t\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_2^t\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Lambda}_2\boldsymbol{\Phi}_{22}$

expectation does not mean unbiasedness, but consistency (see proof of Proposition 1). As seen in Table 1, the bias does not disappear even when the number of subjects goes infinity. See simulation studies in Section 5.

The regression method does not always produce underestimated regression coefficients or correlations, but it usually does, For example, suppose $\boldsymbol{\Phi}_{11} = \boldsymbol{I}, \boldsymbol{\Phi}_{22} = \boldsymbol{I}$. Then, the expectation of the estimated regression coefficient

$$Cov(\hat{\boldsymbol{f}}_1, \hat{\boldsymbol{f}}_2) \times V(\hat{\boldsymbol{f}}_2)^{-1} = (\boldsymbol{I} - (\boldsymbol{I} + \boldsymbol{\Lambda}_1^t\boldsymbol{\Psi}_1\boldsymbol{\Lambda}_1)^{-1})\boldsymbol{\Phi}_{12}, \tag{14}$$

is not greater than the true regression coefficient, $\boldsymbol{\Phi}_{12}$.

This section addressed only Bartlett's method, the regression method and the method by Skrondal and Laake, but it should be noted that estimated factor scores using the other methods and partial least squares yield inconsistent estimators (for inconsistency of the estimators of partial least squares, see Areskoug, 1982) .

The estimated factor score is the sum of the true factor score and the error due to estimation, no matter what estimation method for factor score we use. As shown in this section, the sample variance matrix of estimated factor scores is the biased estimate of the true variance matrix of factors. This fact results in the bias in factor score regression.

In this section, we explained why the bias occurs when estimated factor scores are used, but the degree of bias must be investigated to know whether this problem is of practical importance. The degree of bias due to the familiar three step estimation method using factor score will be examined by simulation studies in section 5.

4 The Proposed Estimation Method

To resolve the problem mentioned in the previous section, we propose a modified stepwise estimation method using estimated factor scores. The method is divided into four steps:

Step 1 Employ the random effect model and estimate the parameters in each measurement model (Eqn.(6)). Then, obtain the ML estimators of the structural parameters $\boldsymbol{\xi}_j$, and $\boldsymbol{\xi}_{\boldsymbol{F}_j}$, $\tilde{\boldsymbol{\xi}}_j$ and $\tilde{\boldsymbol{\xi}}_{\boldsymbol{F}_j}$. The parameter estimation in the random effect model in this step is usually called marginal ML estimation in psychometrics (Bock & Aitkin, 1981). Henceforth we term these "marginal ML estimators." If factor rotation is necessary, it also should be executed in this step.

Step 2 Employ the fixed effect model (Eqn.(7)) for each measurement equation and fix $\boldsymbol{\xi}_j$ at $\tilde{\boldsymbol{\xi}}_j$ obtained in the first step. Subsequently, calculate factor scores $\hat{\boldsymbol{f}}_j^B$ by maximum likelihood (if observed variables are continuous, the method is simply Bartlett's method.)

Step 3 Estimate the factor mean vector and factor covariance matrix in the following manner. Let $\hat{\boldsymbol{\nu}} = \frac{1}{N}\sum_{i=1}^N \hat{\boldsymbol{f}}_i^B$ be the sample mean of the estimated factor scores where $\boldsymbol{f}_i$ is the latent variable vector $\boldsymbol{f} = (\boldsymbol{f}_1^t, \cdots, \boldsymbol{f}_J^t)^t$ for the i-th subject. Let also $\boldsymbol{f}_{ij}$ be the value of $\boldsymbol{f}_j$ for the i-th subject. Let $\hat{\boldsymbol{\Phi}}$ be the sample covariance matrix of the estimated factor scores:

$$\hat{\boldsymbol{\Phi}} = \frac{1}{N}\sum_{i=1}^N (\hat{\boldsymbol{f}}_i^B - \hat{\boldsymbol{\nu}})(\hat{\boldsymbol{f}}_i^B - \hat{\boldsymbol{\nu}})^t. \tag{15}$$

Further, use the sample mean of the estimated factor scores in the second step as the estimator of the factor mean vector, $\hat{\boldsymbol{\nu}}$. Use the estimator $\tilde{\boldsymbol{\Phi}}_{jj}$ in the first step as the estimator of $\boldsymbol{\Phi}_{jj}$ instead of the sample covariance matrix of the estimated factor scores. Use the sample covariance matrix between $\boldsymbol{f}_j$ and $\boldsymbol{f}_k$ as the estimator of factor covariance $\boldsymbol{\Phi}_{jk}$ for $(j \neq k)$.

Subsequently, the resulting estimator of $\mathbf{\Phi}$, $\mathbf{\Phi}^P$ can be expressed as follows:

$$\hat{\mathbf{\Phi}}^P = \begin{pmatrix} \tilde{\mathbf{\Phi}}_{11} & \hat{\mathbf{\Phi}}_{12} & \cdots & \hat{\mathbf{\Phi}}_{1J} \\ \hat{\mathbf{\Phi}}_{21} & \tilde{\mathbf{\Phi}}_{22} & \cdots & \hat{\mathbf{\Phi}}_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{\Phi}}_{J1} & \hat{\mathbf{\Phi}}_{J2} & \cdots & \tilde{\mathbf{\Phi}}_{JJ} \end{pmatrix} \tag{16}$$

Step 4 Estimate the parameters of the structural part in the model τ using the generalized least squares method. Let $\boldsymbol{\tau}$ be the structural parameter vector of the structural equation part; $\boldsymbol{\sigma}$, the vector of the non-redundant element of $\boldsymbol{\nu}$, $\mathbf{\Phi}$; and $\sigma(\boldsymbol{\tau})$, the function of $\boldsymbol{\tau}$.

Moreover, let $\hat{\boldsymbol{\sigma}}$ be the corresponding estimator obtained in the third step, and let the estimate of $\boldsymbol{\tau}$ be the value that minimizes the following generalized least squared error function:

$$Q(\boldsymbol{\tau}) = \frac{1}{2}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}(\boldsymbol{\tau}))^t \boldsymbol{W}^{-1}(\hat{\boldsymbol{\sigma}} - \sigma(\boldsymbol{\tau})), \tag{17}$$

where W is the covariance matrix of $\hat{\boldsymbol{\sigma}}$. See the appendix for detail.

Theoretical Justification of the Proposed Method

There are several criteria for evaluating an estimation method in mathematical statistics, such as invariance, unbiasedness, efficiency, and so on. The most important issue is the "consistency" of the estimator, which implies that as the number of observations increases, the estimator converges to the true value of the parameter. We will investigate the consistency of the proposed estimator.

When the observed variables are continuous

We show the consistency of the proposed estimator when the observed variables are continuous. Without loss of generality, we can restrict the case when the number of measurement models is two.

We use $\tilde{\mathbf{\Phi}}_{jj}$ ($j = 1, 2$) obtained in Step 1 as the consistent estimators of $\mathbf{\Phi}_{jj}$ instead of the sample variance matrices of the estimated factor scores. We can also show the following result.

