Chapter 8

Multilevel Structural Equation Models: the Limited Information Approach and the Multivariate Multilevel Approach

Joop Hox and Cora Maas

University of Utrecht, Department of Methodology and Statistics. Utrecht, The Netherlands.

8.1 Introduction

Social and behavioral research often involves problems and data that are viewed as a hierarchical system, with individuals and groups defined at separate levels of this hierarchical system. Standard multivariate models are not appropriate for the analysis of such hierarchical systems, even if the analysis includes only variables at the lowest (individual) level, because the standard assumption of independent and identically distributed observations is generally not valid. The consequences of using uni-level analysis methods on multilevel data are well known: the parameter estimates are unbiased but inefficient, and the standard errors are negatively biased, which results in spuriously 'significant' effects (cf. Snijders and Bosker, 1999; Hox, 2002). Multilevel analysis techniques for the linear multiple regression model are well developed and the required software is widely available (Raudenbush and Bryk, 2002; Goldstein, 1995).

Structural equation modeling, or *SEM*, is a very general framework for statistical modeling that includes as special cases several traditional multivariate procedures, such as factor analysis, multiple regression analysis,

K. van Montfort et al. (eds.), Recent Developments on Structural Equation Models, 135-149.

discriminant analysis, and canonical correlation. Structural equation models for multilevel data have been formulated by, among others, Goldstein and McDonald (Goldstein & McDonald, 1988; McDonald & Goldstein, 1989, McDonald, 1994), Muthén and Satorra (Muthén, 1989; Muthén & Satorra, 1989) and Longford and Muthén (Longford & Muthén, 1992). We refer to McArdle and Hamagami (1996) for a comparison between multilevel regression techniques and standard multigroup SEM.

The approach to multilevel SEM outlined by Muthén (1989a, 1994) is particularly interesting, because he shows that structural equation modeling of multilevel data is possible using available standard SEM software, such as LISREL (Jöreskog & Sörbom, 1996), EQS (Bentler, 1995), or AMOS (Arbucle & Wothke, 1999). For an introductory exposition of Muthén's method, see Muthén (1994), Kaplan and Elliot (1997) and Hox (2002). Heck and Thomas (2000) present an extended example of multilevel SEM, using Muthén's method and discusses the implementation details for the programs LISREL, STREAMS, and M*PLUS*.

A different approach to estimating multilevel SEM is to estimate the covariance matrices at the distinct levels directly, using standard multilevel regression software, as proposed by Goldstein (1987, 1995) and applied, e.g., by Rowe and Hill (1998) and Rowe (2002). This approach has the advantage that it also uses standard SEM software, but the models and hence the program setups are far less complicated than the models and setups implied by the Muthén approach. For details, we refer to Hox (2002).

This chapter summarizes both approaches in some detail, and discusses their theoretical advantages. Next, both are compared on an exemplary data set. The results for both approaches are very similar.

8.2 The Muthén Approach: Decomposing Multilevel Variables

Multilevel structural models assume that we have a population of individuals that are divided into groups. The individual data are collected in a p-variate vector $\mathbf{Y}_{ig}$ (subscript i for individuals, g for groups). The variates $\mathbf{Y}_{ig}$ can be decomposed into a between groups component $\mathbf{Y}_B = \overline{\mathbf{Y}}_g$, and a within groups component $\mathbf{Y}_W = \mathbf{Y}_{ig} - \overline{\mathbf{Y}}_g$. In other words, for each individual we replace the observed *Total* score $\mathbf{Y}_T = \mathbf{Y}_{ig}$ by its components: the group component $\mathbf{Y}_B$ (the disaggregated group mean) and the individual component $\mathbf{Y}_W$ (the individual deviation from the group mean.) These two components have the attractive property that they are orthogonal and additive (cf. Searle, Casella & McCulloch, 1992):

$$\mathbf{Y}_T = \mathbf{Y}_B + \mathbf{Y}_W \tag{1}$$

This decomposition can be used to compute a between groups covariance matrix $\mathbf{\Sigma}_B$ (the population covariance matrix of the disaggregated group means $\mathbf{Y}_B$) and a within groups covariance matrix $\mathbf{\Sigma}_W$ (the population covariance matrix of the individual deviations from the group means $\mathbf{Y}_W$). These covariance matrices are also orthogonal and additive:

