Diskussionspapiere
Discussion Papers

Discussion Paper No. 218

A Generalized Estimating/Pseudo-Score Equations Approach for the Estimation of Structural Equation Models

by
Martin Spiess

Berlin, May 2000

Deutsches Institut für Wirtschaftsforschung, Berlin
Königin-Luise-Str. 5, 14195 Berlin
Phone: +49-30-89789- 0
Fax: +49-30-89789- 200
Internet: http://www.diw.de
ISSN 1433-0210

A Generalized Estimating/Pseudo-Score Equations Approach for the Estimation of Structural Equation Models

Martin Spiess[a]

Abstract

The results of two simulation studies suggest a mixed 'generalized estimating/pseudo-score equations' approach to lead to more efficient estimators than a GEE approach proposed by Qu, Williams, Beck and Medendorp (1992) or a three-stage approach as proposed e.g. by Scheppers, Arminger and Küsters (1991) in panel probit models with binary responses. Furthermore, the mixed approach led to very efficient estimators of regression and correlation structure parameter estimators in an assumed underlying model relative to the ML estimator for an equicorrelation structure. Using the mixed approach, the regression parameters are estimated using generalized estimating equations and the correlation structure parameters are simultaneously estimated using pseudo-score equations. Both sets of parameters are calculated as if they were orthogonal, thereby preserving the robustness of the regression parameter estimators with respect to misspecification of the correlation matrix. Based on the above simulation results, the mixed approach is extended for the estimation of more general structural equation models with ordered categorical or mixed continuous/ordered categorical responses.

Key words: Multivariate probit model; generalized estimating equations; pseudo-score equations; correlated categorical and continuous responses; structural equation models

[a]German Institute for Economic Research (GSOEP), Königin–Luise–Str. 5, D–14195 Berlin, Germany

1 Introduction

Maximum likelihood (ML) estimation of multivariate probit models with binary and/or ordinal responses is hampered by the computational intractability of multidimensional integrals. Therefore, in the last 15 years a lot of work has been devoted to deriving non-ML approaches for the estimation of these models. Today there are several non-ML approaches available which allow or could easily be extended to estimate multivariate probit models with binary and/or ordered categorical responses. However, a desirable property of estimators is a mean squared error which is as small as possible, not only asymptotically but also in finite samples. Since only a few simulation studies comparing the different approaches are available, not much is known about the difference of the various non-ML estimators with respect to their bias and relative efficiency in finite samples. The estimation of mean and covariance structure probit models, although more complex, is based on the same estimation methods as for the estimation of multivariate probit models. Therefore, it is desirable to develop approaches that yield estimators which are — in the above sense — optimal for the estimation of the simpler models.

Estimation of structural equation models or mean and covariance structure models with continuous and/or categorical responses generally is done in several steps. For example, Muthén (1984) proposed a three-stage procedure for the estimation of structural equation models (see also Muthén and Satorra, 1995). A three-stage approach is also proposed by Küsters (1987) (see also Schepers, Arminger and Küsters, 1991) for the estimation of slightly more general mean and covariance structure models. A three-stage estimation procedure for the estimation of structural equation models is also used by Lee, Poon and Bentler (1990). In another paper by Lee, Poon and Bentler (1992) they propose a two-stage approach for the estimation of their models (see also Lee, Poon and Bentler, 1995). However, their models are not as general as the models considered by Muthén (1984) or Küsters (1987) (cf. Lee, Poon and Bentler, 1992, p. 91).

Instead of starting with the development of estimation approaches of very general and complex models, however, a first step in the development of an optimal estimation approach could be to compare different available approaches in simpler multivariate models via simulated data sets, where all assumptions are controlled for and model specifications can systematically be variied. Then, learning from the results of those studies new estimation approaches may be developed using the most promising estimation methods.

