UCLA Statistics Series #62 August 1990

Mean and Covariance Structure Analysis of Hierarchical Data

Bengt O. Muthén
Graduate School of Education
UCLA

May 1990
Submitted for Publication

Presented at the Psychometric Society meeting in Princeton, N.J., June 1990. An earlier version of this paper was presented at the American Statistical Association meeting in Washington, D.C., August, 1989. This research was supported by Grant No. SES-8821668 from the National Science Foundation. I would like to thank Leigh Burstein, Mike Hollis, Linda Muthén, and Albert Satorra for helpful discussions and Gilbert FitzGerald, Karen Gold, Suk-Woo Kim, Linda Muthén, Ginger Nelson, and Tammy Tam for helpful research assistance.

UCLA Classification: Theory and Methods. Social Sciences-Education

#32

Abstract

Conventional covariance structure analysis, such as factor analysis, is often applied to data that are obtained in a hierarchical fashion, such as students observed within classrooms. A more appropriate specification is presented which explicitly models the within–level and between–level covariance matrices. The likelihood expression under multivariate normality is studied and related to that of conventional covariance structure modeling. It is shown that conventional covariance structure software can be easily adapted to handle hierarchical models. Using this framework, several hierarchical data models are outlined. An example of mathematics achievement testing is analyzed in detail.

Keywords: multilevel, structural equation modeling, latent variables.

1. Introduction

This paper considers data that are collected in a hierarchical fashion, such as students sampled within schools. This is a frequently used design in large–scale surveys, such as The National Longitudinal Study with data gathered regarding the educational aspirations and attainment of high school seniors of 1972 or The National Education Longitudinal Study of 1988 with data on eighth grade students. Multivariate modeling of such data are most frequently done as if the data were obtained as a simple random sample from a single population. Hence, the standard assumption of i.i.d. observations is made. New analysis techniques that are more suited to the hierarchial data structure have recently emerged under the labels of hierarchical, or multilevel models (see e.g. Aitkin & Longford, 1986; Burstein, 1980; De Leeuw & Kreft, 1986; Goldstein 1986, 1987; Longford, 1987; Mason, Wong, & Entwistle, 1984; Raudenbush & Bryk, 1988). As pointed out by Muthen and Satorra (1989), such modeling takes into account that hierarchically gathered data both gives rise to correlated observations and observations from heterogeneous populations with varying parameter values. While appropriate analysis techniques of this kind are now available for standard regressions and analysis of variance situations, Muthen and Satorra emphasized the lack of techniques for covariance structure models, such as factor analysis, path analysis, and structural eqation modeling.

Muthen and Satorra (1989) outlined several possible covariance structure models and their maximum–likelihood estimation. However, they did not discuss how such estimation should be carried out in practice nor did they provide examples of such modeling. A companion article, Muthen (1989), discussed the relationships of multilevel structural equation modeling to conventional structural equation modeling and showed that conventional structural equation modeling software could be used for multilevel structural equation modeling in the special case of balanced data. The aim of this paper is both to further develop the Muthen–Satorra modeling and to present a general, practical estimation scheme in the general unbalanced case. Recent, related work is that of De Leeuw (1985), Goldstein and McDonald (1988), McDonald and Goldstein (1988), and Longford and Muthen (1990). The McDonald and Goldstein paper outlines maximum–likelihood estimation of general multilevel structural model with the aim of developing specially designed software to carry out the complex computational tasks. The Longford and Muthen paper focuses on efficient computation for hierarchical data factor analysis models. It is interesting to note, however, that a good part of these tasks was already carried out more than 20 years ago in the dissertation of Schmidt (1969), see also Schmidt and Wisenbaker (1986), using specially designed software. While the development of special software is of great value for these situations, the present paper shows that quite general models can be fitted by maximum likelihood using only slight modifications of already existing structural equation modeling software. Drawing on the maximum likelihood solution, a simpler to use ad hoc estimator will also be proposed. Examples are given to illustrate new analysis possibilities. Taken together, these developments should enable a more rapid spread of these important modeling schemes.

To develop the framework for specific hierarchical models, Section 2 presents the likelihood for two–level hierarchical data. This development leads up to the presentation of useful relations between the hierarchical data likelihood and the maximum–likelihood fitting function employed in conventional structural equation modeling software. Section 3 proposes an estimator that is simpler to compute than maximum likelihood while still using the structural equation modeling framework. Section 4 proposes specific models and show how they fit into the general structural equation modeling framework. Section 5 uses some of these models and structural equation software to analyze data on mathematics achievement for students observed within classrooms.

2. The likelihood for hierarchical data

Consider for simplicity a two–level structure where data are gathered on individual units obtained within groups of units. Let $\mathbf{y}_{gi}$ denote the p–vector of observed y variables for individual unit i within group g. Also assume the availability of group–level observation on the q–vector $\mathbf{z}_g$. Let

(1) $E(\mathbf{y}_{gi}) = \boldsymbol{\mu}_y$,

(2) $E(\mathbf{z}_g) = \boldsymbol{\mu}_z$,

(3) $V(\mathbf{y}_{gi}) = \boldsymbol{\Sigma}_W + \boldsymbol{\Sigma}_B$,

(4) $V(\mathbf{z}_g) = \boldsymbol{\Sigma}_{zz}$,

(5) $Cov(\mathbf{y}_{gi}, \mathbf{z}_g) = \boldsymbol{\Sigma}_{yz}$,

where $\boldsymbol{\Sigma}_W$ denotes individual–level (within) variation and $\boldsymbol{\Sigma}_B$ denotes group–level (between) variation. With more than two levels of hierarchical nesting more covariance matrix components would be added. It will be assumed that a particular hypothesized covariance structure model, H_0 say, expresses the distinct elements of $\boldsymbol{\mu}_z$, $\boldsymbol{\mu}_y$, $\boldsymbol{\Sigma}_W$, $\boldsymbol{\Sigma}_B$, $\boldsymbol{\Sigma}_{zz}$, $\boldsymbol{\Sigma}_{yz}$ in terms of a smaller number of parameters, while the unrestricted model, H_1, places no restrictions on these elements. Specific models will be discussed in section 3.

Assume g = 1, 2, ..., G independently observed groups with i = 1, 2, ..., N_g individual observations within group g. Arrange the data vector for which independent observations are obtained as

$$(6)\quad \mathbf{d}_g' = (\mathbf{z}_g', \mathbf{y}_{g1}', \mathbf{y}_{g2}', \ldots, \mathbf{y}_{gN_g}'),$$

where we note that the length of $\mathbf{d}_g$ varies across groups. The mean vector and covariance matrix of $\mathbf{d}_g$ are

$$(7)\quad \boldsymbol{\mu}_d' = [\, \boldsymbol{\mu}_z', \mathbf{1}_{Ng}' \otimes \boldsymbol{\mu}_y' \,]$$

$$(8)\quad \Sigma_d = \begin{bmatrix} \Sigma_{zz} & \text{symmetric} \\ 1_{Ng} \otimes \Sigma_{yz} & I_{Ng} \otimes \Sigma_W + 1_{Ng}\, 1_{Ng}' \otimes \Sigma_B \end{bmatrix}$$

where $\mathbf{I}_{Ng}$ is an identity matrix of dimension N_g, $\mathbf{1}_{Ng}$ is a unit vector of length N_g and the symbol $\otimes$ denotes the Kronecker product.

Assuming multivariate normality of $\mathbf{d}_g$, the maximization of the likelihood for G independent draws of $\mathbf{d}_g$ leads to the minimization of the maximum–likelihood (ML) fitting function

$$(9) \quad \sum_{g=1}^{G} \{ \log | \Sigma_{dg} | + (\mathbf{d}_g - \mu_{dg})' \Sigma_{dg}^{-1} (\mathbf{d}_g - \mu_{dg}) \}$$

In simplifying (9) we will make repeated use of the standard matrix algebra results

(10)

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = | A_{11} | \, | A_{22} - A_{21} A_{11}^{-1} A_{12} |$$

(11)

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} .$$

where

$$(12) \quad A^{11} = (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} ,$$

(13) $A^{22} = (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}$,

(14) $A^{12} = - A_{11}^{-1} A_{12} A^{22}$,

and due to symmetry of Σ_{dg}

(15) $A^{21} = (A^{12})'$.

Using (10), and Kronecker product algebra we find

(16) $| \Sigma_{dg} | = | \Sigma_{zz} | \, | I_{Ng} \otimes \Sigma_W + 1_{Ng} 1_{Ng}' \otimes (\Sigma_B - \Sigma_{yz} \Sigma_{zz}^{-1} \Sigma_{yz}') |$.

By standard results, this simplifies as

(17) $| \Sigma_{dg} | = | \Sigma_{zz} | \; | \Sigma_W |^{Ng-1} | \Sigma_W + N_g (\Sigma_B - \Sigma_{yz} \Sigma_{zz}^{-1} \Sigma_{yz}') |$

This expression may be restated using (10) with

$$
(18)\quad \begin{bmatrix} A_{11} & \text{symmetric} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} \Sigma_{zz} & \text{symmetric} \\ \Sigma_{yz} & N_g^{-1}\,\Sigma_W + \Sigma_B \end{bmatrix}
$$

$$
= \Sigma_{vg} ,
$$

say. This reexpresses (17) as

$$
(19)\quad | \Sigma_{dg} | = \quad | \Sigma_{vg} | \; | \Sigma_W |^{Ng-1} \; | N_g | .
$$

In (18) we recognize that A_{22} is the covariance matrix for the sample mean vector $\bar{y}_g$. This may be shown by applying Kronecker product algebra to the y part of (7) and (8). In this way, Σ_{vg} is the covariance matrix for variables that vary across groups, not across individuals. Hence, the determinant in (19) is expressed in terms of that of a between matrix for z and $\bar{y}$ variables and that of a within matrix, Σ_W, for y variables.

Consider now the covariance matrix Σ_{dg} in (8). Using the inverse matrix formulas (11) – (15) together with the definition of A in (18) we may express the inverse as

(inversion trick twice for Σ_s^{22})

(20) $$\left(\Sigma_{dg}\right)^{-1} = \begin{bmatrix} \Sigma_g^{11} & \text{symmetric} \\ \Sigma_g^{21} & \Sigma_g^{22} \end{bmatrix}$$

where

(21) $$\Sigma_g{}^{11} = A^{11} ,$$

$= 1_{N_g} \otimes N_g^{-1} (A^{12})'$

(22) $$\Sigma_g{}^{21} = (\Sigma_g{}^{12})' = - 1_{Ng} \otimes N_g{}^{-1} A^{22} A_{21} A_{11}{}^{-1} ,$$

(23) $$\Sigma_g{}^{22} = 1_{Ng} 1_{Ng}' \otimes N_g{}^{-2} A^{22} + I_{Ng} \otimes \Sigma_W{}^{-1} - 1_{Ng} 1_{Ng} \otimes N_g{}^{-1} \Sigma_W{}^{-1} .$$

Returning to the ML fitting function in (9), and considering the quadratic form for a certain group,

(24) $$q_g = (z_g^{*}{}' \; y_g^{*}{}') \begin{bmatrix} \Sigma_g^{11} & \text{symmetric} \\ \Sigma_g^{21} & \Sigma_g^{22} \end{bmatrix} \begin{pmatrix} z_g^{*} \\ y_g^{*} \end{pmatrix} ,$$

with

$$(25)\qquad \begin{pmatrix} z_g^* \\ y_g^* \end{pmatrix} = \begin{pmatrix} z_g - \mu_z \\ y_{g1} - \mu_y \\ \vdots \\ y_{gN_g} - \mu_y \end{pmatrix},$$

we obtain

$$(26)\qquad q_g = z^*_g{}' \, \Sigma_g^{11} \, z^*_g + 2\, y^*_g{}' \, \Sigma_g^{21} \, z^*_g + y^*_g{}' \, \Sigma_g^{22} \, y^*_g \, .$$

Kronecker algebra gives

$2(\bar{y}_g - \mu_y)' A^{21} (z_g -$

$$(27)\qquad 2\, y^*_g{}' \, \Sigma_g^{21} \, z^*_g = -\,2\, (\bar{y}_g - \mu_y)' \, A^{22} \, A_{21} \, A_{11}^{-1} \, (z_g - \mu_z) \, ,$$

$$(28)\qquad y^*_g{}' \, \Sigma_g^{22} \, y^*_g = (\bar{y}_g - \mu_y)' \, A^{22} \, (\bar{y}_g - \mu_y) + \\ + \sum_{i=1}^{N_g} (y_{gi} - \mu_y)' \, \Sigma_W^{-1} \, (y_{gi} - \mu_y) \\ - N_g \, (\bar{y}_g - \mu_y)' \, \Sigma_W^{-1} \, (\bar{y}_g - \mu_y) \, .$$

These expressions agree with those of McDonald and Goldstein (1988); see also Muthen (1989, (33)).