$$\mathbf{\Sigma}_T = \mathbf{\Sigma}_B + \mathbf{\Sigma}_W \tag{2}$$

Following the same logic, we can also decompose the sample data. Suppose we have data from N individuals, divided into G groups of equal size n (subscript i for individuals, $i=1...N$; subscript g for groups, $g=1...G$). If we decompose the sample data, we have for the sample covariance matrices:

$$\mathbf{S}_T = \mathbf{S}_B + \mathbf{S}_W \tag{3}$$

An unbiased estimate of the population within groups covariance matrix $\mathbf{\Sigma}_W$ is given by the pooled within groups covariance matrix $\mathbf{S}_{PW}$, calculated in the sample by:

$$\mathbf{S}_{PW} = \frac{\sum_g^G \sum_i^n \left(Y_{ig} - \overline{Y}_g\right)\left(Y_{ig} - \overline{Y}_g\right)'}{N-G} \tag{4}$$

Equation (4) corresponds to the conventional equation for the covariance matrix of the individual group-mean deviation scores, with N-G in the denominator instead of the usual N-1.

For computational reasons it is convenient to calculate not the between groups covariance matrix $\mathbf{S}_B$ itself but the *scaled between groups covariance matrix* for the disaggregated group means $\mathbf{S^*}_B$, given by:

$$\mathbf{S}_B^* = \frac{\sum_g^G n\left(\overline{Y} - \overline{Y}_g\right)\left(\overline{Y} - \overline{Y}_g\right)'}{G-1} \tag{5}$$

In equation (5), $\overline{Y}$ is the vector of overall means, and $\overline{Y}_g$ is the vector of group means. Muthén (1989a, 1990) shows that $\mathbf{S}_{PW}$ is the maximum likelihood estimator of $\mathbf{\Sigma}_W$, with sample size N-G, and $\mathbf{S}^*_B$ is the maximum likelihood estimator of the composite $\mathbf{\Sigma}_W + c\mathbf{\Sigma}_B$, with sample size G, and c depending on the group sizes:

$$\mathbf{S}_{\mathrm{PW}} = \hat{\Sigma}_{\mathrm{W}} \tag{6}$$

and

$$\mathbf{S}^{*}_{\mathrm{B}} = \hat{\Sigma}_{\mathrm{W}} + c\hat{\Sigma}_{\mathrm{B}} \tag{7}$$

Equations 6 and 7 suggest using the multi-group option of conventional SEM software for a simultaneous analysis at both levels. If we model the between groups structure, we cannot simply construct and test a model for Σ_B, because S^*_B estimates a combination of Σ_W and Σ_B. Instead, we specify for the S^*_B 'group' two models: one for the within groups structure and one for the between groups structure, with covariance matrices S_{PW} and S^*_B (based on *N-G* and *G* observations). The model for Σ_W is specified for both S_{PW} and S^*_B, with equality restrictions between both 'groups' to estimate the same model in both covariance matrices, and the model for Σ_B is specified for S^*_B only, with the scale factor c built into the model.

This reasoning strictly applies only to the *balanced* case, that is, if all groups have the same group size n. In the balanced case, the scale factor c is equal to the common group size n. The unbalanced case, where the group sizes differ, with G groups of unequal sizes n_g, is more complicated. In this case, S_{PW} is still the maximum likelihood estimator of Σ_W, but S^*_B now estimates a different expression for each set of groups with distinct group size $d=n_d$:

$$\mathbf{S}^{*}_{\mathrm{B}d} = \hat{\Sigma}_{\mathrm{W}} + c_d\hat{\Sigma}_{\mathrm{B}} \tag{8}$$

where equation 8 holds for each distinct set of groups with a common group size equal to n_d, and $c_d=n_d$ (Muthén, 1990, 1994). Full Information Maximum Likelihood (FIML) estimation for unbalanced groups implies specifying a separate between-group model for each distinct group size. These between groups models have different scaling parameters c_d, and require equality constraints across all other parameters and inclusion of a mean structure (Muthén, 1994, p. 385). Thus, using conventional SEM software for the unbalanced case requires a complicated modeling scheme that creates a different 'group' for each set of groups with the same group size. This results in large and complex models, with possibly groups with a sample size less than the number of elements in the corresponding covariance matrix. This makes full Maximum Likelihood estimation problematic, and therefore Muthén (1989a, 1990) proposed to ignore the unbalance, and to compute a single S^*_B. The model for S^*_B includes an ad hoc estimator c^* for the scaling parameter, which is close to the average sample size:

$$c^* = \frac{N^2 - \sum_{g}^{G} n_g^2}{N(G-1)} \tag{9}$$

The result is a Limited Information Maximum Likelihood (LIML) solution, which McDonald (1994) calls a pseudobalanced solution, and Muthén (1989a, 1994) the MUML (for MUthén's ML) solution. Muthén (1989a, 1990) shows that S^*_B is a consistent and unbiased estimator of the composite $\Sigma_W + c\Sigma_B$. This means that with large samples (of both individuals *and* groups!) S^*_B becomes a close estimate of Σ_B, and the pseudobalanced solution should produce a good approximation given adequate sample sizes.

Since S^*_B is not a maximum likelihood estimator, the analysis produces only approximate parameter estimates and standard errors. However, when the group sizes are not extremely different, the pseudobalanced estimates are close enough to the full maximum likelihood estimates to be useful in their own right. Comparisons of pseudobalanced estimates with full maximum likelihood estimates or with known population values have been made by Muthén (1990, 1994), Hox (1993), and McDonald (1994). Their main conclusion is that the pseudobalanced parameter estimates and the standard errors are fairly accurate and useful for a variety of multilevel problems. A large simulation study by Hox and Maas (2001) assesses the robustness of the pseudobalanced method against unequal groups and small sample sizes at both the individual and the group level, in the presence of a low or a high intraclass correlation (ICC). In this study, the within groups part of the model poses no problems in any of the simulated conditions. The most important problem in the between groups part of the model is the occurrence of inadmissible estimates when the group level sample size is small (50) and the intracluster correlation is low. When an admissible solution is found, the factor loadings are generally accurate. However, the residual variances are underestimated, and the standard errors are generally too small. Having more or larger groups or a higher ICC does not effectively compensate this. Therefore, while the nominal alpha level is 5%, the operating alpha level is about 8% in all simulated conditions with unbalanced groups. The strongest contributing factor is an inadequate sample size at the group level. Imbalance is also a problem for the overall goodness-of-fit test. For balanced data, the chi-square test for goodness-of-fit is accurate. For unbalanced data, the model is rejected too often, which again results in an operating alpha level of about 8%. The size of the biases is comparable to the effect of moderate non-normality in ordinary modeling. Hox and Maas conclude that the approximate solution is useful, if the group level sample size is at least 100, and keeping in mind that the operating alpha level is somewhat higher than the nominal alpha level.

The multilevel part of the structural equation model outlined above is simpler than that of the multilevel regression model. It is comparable to the multilevel regression model with random variation of the intercepts. There is no provision for randomly varying slopes (factor loadings and path coefficients). Although it would be possible to include cross-level interactions, introducing interaction variables of any kind in structural equation models is neither simple nor elegant (cf. Bollen, 1989; Marcoulides & Schumacker, 2001). An interesting approach is allowing different within groups covariance matrices in different subsamples.

Since the pseudobalanced approach needs the within groups model both for the pooled within groups and the scaled between groups model, and needs to incorporate the scaling factor for the between groups model, the actual model can become quite complicated. In addition, some software has difficulties finding good starting values. Several software writers have addressed these problems. The program STREAMS (Gustafsson & Stahl, 1999) acts as a preprocessor for standard SEM software. For two-level SEM, it calculates the pooled within and scaled between matrices, and writes the complicated setup, including starting values based on previous analyses. The program *MPLUS* (Muthén & Muthén, 1998) hides all the complications of the pseudobalanced approach from the user. It also uses by default robust estimators for the standard errors and adjusts the chi-square test statistic for the heterogeneity that results from mixing groups of different sizes (cf. Muthén & Satorra, 1995).

8.3 The Multivariate Multilevel Approach: Direct Estimation of the Covariance Matrix at Each Level

Goldstein (1987, 1995) suggested using a multivariate multilevel (MVML) regression model to produce a covariance matrix at the different levels, and to input these in a second step into a standard SEM program for further analysis. Multivariate multilevel regression models are multilevel regression models that contain more than one response variable.