Proposition 1. *The sample covariance matrix of the estimated factor scores,* $\hat{\mathbf{\Phi}}_{12} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{f}}_{i1}^B - \hat{\boldsymbol{\nu}}_1)(\hat{\boldsymbol{f}}_{i2}^B - \hat{\boldsymbol{\nu}}_2)^t$, *is the consistent estimator of* $\mathbf{\Phi}_{12}$, *considering the effect of estimation of structural parameters.*

Proof. Without loss of generality, we consider the case where the means of factors are zero. Using the Bartlett method to estimate the factor scores, we obtain

$$\frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{f}}_{i1}^{B}-\hat{\boldsymbol{\nu}}_1)(\hat{\boldsymbol{f}}_{i2}^{B}-\hat{\boldsymbol{\nu}}_2)^t=\hat{\boldsymbol{A}}_1\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_{1i}\boldsymbol{x}_{2i}^t\}\hat{\boldsymbol{A}}_2^t, \tag{18}$$

where $\boldsymbol{x}_{ji}$ is the observed value vector of $\boldsymbol{x}_j$ for the i-th subject. From invariance and consistency of the ML estimators, $\hat{\boldsymbol{A}}_j \xrightarrow{p} \boldsymbol{A}_j = (\boldsymbol{\Lambda}_j^t\boldsymbol{\Psi}_j^{-1}\boldsymbol{\Lambda}_j)^{-1}\boldsymbol{\Lambda}_j^t\boldsymbol{\Psi}_j^{-1}$, where p stands for "convergence in probability". From the law of large numbers, $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_{1i}\boldsymbol{x}_{2i}^t \xrightarrow{p} \boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_2^t$.

From the above results and by applying Slutsky Theorem, we observe

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{f}}_{i1}^{B}-\hat{\boldsymbol{\nu}}_1)(\hat{\boldsymbol{f}}_{i2}^{B}-\hat{\boldsymbol{\nu}}_2)^t &\xrightarrow{p} (\boldsymbol{\Lambda}_1^t\boldsymbol{\Psi}_1^{-1}\boldsymbol{\Lambda}_1)^{-1}\boldsymbol{\Lambda}_1^t\boldsymbol{\Psi}_1^{-1}\boldsymbol{\Lambda}_1\boldsymbol{\Phi}_{12}\boldsymbol{\Lambda}_2^t\boldsymbol{\Psi}_2^{-1}\boldsymbol{\Lambda}_2(\boldsymbol{\Lambda}_2^t\boldsymbol{\Psi}_2^{-1}\boldsymbol{\Lambda}_2)^{-1} \\ &= \boldsymbol{\Phi}_{12}\end{aligned} \tag{19}$$

□

Therefore, $\hat{\boldsymbol{\Phi}}^P$ in Step 3 is the consistent estimator of $\boldsymbol{\Phi}$. Following the properties of the GLS estimator and the consistency of $\hat{\boldsymbol{\sigma}}$, the proposed estimator of the paramaters in structural equation obtained in Step 4 is consistent.

General case

We show that the proposed method has a kind of "consistency" (i.e., consistency at large) for general case other than the continuous observable variables. By "consistent at large", we mean that at probability one, the estimation method finds the true values of structural parameters when the number of subjects and the observed variables in each measurement model goes to infinity ($N \to \infty$ and $P_j \to \infty$). To prove the consistency of the proposed stepwise estimation method under general model setup, the following additional propositions must be proved to be true.

Proposition 2. *The estimator of $\boldsymbol{\xi}_j$ and $\boldsymbol{\xi}_{F_j}$ in the first step, $\tilde{\boldsymbol{\xi}}_j$ and $\tilde{\boldsymbol{\xi}}_{F_j}$, which are obtained under the random effect measurement model (Eqn.(6)), are consistent although we employ the fixed effect Kiefer & Wolfowitz type model.*

Proposition 3. *Estimator $\hat{\boldsymbol{\Phi}}^P$ is consistent at large.*

In the fixed effect model, factor scores are the incidental parameters in that the number of factor scores increase with the increase in the number of subjects. Further, the joint ML estimators of $\boldsymbol{\xi}_j$ and $\boldsymbol{\xi}_{F_j}$ are not consistent (Neyman & Scott,1948). However, as Kiefer & Wolfowitz (1956) pointed out, marginal ML estimators are shown to be "consistent" in that the estimators converge to the true value when the number of the indicator variables is large. Hence, Proposition 2 is true in the model setup. Then, it is sufficient to prove Proposition 3. The sketch of the proof of Proposition 3 is given in the appendix.

In the next section, by using simulation studies, we will consider the influence of the number of subjects and variables.

An Alternative Estimation Method

Kano (1983) and Shapiro (1984) showed that when the variables are continuous, the estimator of $\boldsymbol{\tau}$, that minimizes the ML discrepancy function where the mean vector and the covariance matrix are replaced by their consistent estimators, is consistent.

In this section, it is also shown that $\hat{\boldsymbol{\nu}}$ and $\hat{\boldsymbol{\Phi}}^P$ are consistent estimators of $\boldsymbol{\nu}$ and $\boldsymbol{\Phi}$, respectively. This indicates that the estimator of $\boldsymbol{\tau}$ that minimizes the ML discrepancy function where the mean and covariance are replaced by $\hat{\nu}$ and $\hat{\boldsymbol{\Phi}}^P$, is also proven to be consistent for $\boldsymbol{\tau}$. Therefore, instead of the proposed fourth step, we can consistently estimate the structural parameter vector τ using prevailing softwares such as SAS/STAT (SAS Institute, 1999) by considering $\hat{\boldsymbol{\Phi}}^P$ as the sample covariance matrix of factors and modeling the structural equation part. It should be noted that the correct standard errors cannot be evaluated by this method.

5 Simulation Studies

In the previous section, the consistency of the proposed stepwise estimation method was proved. However, in order to show the validity of the proposed method in a moderate sample size and in a moderate number of indicators, the degree of bias must be investigated. In this section, some simulation studies that compare the proposed method with familiar methods are discussed.

5.1 Study 1: Two-dimensional factor analysis model

We modeled a situation in which two factors are assumed and each factor is measured by four observed indicators. The model we considered here is the factor analysis model:

$$\boldsymbol{x} = \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 & \lambda_6 & \lambda_7 & \lambda_8 \end{pmatrix}^t \boldsymbol{f} + \boldsymbol{e}, \tag{20}$$

where x (8×1),f (2×1) and e (8×1) are the observable variable vector, latent variable vector, and error vector, respectively. The first element of f is f_X, while the second one is labeled f_Y.