Reassembling these terms, using the definition of Σ_{vg} in (18), and defining

$$(29) \qquad (v_g - \mu) = \begin{pmatrix} z_g - \mu_z \\ \bar{y}_g - \mu_y \end{pmatrix}$$

it follows that we may write the quadratic form part of the ML fitting function of (9) as

(30)

$$\sum_{g=1}^{G} q_g = \mathrm{tr} \sum_{g=1}^{G} q_g$$

$$= \mathrm{tr} \{ \sum_{g}^{G} \Sigma_{vg}^{-1} (v_g - \mu) (v_g - \mu)'$$

$$+ \Sigma_W^{-1} \sum_g \sum_i^{N_g} (y_{gi} - \mu_y) (y_{gi} - \mu_y)'$$

$$- \Sigma_W^{-1} \sum_g N_g (\overline{y}_g - \mu_y) (\overline{y}_g - \mu_y)' \} ,$$

where simplification of the last two terms by centering at $\overline{y}_g$ gives

(31)

$$\mathrm{tr} \{ \sum_g \Sigma_{vg}^{-1} (v_g - \mu) (v_g - \mu)' +$$

$$+ \Sigma_W^{-1} (N - G) S_{PW} \}$$

defining S_{PW} as the usual pooled–within sample covariance matrix

(32)

$$S_{PW} = (N - G)^{-1} \sum_{g=1}^{G} \sum_{i=1}^{N_g} (y_{gi} - \bar{y}_g) (y_{gi} - \bar{y}_g)' .$$

The expression in (31) can be further developed to show a key connection with conventional structural equation modeling. The first term, containing a summation over the groups, can be rearranged as a double sum over distinct group sizes and groups of a particular size. Letting D stand for the number of distinct groups, N_d stand for the d^{th} group size, and G_d stand for the number of groups of the d^{th} size, we obtain

$$(33) \quad \sum_{g} \Sigma_{vg}^{-1} (v_g - \mu) (v_g - \mu)' = \sum_{d=1}^{D} \Sigma_{vd}^{-1} \sum_{k=1}^{G_d} (v_{dk} - \mu) (v_{dk} - \mu)'$$

$$= \sum_{d=1}^{D} G_d \Sigma_{vd}^{-1} \{ G_d^{-1} \sum_{k} (v_{dk} - \bar{v}_d) (v_{dk} - \bar{v}_d)' + (\bar{v}_d - \mu) (\bar{v}_d - \mu)' \}$$

where $\bar{v}_d$ represents the mean vector for z and y in group category d.

Collecting terms, the ML fitting function of (9) can now be written as

(34)

$$\sum_{d}^{D} G_d \ln \begin{vmatrix} \Sigma_{zz} & \text{symmetric} \\ \Sigma_{yz} & N_d^{-1}\Sigma_W + \Sigma_B \end{vmatrix} + (N - G) \ln |\Sigma_W| + \sum_{d} G_d \ln |N_d| +$$

$$\sum_{d=1}^{D} G_d \operatorname{tr} \{ \begin{bmatrix} \Sigma_{zz} & \text{symmetric} \\ \Sigma_{yz} & N_d^{-1}\Sigma_W + \Sigma_B \end{bmatrix}^{-1} [G_d^{-1} \sum_{k} (v_{dk} - \bar{v}_d)(v_{dk} - \bar{v}_d)' + (\bar{v}_d - \mu)(\bar{v}_d - \mu)'] \} +$$

$$+ \operatorname{tr} \{ \Sigma_W^{-1} (N - G) S_{PW} \} .$$

In terms of the software that will be used it is sligthly more convenient to consider the corresponding ML fitting function where we multiply Σ_{vd} by N_d. The ML fitting function of (34) may then be rewritten as

(35)

$$\sum_{d}^{D} G_d \{ \ln \begin{vmatrix} N_d \Sigma_{zz} & \text{symmetric} \\ N_d \Sigma_{yz} & \Sigma_W + N_d \Sigma_B \end{vmatrix} +$$

$$+ \text{tr} [\begin{bmatrix} N_d \Sigma_{zz} & \text{symmetric} \\ N_d \Sigma_{yz} & \Sigma_W + N_d \Sigma_B \end{bmatrix}^{-1} (S_{Bd} + N_d (\overline{v}_d - \mu) (\overline{v}_d - \mu)')] \} +$$

$$+ (N - G) \{ \ln |\Sigma_W| + \text{tr} [\Sigma_W^{-1} S_{PW}] \} ,$$

where S_{Bd} denotes the between–group matrix for groups of distinct size category d,

(36)

$$S_{Bd} = N_d G_d^{-1} \sum_{k=1}^{G_d} \begin{bmatrix} z_{dk} - \overline{z}_d \\ \overline{y}_{dk} - \overline{y}_d \end{bmatrix} [(z_{dk} - \overline{z}_d)' (\overline{y}_{dk} - \overline{y}_d)']$$

and S_{PW} is defined as in (32).

The way the log likelihood function of (35) is written shows that it may be optimized with respect to the model parameters using conventional structural equation modeling software that allows for the simultaneous analysis of multiple populations with structured means (see for example Muthen, 1983, 1984, 1989). For example, the maximum–likelihood (ML) fitting function in the generally available LISCOMP program (Muthen, 1987) is expressed as

$$(37) \quad \sum_{p=1}^{P} \{ N_p [\ln | \Sigma_p | + \mathrm{tr} (\Sigma_p^{-1} T_p) - \ln | S_p | - r] \} N^{-1} ,$$

$$(38) \quad T_p = S_p + (\bar{x}_p - \mu_p) (\bar{x}_p - \mu_p)' .$$

In (37), independent random samples from P populations with sample sizes N_p and total sample size N are considered. Here, an r–dimensional vector x, say, is observed with sample covariance matrix S_p, sample mean vector $\bar{x}_p$, population covariance matrix Σ_p, and population mean vector μ_p. The terms containing $\ln | S_p | - r$ are offsets so that a perfectly fitting model has the function value of zero. The sample covariance matrices S_p are the ML estimates of the unrestricted Σ_p matrices and are therefore divided by N_p, not $N_p - 1$. Multiplying the minimum value for any model by $2 \times N$ then gives the value of the likelihood–ratio chi–square test of the H_0 model against the H_1 model of unrestricted mean vectors μ_p and covariance matrices Σ_p.

The structural modeling ML expression of (37) shows that from a structural modeling point of view, the hierarchical data ML fitting function of (35) can be viewed as that corresponding to a simultaneous analysis of independent observations from D + 1 populations. The D between–group populations are viewed as having sample sizes G_d with sample covariance matrices S_{Bd} fitted to the population covariance structures

(39)

$$\begin{bmatrix} N_d \Sigma_{zz} & \text{symmetric} \\ N_d \Sigma_{yz} & \Sigma_W + N_d \Sigma_B \end{bmatrix}$$

while the within–group population has sample size N – G with the sample covariance matrix S_{PW} fitted to the covariance structure Σ_W. The mean vectors of the D between–group populations are viewed as all having the form $\sqrt{N_d} \times \mu$ with sample counterparts $\sqrt{\bar{v}_d}$, while the mean vector for the within population is viewed as being zero in the population and in the sample. A dummy variable arrangement can accommodate the fact that the first D populations have observations on q + p variables, whereas the last population has observations on only p variables. In the balanced case the analysis is particularly simple in that only two populations would be needed, but even in the unbalanced case the number of groups with distinct group sizes may be rather small in any given application.

Since the hierarchical data structure fits into the framework of maximum likelihood estimation with conventional structural modeling software, it follows that there is no need to specifically derive and program first– and second–order derivatives for getting estimates or expressions for estimated standard errors. In addition to the H_0 model, a separate analysis of an H_1 model is needed to obtain a likelihood–ratio chi–square test of model fit. Usually, the H_1 model would be the model where no restrictions are placed on the mean vectors and covariance matrices of (1) – (5). A separate analysis is needed because the corresponding H_1 model assumed in the use of conventional software would incorrectly neglect to impose equality restrictions on common mean vector and covariance matrix parameters across groups as indicated in (35). Longford and Muthen (1990) considers a simplification to the calculation of the H_1 likelihood value.

The use of conventional structural models in capturing mean and covariance structures for hierarchical data will be discussed in Section 4, showing that the structural modeling framework enables very general hierarchical data models to be specified. In fact, multiple populations of group units may be studied in a simultaneous analysis where each (factual) population is represented by the D + 1 imaginary populations just discussed. Structured means models can be included.

The analyses to be presented below in Section 6 were computed by a slightly revised version of LISCOMP (Muthen, 1987). Conventional structural modeling software usually applies offsets as in (37). In (35) it should be noted that the between sample covariance matrices may be singular due to being created by summation over fewer units than variables. This may prevent the use of certain conventional structural modeling software where positive definite matrices are assumed. Relatively modest revisions of conventional software could simplify the software use for hierarchical data. For example, by making the equation (35) division into D parts of the between structure of the model internal to the program, the user would only have to specify one between structure in addition to the within structure.

3. A simple ad hoc estimator

Consider balanced data as a special case of the ML fitting function of (35). Here, $D = 1$, $G_d = G$, and $N_d = N/G$ for all groups. When $\mathbf{z}_g$'s are absent and $\mathbf{y}_{gi}$'s are the only observed variables, $\mathbf{S}_B$ is the usual between–group covariance matrix, apart from dividing by G_d and not $G_d - 1$. For the special case of z variables being absent and the case of a model with no mean structure, (35) reduces to the maximum–likelihood fitting function considered in Schmidt (1969) and Schmidt and Wisenbaker (1986).

Consider again the balanced case and the common special case of a model with no mean structure, including both z and y variables. Here, the ML fitting function of (35) simplifies as

(40)

$$G \left\{ \ln \begin{vmatrix} c\,\Sigma_{zz} & \text{symmetric} \\ c\,\Sigma_{yz} & \Sigma_W + c\,\Sigma_B \end{vmatrix} + \text{tr} \begin{bmatrix} c\Sigma_{zz} & \text{symmetric} \\ c\,\Sigma_{yz} & \Sigma_W + c\,\Sigma_B \end{bmatrix}^{-1} S_B \right\} +$$
$$+ (N - G) \left\{ \ln | \Sigma_W | + \text{tr} [\Sigma_W^{-1} S_{PW}] \right\} ,$$

where c is the common group size, S_B is the single between-group covariance matrix obtained as the special balanced case of (36), and S_{PW} is as in (32). In the balanced case the ML estimates for the unrestricted model are according to standard ML theory

(41)

$$\widehat{\begin{bmatrix} c\Sigma_{zz} & \text{symmetric} \\ c\,\Sigma_{yz} & \Sigma_W + c\,\Sigma_B \end{bmatrix}} = S_B \,,$$

$$\widehat{\Sigma_W} = S_{PW} \,,$$

giving the ML estimate for the between matrix

(42)

$$\widehat{\begin{bmatrix} \Sigma_{zz} & \text{symmetric} \\ \Sigma_{yz} & \Sigma_B \end{bmatrix}} = c^{-1} \left(S_B - \begin{bmatrix} 0 & \text{symmetric} \\ 0 & S_{PW} \end{bmatrix} \right)$$

Given the ML estimators in (41) the customary test of H_0 against an unrestricted H_1 model is automatically carried out by conventional software using the standard offsets shown in (37). Hence, in this case, an additional analysis of H_1 is not needed to obtain a chi–square test of model fit.

The expression of (40) suggests a simple ad hoc estimator for the general unbalanced case with z and y variables and a model without mean structure. As indicated in (40), this estimator would use only one between model part in addition to the within model part. We may define a new S_B matrix as

(43)

$$S_B = (G-1)^{-1} \begin{bmatrix} c \sum_g (z_g - \bar{z})(z_g - \bar{z})' & \text{symmetric} \\ c \times G/N \sum_g N_g (\bar{y}_g - \bar{y})(z_g - \bar{z})' & \sum_g N_g (\bar{y}_g - \bar{y})(\bar{y}_g - \bar{y})' \end{bmatrix}$$

where the lower right–hand part of the partitioned matrix is the regular between matrix for unbalanced data on y. Letting S_{PW} be the regular pooled–within sample covariance matrix as before, it may be shown in line with Muirhead (1982, pp. 16 – 17) that the expected values are obtained as

(44)

$$E(S_B) = \begin{bmatrix} c\Sigma_{zz} & \text{symmetric} \\ c\,\Sigma_{yz} & \Sigma_W + c\,\Sigma_B \end{bmatrix}$$

(45) $E(S_{PW}) = \Sigma_W$

where in this general unbalanced case the constant c is expressed as (see also Graybill, 1961, p. 354)

(46)

$$c = [\,N^2 - \sum_{g=1}^{G} N_g^2\,]\;[\,N(G-1)\,]^{-1}.$$

In the balanced case, c reduces to the common group size.

Since S_B and S_{PW} tend to the corresponding population covariance matrices in (40) as the sample size increases, this means that using c of (46) and S_B of (43) in the fitting function of (40) gives a consistent estimator. When the data are balanced it is the ML estimator. When the data are in some sense not too far from being balanced, the estimator may not only give estimates close to those of ML but also approximations to the ML parameter estimate standard errors and the ML chi–square test of model fit. The estimator is simple to use. The sample covariance matrices S_B and S_{PW} can be obtained from standard software packages. The analysis of H_0 can be carried out in generally available conventional structural modeling software without modifications. And the analysis also provides a quasi chi–square test of H_0 against the unrestricted model of H_1. In the examples presented below we will use both the ML estimator and this ad hoc estimator for comparison purposes. While these results are encouraging, a more thorough study of the estimator seems warranted.

4. Modeling

A general covariance structure model will first be presented as an abstract framework within which particular models for hierarchical data will be discussed.

4.1 A general covariance structure framework

In line with the conventional structural equation modeling framework of Muthen (1984, 1987), consider the measurement model for the vector $\mathbf{x}' = (\mathbf{z}', \mathbf{y}')$ of length $q + p$,

(47) $\quad \mathbf{x} = \boldsymbol{\upsilon} + \boldsymbol{\Lambda}\,\boldsymbol{\eta} + \boldsymbol{\varepsilon}$,

together with the structural equations

(48) $\quad \boldsymbol{\eta} = \boldsymbol{\alpha} + \mathbf{B}\,\boldsymbol{\eta} + \boldsymbol{\zeta}$,

such that with $E(\boldsymbol{\eta}) = \boldsymbol{\alpha}$, $V(\boldsymbol{\varepsilon}) = \boldsymbol{\Theta}$, $V(\boldsymbol{\zeta}) = \boldsymbol{\Psi}$, usual assumptions give

(49) $\quad E(\mathbf{x}) = \boldsymbol{\nu} + \boldsymbol{\Lambda}(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\alpha}$,

(50) $\quad V(\mathbf{x}) = \boldsymbol{\Lambda}(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B})'^{-1}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}$.