In multivariate multilevel models, the variables define a separate level. In most applications, the variables would be the first level, the individuals the second level, and if there are groups, these form the third level. If we have P response variables, Y_{pig} is the response on measure p of individual i in group g. We define a total of P dummy variables scored 0/1, one for each response variable. To use these P dummy variables in a model, we exclude the usual intercept from the model. Hence, on the lowest level we have

$$Y_{pig} = \pi_{1ig} d_{1ig} + \pi_{2ig} d_{2ig} + \ldots + \pi_{pig} d_{pig} \quad (10)$$

The extra level, the *dummy-variable* level, exists solely to specify a multivariate response structure using software that is essentially developed for

univariate analyses. Therefore, there is no lowest-level error term in equation (10). If there are no explanatory variables, we have at the individual level:

$$\pi_{pig} = \beta_{pg} + u_{pig}, \quad (11)$$

where the u_{pig} terms are the residual errors for each variable at the individual level.

At the group level (the third level in the multivariate model), we have

$$\beta_{pg} = \gamma_p + u_{pg}, \quad (12)$$

where the u_{pg} terms are the residual errors for each variable at the group level.

By substitution we obtain

$$\begin{aligned} Y_{pig} &= \gamma_1 d_{1ig} + \gamma_2 d_{2ig} + \ldots + \gamma_p d_{pig} \\ &+ u_{1ig} d_{1ig} + u_{2ig} d_{2ig} + \ldots + u_{pig} d_{pig} + u_{1g} d_{1ig} + u_{2g} d_{2ig} + \ldots + u_{pg} d_{pig} \end{aligned} \quad (13)$$

Equation (13) is more conveniently expressed using sum notation:

$$Y_{pig} = \sum_{p=1}^{P} \gamma_p d_{pig} + \sum_{p=1}^{P} u_{pig} d_{pig} + \sum_{p=1}^{P} u_{pg} d_{pig} \quad (14)$$

In this multivariate multilevel model, the fixed part contains P regression coefficients for the dummy variables, which are the P overall means for the P outcome variables. The random part contains two covariance matrices, Σ_{ig} and Σ_g, which contain the variances and the covariances of the regression slopes for the dummies on the individual and the group level. Since individual level and group level covariances are estimated directly, they can be modeled directly and separately by any SEM program, using standard Maximum Likelihood estimation techniques. As a result, we get separate model tests and fit indices at all levels. The straightforward approach is a distinct advantage of the multivariate multilevel approach.

There are some disadvantages to the multivariate multilevel approach. An potential problem is that the covariances produced by the multivariate multilevel approach are themselves estimated values. They are not directly calculated, as the pooled within groups and scaled between groups covariances are, but they are estimates produced by a complex statistical procedure. If the data have a multivariate normal distribution, the pooled within groups and scaled between groups covariances can be viewed as observed values, which have a known sampling distribution. This sampling distribution is used by

SEM programs for the chi-square model test and the standard errors of the parameter estimates. It is unknown how well the sampling distribution of the multivariate multilevel covariance estimates follows the sampling distribution of the observed covariances.