The structural equation is assumed as follows:

$$f_X = \beta f_Y + \zeta, \tag{21}$$

and the main point of interest in this study is β. The model we assume here is depicted in Figure 1. The true values of the factor loadings are shown in Figure 1. The variances of the errors for the observed variables were 0.5. The true value of β was 0.7. $V(f_Y)$ and $V(\zeta)$ were fixed at 1 and 0.51, respectively (so that $V(f_X) = 1$.) Five estimation methods were compared in this study: (1) factor score regression using the estimated factor scores by Bartlett's method (Bart), (2) factor score regression using the estimated factor scores by the regression method (Reg), (3) the modified estimation proposed by Skrondal & Laake (Skro), (4) the proposed method (Prop) and (5) the ML estimation method in SEM. It should be noted that in methods (1)-(4), the factor scores of f_X and f_Y are independently estimated. Moreover, the factor variances were fixed at 1 in each method. The variances of the estimated factor scores in methods (1)-(4) are not equal to one (see section 3).

Data generation and estimation were performed using the SAS package by combining SAS/IML program with proc CALIS in SAS/STAT. To calculate the mean and the mean of squared error (MSE) of estimates of β, 10000 data sets were generated; the results are reported in Table 2.

These results indicate that even when the number of subjects is very large ($N = 10000$), the factor score regression by Bartlett's method and the regression method underestimate β to a large extent. On the other hand, the proposed method and the ML method produce valid estimates, even for relatively small sample sizes ($N = 300$).

5.2 Study 2: Two-level factor analysis model

In study 2, we assumed that three factors (level-one factor f_1, f_2, f_3) exist, and each factor is measured by four observed continuous indicators; hence, the total number of observed variables is 12. Moreover, the three factors are assumed to measure a common factor (level-two factor f_4). Therefore, the structural equation is assumed to be as follows:

$$f_j = \gamma_j f_4 + \zeta_j \quad (j = 1, 2, 3). \tag{22}$$

The model assumed here is shown in Figure 2.

The true values of factor loadings are also shown in the figure. The variance of errors for the observed variables was 0.5. $V(f_4)$ was fixed at 1 and $V(\zeta_1) = 0.51, V(\zeta_2) = 0.64, V(\zeta_3) = 0.75$, respectively (so that $V(f_1) = V(f_2) = V(f_3) = 1$.)

In this model setup, the estimation method using the Bartlett factor scores, the method using Regression factor scores, and the method proposed by Skrondal & Laake produce the same estimates, then the following three methods were compared: (1) factor analysis using estimated factor scores by Bartlett method (Bart), (2) the proposed method (Prop) and (3) the ML method in SEM.

To calculate the mean and the mean of squared error ("MSE") of estimates of λ_j ($j = 1, 2, 3$), 10000 data sets were generated; the true values and the corresponding results are reported in Table 3.

These results indicate that even when the number of subjects is very large ($N = 10000$), the factor score regression by the Bartlett method (or the regression method, Skrondal & Laake) underestimate γ_s to a large extent. On the other hand, the proposed method and ML method produce valid estimates even in relatively small sample sizes ($N = 300$).

5.3 Study 3: Factor analysis model and two-parameter logistic model

In study 1, each measurement model was a factor analysis model. In this study, we assumed that one factor was measured by continuous variables (In this case, the measurement model is factor analysis), and the indicators of the other factor are binomial (In this case, the measurement model is the two-parameter logistic item response model (Eqn.(3))).

The true value of β was changed from -0.8 to 0.8 with increments of 0.2, and the number of subjects was 200 or 500. The number of items (indicators) was set to 20 or 50. For each model setup, 1000 data sets were generated; hence, the resulting number of data sets was 36000. The data sets that yield improper solutions are discarded and not included in the study. Therefore, the total number of data sets included in the study was 33683.

The following three methods are carried out: (i) the proposed method ("Prop"), (ii) the ordinary factor score regression ("Old"), and (iii) the generalized least square estimation using estimated polychoric and polyserial correlations ("poly-GLS").

Data generation was carried out by SAS/IML, and the estimation was performed using SAS/IML, Bilog-MG, and M-plus.

The resulting mean and the MSE of each estimate calculated using the 1000 data sets are listed in Table 4.

There are five points that should be noted. (i) The estimates of the proposed method are more accurate than the ordinary factor score regression in most of the model setup. (ii) As the number of subjects increases, the estimates of the proposed method move closer to the true value. This is not the case in ordinary factor score regression. (iii) As the number of subjects increases, both the estimates of the proposed method and those of the ordinary factor score regression move closer to the true value. However, even when the number of items is 50, the squared errors of the ordinary method is very large as compared to the proposed method. This result is consistent with the theoretical investigation in the appendix (Eqn. (29)), (iv) The proposed method is sometimes biased as compared to GLS using polyserial/polychoric correlations; however, it is not always biased, and (v) GLS using polyserial/polychoric correlations yields improper solutions at a high rate, while the proposed method scarcely yields improper solutions (see Table 5).

5.4 Study 4: Factor analysis model and nominal response model

In study 1, each measurement model was a factor analysis model. In this study, we assumed that one factor was measured by continuous variables (In this case, the measurement model is factor analysis.) and the indicators of the other factor are nominal responses (Bock, 1972). In theory, simultaneous estimation such as that based on ML estimation is applicable to SEM with mixed continuous and nominal observed variables (Moustaki & Knott, 2000; Sammel, Ryan & Legler, 1997). However, simultaneous estimation methods are generally very difficult for researchers to use in applied areas because the evaluation of the likelihood requires numerical integration which becomes impractical in reasonably sized models. There is no prevailing software to deal with this model.

The true value of β was changed from -0.8 to 0.8 with increments of 0.2, and the number of subjects was 200 or 500. The number of items (indicators) was set to 25.

For each model setup, 1000 data sets were generated; hence, the resulting number of data sets was 18000. (The data sets that yield improper solutions are discarded and not included in this study.) The following two methods are carried out: (i) the proposed method ("Prop") and (ii) the ordinary factor score regression ("Old"). Data generation was carried out by SAS/IML, and the estimation was performed using SAS/IML and MULTILOG.

The resulting mean and the MSE of each estimate calculated using the 1000 data sets are listed in Table 6. The simulation study shows that the estimates of the proposed method are finer than the ordinary factor score regression in most of the model setup. It is also observed that as the number of subjects increases, the estimates of the proposed method move closer to the true value. This is not the case in ordinary factor score regression.

6 Discussion

In this paper, we resolved the reason why the estimators of the parameters related to the relationship between the latent variables using estimated factor scores are biased, even when the number of the subject is large.

We also proposed a modified stepwise estimation method using estimated factor scores and showed some asymptotic properties. The simulation studies indicated that the proposed method is valid under several model setups, while the ordinary factor score regression is biased even for very large sample sizes.

The proposed method has the following advantages:

(1) The analysis becomes practical when the level of variables differs with the measurement models.

For example, assume that the indicators of factor A in the first measurement model are continuous, and the indicators of factor B in the second measurement model are nominal. The ML estimation in such a model is impracticable; however, it is very easy to infer the relationship between factors A and B by the proposed method, because most of prevailing programs such as SAS/STAT solve the first measurement model. The second measurement model can be solved by using software such as MULTILOG. We can estimate the relationships between the factors using the outputs of the prevailing softwares.