In (47) and (48) $\boldsymbol{\eta}$ is a vector of random latent variables (factors), $\boldsymbol{\varepsilon}$ is a vector of random measurement errors, and $\boldsymbol{\zeta}$ is a vector of random residuals in the structural equations. The other arrays contain (non–random) parameters. This model framework is considered for several populations, where the populations may have some parameters in common, a feature which is captured by equality constraints on the parameters. This is the structural modeling framework for continuous variables used in the LISCOMP computer program. The ML fitting function was given in (37). The LISREL program can also fit such structures (see e.g. Joreskog, 1977; Muthen, 1983).

For purposes of modeling hierarchical data it is convenient to denote the parts of the parameter arrays according to three parts of η ,

(51) $\eta' = (\eta_Z' , \eta_B' , \eta_W') ,$

where η_Z is related to the between–group variation in **z** and η_B and η_W are related to between–group and within–group variation in **y**, respectively. In this way, the summation of Σ_W and Σ_B in the between covariance structure terms (populations) of the hierarchical data ML fitting function of (35) can be accomodated. The ad hoc estimation is approached in the same way. The parameter arrays are then partitioned as

$$(52)\quad \Lambda = \begin{bmatrix} \Lambda_{zz} & \Lambda_{zB} & \Lambda_{zW} \\ \Lambda_{yz} & \Lambda_{yB} & \Lambda_{yW} \end{bmatrix},$$

$$(53)\quad B = \begin{bmatrix} B_{zz} & B_{zB} & B_{zW} \\ B_{Bz} & B_{BB} & B_{BW} \\ B_{Wz} & B_{WB} & B_{WW} \end{bmatrix},$$

$$(54)\quad \Psi = \begin{bmatrix} \Psi_{zz} & \text{symmetric} & \\ \Psi_{Bz} & \Psi_{BB} & \\ \Psi_{Wz} & \Psi_{WB} & \Psi_{WW} \end{bmatrix},$$

and similarly for Θ. For the within covariance structure term, arrays with subscript B are set to zero. Section 5 contains a graphic display of this model structure as applied to a particular data set and Appendix 2 gives examples of array specifications.

In the model types to be considered below, the following parameter arrays are not used for the between or within structures, but are assumed to be set at zero:

(55) $\Lambda_{zB}, \Lambda_{zW}, \Lambda_{yz}, B_{zz}, B_{zB}, B_{zW}, B_{BW}, B_{Wz}, B_{WB}$.

The fact that these parameter arrays are also available for special applications reflects the richness of the possibilities of hierarchical data modeling in the structural modeling framework.

The hierarchical data ML fitting function of (35) shows that in the D first between terms the between parameter matrices Σ_{zz}, Σ_{yz}, Σ_B, are scaled by the distinct group size constants N_d, while μ_z and μ_y are scaled by $\sqrt{N_d}$. The last term only involves the within matrix Σ_W without scaling. The ad hoc estimator has analogous features. This scaling can be taken into account in the specification of Λ_{zz} and Λ_{yB} for the between terms. As an example, consider the scaling for matrices related to the y variables. With m variables in η_B, and letting c represent the group size constant,

(56) $\Lambda_{yB} = [\, c^{1/2}\, I_{p \times p} \quad O_{p \times m} \,]$,

$$(57) \quad B_{BB} = \begin{bmatrix} 0_{p\times p} & \Lambda_{B_{p\times m}} \\ 0_{m\times p} & B_{B_{m\times m}} \end{bmatrix},$$

$$(58) \quad B_{Bz} = \begin{bmatrix} B_{Byz_{p\times q}} \\ B_{B\eta z_{m\times q}} \end{bmatrix},$$

$$(59) \quad \Psi_{BB} = \begin{bmatrix} \Psi_{By_{p\times p}} & \text{symmetric} \\ 0_{m\times m} & \Psi_{B\eta} \end{bmatrix}.$$

The unrestriced H_1 model can be simply captured by letting $\boldsymbol{\Sigma}_B$ be represented by the $\boldsymbol{\Psi}$ matrix and $\boldsymbol{\Sigma}_W$ by the $\boldsymbol{\Theta}$ matrix. The scaling by the constant is handled by diagonal loading matrices for the between parts.

Identification issues will not be considered in this paper. We note however, that a model is identified if within–level parameters can be identified in terms of the Σ_W elements and between–level parameters can be identified in terms of Σ_B elements. This is a helpful observation since it brings the identification issue into the framework of conventional covariance structure analysis.

Some examples of model specifications are given in the appendices.

4.2 Specific models for hierarchical data

In this section we present some basic model variations, stressing their interpretation. Section 4.3 gives a discussion of further, more complex variations.

4.2.1 The Muthen–Satorra model for varying factor means

Muthen and Satorra (1989) and Muthen (1989) proposed a factor analysis model for hierarchical data where variations in factor means across groups was accounted for. Their motivating example was the same as the one to be analyzed in Section 5. Responses to achievement items are modeled as a factor analysis model. It is reasonable to assume that there is considerable heterogeneity in the achievement factor level since students come from classes with a variety of instructional histories. The model may be expressed as

$$(60)\quad \mathbf{y}_{gi} = \boldsymbol{\nu} + \boldsymbol{\Lambda}\, \boldsymbol{\eta}_{gi} + \boldsymbol{\varepsilon}_{gi}\,,$$

(61) $\boldsymbol{\eta}_{gi} = \boldsymbol{\alpha}_g + \boldsymbol{\eta}_{Wgi}$,

(62) $\boldsymbol{\alpha}_g = \boldsymbol{\alpha} + \mathbf{B}_{\alpha} \mathbf{z}_g + \boldsymbol{\eta}_{B\alpha g}$,

where α_g is a group–level random component, $\boldsymbol{\eta}_{Wgi}$ is a random vector with zero expectation and constant covariance matrix $\boldsymbol{\Psi}_{WW}$ across g and i, $\boldsymbol{\eta}_{B\alpha g}$ is a random vector with zero expectation and constant covariance matrix $\boldsymbol{\Psi}_{\alpha\alpha}$ across g and uncorrelated with $\boldsymbol{\eta}_{Wgi}$ and $\mathbf{z}_g$. Expressing $\boldsymbol{\eta}_{gi}$ as in (61) and (62) emphasizes that the model may be viewed as a random intercept (mean) model for the factors of $\boldsymbol{\eta}$. Conditional on group membership, expected factor values for an individual vary across groups. Consider first the special case of the model where there is no $\mathbf{z}$. Group (classroom) membership is assumed to bring with it a latent class–level (between) component $\boldsymbol{\eta}_{B\alpha g}$ which influences the outcome measures of $\mathbf{y}_{gi}$ through an influence on the individual factor values of $\boldsymbol{\eta}_{gi}$. For example, $\boldsymbol{\eta}_{B\alpha g}$ may reflect the part of the achievement factor level which is due to class–specific subject–matter training. Adding $\mathbf{z}$, we may bring in explanatory variables describing part of the class–specific variation, such as class records of instruction.

The model specifies

(63) $V(\mathbf{y}) = \boldsymbol{\Sigma}_W + \boldsymbol{\Sigma}_B$,

where

(64) $\boldsymbol{\Sigma}_W = \boldsymbol{\Lambda}\, \boldsymbol{\Psi}_{WW}\, \boldsymbol{\Lambda}' + \boldsymbol{\Theta}$,

while the between parts contain the covariance matrix for z, $\boldsymbol{\Sigma}_{zz}$, the covariance matrix for y,

(65) $\boldsymbol{\Sigma}_B = \boldsymbol{\Lambda}\, \mathbf{B}_\alpha\, \boldsymbol{\Sigma}_{zz}\, \mathbf{B}_\alpha{}'\, \boldsymbol{\Lambda}' + \boldsymbol{\Lambda}\, \boldsymbol{\Psi}_{\alpha\alpha}\, \boldsymbol{\Lambda}'$,

and the covariances between y and z variables

(66) $\boldsymbol{\Sigma}_{yz} = \boldsymbol{\Lambda}\;\; \mathbf{B}_\alpha\, \boldsymbol{\Sigma}_{zz}$.

Let us now express this model in terms of the general framework of Section 4.1, particularly (52) – (55) and (56) – (59). It suffices to consider the specification for the between terms. The between–group covariance matrices are parameterized as $\Lambda_{zz} = c^{1/2}\,\mathbf{I}$, $\Lambda_B = \Lambda_{yW} = \Lambda$, $\mathbf{B}_{Bz} = \mathbf{B}$, $\mathbf{B}_{BB} = \mathbf{0}$, $\mathbf{B}_{WW} = \mathbf{0}$, $\Psi_{Bz} = \mathbf{0}$, $\Psi_{Wz} = 0$, $\Psi_{Wz} = 0$, $\Psi_{WB} = 0$, $\Psi_{zz} = \Sigma_{zz}$, $\Psi_{By} = 0$, $\Psi_{B\eta} = \Psi_{\alpha\alpha}$, $\Psi_{WW} = \Psi_{WW}$, and as is customary in factor analysis, taking Θ to be diagonal with zeros in the first q diagonal entries and containing the variances of the ε's in the next p elements.

4.2.2 The factor model with varying factor means and measurement intercepts.

Assume the model

$$(67) \quad \mathbf{y}_{gi} = \boldsymbol{\nu}_g + \Lambda\, \boldsymbol{\eta}_{gi} + \boldsymbol{\varepsilon}_{gi}\,,$$

$$(68) \quad \boldsymbol{\nu}_g = \boldsymbol{\nu} + \mathbf{B}_\nu\, \mathbf{z}_g + \boldsymbol{\eta}_{\nu g}\,,$$

and $\boldsymbol{\eta}_{gi}$ defined as in our first model, Section 4.2.1. Here, the measurement intercepts of $\boldsymbol{\nu}_g$ are allowed to vary randomly across the groups. To continue the achievement example, certain of the **y** variables may correspond to subject–matter content which is relatively easier for students from classes which have obtained more training on these topics. We may assume that $V(\boldsymbol{\eta}_\nu) = \Psi_{\nu\nu}$ is a diagonal matrix and that η_{vg} is uncorrelated with $\eta_{\alpha g}$ and $\mathbf{z}_g$. While the within–group covariance matrix for $\mathbf{y}$ is still

(69) $\boldsymbol{\Sigma}_W = \boldsymbol{\Lambda}\, \boldsymbol{\Psi}_{WW}\, \boldsymbol{\Lambda}' + \boldsymbol{\Theta}$,

the between group covariance matrix for $\mathbf{y}$ is now

(70) $$\boldsymbol{\Sigma}_B = \mathbf{B}_v\, \boldsymbol{\Sigma}_{zz}\, \mathbf{B}_v' + \mathbf{B}_v\, \boldsymbol{\Sigma}_{zz}\, \mathbf{B}_\alpha'\, \boldsymbol{\Lambda}' + \boldsymbol{\Lambda}\, \mathbf{B}_\alpha\, \boldsymbol{\Sigma}_{zz}\, \mathbf{B}_v' +$$

$$+ \boldsymbol{\Lambda}\, \mathbf{B}_\alpha\, \boldsymbol{\Sigma}_{zz}\, \mathbf{B}_\alpha'\, \boldsymbol{\Lambda}' + \boldsymbol{\Lambda}\, \boldsymbol{\Psi}_{\alpha\alpha}\, \boldsymbol{\Lambda}' + \boldsymbol{\Psi}_{vv} .$$

If there are no $\mathbf{z}$ variables, only the last two terms of $\boldsymbol{\Sigma}_B$ remain and we note that the model implies that $\boldsymbol{\Sigma}_W$ and $\boldsymbol{\Sigma}_B$ follow the same type of factor model with different factor covariance matrix, different measurement error covariance matrix, but the same factor loading matrix.

This type of model is a natural extension of some work on measurement modeling in heterogeneous populations discussed in Muthen (1989), where it was argued that failure to take into account group–differences in measurement intercepts may lead to a distorted factor analysis. The model may be fitted into the general framework by letting $\mathbf{B}_{\mathrm{Byz}} = \mathbf{B}_{\nu}$, $\mathbf{B}_{\mathrm{B}\eta\mathrm{z}} = \mathbf{B}_{\alpha}$, and $\mathbf{\Psi}_{\mathrm{By}} = \mathbf{\Psi}_{\nu\nu}$, whereas other matrices are as for the model of varying factor means only in Section 4.2.1.

4.2.3. Structural equation modeling with varying levels

The model of the previous section may be directly generalized to include not only the measurement part relating $\mathbf{y}$ to $\boldsymbol{\eta}$, but also permitting linear structural relations among the variables of $\boldsymbol{\eta}$. The model is then as above, except that we write

$$(71) \quad \boldsymbol{\eta}_{gi} = \boldsymbol{\alpha}_g + \mathbf{B}\,\boldsymbol{\eta}_{gi} + \boldsymbol{\zeta}_{gi}\,,$$

where $\mathbf{B}$ has zero diagonal elements and $(\mathbf{I} - \mathbf{B})^{-1}$ is nonsingular. This permits randomly varying means for both independent (exogenous) and dependent latent variable constructs.

In terms of the general framework, we have the same model as in 4.2.2., except that we now have $\mathbf{B}_{\mathrm{B}} = \mathbf{B}$ and $\mathbf{B}_{\mathrm{ww}} = \mathbf{B}$.