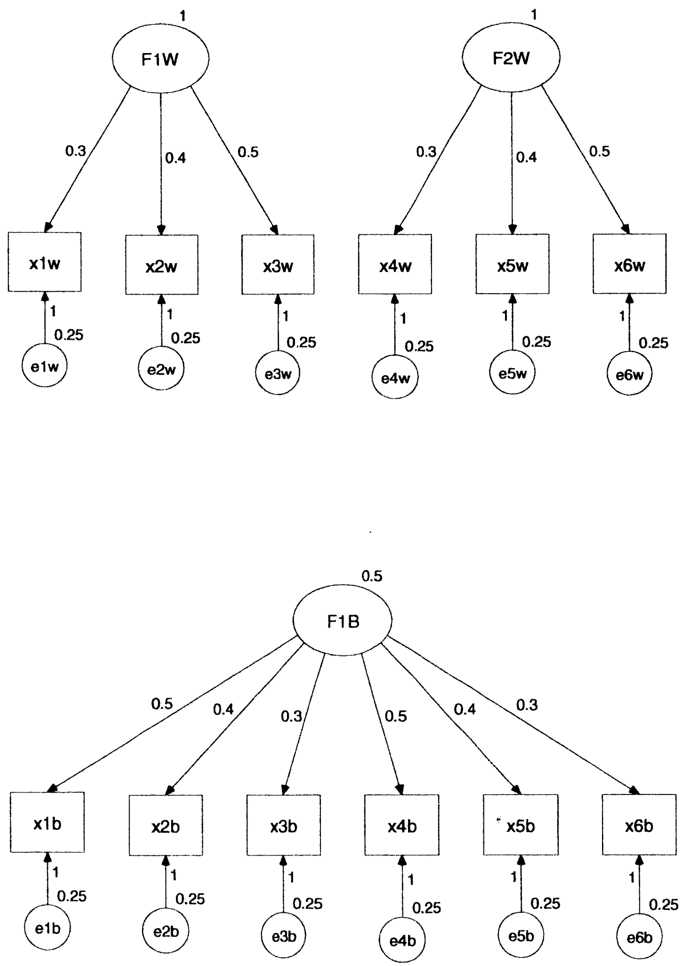

Figure 8.1. Individual level (within) and group level (between) population model.

8.4 Comparing the Two Approaches

We compare the Muthén approach and the multivariate multilevel approach to multilevel SEM. Our benchmark model is a two-level factor model with six observed variables, one factor at the group level, and two factors at the individual level.

The model and population parameter values are presented in Figure 8.1. The covariance matrices are given in Appendix 8.1. Using procedures outlined by Bollen and Stine (1993), we have constructed a two-level data set that *exactly* reproduces the benchmark model. The multilevel structure has 100 groups all of size 50. The group size and the number of groups have been chosen to be both large enough to ensure accurate estimation of both parameters and standard errors at all levels (cf. Hox & Maas, 2001).

Since the data are balanced, Muthén's method in this case is a full information Maximum Likelihood method. The multivariate multilevel approach produces Maximum Likelihood estimates, and since the input are two covariance matrices estimated by Maximum Likelihood methods the results should be comparable to the estimates produced by the Muthén method.

8.5 Results

Both methods lead to a model with an almost perfect fit. The Muthén approach leads to a chi-square of 0.059 (df=18, p=1.000); the multivariate multilevel approach leads to a chi-square of 0.000 for the within model (df=0, p=1.000) and a chi-square of 0.063 for the between model (df=9, p=1.000). The difference between the two approaches is tiny and totally immaterial. Table 8.1 shows the population values of the factor loadings and variances in the two-level factor model. In addition, it shows the estimates produced by Muthén's method and by the multivariate multilevel (MVML) method.

It is clear from Table 8.1 that the individual level loadings are estimated with total accuracy. Since the individual-level sample size is N-G = 4900 this is not surprising. At the group level, where the sample size is 100, the loading estimates are very close to the population values. The estimates produced by the MVML method are somewhat closer to the known population values, but both sets of estimates are so close to the true values that this difference is utterly trivial.

The doubts expressed in section 8.3 about the MVML approach concern mainly the effect on the standard errors. Table 8.2 shows asymptotic values of the standard errors for the factor loadings and variances, and the standard errors produced by Muthén's method and by the multivariate multilevel (MVML) method.

Table 8.1. Population and estimated values of the model parameters.

Loadings Variables	Population WF1	WF2	BF1	Muthén method WF1	WF2	BF1	MVML method WF1	WF2	BF1
X1	.3		.5	.300		.497	.300		.499
X2	.4		.4	.400		.395	.400		.397
X3	.5		.3	.500		.293	.500		.295
X4		.3	.5		.300	.497		300	.499
X5		.4	.4		.400	.395		400	.397
X6		.5	.3		.500	.293		500	.295

Res. var. method Variables	Population With./Betw.	Muthén method With.	Betw.	MVML With.	Betw.
X1	.25	.250	.241	.250	.244
X2	.25	.250	.241	.250	.243
X3	.25	.250	.239	.250	.242
X4	.25	.250	.241	.250	.244
X5	.25	.250	.241	.250	.243
X6	.25	.250	.239	.250	.242

It is clear from Table 8.2 that the standard errors of the individual level loadings are estimated with total accuracy. Since the individual-level sample size is N-G = 4900 this is not surprising. At the group level, where the sample size is 100, the standard errors produced by the Muthén and MVML method are very similar. To check the standard errors, we carried out a parametric bootstrap (10,000 bootstrap samples) on the covariance matrices produced by the MVML method. The bootstrapped standard errors are very close (all within |0.002|) to the asymptotic standard errors. Interestingly enough, the bias-correction based on the bootstrap (Stine, 1989) almost totally removes the tiny bias in the MVML-based estimates.