(2) The proposed method can eliminate the problem in which the model setup of the structural equation part affects the estimation of the measurement part because the proposed method separates the estimation of parameters in the measurement part (the first step) from the parameters in the structural part (the third and fourth parts).

(3) Raw data is not necessary to estimate the parameters in the structural part. The estimated factor scores, the estimates of the measurement parts, and their variance matrices are sufficient. Hence, the proposed method enables a secondary analysis for relationships between factors to be easily executed.

(4) The proposed method can include the variations in the estimators caused by factor rotation. Simultaneous estimation methods fail to incorporate factor rotation into the whole analysis, which leads them to underestimate the standard errors of the parameter estimators in the structural part if the rotated factor loadings are set as constants. The proposed method can evaluate the variation due to factor rotation by calculating the correct variance matrix of the rotated factor loadings (e.g., Ogasawara, 1998).

(5) The proposed method diminishes improper solutions as compared to the previously proposed simultaneous estimation methods.

As discussed in the previous sections, factor score regression is frequently used not only in psychology but also in other social sciences. The "item parceling" method that is more frequently used in psychology, social sciences and behavioral sciences can be considered as a coarse variant of factor score regression. Several simulation studies have shown that the item parceling causes the parameters in the structural part to be underestimated (Bandalos, 2002; MacCallum, Widamam. Zhang & Hong, 1999; Marsh, Hau, Balla & Grayson, 1998).The theoretical investigations in section 3 and the appendices also show that the correlation or regression coefficients are underestimated in the studies in which scale scores (parcel scores) are used to estimate the factor scores. The item parceling method using scale scores can be expected to produce a coarser estimate than the factor score regression using Bartlett's estimates or regression estimates; thus, they are also expected to cause a more serious bias. It is clear that this issue requires further research.

APPENDIX

Generalized least squares as the fourth step

Following results shown in Lee, Poon & Bentler (1990), the resulting estimator of the fourth step $\boldsymbol{\tau}$ can be shown to follow multivariate normal distribution with the mean $\boldsymbol{\tau}$ and covariance matrix $((\partial\sigma(\boldsymbol{\tau})/\partial\boldsymbol{\tau})^t\boldsymbol{W}^{-1}(\partial\sigma(\boldsymbol{\tau})/\partial\boldsymbol{\tau}))^{-1}$ asymptotically. Also, since it has been shown that $2Q(\boldsymbol{\tau})$ asymptotically follows a χ^2 distribution, we can execute hypothesis testing (see Lee, Poon, & Bentler (1990) for detail).

To use the above results, we must calculate the covariance matrix of $\hat{\boldsymbol{\sigma}}$, W. Let $\hat{\boldsymbol{\phi}} = (\hat{\boldsymbol{\nu}}^t, vec(\hat{\boldsymbol{\Phi}})^t)^t$ and $\tilde{\boldsymbol{\phi}_D} = vec(\tilde{\boldsymbol{\Phi}_D})$ where $\boldsymbol{\phi}_D$ is the vector containing all elements of $\boldsymbol{\Phi}_{jj}$ $(j = 1 \sim J)$. Let also $vec(\boldsymbol{f}, \boldsymbol{\phi}_D) = (\boldsymbol{f}^t, \boldsymbol{\phi}_D^t)^t$ and $vec(\hat{\boldsymbol{\Phi}}, \tilde{\boldsymbol{\phi}_D}) = (\hat{\boldsymbol{\phi}}^t, \tilde{\boldsymbol{\phi}_D}^t)^t$.

Following pseudo maximum likelihood estimation theory (Gong & Samaniego, 1981; Parke, 1986),

$$vec(\hat{\boldsymbol{f}}^B, \tilde{\boldsymbol{\phi}_D}) \sim N(vec(\boldsymbol{f}, \boldsymbol{\phi}_D), \boldsymbol{\Sigma}_{\boldsymbol{f},\boldsymbol{\phi}_D}), \tag{23}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{f},\boldsymbol{\phi}_D}$ is estimated as the inverse of the Fisher information matrix when the fixed effect model is employed (Eqn.(7)).

Then we obtain,

$$vec(\hat{\boldsymbol{\Phi}}, \tilde{\boldsymbol{\phi}_D}) \sim N(vec(\boldsymbol{\Phi}, \boldsymbol{\phi}_D), \boldsymbol{A}\boldsymbol{\Sigma}_{\boldsymbol{f},\boldsymbol{\phi}_D}\boldsymbol{A}^t), \tag{24}$$

where $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I} \otimes \boldsymbol{a}^t & 2N_p(\boldsymbol{F}^t\boldsymbol{R} \otimes \boldsymbol{I}_p) & 0 \\ 0 & 0 & \boldsymbol{I} \end{pmatrix}$, $N_p = \frac{1}{2}(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{pp})$, $\boldsymbol{a} = \frac{1}{N}1$ and $\boldsymbol{K}_{pp}$ is a commutation matrix (Magnus & Neudecker (1999)).

Then $\boldsymbol{W}$ is expressed as $\boldsymbol{W} = \boldsymbol{P}\boldsymbol{A}\boldsymbol{\Sigma}_{\boldsymbol{f},\boldsymbol{\phi}_D}\boldsymbol{A}^t\boldsymbol{P}^t$ where $\boldsymbol{P}$ is the appropriate permutation matrix.

Sketch of Proof of Proposition 3

Let $\hat{\boldsymbol{f}}_j^B$ be the estimator of the factor score vector in the j-th measurement model. Let also $\begin{pmatrix} I_{\boldsymbol{f}_j} & I_{\boldsymbol{f}_j\boldsymbol{\xi}_j} \\ I_{\boldsymbol{\xi}_j\boldsymbol{f}_j} & I_{\boldsymbol{\xi}_j} \end{pmatrix}$ be the Fisher information matrix for factor scores and structural parameters of the fixed effect model expressed by Eqn.(7). Using the results of pseudo maximum likelihood estimation (Gong & Samaniego, 1981: Yuan & Jennrich,2000),

$$\begin{pmatrix} \hat{\boldsymbol{f}}_j^B(\tilde{\boldsymbol{\xi}}_j) \\ \tilde{\boldsymbol{\xi}}_j \end{pmatrix} \sim N(\begin{pmatrix} \boldsymbol{f}_j \\ \boldsymbol{\xi}_j \end{pmatrix}, \begin{pmatrix} I_{\boldsymbol{f}_j}^{-1} + I_{\boldsymbol{f}_j}^{-1} I_{\boldsymbol{f}_j \boldsymbol{\xi}_j} \boldsymbol{\Sigma}_{\boldsymbol{\xi}_j} I_{\boldsymbol{\xi}_j \boldsymbol{f}_j} I_{\boldsymbol{f}_j}^{-1} & -I_{\boldsymbol{f}_j}^{-1} I_{\boldsymbol{f}_j \boldsymbol{\xi}_j} \boldsymbol{\Sigma}_{\boldsymbol{\xi}_j} \\ -\boldsymbol{\Sigma}_{\boldsymbol{\xi}_j} I_{\boldsymbol{\xi}_j \boldsymbol{f}_j} I_{\boldsymbol{f}_j}^{-1} & \boldsymbol{\Sigma}_{\boldsymbol{\xi}_j} \end{pmatrix}), \tag{25}$$

where $\tilde{\boldsymbol{\xi}}_j$ is the marginal ML estimator in the first step with the variance $\boldsymbol{\Sigma}_{\boldsymbol{\xi}_j}$ (calculated by the inverse of the Fisher information matrix in the random effect model), and $\hat{\boldsymbol{f}}_j^B(\tilde{\boldsymbol{\xi}}_j)$ is the estimator of factor scores in the second step with $\tilde{\boldsymbol{\xi}}_j$ given.