4.2.4 Structural equation modeling with varying levels and different between– and within–group slopes.

The structural eqation model of 4.2.3 may be written as

$$(72) \quad \mathbf{y}_{gi} = \boldsymbol{\nu}_g + \boldsymbol{\Lambda}\,(\mathbf{I} - \mathbf{B})^{-1}\,\boldsymbol{\alpha}_g + \boldsymbol{\Lambda}\,(\mathbf{I} - \mathbf{B})^{-1}\,\boldsymbol{\zeta}_{gi} + \boldsymbol{\varepsilon}_{gi}\,.$$

This expression shows that the slopes of $\boldsymbol{\Lambda}$ and $\mathbf{B}$ are taken to be the same for group– and individual–level variation in $\boldsymbol{\eta}$ and $\mathbf{y}$. If one hypothesizes that group components may have different influence than individual components, one should instead consider the model with

$$(73) \quad \mathbf{y}_{gi} = \boldsymbol{\nu}_g + \boldsymbol{\Lambda}_B\,(\mathbf{I} - \mathbf{B}_B)^{-1}\,\boldsymbol{\alpha}_g +$$

$$+ \boldsymbol{\Lambda}_W\,(\mathbf{I} - \mathbf{B}_W)^{-1}\,\boldsymbol{\zeta}_{gi} + \boldsymbol{\varepsilon}_{gi}\,.$$

The model may be fitted into the general framework in an obvious way by relaxing equality restrictions related to $\boldsymbol{\Lambda}$ and $\mathbf{B}$.

This was the model used by Schmidt (1969), Schmidt and Wisenbaker (1986), and also considered by McDonald and Goldstein (1988). These authors were particularly interested in studying a multilevel model (Burstein, 1980) with separate relationships such that the full $\mathbf{\Lambda}$ and $\mathbf{B}$ matrices of (52) and (53) may be used.

Instead of viewing the model of (73) as having randomly varying parameters of factor means and measurement intercepts, one may in line with Schmidt (1969) simply view the model as having group– and individual variation in the random variables of the factors and the measurement errors. Excluding z variables for simplicity, we may specify a measurement and a structural model for both the between and the within level

(74)
$$\mathbf{y}_{Bg} = \boldsymbol{\nu} + \boldsymbol{\Lambda}_B \, \boldsymbol{\eta}_{Bg} + \boldsymbol{\varepsilon}_{Bg},$$
$$\boldsymbol{\eta}_B = \boldsymbol{\alpha} + \mathbf{B}_B \, \boldsymbol{\eta}_{Bg} + \boldsymbol{\zeta}_{Bg} \, ,$$
$$\mathbf{y}_{Wgi} = \boldsymbol{\Lambda}_W \, \boldsymbol{\eta}_{Wgi} + \boldsymbol{\varepsilon}_{Wgi},$$
$$\boldsymbol{\eta}_{Wgi} = \mathbf{B}_W \, \boldsymbol{\eta}_{Wgi} + \boldsymbol{\zeta}_{Wgi}.$$

4.3 Model extensions

In section 4.2.1 we considered the varying factor means model

$$(75) \quad y_{gi} = \nu + \Lambda\, \eta_{B\alpha g} + \Lambda\, \eta_{Wgi} + \varepsilon_{gi} ,$$

whereas in section 4.2.4 we used the measurement model

$$(76) \quad y_{gi} = \nu + \Lambda_B\, \eta_{Bg} + \Lambda_W\, \eta_{Wgi} + \varepsilon_{gi} .$$

In (75) we view the group–level factor component $\eta_{B\alpha g}$, or η_{Bg}, as influencing the y variables only indirectly through a change in η_{gi}. Explicitly allowing for a direct effect of η_{Bg} on the y variables leads to an alternative way to view the model of (76),

$$(77) \quad y_{gi} = \nu + \Lambda\, \eta_{gi} + \Lambda_D\, \eta_{Bg} + \varepsilon_{gi} ,$$

$$(78) \quad \eta_{gi} = \eta_{Bg} + \eta_{Wgi} ,$$

where we note that $\Lambda_W = \Lambda$, $\Lambda_B = (\Lambda + \Lambda_D)$.

These model variations show that one issue to consider is if and how the measurement parameters of the loadings (slopes) vary. Note, however, that none of the above models specifies a random parameter model for the loadings as was done for the measurement parameters of the intercepts. For certain applications randomly varying loadings could be of importance. This may be viewed as analogous to random slopes in regular multilevel regression models. However, randomly varying loadings lead to considerably more complex modeling than in the regression case since the independent variables of the factors are random, whereas in regression they are fixed, or conditioned upon.

Let us briefly outline how a randomly varying loadings model could be specified. Consider for simplicity a one–factor model with varying loadings and factor means,

(79) $$y_{gi} = \nu + \lambda_{gi}\, \eta_{gi} + \varepsilon_{gi},$$

(80) $$\eta_{gi} = \eta_{Bg} + \eta_{Wgi},$$

(81) $$\lambda_{gi} = \lambda + \zeta_g,$$

where all latent variable variance components may be taken to be uncorrelated. It may be shown that this model gives the covariance matrix components for **y**

(82) $\mathbf{\Sigma}_W = \boldsymbol{\lambda}\, V(\eta_w)\, \boldsymbol{\lambda}' + V(\eta_w)\, V(\boldsymbol{\zeta}) + \mathbf{\Theta},$

(82) $\mathbf{\Sigma}_B = \boldsymbol{\lambda}\, V(\eta_B)\, \boldsymbol{\lambda}' + V(\boldsymbol{\zeta})\, V(\eta_B)\ .$

Such models need special consideration in terms of parameterization, identification, and estimation. In terms of parameterization, it appears that at least some such covariance structures could be fitted in general structural equation modeling framework. We note, however, that if the $\boldsymbol{\varepsilon}_{gi}$, η_{Bg}, η_{wgi}, and $\boldsymbol{\zeta}_g$ terms are taked to be normally distributed the resulting distribution for **y** is not normal. The robustness of nevertheless using the ML fitting function for estimation and testing is of interest. Nonnormal variable estimators such as ADF (see e.g. Browne, 1984) could be attempted, fitting the model to two types of sample covariance matrices, between and within groups, in line with the ad hoc estimator. However, small numbers of groups may be an obstacle to good performance. One may also attempt true maximum likelihood estimation, e.g. by the EM approach. These and other issues will be considered in the author's future research.

5. An achievement example

To illustrate some of the possibilities of the new methodology for hierarchical data analysis, we will use data from the Second International Mathematics Study (Crosswhite, Dossey, Swafford, McKnight, & Cooney, 1985). We will be concerned with a subset of data from the population of U.S eighth–grade students./ A national probability sample of school districts was selected proportional to size within school district; two classes were randomly selected within each school. An achievement test was administered at the end of Spring 1982 containing a total of 180 items in the areas of arithmetic, algebra, geometry, and measurement. There were four test forms. Each student responded to a core test (40 items) and one of four randomly assigned rotated forms (35 items). Our analyses will consider the core and rotated form A items taken by 819 students from 179 different classrooms./ Data of this kind are typically analyzed as a simple random sample. In our analyses we will take into account that students are hierarchically observed within classrooms. For simplicity, however, we will ignore the level of hierarchies corresponding to school districts and schools, and assume simple random sampling of classrooms.

Eight achievement variable sums were created from the 75 right–wrong scored items corresponding to two sums for each of arithmetic, algebra, measurement, and geometry with the topic classifications: Ratio, proportion, percent (11 items); Common and decimal fractions (12 items); Equations and expressions (10 items); Integers, numbers (9 items); Standard units, estimation (6 items); Area, volume (5 items); Coordinates, vizualization (9 items); Plane figures (8 items).

The mathematics curriculum is very varied for U.S. 13–year olds. Comparing individuals, classrooms, schools, or school districts with respect to the total sum of correct answers may to a large extent reflect differences in opportunity to learn rather than achievement in areas taught. In this sense there is a risk of test bias due to instructional differences. From this point of view it is interesting that the data also contains information on such instructional differences across classrooms. A class–level variable categorizes the math classes into four types: basic or remedial; general or typical; pre–algebra or enriched; and algebra. Enriched and algebra classes are taught much more advanced topics than typical classes. In particular, more advanced topics related to algebra and geometry are introduced. The classification into class types was made based on both the text books reportedly used and by teacher–reported opportunity–to–learn (OTL). OTL was recorded for each item. For the purposes of the present analyses OTL was scored 1 if the teacher stated that the mathematics needed to answer the item correctly had been taught during eighth grade. For typical classes this means that almost all 0's imply that the topic had never been taught, while for enriched and algebra classes almost all 0's correspond to coverage during previous years. These dichotomous variables were summed up into eight OTL scores corresponding to the summing of the achievement items. For typical classes we expect positive effects on achievement from each of these OTL sums. However, for enriched and algebra classes we might obtain negative effects for less advanced topics since the fact that the topic was not sufficiently covered already before may indicate a low–achieving class.

Based on previous item–level analyses (see, e.g. Muthen, 1988), a one–factor model is expected for the eight student–level achievement variables. While several minor, topic–specific factors have been identified, the responses are dominated by a single general factor and the effect of this dominant factor is strengthened by the aggregation into eight scores. A conventional ML factor analysis of all 819 students ignoring class membership and the hierarchical nature of the data gave a chi–square test of fit of 35.72 with 20 degrees of freedom. In addition to this analysis of the total covariance matrix, we also analyzed the regular pooled–within classroom covariance matrix line with the suggestion in Muthen (1989). Here, the one–factor model obtained a chi–square of 26.28, specifying the sample size of N – G = 819 – 179 = 640. The corresponding analysis of the regular between matrix, specifying a sample size of 179, gave a chi–square of 37.50. We regard this fit as reasonably good, although there is some room for improvement. Although these conventional chi–square values are not completely trustworthy they do appear to indicate a relatively better fit on the within level than the between level. The corresponding three chi–square values for the 474 students in the 105 typical classes were 30.80 (sample size 474), 28.38 (sample size 369), and 36.09 (sample size 105).

Turning to hierarchical data modeling of the eight student–level achievement variables, we will specify the following model in line with the model variation in (74),

$$\mathbf{y}_{gi} = \boldsymbol{\nu} + \boldsymbol{\lambda}_B \, \eta_{Bg} + \boldsymbol{\lambda}_W \, \eta_{Wgi} + \boldsymbol{\varepsilon}_{Bg} + \boldsymbol{\varepsilon}_{Wgi} \, . \qquad (83)$$

Here, a single factor η is specified on both the group level and the individual level, where the group component captures contributions to an individual's achievement level from the class. The two components are allowed to have different loadings. As in conventional factor analysis, the metric of the factor can be set by fixing a loading to unity. This will be done for the first variable for both the between and within loadings. A test of equality across levels of the remaining loadings will be performed. If equality cannot be rejected, the factor part of the model can be interpreted as a varying factor means model in line with the Muthen–Satorra model of Section 4.2.1. In this case it is of interest to compare the between and within variance contributions to η. In particular, we will focus on the ratio of the between factor variance relative to the sum of between and within factor variances. The size of this ratio is an interesting indicator of the heterogeneity in achievement level across classrooms. A group–level component is also specified for the variable–specific residuals of $\boldsymbol{\varepsilon}$ to capture the anticipated classroom variation in variable–specific topic coverage. As discussed in Section 4, this variation may alternatively be seen as classroom variation in measurement intercepts corresponding to the perceived difficulty level of the respective set of items.

The specification of this model can be illustrated by a diagram such as that in Figure 1. In line with structural modeling conventions, squares denote observed variables and circles denote latent variables. The part of the model below the squares refer to the within structure, while the part above refers to the between structure. For the between terms of the hierarchical data ML fitting function of (35), both parts are involved, while for the within term only the bottom part is involved. This is in line with the parameter array arrangement discussed in Section 4.1.

Insert Figure 1

The LISCOMP input for the ad hoc estimator is given in Appendix 1. The setup for the ML estimator generalizes in a relatively straightforward fashion, adding the means. In the case of analyzing all 179 classrooms, there are 10 distinct class sizes, ranging from 1 to 10. Typical and Algebra/enriched classes are similar in this respect.

Table 1 gives the estimated model using both the ML estimator and the ad hoc estimator. The ML chi–square is 62.42 with 40 degrees of freedom and the ad hoc estimator quasi chi–square is 62.47. Adding the restrictions of equal loadings across levels gives a chi–square of 92.31 (93.90) with 47 degrees of freedom for the ML (ad hoc) estimator. Considering the chi–square difference, equality is rejected in this case. Table 1 presents estimates for the model with unequal loadings. The ML and ad hoc estimates and standard errors are close. We note that the between factor variance is strongly significant.

Insert Table 1

As in the conventional analysis, the model fit is marginal since the p value is just less than .01. As we will see in subsequent analyses, this may be due to the need for a more complex between structure. This is also in line with the relatively better fit to the pooled–within covariance matrix than the total or between matrices seen in the conventional analyses.

It is now of interest to bring in the available information on classroom differences in instruction. The class type categorization classified the 179 classrooms (819 students) into 105 typical and 56 algebra/enriched classes (474 and 285 students, respectively). A hierarchical data factor analysis by the same model was also carried out for each of these two types of classrooms separately. with about the same model fit as in the total set of classes.

For typical classes the ML chi–square was 65.389 for the 40 degree of freedom model with unequal loadings across groups and 72.30 with equality of loadings. As opposed to the analysis of the total set of classes, equality of loadings appears to fit very well, perhaps due to the larger degree of homogeneity of classes. It is then relevant to consider the between to total factor variance ratio. The between and within factor variances are estimated as 2.09 and 2.77, respectively, giving a ratio of 0.43. This indicates that a large degree of heterogeneity exists also among typical classes.