In this paper, results from simulation studies are presented where different estimation approaches are compared with each other in the case of binary probit models with correlated responses. Furthermore, a mixed generalized estimating/pseudo-score equations approach is proposed and compared to an existing approach proposed by Qu, Williams, Beck and Medendorp (1992) and Qu, Piedmonte and Williams (1994) in a second simulation study for the simple binary panel model. Since the results of the simulation studies suggest that the mixed approach is a very promising candidate for the estimation of more general structural equation models with ordered categorical or mixed continuous/ordered categorical responses, an extension of the mixed approach to the estimation of the more general structural equation models is outlined.

2 Model and Notation

The general model is based on the assumption of a hierarchical latent model with $H+1$ levels (e.g. Küsters, 1987; Schepers, Arminger and Küsters, 1991). For simplicity, however, only two hierarchical levels will be considered. Omitting the index i, where $i = 1, \ldots, n$ and n is the number of clusters (e.g. subjects), the latent model is given by

$$\mathrm{B}_1\eta_0 = \mu_1 + \Lambda_1\eta_1 + \Gamma_1 x_1 + \epsilon_1 \qquad \text{(first level)}$$

and

$$\mathrm{B}_2\eta_1 = \mu_2 + \Lambda_2\eta_2 + \Gamma_2 x_2 + \epsilon_2 \qquad \text{(second level)},$$

or, given the necessary inverse matrices exist, more compactly by

$$\eta_0 = \mathrm{B}_1^{-1}(\mu_1 + \Lambda_1\mathrm{B}_2^{-1}(\mu_2 + \Lambda_2\eta_2 + \Gamma_2 x_2 + \epsilon_2) + \Gamma_1 x_1 + \epsilon_1) \tag{1}$$

and the assumption that the random vector $\eta_3 = (\epsilon_1', \epsilon_2', \eta_2')'$ is — independently of the exogenous vector variables x_1 and x_2 — normally distributed with expected value zero and block diagonal covariance matrix $\Omega(\vartheta)$, which is a function of a structural parameter vector ϑ. The matrices on the diagonal of $\Omega(\vartheta)$ are the covariance matrices of the random vector variables ϵ_1, ϵ_2 and η_2 respectively. The non-singular matrices B_1 and B_2 represent, for example, relations between the components of the latent vector variables η_0 and η_1. The matrices Γ_1 and Γ_2 consist of regression parameters, Λ_1 and Λ_2 may be considered, for example, as matrices of factor loadings and the vectors μ_1 and μ_2 are the general means of the latent variables η_0 and η_1. In general, the elements of these matrices are either fixed variables or elements of the structural parameter ϑ.

Defining the matrices $\mathrm{F}_1 = \mathrm{B}_1^{-1}$, $\mathrm{G}_1 = \begin{pmatrix} \mathrm{I} & \Lambda_1 \end{pmatrix}$,

$$\mathrm{F}_2 = \begin{pmatrix} \mathrm{I} & 0 \\ 0 & \mathrm{B}_2^{-1} \end{pmatrix}, \qquad \mathrm{G}_2 = \begin{pmatrix} \mathrm{I} & 0 & 0 \\ 0 & \mathrm{I} & \Lambda_2 \end{pmatrix}, \qquad \mathrm{K}_2 = \begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix},$$

where I is the identity matrix and 0 denotes matrices with all elements equal to zero, and

$$\gamma(\vartheta) = \mathrm{F}_1\mathrm{G}_1\mathrm{F}_2 \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \qquad \Pi(\vartheta) = \mathrm{F}_1\mathrm{G}_1\mathrm{F}_2\mathrm{K}_2,$$

$$v = \mathrm{F}_1\mathrm{G}_1\mathrm{F}_2\mathrm{G}_2 \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \eta_2 \end{pmatrix}, \qquad y^\star = \eta_0,$$

the reduced form of the general model (1) may written as

$$y^\star = \gamma(\vartheta) + \Pi(\vartheta)x + v, \tag{2}$$

where v is normally distributed with expected value zero and covariance matrix $\Sigma(\vartheta) = (\mathrm{F}_1\mathrm{G}_1\mathrm{F}_2\mathrm{G}_2)\,\Omega(\vartheta)\,(\mathrm{F}_1\mathrm{G}_1\mathrm{F}_2\mathrm{G}_2)'$, $y^\star = (y_1^\star, \ldots, y_t^\star)'$ is a $(t \times 1)$-latent response variable $(j = 1, \ldots, t)$ and $x = (x_1', x_2')'$.