Because the mean of the marginal distribution is equal to the mean of the conditional distribution:

$$\operatorname*{plim}_{P_j \to \infty} \mathcal{E}(\hat{\boldsymbol{f}}_j^B) = \operatorname*{plim}_{P_j \to \infty} \mathcal{E}(\mathcal{E}(\hat{\boldsymbol{f}}_j^B | \boldsymbol{f}_j)) = \boldsymbol{\nu}_j. \tag{26}$$

From the relationship between the variance of the marginal distribution and that of the conditional distribution, it is also shown that

$$\operatorname*{plim}_{P_j \to \infty} V(\hat{\boldsymbol{f}}_j^B) = \operatorname*{plim}_{P_j \to \infty} \{V(\mathcal{E}(\hat{\boldsymbol{f}}_j^B | \boldsymbol{f}_j)) + \mathcal{E}(V(\hat{\boldsymbol{f}}_j^B \mid \boldsymbol{f}_j))\} = V(\boldsymbol{f}_j) + \operatorname*{plim}_{P_j \to \infty} \mathcal{E}(V(\hat{\boldsymbol{f}}_j^B \mid \boldsymbol{f}_j)). \tag{27}$$

From the definition, the variance matrix of $\boldsymbol{f}_j$ is the true variance of factor, $\boldsymbol{\Phi}_j$. $\mathcal{E}(V(\hat{\boldsymbol{f}}_j^B \mid \boldsymbol{f}_j))$ is the variance of the factor score estimator $\hat{\boldsymbol{f}}_j^{\,B}$ with the true factor score f_j given.

If the structural parameters in the measurement models are known, $\mathcal{E}(V(\hat{\boldsymbol{f}}_j^B \mid \boldsymbol{f}_j))$ is the inverse matrix of the Fisher information matrix $I_{\boldsymbol{f}_j}$ (or "test information matrix" in the area of educational statistics) with the structural parameters given in the fixed effect j-th measurement model (Eqn.(7).).

However, if we do not know the true values of structural parameters and if we evaluate the influence of the estimation in the first step, the second term in the right side of Eqn.(27) can be expressed as follows in the large number of indicators:

$$\operatorname*{plim}_{P_j \to \infty} \mathcal{E}(V(\hat{\boldsymbol{f}}_j^B \mid \boldsymbol{f}_j)) = I_{\boldsymbol{f}_j}^{-1} + I_{\boldsymbol{f}_j}^{-1} I_{\boldsymbol{f}_j \boldsymbol{\xi}_j} \boldsymbol{\Sigma}_{\boldsymbol{\xi}_j} I_{\boldsymbol{\xi}_j \boldsymbol{f}_j} I_{\boldsymbol{f}_j}^{-1}, \tag{28}$$

See Hoshino & Shigemasu (2008) for concrete expression.

It is shown that the sample covariance matrix of f_j using the factor scores (calculated in the third step) is inconsistent for estimating $\boldsymbol{\Psi}_{jj}$, because the following equation holds:

$$\operatorname*{plim}_{N,P_j\to\infty}\frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{f}}_{ji}^{B}-\bar{\boldsymbol{f}}_{j}^{B})(\hat{\boldsymbol{f}}_{ji}^{B}-\bar{\boldsymbol{f}}_{j}^{B})^{t}=\boldsymbol{\Phi}_{jj}+\boldsymbol{G}, \tag{29}$$

where $\bar{\boldsymbol{f}}_j^B=\frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{f}}_j^B$ and $\boldsymbol{G}=\mathrm{plim}_{N,\ P_j\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left[I_{\boldsymbol{f}_j}^{-1}+I_{\boldsymbol{f}_j}^{-1}I_{\boldsymbol{f}_j\boldsymbol{\xi}_j}\boldsymbol{\Sigma}_{\boldsymbol{\xi}_j}I_{\boldsymbol{\xi}_j\boldsymbol{f}_j}I_{\boldsymbol{f}_j}^{-1}\right]$.

On the other hand, the sample covariance matrix between f_j and f_k using estimated factor scores is consistent for estimating $\boldsymbol{\Phi}_{jk}$:

$$\operatorname*{plim}_{N,P_j\to\infty}\hat{\boldsymbol{\Phi}}_{jk}=\operatorname*{plim}_{N,P_j\to\infty}\frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{f}}_{ji}^{B}-\bar{\boldsymbol{f}}_{j}^{B})(\hat{\boldsymbol{f}}_{ki}^{B}-\bar{\boldsymbol{f}}_{k}^{B})^{t}=Cov(\hat{\boldsymbol{f}}_{ji}^{B},\hat{\boldsymbol{f}}_{ki}^{B})=Cov(\boldsymbol{f}_{ji},\boldsymbol{f}_{ki})=\boldsymbol{\Phi}_{jk}. \tag{30}$$

It follows from that $\tilde{\xi}_j$ and $\tilde{\xi}_k$ are mutually independent for $j\neq k$ because each ξ_s is estimated for each measurement equation, Therefore we obtain that $Cov(\hat{\boldsymbol{f}}_j^{\ B},\hat{\boldsymbol{f}}_k^{\ B}\mid\boldsymbol{f}_j,\boldsymbol{f}_k)=0$ for $\forall j\neq l$.

Therefore we observe that $\hat{\boldsymbol{\Phi}}^{P}$ is a consistent estimator of $\boldsymbol{\Phi}$ when the number of indicators is large.

References

[1] Anderson, J. C., & Gerbing, D. W. (1984). The effects of sampling error on convergence, improper solutions and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*,**49**, 155–173.

[2] Anderson, J.C., & Gerbing, D.W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin,* **103**, 411–423.

[3] Areskoug, B. (1982). The first canonical correlation: Theoretical PLS analysis and simulation experiments. In K. G. Joreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction. Part I* (pp. 263-270). Amsterdam: North-Holland.

[4] Ashton, M.C., Lee, K., & Goldberg, L.R. (2004). A hierarchical analysis of 1,710 English personality-descriptive adjectives. *Journal of Personality and Social Psychology*,**87**, 707–721.

[5] Baker, F.B. & Kim, S-H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York: Marcel Dekker.