The analysis of algebra/enriched classes gave a ML chi–square of 48.66 with 40 degrees of freedom and adding equality of loadings gave a value of 65.98. Here, the seven–degree of freedom difference is significant, perhaps due to the larger differences among classes than in the typical category. The model with unequal loadings estimated the between and within factor variances as 2.01 and 2.98, respectively.

These analyses indicate that particularly for typical classes there is a large amount of heterogeneity in class achievement level. It is therefore of interest to investigate how much of this heterogeneity can be further accounted for by the teacher–reported OTL variables.

The class–level OTL variables may be seen as influencing the class–level part of the students' achievement scores indirectly through the class–level factor component in (84), η_{Bg}. For example, for typical classes we expect that the OTL scores increase the class level of overall achievement. However, as indicated in Table 1, the variable–specific variation in students' achievement across classrooms is strong and it is quite possible that each OTL variable has a direct effect on its corresponding achievement variable beyond any indirect effect through the factor η_B. In related, item–level analyses of the same data set Muthen, Kao and Burstein (1990) and Kao (1990) have investigated the extent of such direct effects, viewed as evidence of instructional item sensitivity. Muthen et al used dichotomous variables for both achievement items and OTL, while Kao aggregated topic–specific OTL variables. It is of interest to see if the present analysis, aggregating both achievement items and OTL items gives a different picture of instructional sensitivity. If there are large direct effects the use of the total test score for comparisons across groups of students with different amount of OTL may give a misleading achievement comparison.

The hierarchical data model with both student–level achievement variables and class–level OTL variables is shown in Figure 2. As for Figure 1, Figure 2 can be directly translated into the general modeling framework of Section 4.1.

Insert Figure 2

To aid in understanding the LISCOMP input for the ad hoc estimator, Appendix 2 gives the arrangement of the parameter arrays. The LISCOMP input is given in Appendix 3. Since the OTL variable part of the within covariance matrix has fixed values, the degrees of freedom given by LISCOMP will be too large. In this example, there are 100 such artificial restrictions. From a structural modeling point of view, the between part (top part) of Figure 2 corresponds to that of a MIMIC (multiple indicators, multiple causes) model (see also Muthen, 1989). While the within part of the model is that of a regular one–factor model for the achievement variables, the between part allows for correlations among the between components of the achievement variables that are explained not only by a single between factor, but also by correlations via the OTL variables and their direct effects. Given that the OTL variables are correlated, the existence of two or more direct OTL effects implies that a single factor does not explain all of the between correlation among the achievement variables, which could in turn contribute to the marginal fit observed in earlier analyses.

The hierarchical data model of Figure 2 was applied to the students of both typical and enriched/algebra classes. For simplicity, model fit is gauged approximately by the ad hoc estimator, alleviating the H_1 ML analysis. In both types of classes reasonably well-fitting models appear to have been found. For typical classes the ad hoc estimator quasi chi-square tests of fit were 143.76 and 137.38 for the models with and without loading constraints. The model with loading constraints does not fit significantly worse and will be presented (the ML chi-square difference from these two H_0 models was 7.33 and so agrees with this decision). The model has 95 degrees of freedom. The corresponding quasi chi-square values for algebra/enriched classes were 163.01 and 142.77, still indicating loading inequality.

We will now present the estimates for the model for typical classes. Table 2 shows the ML and ad hoc estimates of central parameters and their (quasi) significance. For ease of interpretation, the between-level structural estimates are given in standardized form corresponding to unit variances for the between-level variables (the circles above the squares in Figure 2).

Insert Table 2

Table 2 shows that when holding class type constant, only four of the eight OTL variables have any explanatory power. The variable OTL (Integers, Numbers) is the only OTL variable that has an effect on the between–level factor, while the variables OTL (Equations, expression), OTL (Area, volume), and OTL (Plane figures) have direct effects on their corresponding between–level achievement components. The estimated model says that OTL (Integers, Numbers) influences all the achievement variables indirectly through the class–level factor component. This OTL variable represents standard algebra topics for eighth grade students typical classes. The estimated R^2 for the class–level factor is 0.14. The more interesting OTL effects are the three direct ones; all three of these OTL variables correspond to topics that are usually new to eighth grade students in typical classes. OTL (Plane figures) correlates 0.51 and 0.54 with OTL (Equations, expression) and OTL (Area, volume), respectively. The existence of these three direct effects shows that the between–level variation cannot be properly explained by a single–factor construct. And it suggests that using a single total test score for comparison of students across classrooms is inappropriate if one wants to disentangle exposure from achievement on studied topics.

The between–level R^2's for the achievement variables are high, ranging from 0.78 for Coordinate, vizualization to 0.98 for Common and decimal fractions. The loadings indicate strong effects on the between components from the class–level factor, except on the Area, volume component. Arithmetic topics, followed by algebra topics, seem to dominate the class–level factor. However, as we have seen the OTL variables explain rather little of the class–level factor variation; other class–level characteristics will have to be sought for typical classes. The previously analyzed model without OTL variables obtained the between to total factor variance ratio of 0.43. In this model the between and within factor variances obtain ML estimates of 2.06 and 2.78, respectively, maintaining the ratio of 0.43. The corresponding residual variance ratio, holding OTL constant, is only reduced to 0.39.

The marginal fit of the hierarchical MIMIC model for typical classes was further investigated and this will be outlined since the strategy is of more general interest for hierarchical modeling. Including further direct effects did not improve the fit in any major way. Neither did the respecification allowing the influence of all the OTL variables to all the achievement variable–specific between components, excluding the influence on the factor. This latter respecification relaxes all between–level restrictions on correlations between OTL variables and achievement variable–specific between components. Since we have found the one–factor structure on the within level to fit well, the remaining model part to investigate is the one–factor specification for the between components of the achievement variables. Already our conventional between analysis pointed to the possibility of one–factor misfit, although the first eigenvalue was clearly dominant. It may be the case that a more complex model structure is warranted for this between model part. This question will not be pursued here, however.

The hierarchical MIMIC analysis of algebra/enriched classes, using the ad hoc estimator for simplicity, pointed to both positive and negative OTL effects on the class–level components. Five out of the eight OTL variables had strong positive direct effects on its corresponding achievement variable, OTL (Equations, expression), OTL (Standard units, estimation), OTL (Area, volume), OTL (Coordinate, vizualization), and OTL (Plane figures). OTL (Integers, numbers) was again found to have an effect on the class–level factor and in addition OTL (Common and decimal fractions) had such an effect. Both of these latter two effects, however, were negative. This is presumable because these topics are of rather low level for these types of classes and coverage during eighth grade and not earlier may reflect lower student achievement. In these classes the class–level factor R^2 obtains a high value of 0.78.

In summary, the hierarchical data analysis of this example uses a latent variable model that not only takes the hierarchical data nature into account, but is able to uncover class–level relationships that would be ignored by conventional analyses.

6. Conclusion

This paper has shown how covariance structure models can be formulated for hierarchical data and how they can be analyzed by maximum likelihood using a simple adaptation of conventional covariance structure software. A simple ad hoc estimator which can be carried out without change of existing software was also presented. This estimator gave results that were very close to those of the maximum likelihood estimator, but further research is needed before the estimator can be adapted for general use. The maximum likelihood estimator was applied to an educational example with students nested within classrooms. This analysis shows that new information is uncovered that conventional analysis would not be able to properly describe.

REFERENCES

Aitkin, M., & Longford, N. (1986). Statistical modeling issues in school effectiveness studies. Journal of the Royal Statistical Society, Series A, 149, 1, 1-43.

Browne, M.W. (1984). Asymptotically distribution free methods for the analysis of covariance structures. British Journal of Mathematical and Statistical Psychology, 37, 62-83.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. Review of Research in Education, 8, 158-233.

Crosswhite, F.J., Dossey, J. A., Swafford, J. O., McKnight, C. C., & Cooney, T. J. (1985). Second international mathematics study: Summary report for the United States. Champaign, IL: Stipes.

de Leeuw, J. (1985). Path models with random coefficients. Leiden, The Netherlands: University of Leiden, Department of Data Theory.

de Leeuw, J., & I. Kreft (1986). Random coefficient models for multilevel analysis. Journal of Educational Statistics, 11(1), 57-85.

Goldstein, H. I. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. Biometrika, 73, 43-56.

Goldstein, H. I. (1987). Multilevel models in educational and social research. London: Oxford University Press.

Goldstein, H. I., & McDonald, R. P., (1988). A general model for the analysis of multilevel data. Psychometrika, 53, 455-467.

Graybill, F.A. (1961) An Introduction to Linear Statistical Models Vol I. McGraw-Hill, New York.

Joreskog, K.G. (1977). Structural equation models in the social sciences: Specification, estimation and testing. In P.R. Krishnaiah (ed.) Applications of Statistics. Amsterdam: North-Holland.

Kao, C. F. (1990). An investigation of instructional sensitivity in mathematics achievements test items for U. S, eighth grade students. Unpublished doctoral dissertation, University of California, Los Angeles.

Longford, N.T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested effects. Biometrika, 74, 817-827.

Longford, N.T., Muthén, B. (1990). In preparation. University of California, Los Angeles.

Mason, W.M., Wong, G.Y., & Entwistle, B. (1984). Contextual analysis through the multi-level linear model. Sociological Methodology, San Francisco, CA: Jossey-Bass.

McDonald, R.P., & Goldstein, H. (1988). Balanced versus unbalanced designs for linear structural relations in two-level data. Paper presented at the Annual meeting of the Psychometric Society, Los Angeles.

Muirhead, R.J. Aspects of Multivariate Statistical Theory. (1982). John Wiley & Sons, New York.

Muthén, B. (1983). Latent variable structural equation modeling with categorical data. Journal of Econometrics, 22, 43-65.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika, 49, 115-132.

Muthén, B. (1987). LISCOMP. Analysis of linear structural equations with a comprehensive measurement model. Theoretical integration and user's guide. Mooresville, IN: Scientific Software.

Muthén, B. (1988). Some uses of structural equation modeling in validity studies: Extending IRT to external variables. In H. Wainer & H. Braun (Eds.), Test Validity (pp. 213-238). Hillsdale, NJ: Lawrence Erlbaum Associates.

Muthén, B. (1989). Latent variable modeling in heterogeneous populations. Presidential address to the Psychometric Society, July, 1989. Psychometrika, 54, 557-585.

Muthén, B. Kao, Chih-fen, & Burstein, L. (1990). Instructional sensitivity in mathematics achievement test items: Applications of a new IRT-based detection technique. Forthcoming in Journal of Educational Measurement.

Muthén, B., & Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. In Bock (Ed.), Multilevel Analysis of Educational Data (pp. 87-99). San Diego: Academic Press. (Invited paper for the conference on Multilevel analysis of educational data, Princeton, NJ, April 1987).

Raudenbush, S., & Bryk, A. (1988). Methodological advances in studying effects of schools and classrooms on student learning. Review of Research in Education, 1988.

Schmidt, W. H. (1969). Covariance structure analysis of the multivariate random effects model. Unpublished doctoral dissertation, University of Chicago.

Schmidt, W., & Wisenbaker, J. (1986). Hierarchical data analysis: an approach based on structural equations. (CEPSE, No.4., Research Series). University of Michigan, Department of Counseling Educational Psychology and Special Education.

Table 1
Estimated Model For Achievement Variables
All classes (z values in parenthesis)

Variable	ML				Ad Hoc			
	Loadings		Residual variances		Loadings		Residual variances	
	Between	Within	Between	Within	Between	Within	Between	Within
Ratio, proportion, percent	1. (-)	1. (-)	.10 (1.2)	2.10 (14.1)	1. (-)	1. (-)	.11 (1.3)	2.08 (13.9)
Common and decimal fractions	1.13 (18.9)	.95 (17.9)	.23 (2.2)	2.50 (15.1)	1.13 (18.4)	.94 (17.8)	.23 (2.1)	2.51 (15.0)
Equations, expressions	1.03 (18.0)	.66 (16.3)	.30 (3.2)	1.67 (15.9)	1.03 (17.5)	.65 (16.2)	.30 (3.2)	1.67 (15.8)
Integers, numbers	.90 (19.5)	.72 (17.3)	.08 (1.3)	1.63 (15.5)	.90 (18.7)	.71 (17.3)	.09 (1.4)	1.61 (15.3)
Standard units, estimation	.52 (16.1)	.46 (15.3)	.07 (2.1)	1.00 (16.4)	.52 (15.5)	.45 (15.3)	.09 (2.3)	.99 (16.2)
Area, volume	.41 (13.4)	.42 (15.0)	.11 (3.0)	0.89 (16.4)	.41 (13.2)	.42 (14.9)	.10 (2.9)	.89 (16.3)
Coordinate, visualization	.51 (13.7)	.50 (14.2)	.11 (2.2)	1.50 (16.7)	.51 (13.5)	.50 (14.1)	.11 (2.0)	1.51 (16.5)
Plane figures	.65 (15.1)	.57 (14.9)	.18 (2.7)	1.67 (16.5)	.65 (14.7)	.57 (14.9)	.18 (2.7)	1.67 (16.3)

	ML FACTOR VARIANCE		Ad Hoc FACTOR VARIANCE	
	Between	Within	Between	Within
	2.98 (6.6)	2.98 (10.9)	2.78 (6.6)	2.99 (10.9)