Table 8.2. Asymptotic and estimated values of the standard errors.

Loadings Variables	Asymptotic WF1	WF2	BF1	Muthén method WF1	WF2	BF1	MVML method WF1	WF2	BF1
X1	.010		.102	.010		.102	.010		.101
X2	.012		.096	.012		.096	.012		.095
X3	.014		.092	.014		.092	.014		.090
X4		.010	.102		.010	.102		.010	.101
X5		.012	.096		.012	.096		.012	.095
X6		.014	.092		.014	.092		.014	.090

Res. var. method Variables	Asymptotic With.	Betw.	Muthén method With.	Betw	MVML With.	Betw
X1	.006	.047	.006	.047	.006	.047
X2	.008	.042	.008	.042	.008	.041
X3	.012	.039	.012	.038	.012	.038
X4	.006	.047	.006	.047	.006	.047
X5	.008	.042	.008	.042	.008	.041
X6	.012	.039	.012	.038	.012	.038

8.6 Discussion

When the individual level and group level sample sizes are adequate, both approaches appear to produce accurate parameter estimates. The standard errors are also very close to their asymptotic values. The conclusion is that the multivariate multilevel approach to multilevel SEM is a viable method. It's simplicity compared to the Muthén pseudobalanced approach is an advantage. It does require that analysts have access to a multilevel regression program that can produce the required individual level (within groups) and group level (between groups) covariance matrices. The procedure described in this paper requires a three-level program that can constrain the lowest level variance to zero. Currently, the programs aML, HLM, MLwiN, Lisrel/Prelis and SAS Proc Mixed all can do this. The multivariate multilevel approach to multilevel SEM also generalizes straightforwardly to more than two levels, but this requires a multilevel regression program that can handle more than three levels. Currently, only aML and MLwiN have this capability. MLwiN can handle up to 50 levels, but the requirement that the highest level has an adequate sample size (Hox and Maas, 2001, suggest a lower limit of 100 for the FIML approach) puts strong practical limits to such endeavors.

There are other advantages to the MVML approach. First, it does not assume that the group sizes are all equal, an assumption made in Muthén's approach. If the group sizes are unequal, Muthén's LIML approach produces unbiased and consistent parameter estimates, but the standard errors are underestimated. As Hox and Maas (2001) show, this bias is relatively small, which justifies using the LIML approach with unequal group sizes. A second advantage of the multilevel multivariate model is that it does not assume having a complete set of variables for each individual; incomplete data are accommodated without special effort. Finally, if we have dichotomous or ordinal categorical variables, we can use the multilevel generalized linear model to produce the covariance matrices, again without special effort. The problem here is, of course, that in this case we know for certain that the covariance matrices so produced do not follow a Wishart distribution. In the incomplete data case, different parts of the covariance matrix will be based on different subsets of cases, a fact that is not reflected in the covariance matrices themselves. If we analyze dichotomous or ordered categorical variables using the probit link-function instead of the more usual logit link-function (cf. Hox, 2002), the covariances produced are polychoric correlations (cf. Olsson, 1979). Polychoric correlations have larger sampling variances than Pearson correlations (Muthén, 1989b) and require special estimation methods. The appropriate estimation method here is weighted least squares, but this requires very large sample sizes (1000 or more) for accurate estimation (Muthén & Kaplan, 1985). In multilevel SEM, this means that the highest-level sample size must be very large, a requirement that is in practice difficult to meet.

References

Arbuckle, J.L., & Wothke, W. (1999). *Amos 4.0 user's guide*. Chicago: SmallWaters.

Bentler, P.M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*, 238-246.

Bentler, P.M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

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

Bollen, K.A. & Stine, R.A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. Pp. 111-135 in K.A. Bollen & J.S. Long (Eds.). Testing structural equation models. Newbury Park, CA: Sage.

Goldstein, H. (1987). Multilevel Models in Educational and Social Research. London: Griffin.

Goldstein, H. (1995). Multilevel Statistical Models. London: Edward Arnold/New York: Halsted.

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

Gustafsson, J.-E., & Stahl, P.E. (1999). Streams user's guide, Vs. 2.0. Mölndal, Sweden: MultivariateWare.

Heck, R.H., & Thomas, S.L. (2000). An introduction to multilevel modeling techniques. Mahwah, NJ: Lawrence Erlbaum Associates.