[6] Bandalos, D.L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling,* **9**, 78–102.

[7] Bartholomew, D.J. (1981). Posterior analysis of the factor model. *British Journal of Mathematical and Statistical Psychology,* **34**, 93–99.

[8] Bartholomew, D. (2007). Three faces of factor analysis. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* . Wahmah NJ: Erlbaum.

[9] Bauer, D.J. & Curran, P.J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*,**9**, 3–29.

[10] Bentler, P.M. (1983). Some contributions to efficient statistics in structural equation models: Specification and estimation of moment structure. *Psychometrika,* **48**, 493–517.

[11] Bentler, P. M. & Weeks, DG. (1980). Linear structural equations with latent variables. *Psychometrika,* **45**, 289–308.

[12] Bock, R.D. (1972). Estimating item parameters and latent ability when responses are scored in two or morenominal categories. *Psychometrika*, **37,** 29–51.

[13] Bock, R.D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika,* **46**, 443–459.

[14] Bollen, K.A. (1989). *Structural equations with latent variables.* New York: Wiley.

[15] Bollen, K. A. (1996). An alternative two stage least squares (2SLS) estimator for latent variable equations. *Psychometrika*,**61**, 109–121.

[16] Bollen, K.A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*,**53**, 605–634.

[17] Bollen, K.A. & Curran, P.J. (2006). *Latent curve models: A structural equation approach*. New York: Wiley.

[18] Burt, R.S. (1973). Confirmatory factor analytic structures and the theory construction process. *Sociological Methods and Research*, **2**, 131–190.

[19] Burt, R.S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods and Research,* **5**, 3–52.

[20] Cattell, R. B., & Burdsall, C. A. (1975). The radial parcel double factor design: A solution to the item-versus-parcel controversy. *Multivariate Behavioral Research,* **10**, 165–179.

[21] Chen, F., Bollen, K.A., Paxton, P., Curran, P.J., & Kirby, J.B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods & Research*,**29**, 468–508.

[22] De Boeck, P., & Wilson, M. (Eds.) (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York: Springer.

[23] Dijkstra, T. (1983). Some comments on maximum likelihood and partial least squares methods. *Journal of Econometrics*,**22**, 67–90.

[24] Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

[25] Gass, C.S. (1996). MMPI-2 Variables in Attention and Memory Test Performance. *Psychological Assesment*, **8**, 135–138.

[26] Gong, G., & Samaniego, F.J. (1981). Pseudo Maximum Likelihood Estimation: Theory and Applications. *Annals of Statistics,* **9**, 861–869.

[27] Hays, R. D., Revicki, D., & Coyne, K. (2005). Application of structural equation modeling to health outcomes research. *Evaluation and the Health Professions*,**28**, 295–309.

[28] Hoshino, T., & Shigemasu, K. (2008). Standard errors of estimated latent variables with estimated structural parameters. *Applied Psychological Measurement*, **32**, 181–189.

[29] Hunter, J.E., & Garbing, D.W. (1982). Unidimensional measurement, second order factor analysis, and causal models. *Research in Organizational Behavior,* **4**, 267–320.

[30] Hwang, H. & Takane, Y. (2004). Generalized structured components analysis. *Psychometrika*, **69**, 81–99.

[31] Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, **32**, 443–482.

[32] Jöreskog, K.G. (1970). A general method for the analysis of covariance structure. *Biometrika*, **57**, 239–251.

[33] Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, **36**, 409–426.

[34] Jöreskog, K. G. (1977). Structural equation models in the social sciences: Specification, estimation and testing. In P. R. Krishnaiah (Ed.), *Applications of statistics* (pp. 265-287). Amsterdam: North-Holland.

[35] Jöreskog, K. G., & Wold, H. (1982). The ML and PLS techniques for modeling with latent variables: Historical and comparative aspects. In K. G. Joreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction. Part I* (pp. 263-270). Amsterdam: North-Holland.

[36] Kano, Y. (1983). Consistency of estimators in factor analysis. *Journal of the Japan Statistical Society*, **13**, 137–144.

[37] Kiefer, J., & Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Annals of Mathematical Statistics*, **27**, 887–906.

[38] Kishton, J.M., & Widaman, K.F. (1994). Unidimensional versus domain representative parceling of questionnaire items: An empirical example. *Educational & Psychological Measurement*, **54**, 757–765.

[39] Kumar, A, & Dillon, W.R. (1987), The Interaction of Measurement and Structure in Simultaneous Equation Models with Unobservable Variables. *Journal of Marketing Research*, **24**, 98–105.

[40] Lance, C.E., Cornwell, J.M., & Mulaik, S.A. (1988). Limited Information parameter Estimates for Latent or Mixed Manifest and Latent Variable Models. *Multivariate Behavioral Research*, **23**, 171–187.

[41] Lee, S. -Y. (Ed.) (2007). *Handbook of latent variable and related models.* Amsterdam: North-Holland.

[42] Lee, S-Y., Poon, W-Y., & Bentler, P.M. (1990). A three-stage estimation procedure for structural equation models with polytomous variables. *Psychometrika,* **55**, 45–51.

[43] Lee, S. -Y., Poon, W. -Y., & Bentler, P. M. (1995). A two-stage estimation of structural equation models with continuous and polytomous variables. *British Journal of Mathematical and Statistical Psychology*,**48**, 339–358.

[44] Lee, S. -Y., & Song, X. -Y. (2003). Bayesian analysis of structural equation models with dichotomous variables. *Statistics in Medicine*,**22**, 3073–3088.

[45] Lindsay, B., Clogg, C.C., & Grego, J. (1991). Semparametric Estimation in the Rasch Model and Related Exponential Response Models, Including a Simple Latent Class Model for Item Analysis, *Journal of the American Statistical Association,* **86**, 96–107.

[46] Lord, F.M. & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

[47] Magnus, J.R. & Neudecker, H. (1999). *Matrix Differential Cauculus with Applications in Statistics and Econometrics. 2nd ed.* Chichester: John Wiley & Sons.

[48] MacCallum, R.C. & Austin, J.T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*,**51**, 201–226.

[49] MacCallum, R.C., Widaman, K.F., Zhang,S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods,* **4**, 84–99.

[50] Marcoulides, G.A., & Schumacker, R.E. (1996). *Advanced structural equation modeling: issue and techniques*. Mahwah: Erlbaum Associates.

[51] Marsh, H.W., Hau, K,T,, Balla, J.R., & Grayson, D. (1998) Is more ever too much?: The number of indicators per factors in confirmatory factor analysis. *Multivariate Behavioral Research*, **33**, 181–222.

[52] McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research*, **31**, 239–270.

[53] Moustaki, I., Jöreskog, K. G., & and Mavridis, D. (2004). Factor models for ordinal variables with covariate effects on the manifest and latent variables: A comparison of LISREL and IRT approaches. *Structural Equation Modeling*,**11**, 487–513.

[54] Moustaki, I., & Knott, M. (2000). Generalized latent trait models. *Psychometrika*, **65**, 391–411.