Achievement Variable	Loadings	R^2	1	2	3	4	5	6	7	
Ratio, proportion, percent	1. (1.)	.96 (.96)	0.09 (0.09)							
Common and decimal fractions	1.06 (1.06)	.98 (.97)		0.01 (0.01)						
Equations, expressions	.74 (.75)	.80 (.81)			0.22* (0.23*)					
Integers, numbers	.79 (.80)	.92 (.91)				0.08 (0.09)				
Standard units, estimation	.51 (.51)	.93 (.88)					-0.03 (-0.04)			
Area, volume	.41 (.41)	.82 (.82)						0.19* (0.18*)		
Coordinate, visualization	.48 (.49)	.78 (.77)							0.14 (0.13)	
Plane figures	.60 (.60)	.87 (.86)								0.22* (0.22*)
	Between-level Factor	.15 (0.14)	-0.07 (-0.10)	0.01 (-0.06)	-0.01 (0.05)	0.42* (0.45*)	-0.17 (-0.10)	0.06 (-0.02)	-0.20 (-0.20)	0.14 (0.07)

* Significant at the 5% level

nent variables

```
//IYX4BEN JOB
// EXEC LISCOMP
//SYSIN DD *
TI MULTILEVEL FACTOR ANALYSIS ONE FACTOR SIMS EIGHT VARIABLES AD HOC
TI
TI GROUP 1 BETWEEN
TI
DA IY=8 NG=2 NO=179
MO MO=SE P3 NE=10 LY=FI BE=FI PS=FI TE=FI
VA 2.1381 LY(1,1) LY(2,2) LY(3,3) LY(4,4) LY(5,5) LY(6,6) LY(7,7) LY(8,8)
FR  LY(2,10) LY(3,10) LY(4,10) LY(5,10) LY(6,10) LY(7,10)
   LY(8,10)
VA 1 LY(1,10) LY(2,10) LY(3,10) LY(4,10) LY(5,10) LY(6,10) LY(7,10)
   LY(8,10)
EQ BE(2,9) LY(2,10)
EQ BE(3,9) LY(3,10)
EQ BE(4,9) LY(4,10)
EQ BE(5,9) LY(5,10)
EQ BE(6,9) LY(6,10)
EQ BE(7,9) LY(7,10)
EQ BE(8,9) LY(8,10)
VA 1 BE(1,9) BE(2,9) BE(3,9) BE(4,9) BE(5,9) BE(6,9) BE(7,9) BE(8,9)
FR TE(1,1) TE(2,2) TE(3,3) TE(4,4) TE(5,5) TE(6,6) TE(7,7) TE(8,8)
VA 1  TE(1,1) TE(2,2) TE(3,3) TE(4,4) TE(5,5) TE(6,6) TE(7,7) TE(8,8)
FR PS(1,1) PS(2,2) PS(3,3) PS(4,4) PS(5,5) PS(6,6) PS(7,7) PS(8,8)
   PS(9,9) PS(10,10)
VA 1  PS(1,1) PS(2,2) PS(3,3) PS(4,4) PS(5,5) PS(6,6) PS(7,7) PS(8,8)
   PS(9,9) PS(10,10)
OU MN AL
SC
 18.5417
 17.7991    22.6345
 14.7456    16.4805    17.8631
 13.5260    14.9769    13.5230    13.8207
  8.1677     9.0267     7.7566     6.8916     5.5128
  6.4277     7.0745     6.2534     5.6848     3.1125     4.0353
  8.0290     8.3090     7.9135     7.0517     4.1459     3.4005     6.0943
 10.3805    11.1561     9.7292     8.2971     5.1475     4.2677     5.1598
  8.9307
TI
TI GROUP 2 WITHIN
TI
DA IY=8 NG=2 NO=640
MO MO=SE P3 NE=10 LY=FI BE=FI PS=FI TE=FI
EQ LY(2,2,10) LY(1,2,10)
EQ LY(2,3,10) LY(1,3,10)
EQ LY(2,4,10) LY(1,4,10)
EQ LY(2,5,10) LY(1,5,10)
EQ LY(2,6,10) LY(1,6,10)
EQ LY(2,7,10) LY(1,7,10)
EQ LY(2,8,10) LY(1,8,10)
VA 1 LY(1,10) LY(2,10) LY(3,10) LY(4,10) LY(5,10) LY(6,10) LY(7,10)
   LY(8,10)
EQ PS(2,10,10) PS(1,10,10)
VA 1  PS(10,10)
EQ TE(2,1,1) TE(1,1,1)
EQ TE(2,2,2) TE(1,2,2)
EQ TE(2,3,3) TE(1,3,3)
EQ TE(2,4,4) TE(1,4,4)
EQ TE(2,5,5) TE(1,5,5)
EQ TE(2,6,6) TE(1,6,6)
EQ TE(2,7,7) TE(1,7,7)
EQ TE(2,8,8) TE(1,8,8)
```

```
VA 1  TE(1,1) TE(2,2) TE(3,3) TE(4,4) TE(5,5) TE(6,6) TE(7,7) TE(8,8)
OU MN AL
SC
  5.0344
  2.6302     5.1245
  2.0000     2.0195     2.9660
  2.1942     1.9509     1.3970     3.1538
  1.3755     1.3266     0.8726     0.9389     1.6033
  1.2464     1.2065     0.7827     0.9302     0.5067     1.4182
  1.4512     1.3217     0.9370     1.0984     0.7647     0.6709     2.2655
  1.7001     1.5941     1.0003     1.2077     0.7407     0.7407     0.8900
  2.6232
```

Appendix 2

LISCOMP parameter array arrangement for the ad hoc estimator using eight achievement variables and eight OTL variables

Note:

The eight OTL variables are given as the first eight y variables. The latent variables are arranged in line with Figure 2. The matrix $\mathbf{B}_W = \mathbf{0}$.

$\Lambda_B =$

	η_1	η_2	η_3	η_4	η_5	η_6	η_7	η_8	η_9	η_{10}	η_{11}	η_{12}	η_{13}	η_{14}	η_{15}	η_{16}	η_{17}	η_{18}
y_1	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_2	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_3	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_4	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_5	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0	0	0	0
y_6	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0	0	0
y_7	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0	0
y_8	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	0	0
y_9	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	0	λ_W
y_{10}	0	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	0	λ_W
y_{11}	0	0	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	0	λ_W
y_{12}	0	0	0	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	0	λ_W
y_{13}	0	0	0	0	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	0	λ_W
y_{14}	0	0	0	0	0	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	0	λ_W
y_{15}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	0	λ_W
y_{16}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\sqrt{C}$	0	λ_W

$B_B =$

	η_1	η_2	η_3	η_4	η_5	η_6	η_7	η_8	η_9	η_{10}	η_{11}	η_{12}	η_{13}	η_{14}	η_{15}	η_{16}	η_{17}	η_{18}
η_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
η_9	x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_B	0
η_{10}	0	x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_B	0
η_{11}	0	0	x	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_B	0
η_{12}	0	0	0	x	0	0	0	0	0	0	0	0	0	0	0	0	λ_B	0
η_{13}	0	0	0	0	x	0	0	0	0	0	0	0	0	0	0	0	λ_B	0
η_{14}	0	0	0	0	0	x	0	0	0	0	0	0	0	0	0	0	λ_B	0
η_{15}	0	0	0	0	0	0	x	0	0	0	0	0	0	0	0	0	λ_B	0
η_{16}	0	0	0	0	0	0	0	x	0	0	0	0	0	0	0	0	λ_B	0
η_{17}	x	x	x	x	x	x	x	x	0	0	0	0	0	0	0	0	0	0
η_{18}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

$$
\Psi_B = \begin{array}{c|cccccccccccccccc|}
 & \zeta_1 & \zeta_2 & \zeta_3 & \zeta_4 & \zeta_5 & \zeta_6 & \zeta_7 & \zeta_8 & \zeta_9 & \zeta_{10} & \zeta_{11} & \zeta_{12} & \zeta_{13} & \zeta_{14} & \zeta_{15} & \zeta_{16} \\
\hline
\zeta_1 & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_2 & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_3 & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_4 & x & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_5 & x & x & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_6 & x & x & x & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_7 & x & x & x & x & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_8 & x & x & x & x & x & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 & 0 & 0 \\
\zeta_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 & 0 \\
\zeta_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 \\
\zeta_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 \\
\zeta_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 \\
\zeta_{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 \\
\zeta_{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x \\
\end{array}
$$

$\Theta_B =$

	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6	ε_7	ε_8	ε_9	ε_{10}	ε_{11}	ε_{12}	ε_{13}	ε_{14}	ε_{15}	ε_{16}
ε_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ε_9	0	0	0	0	0	0	0	0	x	0	0	0	0	0	0	0
ε_{10}	0	0	0	0	0	0	0	0	0	x	0	0	0	0	0	0
ε_{11}	0	0	0	0	0	0	0	0	0	0	x	0	0	0	0	0
ε_{12}	0	0	0	0	0	0	0	0	0	0	0	x	0	0	0	0
ε_{13}	0	0	0	0	0	0	0	0	0	0	0	0	x	0	0	0
ε_{14}	0	0	0	0	0	0	0	0	0	0	0	0	0	x	0	0
ε_{15}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x	0
ε_{16}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x

$\Lambda_W =$

	η_1	η_2	η_3	η_4	η_5	η_6	η_7	η_8	η_9	η_{10}	η_{11}	η_{12}	η_{13}	η_{14}	η_{15}	η_{16}	η_{17}	η_{18}
y_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y_9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W
y_{10}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W
y_{11}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W
y_{12}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W
y_{13}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W
y_{14}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W
y_{15}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W
y_{16}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	λ_W

$\Psi_W =$

	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5	ζ_6	ζ_7	ζ_8	ζ_9	ζ_{10}	ζ_{11}	ζ_{12}	ζ_{13}	ζ_{14}	ζ_{15}	ζ_{16}
ζ_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_{10}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_{11}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_{12}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_{13}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_{14}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_{15}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ζ_{16}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	x

$$
\Theta W = \begin{array}{c} \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{15} \\ \varepsilon_{16} \end{array}
\begin{array}{cccccccccccccccc}
\varepsilon_1 & \varepsilon_2 & \varepsilon_3 & \varepsilon_4 & \varepsilon_5 & \varepsilon_6 & \varepsilon_7 & \varepsilon_8 & \varepsilon_9 & \varepsilon_{10} & \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} & \varepsilon_{14} & \varepsilon_{15} & \varepsilon_{16} \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x
\end{array}
$$