Hox, J.J. (1993). Factor analysis of multilevel data: Gauging the Muthén model. In J.H.L. Oud & R.A.W.van Blokland-Vogelesang (Eds.), *Advances in longitudinal and multivariate analysis in the behavioral sciences*. Pp.141-156. Nijmegen: ITS.

Hox, J.J. (2002). *Multilevel Analysis. Techniques and Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hox, J.J., & Maas, C.J.M. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling*, *8*, 157-174.

Jöreskog, K.G., & Sörbom, D. (1996). *Lisrel 8: User's reference guide. 2nd Edition.* Chicago: Scientific Software International Inc.

Kaplan, D., & Elliot, P.R. (1997). A didactic example of multilevel structural equation modeling applicable to the study of organizations. *Structural Equation Modeling, 4*, 1-24.

Longford, N.T., & Muthén, B.O. (1992). Factor analysis for clustered observations. *Psychometrika*, *57*, 581-597.

Marcoulides, G.A., & Schumacker, R.E. (Eds.), (2001). *New developments and techniques in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.

McArdle, J.J., & Hamagami, F. (1996). Multilevel models from a multiple group structural equation perspective. In G.A. Marcoulides & R.E. Schumacker (Eds.), *Advanced Structural Equation Modeling*. Pp. 89-125. Mahwah, NJ: Lawrence Erlbaum Associates.

McDonald, R.P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological Methods & Research*, *22*, 399-413.

McDonald, R., & Goldstein, H. (1989). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical and Statistical Psychology*, *42*, 215-232.

Muthén, B.O. (1989a). Latent variable modeling in heterogeneous populations. *Psychometrika*, *54*, 557-585.

Muthén, B.O. (1989b). Dichotomous factor analysis of of symptom data. *Sociological Methods & Research*, *18*, 19-65.

Muthén, B. (1990). Means and covariance structure analysis of hierarchical data. Los Angeles: UCLA Statistics series, #62.

Muthén, B.O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, *22*, 376-398.

Muthén, B.O. & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, *38*, 171-189.

Muthén, L.K., & Muthén, B.O. (1998). *Mplus user's guide*. Los Angeles: Muthén & Muthén.

Muthén, B.O., & Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. In Bock, R.D. (Ed.), *Multilevel analysis of educational data*. San Diego: Academic Press.

Muthén, B.O., & Satorra, A. (1995). Complex sample data in structural equation modeling. In Marsden, P.V. (Ed.), *Sociological Methodology, 1995*. Oxford: Blackwell.

Olsson, U. (1979). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavior Research*, *14*, 485-500.

Raudenbush, S.W & Bryk, A.S. (2002). *Hierarchical linear models*. Newbury Park, CA: Sage.

Rowe, K.J. (2002). Simultaneous estimation of interdependent effects among multilevel composite variables in psychosocial research: An annotated example of the application of multilevel structural equation modeling. In N. Duan & S. Reise (Eds.), *Multilevel Modeling: Methodological advances, issues and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rowe, K.J., & Hill, P.W. (1998). Modeling educational effectiveness in classrooms: the use of multi-level structural equations to model students' progress. *Educational Research and Evaluation*, *4*, 307-347.

Searle, S.R., Casella, G., & McCulloch, C.E. (1992). *Variance components*. New York: Wiley.

Snijders, T.A.B., & Bosker, R. (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.

Stine, R. (1989). An introduction to bootstrap methods. *Sociological Methods & Research*, *18*, 2-3, 243-291.

Appendix.

Population covariances on within and between level

Covariances, Within Model

Var.	x1w	x2w	x3w	x4w	x5w	x6w
x1w	0.34					
x2w	0.12	0.41				
x3w	0.15	0.20	0.50			
x4w	0.00	0.00	0.00	0.34		
x5w	0.00	0.00	0.00	0.12	0.41	
x6w	0.00	0.00	0.00	0.15	0.20	0.50

Covariances, Between Model

Var.	x1b	x2b	x3b	x4b	x5b	x6b
x1b	0.375					
x2b	0.100	0.330				
x3b	0.075	0.060	0.295			
x4b	0.125	0.100	0.075	0.375		
x5b	0.100	0.080	0.060	0.100	0.330	
x6b	0.075	0.060	0.045	0.075	0.060	0.295