[55] Muthén, B.O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**, 115–132.

[56] Neyman, J., & Scott, E.L. (1948). Consistent estimation from partially consistent observations. *Econometrica*, **16**, 1–32.

[57] Ogasawara, H. (1998). Standard errors for rotation matrices with an application to the promax solution. *British Journal of Mathematical and Statistical Psychology*, **51**, 163–178.

[58] Parke, W.R. (1986). Pseudo maximum likelihood estimation: The asymptotic distribution. *Annals of Statistics*, **14**, 355–357.

[59] Rabe-Hesketh, S., & Skrondal, A. (2005). *Multilevel and longitudinal modeling using Stata*. College Station, TX: Stata Press.

[60] Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). Generalized multilevel structural equation modeling. *Psychometrika*,**69**, 167–190.

[61] Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monographs*, (Whole No. 17).

[62] Sammel, M.D., Ryan, L.M., & Legler, J.M. (1997). Latent variable models for mixed discrete and continuous outcomes. *Journal of the Royal Statistical Society, Series B*, **59**, 667–678.

[63] SAS Institute (1999). *SAS/STAT user's guide*. Cary, NC: SAS Institute.

[64] Shapiro, A. (1984). A note on the consistency of estimators in the analysis of moment structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 84–88.

[65] Skrondal, A., & Laake, P. (2001). Regression among Factor Scores. *Psychometrika*, **66**, 563–576.

[66] Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton: Chapman & Hall/CRC.

[67] Sörbom, D. (1974). A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology*, **27**, 229–239.

[68] Stoel, R.D., Garre, F.G., Dolan, C., & vanden Wittenboer, G. (2006). On the likelihood ratio test in structural equation modeling when parameters are subject to boundary constraints. *Psychological Methods*,**11**, 439–455.

[69] Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis*, **48**, 159–205.

[70] Thompson, B., & Melancon, J. G. (1996). *Using item "Testlest" / "Parcels" in confirmatory factor analysis: An example using the PPSDQ-78*. (ERIC Document No. ED 404 349).

[71] Vinzi, V. E., & Lauro, C. (Eds.) *Partial least squares*. *Computational Statistics & Data Analysis*, **48**, 1–219.

[72] Wold, H. (1973). Non-linear iterative partial least squares (NIPALS) modelling. Some current developments. In P. R. Krishnaiah (Ed.), *Multivariate Analysis, vol. III* (pp. 383-407). New York: Academic Press.

[73] Wold, H. (1982). Soft modeling: the basic design and some extensions. In K. G. Jöreskog, & H. Wold (Eds.), *Systems under indirect observation, part 2* (pp. 1–54). Amsterdam: North-Holland.

[74] Yanai, H., & Ichikawa, M. (2007). Factor analysis. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics 26 Psychometrics* (pp. 257–296). Amsterdam: North-Holland.

[75] Yuan, K.-H., & Bentler, P. M. (2007). Structural equation modeling. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics 26 Psychometrics* (pp. 297–358). Amsterdam: North-Holland.

[76] Yuan, K.-H., Bentler, P. M., & Kano, Y. (1997). On averaging variables in a confirmatory factor analysis model. *Behaviormetrika*,**24**, 71–83.

[77] Yuan, K. -H., & Chan, W. (2002). Fitting structural equation models using estimating equations: A model segregation approach. *British Journal of Mathematical and Statistical Psychology*,**55**, 41–62.

[78] Yuan, K-H., & Jennrich, R.I. (2000). Estimating equations with nuisance parameters: Theory and applications. *Annals of the Institute of Statistical Mathematics*, **52**, 343–350.

[79] Zammuner, V.L. (1998). Concepts of Emotion: “Emotionness”, and Dimensional Ratings of Italian Emotion Words. *Cognition and Emotion*, **12**, 243–272.

Table 2: Resulting estimates of Study 1

N=300	True	Bart	Reg	Skro	Prop	ML
Correlation	.7000	.5190	.5190	.5190	.6953	.6989
MSE of Corr.		.0346	.0346	.0346	.0078	.0028
Reg. coef.	.7000	.4995	.5391	.6977	.6953	.6989
MSE of Reg.		.0425	.0285	.0045	.0078	.0028
N=10000	True	Bart	Reg	Skro	Prop	ML
Correlaton	.7000	.5206	.5206	.5408	.6997	.7000
MSE of Corr.		.0323	.0323	.0254	.0002	.0000
Reg. Coef.	.7000	.5010	.5408	.6998	.6997	.7000
MSE of Reg.		.0397	.0254	.0001	.0002	.0000

Table 3: Resulting estimates of Study 2

N=300	True	Bart	Prop	MLE
γ_1	.7000	.6085	.7038	.7051
γ_2	.6000	.5115	.6036	.6038
γ_3	.5000	.4058	.4969	.4971
MSE		.0544	.0429	.0388
N=10000	True	Bart	Prop	MLE
γ_1	.7000	.6035	.7005	.7005
γ_2	.6000	.5074	.5995	.5998
γ_3	.5000	.4089	.5007	.5008
MSE		.0269	.0012	.0010

Table 4: Resulting estimates of Study 3

N=500							
Items	True	Prop	Prop MSE	Old	Old MSE	poly-GLS	poly-GLS MSE
50	0.8	0.8053	.001517	0.6912	.014058	0.8123	.001156
50	0.6	0.6087	.001606	0.5223	.007783	0.6141	.001477
50	0.4	0.4105	.001652	0.3521	.003720	0.4136	.001622
50	0.2	0.2111	.001655	0.1809	.001621	0.2118	.001650
50	0.0	0.0107	.001625	0.0089	.001303	0.0333	.001722
50	-0.2	-0.1905	.001576	-0.1638	.002631	-0.2044	.001543
50	-0.4	-0.3922	.001526	-0.3369	.005525	-0.4083	.001470
50	-0.6	-0.5941	.001490	-0.5102	.009956	-0.6116	.001349
50	-0.8	-0.7960	.001470	-0.6835	.015941	-0.8127	.001131
20	0.8	0.8171	.002025	0.6501	.026068	0.8062	.001222
20	0.6	0.6175	.002100	0.4912	.014287	0.6096	.001405
20	0.4	0.4162	.002045	0.3310	.006475	0.4105	.001602
20	0.2	0.2136	.001954	0.1698	.002217	0.2099	.001692
20	0.0	0.0101	.001853	0.0082	.001276	0.0337	.001767
20	-0.2	-0.1941	.001780	-0.1545	.003512	-0.2024	.001640
20	-0.4	-0.3988	.001750	-0.3174	.008790	-0.4052	.001571
20	-0.6	-0.6037	.001779	-0.4804	.017086	-0.6072	.001425
20	-0.8	-0.8085	.001859	-0.6433	.028429	-0.8071	.001164
N=200							
Items	True	Prop	Prop MSE	Old	Old MSE	poly-GLS	poly-GLS MSE
50	0.8	0.8078	.004231	0.6935	.014440	0.8337	.004928
50	0.6	0.6101	.004393	0.5238	.008953	0.6333	.006293
50	0.4	0.4096	.004243	0.3516	.005383	0.4272	.005870
50	0.2	0.2075	.004131	0.1781	.003460	0.2176	.004827
50	0.0	0.0043	.004006	0.0036	.002946	0.0541	.004674
50	-0.2	-0.1999	.003911	-0.1717	.003702	-0.2187	.004778
50	-0.4	-0.4028	.004025	-0.3459	.005936	-0.4273	.006250
50	-0.6	-0.6053	.004035	-0.5197	.009468	-0.6344	.006124
50	-0.8	-0.8039	.004119	-0.6901	.015185	-0.8310	.007647
20	0.8	0.8365	.006268	0.6435	.030317	0.8251	.003204
20	0.6	0.6343	.006214	0.4877	.017304	0.6305	.004256
20	0.4	0.4286	.005835	0.3293	.008825	0.4285	.004671
20	0.2	0.2208	.005371	0.1693	.004262	0.2217	.004711
20	0.0	0.0114	.004988	0.0080	.003223	0.0569	.004952
20	-0.2	-0.1988	.004803	-0.1537	.005487	-0.2136	.004250
20	-0.4	-0.4088	.004803	-0.3150	.011122	-0.4208	.004119
20	-0.6	-0.6173	.004995	-0.4755	.020298	-0.6240	.003743
20	-0.8	-0.8239	.005217	-0.6344	.033457	-0.8209	.002896