Appendix 3

LISCOMP input for the ad hoc estimator using eight achievement variables

and eight OTL variables

```
//IYX4BEN JOB
// EXEC LISCOMP
//SYSIN DD *
TI MULTILEVEL FACTOR ANALYSIS ONE FACTOR SIMS EIGHT VARIABLES AD HOC
TI         WITH OTL FOR ALL STUDENTS
TI GROUP 1 BETWEEN
TI
DA IY=16 NG=2 NO=179
MO MO=SE P3 NE=18 LY=FI BE=FI PS=FI TE=FI
VA 2.13804  LY(1,1) LY(2,2) LY(3,3) LY(4.4) LY(5,5) LY(6,6) LY(7,7)
 LY(8,8) LY(9,9) LY(10,10) LY(11,11) LY(12,12) LY(13,13) LY(14,14)
 LY(15,15) LY(16,16)
FR  LY(10,18) LY(11,18) LY(12,18) LY(13,18) LY(14,18) LY(15,18)
   LY(16,18)
VA 1 LY(9,18) LY(10,18) LY(11,18) LY(12,18) LY(13,18) LY(14,18) LY(15,18)
   LY(16,18)
FR  BE(9,1) BE(10,2) BE(11,3) BE(12,4) BE(13,5) BE(14,6) BE(15,7)
  BE(16,8) BE(17,1) BE(17,2) BE(17,3) BE(17,4) BE(17,5) BE(17,6)
  BE(17,7) BE(17,8)          BE(10,17) BE(11,17) BE(12,17) BE(13,17)
   BE(14,17) BE(15,17) BE(16,17)
VA 1 BE(9,1) BE(10,2) BE(11,3) BE(12,4) BE(13,5) BE(14,6) BE(15,7)
  BE(16,8) BE(17,1) BE(17,2) BE(17,3) BE(17,4) BE(17,5) BE(17,6)
  BE(17,7) BE(17,8) BE(9,17) BE(10,17) BE(11,17) BE(12,17) BE(13,17)
   BE(14,17) BE(15,17) BE(16,17)
FR TE(9,9) TE(10,10) TE(11,11) TE(12,12) TE(13,13) TE(14,14)
   TE(15,15) TE(16,16)
VA 1 TE(9,9) TE(10,10) TE(11,11) TE(12,12) TE(13,13) TE(14,14)
   TE(15,15) TE(16,16)
FR PS(1,1) PS(2,2) PS(3,3) PS(4,4) PS(5,5) PS(6,6) PS(7,7) PS(8,8)
   PS(9,9) PS(10,10) PS(11,11) PS(12,12) PS(13,13) PS(14,14)
   PS(15,15) PS(16,16) PS(17,17) PS(18,18)
   PS(2,1) PS(3,1) PS(3,2) PS(4,1) PS(4,2)
   PS(4,3) PS(5,1) PS(5,2) PS(5,3) PS(5,4) PS(6,1) PS(6,2)
   PS(6,3) PS(6,4) PS(6,5) PS(7,1) PS(7,2) PS(7,3) PS(7,4)
   PS(7,5) PS(7,6) PS(8,1) PS(8,2) PS(8,3) PS(8,4) PS(8,5)
   PS(8,6) PS(8,7)
VA 1 PS(1,1) PS(2,2) PS(3,3) PS(4,4) PS(5,5) PS(6,6) PS(7,7) PS(8,8)
   PS(9,9) PS(10,10) PS(11,11) PS(12,12) PS(13,13) PS(14,14)
   PS(15,15) PS(16,16) PS(17,17) PS(18,18)
OU MN AL
SC
 0.328794D+02 0.272382D+02 0.463819D+02 0.159069D+01 -.584805D+01 0.329165D+02
 0.829732D+01 0.665241D+01 0.113146D+02 0.138911D+02 0.133674D+02 0.208024D+02
 0.178946D+01 0.729790D+01 0.209786D+02 0.821101D+01 0.107013D+02 0.131862D+01
 0.472602D+01 0.105862D+02 0.124889D+02 0.452895D+01 0.492802D+01 0.743633D+01
 0.715226D+01 0.829310D+01 0.646744D+01 0.148802D+02 0.120298D+02 0.127573D+02
 0.114615D+02 0.927398D+01 0.134128D+02 0.130632D+02 0.127723D+02 0.313686D+02
 -.303309D+01 -.965940D+01 0.894703D+01 0.361625D+01 -.319898D+01 -.148741D+01
 0.236301D+01 0.280442D+01 0.185418D+02 -.316206D+01 -.115046D+02 0.979588D+01
 0.399496D+01 -.513949D+01 -.239772D+01 0.108811D+01 0.317200D+00 0.177992D+02
 0.226347D+02 -.521663D+01 -.126894D+02 0.112245D+02 0.302305D+01 -.399315D+01
 -.172678D+01 0.167987D+01 0.650265D+00 0.147457D+02 0.164806D+02 0.178632D+02
 -.408291D+01 -.106087D+02 0.888684D+01 0.295462D+01 -.315722D+01 -.108353D+01
 0.176116D+01 0.119107D+01 0.135261D+02 0.149770D+02 0.135232D+02 0.138208D+02
 -.175670D+01 -.465956D+01 0.405993D+01 0.133074D+01 -.175026D+01 -.596946D+00
 0.571977D+00 0.122990D+01 0.816783D+01 0.902683D+01 0.775667D+01 0.689169D+01
 0.551286D+01 -.201000D+01 -.434065D+01 0.345960D+01 0.167257D+01 -.117573D+01
 -.156460D+00 0.110929D+01 0.715582D+00 0.642786D+01 0.707465D+01 0.625355D+01
 0.568493D+01 0.311254D+01 0.403544D+01 -.393684D+01 -.785468D+01 0.586062D+01
 0.139657D+01 -.245435D+01 -.755813D+00 0.137602D+01 0.809051D+00 0.802909D+01
 0.830916D+01 0.791360D+01 0.705185D+01 0.414599D+01 0.340060D+01 0.609444D+01
 -.235702D+01 -.736279D+01 0.648020D+01 0.166407D+01 -.245581D+01 -.724019D+00
 0.171368D+01 0.321947D+01 0.103806D+02 0.111562D+02 0.972928D+01 0.829716D+01
```

```
 0.514755D+01 0.426777D+01 0.515989D+01 0.893079D+01
TI
TI GROUP 2 WITHIN
TI
DA IY=16 NG=2 NO=640
MO MO=SE P3 NE=18 LY=FI BE=FI PS=FI TE=FI
EQ LY(2,10,18) LY(1,10,18)
EQ LY(2,11,18) LY(1,11,18)
EQ LY(2,12,18) LY(1,12,18)
EQ LY(2,13,18) LY(1,13,18)
EQ LY(2,14,18) LY(1,14,18)
EQ LY(2,15,18) LY(1,15,18)
EQ LY(2,16,18) LY(1,16,18)
VA 1 LY(9,18) LY(10,18) LY(11,18) LY(12,18) LY(13,18) LY(14,18) LY(15,18)
   LY(16,18)
EQ PS(2,18,18) PS(1,18,18)
VA 1  PS(18,18)
EQ TE(2,9,9) TE(1,9,9)
EQ TE(2,10,10) TE(1,10,10)
EQ TE(2,11,11) TE(1,11,11)
EQ TE(2,12,12) TE(1,12,12)
EQ TE(2,13,13) TE(1,13,13)
EQ TE(2,14,14) TE(1,14,14)
EQ TE(2,15,15) TE(1,15,15)
EQ TE(2,16,16) TE(1,16,16)
VA 1  TE(1,1) TE(2,2) TE(3,3) TE(4,4) TE(5,5) TE(6,6) TE(7,7) TE(8,8)
      TE(9,9) TE(10,10) TE(11,11) TE(12,12) TE(13,13) TE(14,14)
    TE(15,15) TE(16,16)
OU MN AL
SC
 0.100000D+01 0.000000D+00 0.100000D+01 0.000000D+00 0.000000D+00 0.100000D+01
 0.000000D+00 0.000000D+00 0.000000D+00 0.100000D+01 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.100000D+01 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.100000D+01 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.000000D+00 0.100000D+01 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.100000D+01
 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.503438D+01 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.263023D+01
 0.512456D+01 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.000000D+00 0.200004D+01 0.201947D+01 0.296605D+01
 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.219425D+01 0.195090D+01 0.139700D+01 0.315379D+01
 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.137556D+01 0.132662D+01 0.872592D+00 0.938937D+00
 0.160330D+01 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.000000D+00 0.124644D+01 0.120651D+01 0.782719D+00
 0.930233D+00 0.506747D+00 0.141824D+01 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.145123D+01
 0.132172D+01 0.937018D+00 0.109838D+01 0.764738D+00 0.670944D+00 0.226553D+01
 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00 0.000000D+00
 0.000000D+00 0.000000D+00 0.170008D+01 0.159410D+01 0.100027D+01 0.120771D+01
 0.740667D+00 0.740707D+00 0.889989D+00 0.262322D+01
```

Figure 1

Path diagram for a one–factor model

using within– and between–level components

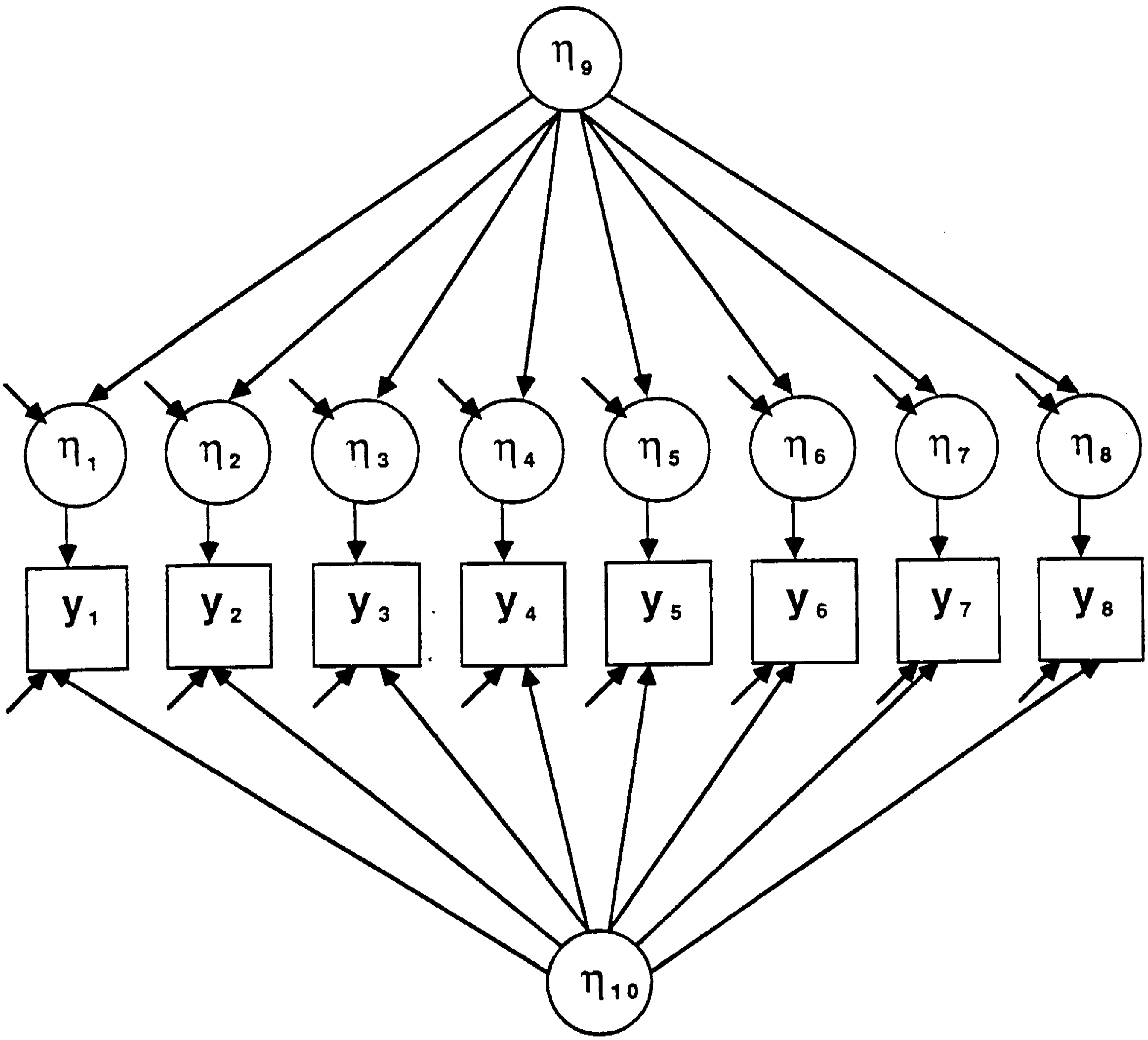
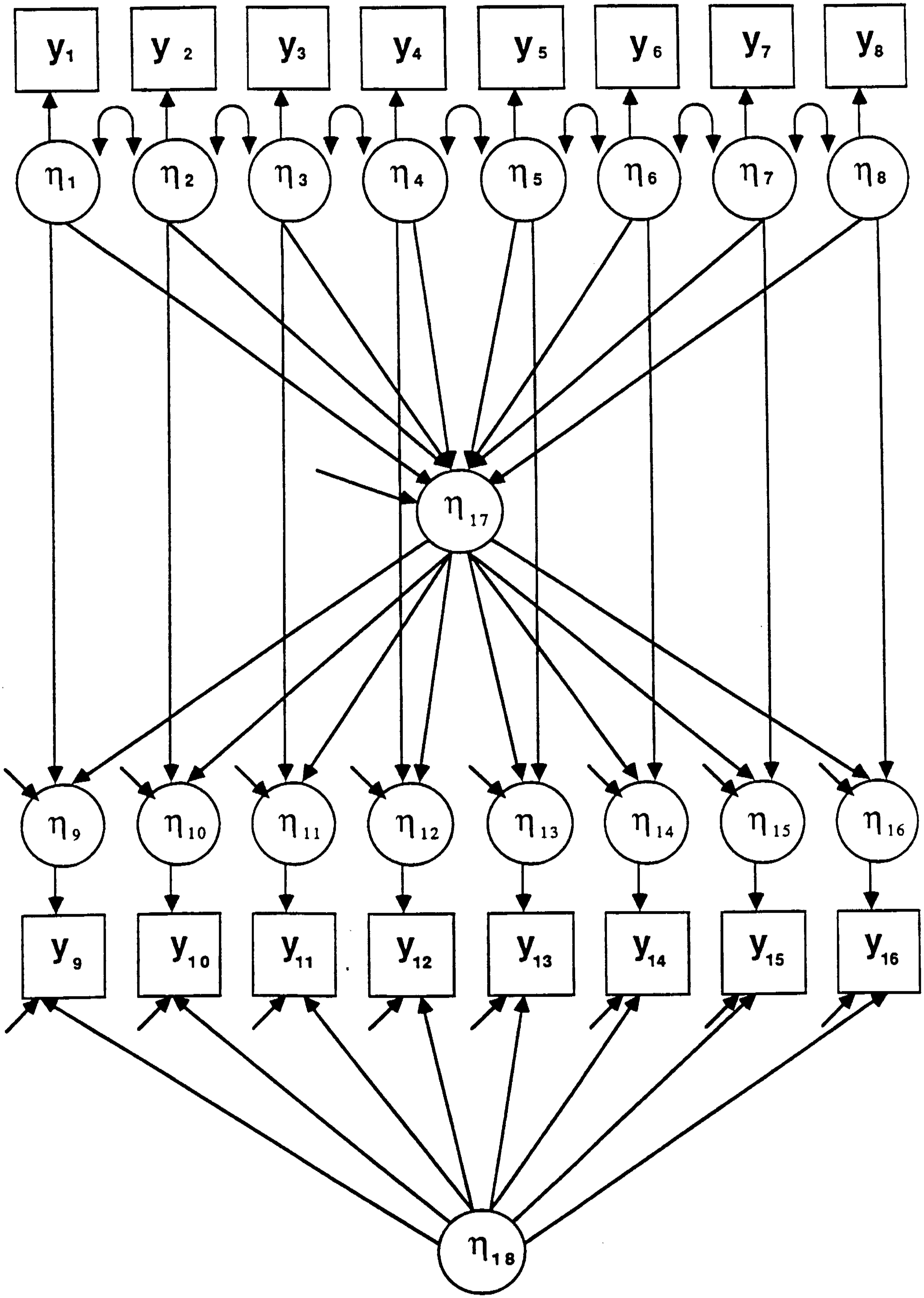

η9
η1
η2
η3
η4
η5
η6
η7
η8
y1
y2
y3
y4
y5
y6
y7
y8
η10

Figure 2

Path diagram for a two–level structural equation model

using individual– and class–level observations



y_1 y_2 y_3 y_4 y_5 y_6 y_7 y_8
η_1 η_2 η_3 η_4 η_5 η_6 η_7 η_8
η_{17}
η_9 η_{10} η_{11} η_{12} η_{13} η_{14} η_{15} η_{16}
y_9 y_{10} y_{11} y_{12} y_{13} y_{14} y_{15} y_{16}
η_{18}

(To obtain a copy of a paper, write to the author at the address given)

01 Multivariate Analysis with Linearizable Regressions - *Jan de Leeuw*
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

02 Scaling Corrections for Statistics in Covariance Structure Analysis - *Albert Satorra and Peter Bentler*
Peter Bentler, Psychology Dept., University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1563

03 Slicing Regression: Dimension Reduction via Inverse Regression - *Ker-Chau Li*
Ker-Chau Li, Mathematics Dept., University of California, 6364 MathSci, 405 Hilgard Ave., Los Angeles, CA 90024-1555

04 Model Conditions for Asymptotic Robustness in the Analysis of Linear Relations - *Albert Satorra and Peter Bentler*
Peter Bentler, Psychology Dept., University of California, Franz Hall 4627, 405 Hilgard Ave., Los Angeles, CA 90024-1563

05 Robust Statistical Modeling Using the t- Distribution - *Kenneth Lange, Rod Little and Jeremy Taylor*
Rod Little, Dept. of Biomathematics, AV-617 CHS, 10833 LeConte Ave., Los Angeles, CA 90024-1766

06 When Random Assignment Fails: Some Lessons from the Minneapolis Spouse Abuse Experiment - *Richard Berk*
Richard Berk, Sociology Dept., University of California, Haines Hall 264, 405 Hilgard Ave., Los Angeles, CA 90024-1551

07 Statistical Consultant for a Day - *Donald Ylvisaker*
Donald Ylvisaker, Mathematics Dept., University of California, MathSci 6364, 405 Hilgard Ave., Los Angeles, CA 90024-1555

08 Conjugate Gradient Methods in Confirmatory Factor Analysis - *Mortaza Jamshidian and Robert Jennrich*
Robert Jennrich, Mathematics Dept., University of California, MathSci 6364, 405 Hilgard Ave., Los Angeles, CA 90024-1555

09 Comparative Fit Indices in Structural Models - *Peter Bentler*
Peter Bentler, Psychology Dept., University of California, Franz Hall 4627, 405 Hilgard Ave., Los Angeles, CA 90024-1563

10 Data Modeling and Theory Construction - *Jan de Leeuw*
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

11 Correspondence Analysis, with Special Attention to the Analysis of Panel Data and Event History Data - *Peter van der Heyden and Jan de Leeuw*
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

12 A Primer on Robust Regression - *Richard Berk*
Richard Berk, Sociology Dept., University of California, Haines Hall 264, 405 Hilgard Ave., Los Angeles, CA 90024-1551

13 Minimax and Maximin Distance Designs - *Mark Johnson, Leslie Moore and Donald Ylvisaker*

Donald Ylvisaker, Mathematics Dept., University of California, MathSci 6364, 405 Hilgard Ave., Los Angeles, CA 90024-1555

14 Testing the Assumptions Underlying Tetrachoric Correlations
- Bengt O. Muthén and Charlie Hofacker
Bengt O. Muthén, Grad School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521

15 Factor Structure in Groups Selected on Observed Scores - *Bengt O. Muthén*
Bengt O. Muthén, Grad School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521

16 Some Models for the Analysis of Asymmetric Association in Square Contingency Tables with Ordered Categories - *Kazuo Yamaguchi*
Kazuo Yamaguchi, Sociology Dept., University of California, 264 Haines Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1551

17 Model Based Rankings of Schools - *Ita Kreft and Jan de Leeuw*
Ita G. G. Kreft, Center for the Study of Evaluation, University of California, 145 Moore Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1522

18 The Strange Case of Surge and Decline - *James Denardo*
James Denardo, Political Science Dept., University of California, 4289 Bunche Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1472

19 Logit and Multinomial Logit Models for Discrete-Time Event-History Analysis: An Exploratory Causal Analysis of Interdependent Discrete-State Processes
- Kazuo Yamaguchi
Kazuo Yamaguchi, Sociology Dept., University of California, 264 Haines Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1551

20 The Relation Between Input and Output Variables in Various Groups of Individuals
- Jan de Leeuw and Ita Kreft
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

21 Estimates of AIDS Incubation Periods from Convenience Samples
-Paul R. Abramson and Richard Berk
Richard Berk, Sociology Dept., University of California, Haines Hall 264, 405 Hilgard Ave., Los Angeles, CA 90024-1551

22 Modeling Prospective Duration: Gender Differences in Worker Intention to Remain in an Organization
- Kazuo Yamaguchi, Lynne G. Zucker, Anna Y. Leon-Guerrero, and Mark W. Plant
Kazuo Yamaguchi, Sociology Dept., University of California, 264 Haines Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1551

23 The Accuracy of the Stability of Perimeter Estimates of the Endogenous Feedback Model for Contextual Analysis - *Chih-Ping Chou and Leigh Burstein*
Chih-Ping Chou, Psychology Dept., University of California, 3558 Franz Hall, 405 Hilgard Ave.,Los Angeles, CA 90024-1563

24 "Optimal" Transformation of the Dependent Variable: Another Look at Slicing Regression - *Ker-Chau Li*
Ker-Chau Li, Mathematics Dept., University of California, 6364 MathSci, 405 Hilgard Ave., Los Angeles, CA 90024-1555

25 Slicing Regression: A Link-free Regression Method - *Naihua Duan and Ker-Chau Li*
Naihua Duan, Rand Corporation, 1700 Main St., Santa Monica, CA 90406

26 A Bias Bound for Applying Linear Regression to a General Linear Model
- *Naihua Duan and Ker-Chau Li*
Naihua Duan, Rand Corporation, 1700 Main St., Santa Monica, CA 90406

27 The Performance of Tests When Observations Have Different Variances
- *Peter A. Lachenbruch*
Peter A. Lachenbruch, Dept. of Biostatistics, School of Public Health, 10833 Le Conte Ave., Los Angeles, CA 90024-1772

28 Exact Analysis for Binomial Data with Covariates - *Karim F. Hirji*
Karim F. Hirji, Dept. of Biomathematics, AV-617 CHS, 10833 LeConte Ave., Los Angeles, CA 90024-1766

29 Regression Analysis with an Unknown Link Function: the Adjoint Projection Pursuit Regression - *Naihua Duan*
Naihua Duan, Rand Corporation, 1700 Main St., Santa Monica, CA 90406

30 Homophily and Social Distance in the Choice of Multiple Friends: An Analysis Based on Conditionally Symmetric Log-Bilinear Association Models - *Kazuo Yamaguchi*
Kazuo Yamaguchi, Sociology Dept., University of California, 264 Haines Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1551

31 A Quasi-Exact Test for Comparing Two Binomial Proportions
- *Karim F. Hirji, Shu-Jane Tau and Robert Elashoff*
Robert Elashoff, Dept. of Biomathematics, AV-617 CHS, 10833 LeConte Ave., Los Angeles, CA 90024-1766

32 Linear Dynamical Systems Analysis by Alternating Least Squares
- *Jan de Leeuw and Catrien Bijleveld*
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

33 Stochastic Microenvironment Models for Air Pollution Exposure - *Naihua Duan*
Naihua Duan, Rand Corporation, 1700 Main St., Santa Monica, CA 90406

34 A Latent Time-Budget Model
- *Jan de Leeuw, Peter G.M. van der Heijden and Peter Verboon*
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

35 Fitting Reduced Rank Regression Models by Alternating Maximum Likelihoods
- *Jan de Leeuw*
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

36 Estimating the Ratio of Parameters in Ordinal Response Regression Models
- *Jeremy M. G. Taylor*
Jeremy Taylor, Dept. of Biostatistics, School of Public Health, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1772

37 Estimating the Distribution of Times from HIV Seroconversion to AIDS Using Multiple Imputation - *Jeremy M.G. Taylor, Alvaro Munoz, Sue M. Bass, Joan S. Chmiel, Lawrence A. Kingsley, and Alfred J. Saah*Jeremy Taylor, Dept. of Biostatistics, School of Public Health, University of California, 10833 LeConte Ave., Los Angeles, CA 90024-1772

38 An Imputation Method for Dealing with Missing Data in Regression
- Lee G. Cooper, Jan de Leeuw, and Aram G. Sogomonian
Lee Cooper, Anderson Graduate School of Management University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1481

39 Analysis of Public-Use Data with Multiply-Imputed Industry and Occupation Codes
-Nathaniel Schenker, Donald J. Treiman, and Lynn Weidman
Nathaniel Schenker, Dept. of Biostatistics, School of Public Health, University of California, 10833 LeConte Ave., Los Angeles, CA 90024-1772

40 Trends in Health Statistics and the Spiegelman Award *- Peter A. Lachenbruch*
Peter A. Lachenbruch, Dept. of Biostatistics, School of Public Health, 10833 LeConte Ave., Los Angeles, CA 90024-1772

41 Cross-sectional and Dynamic Analyses of the Concomitants of Police Behavior
- Catrien C. J. H. Bijleveld and Eric H. Monkkonen
Eric H. Monkkonen, Dept. of History, 6265 Bunche Hall, University of California, Los Angeles, CA 90024-1473

42 Signed-Rank Tests for Censored Matched Pairs *- Dorota M. Dabrowska*
Dorota M. Dabrowska, Dept. of Biostatistics, School of Public Health, 10833 LeConte Ave., Los Angeles, CA 90024-1772

43 An Evaluation of the UCLA AIDS Course
- Paul R. Abramson, Joan C. Sekler, Richard Berk and Monique Clowd
Richard Berk, Sociology Dept., University of California, Haines Hall 264, 405 Hilgard Ave., Los Angeles, CA 90024-1551

44 What Your Mother Never Told You About Randomized Field Experiments
- Richard Berk
Richard Berk, Sociology Dept., University of California, Haines Hall 264, 405 Hilgard Ave., Los Angeles, CA 90024-1551

45 Reduced Rank Models for Contingency Tables
- Jan de Leeuw and Peter van der Heijden
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

46 Latent Variable Modeling in Heterogeneous Populations *- Bengt O. Muthén*
Bengt O. Muthén, Graduate School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521

47 Standard Errors for the Means of Proficiencies in NAEP: Jackknife versus Variance Components *- Nicholas T. Longford*
Nicholas T. Longford, Mathematics Dept., University of California, MathSci 6364, 405 Hilgard Ave., Los Angeles, CA 90024-1555

48 Conjugate Gradient Acceleration of the EM Algorithm
- Mortaza Jamshidian, Robert I. Jennrich
Robert I Jennrich, Department of Mathematics, University of California 405 Hilgard Ave., Los Angeles, CA 90024-1555, U.S.A.

49 Geometrical Aspects of Multiple Correspondence Analysis: Implications for the Coordinate Scaling Debate *- Donna L. Hoffman, Jan de Leeuw*

Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

50 Comparing Four Different Statistical Packages for Hierarchical Linear Regression: GENMOD, HLM, ML2, and VARCL
- Ita G.G. Kreft, Jan de Leeuw, and Kyung-Sung Kim
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

51 Pseudo Maximum Likelihood Estimation and Elliptical Distribution Theory in a Misspecified Model - *Yutaka Kano, Maia Berkane, and Peter M. Bentler*
Peter Bentler, Psychology Dept., University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1563

52 On Principal Hessian Directions for Data Visualization and Dimension Reduction: Another Application of Stein's Lemma - *Ker-Chau Li*
Department of Mathematics, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1555

53 Variable Bandwidth Kaplan-Meier Estimate - *Dorota M. Dabrowska*
Dorota M. Dabrowska, Dept. of Biostatistics, School of Public Health, 10833 LeConte Ave., Los Angeles, CA 90024-1772

54 Equality Constraints in Multiple Correspondence Analysis
-Stef van Buuren and Jan de Leeuw
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

55 Richness Curves for Evaluating Market Segmentation
- Thomas P. Novak, Jan de Leeuw, and Bruce MacEvoy
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

56 Teaching Biostatistics to Postgraduate Students in Public Health
- Peter A. Lachenbruch
Peter A. Lachenbruch, Dept. of Biostatistics, School of Public Health, 10833 Le Conte Ave., Los Angeles, CA 90024-1772

57 Accelerated Failure Time Regression Model with a Regression Model of Surviving Fraction: An Application to the Analysis of "Permanent Employment" in Japan
- Kazuo Yamaguchi
Kazuo Yamaguchi, Sociology Dept., University of California, 264 Haines Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1551

58 Additive Logistic Elliptical Distributions
- Kai-Tai Fang, Peter M. Bentler, Chih-Ping Chou
Peter Bentler, Psychology Dept., University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1563

59 Information Theory and an Extension of the Maximum Likelihood Principle by Hirotogu Akaike - *Jan de Leeuw*
Jan de Leeuw, Program in Social Statistics, Institute for Social Science Research, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1484

60 The Death Penalty and the Role of Chance - *Richard Berk and Robert Weiss*
Richard Berk, Sociology Dept., University of California, Haines Hall 264, 405 Hilgard Ave., Los Angeles, CA 90024-1551

61 Analysis of Longitudinal Data Using Latent Variable Models with Varying Parameters - *Bengt O. Muthén*
Bengt O. Muthén, Grad School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521

62 Mean and Covariance Structure Analysis of Hierarchical Data - *Bengt O. Muthén*
Bengt O. Muthén, Grad School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521

63 Applying Logistic Regression to Choice-Based and Supplementary Samples for the Prediction of MBA Matriculation and Persistence - *Bengt O. Muthén, Michael Hollis, Linda Muthén, Tammy Tan*
Bengt O. Muthén, Grad School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521

64 Selection and Predictive Validity with Latent Variable Structure - *Bengt O. Muthén, Jin-Wen Yang Hsu*
Bengt O. Muthén, Grad School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521

65 A Comparison of Some Methodologies for the Factor Analysis of Non-Normal Likert Variables: A Note on the Size of the Model - *Bengt O. Muthén and David Kaplan*
Bengt O. Muthén, Grad School of Education, SRM, University of California, 405 Hilgard Ave., Los Angeles, CA 90024-1521