Table 5: The number of Improper Solutions

N=500	Items	True value	# of datasets that could not estimate: poly-GLS	# of datasets that could not estimate: Prop. / Old	# of valid datasets
	50	0.8	13	0	1000
	50	0.6	12	0	1000
	50	0.4	12	0	1000
	50	0.2	12	0	1000
	50	0.0	12	0	1000
	50	-0.2	12	0	1000
	50	-0.4	12	0	1000
	50	-0.6	12	0	1000
	50	-0.8	12	0	1000
	20	0.8	0	0	1000
	20	0.6	0	0	1000
	20	0.4	0	0	1000
	20	0.2	0	0	1000
	20	0.0	0	0	1000
	20	-0.2	0	0	1000
	20	-0.4	0	0	1000
	20	-0.6	0	0	1000
	20	-0.8	0	0	1000
N=200					
	50	0.8	360	8	735
	50	0.6	349	7	747
	50	0.4	351	7	745
	50	0.2	350	7	746
	50	0.0	353	7	743
	50	-0.2	353	7	743
	50	-0.4	351	5	746
	50	-0.6	353	4	743
	50	-0.8	362	3	735
	20	0.8	1	7	1000
	20	0.6	0	7	1000
	20	0.4	0	7	1000
	20	0.2	0	7	1000
	20	0.0	3	7	1000
	20	-0.2	1	7	1000
	20	-0.4	0	5	1000
	20	-0.6	0	4	1000
	20	-0.8	1	3	1000

Table 6: Resulting estimates of Study 4

N=500	True value	Prop	Prop MSE	Old	Old MSE
	0.8	0.7645	.003273	0.4936	.094399
	0.6	0.5755	.002449	0.3665	.055074
	0.4	0.3878	.001838	0.2434	.025089
	0.2	0.1993	.001629	0.1238	.006413
	0	0.0045	.001538	0.0028	.000591
	-0.2	-0.1945	.001608	-0.1207	.006886
	-0.4	-0.3954	.001587	-0.2463	.024168
	-0.6	-0.5955	.001790	-0.3709	.053029
	-0.8	-0.7857	.002070	-0.4953	.093305
N=200	True value	Prop	Prop MSE	Old	Old MSE
	0.8	0.7941	.006040	0.5039	.088946
	0.6	0.6017	.005146	0.3760	.051568
	0.4	0.4041	.004720	0.2492	.024174
	0.2	0.2007	.004392	0.1230	.007414
	0	-0.0008	.004018	-0.0009	.001501
	-0.2	-0.2052	.003920	-0.1270	.006792
	-0.4	-0.4091	.004098	-0.2555	.022230
	-0.6	-0.6108	.004652	-0.3825	.048569
	-0.8	-0.7962	.005441	-0.5059	.087619

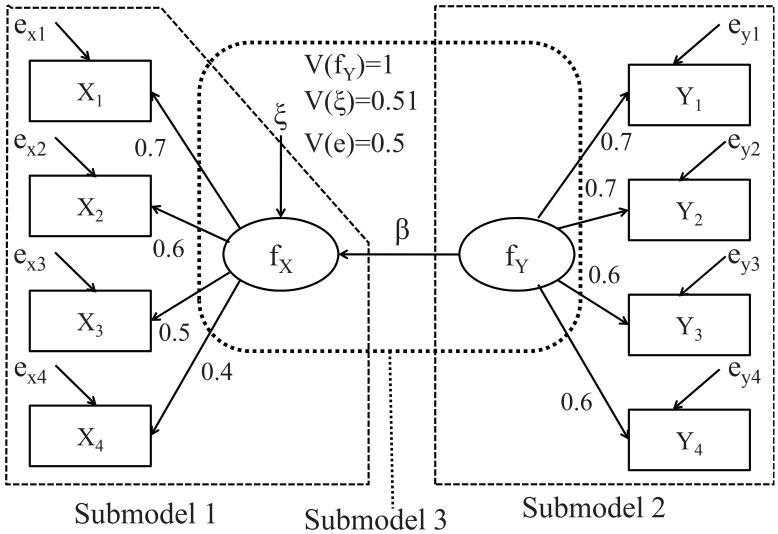

Figure 1: The model for study 1

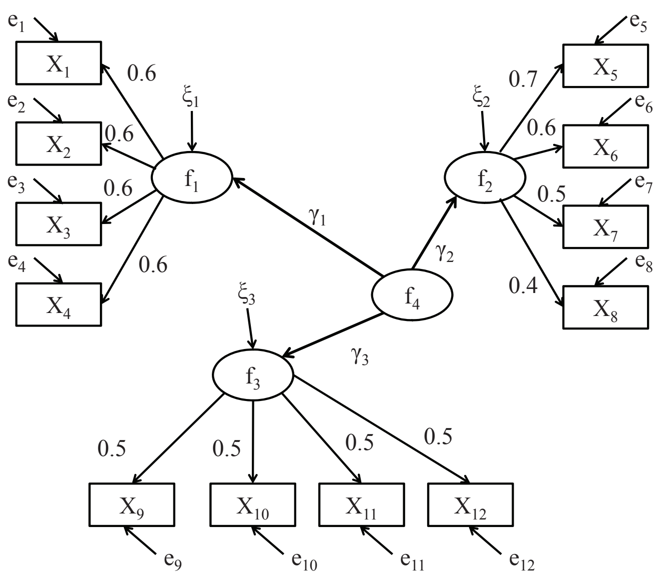

Figure 2: The model for study 2