Depending on the measurement level, each observable response y_j is connected with the latent, not observable response $y_j^\star$ by a corresponding measurement relation. For example, if y_j is metric, then $y_j = y_j^\star$. If y_j is ordered categorical with $q = 1, \ldots, k+1$ categories and k unknown thresholds $\tau_{j,1}, \ldots, \tau_{j,k}$, then $y_j = q \Leftrightarrow \tau_{j,q-1} < y_j^\star \leq \tau_{j,q}$ where $\tau_{j,0} = -\infty$ and $\tau_{j,k+1} = +\infty$.

Many special cases of the general model with $H+1$ levels are given in Küsters (1987).

The binary panel model as a special case The binary panel probit model can be considered as a special case of the general model with only one level. To see this, let B_1 be the identity matrix and Λ_1 be a matrix whose elements are all equal to zero. Then (1) can be written as

$$\eta_0 = \mu_1 + \Gamma_1 x_1 + \epsilon_1$$

and the reduced form is given by

$$y^\star = \mu + \Gamma x + \epsilon,$$

where $\mu \equiv \mu_1$, $\Gamma \equiv \Gamma_1$, $x \equiv x_1$, $\epsilon \equiv \epsilon_1$ and $\epsilon \sim N(0, \Sigma(\vartheta))$. Note that $\Sigma(\vartheta) = \Omega(\vartheta)$.

However, if all regression parameters and all general means are assumed to be equal over the observation points within each cluster, then the different parameters can be collected in a $(p+1)\times 1$ parameter vector β. If, for each cluster, all exogenous variables and a vector of ones are arranged according to the scalar parameters in β in a $(t \times (p+1))$-matrix X, the reduced form can be written in a more familiar way as

$$y^\star = X\beta + \epsilon.$$

Since the observed responses are binary, the measurement relation

$$y_j = \begin{cases} 1 & \text{if } y_j^\star > \tau, \\ 0 & \text{otherwise} \end{cases}$$

is assumed, where τ is an unknown threshold parameter. However, since not all parameters are identifiable in this model, τ is set to zero and the variances are restricted to one.

3 Estimation approaches

3.1 Part I: Estimation of regression parameters

Using the binary panel model, in a first step, two estimation approaches are compared with each other and for the special case of an equicorrelation structure with the maximum likelihood (ML) approach. The approaches are compared using simulated data sets with respect to convergence behavior of the estimates as well as bias and efficiency of the estimators in finite samples. Furthermore, in this first step estimation of the regression parameters is the main interest.

The MECOSA approach ('mean and covariance structure analysis'; Küsters, 1987; Schepers, 1991; Schepers, Arminger and Küsters, 1991) is a three-stage approach for the estimation of general mean and covariance structure models. To estimate the binary panel model, at the first stage, regression parameter estimates are calculated separately for every observation point j using the ML principle, that is, t independent binary probit models are estimated. In the second step, all correlations of the underlying errors are estimated using the t regression parameter estimates from the first stage. This is done by maximizing all possible $t \times (t-1)/2$ different conditional marginal bivariate log-likelihood functions. Then a consistent estimate of the asymptotic covariance matrix of all parameter estimators of the first two stages is calculated. Note, that if the mean structure parameters are of dimension $p+1$, then this covariance matrix has $t \times (p+1) + t \times (t-1)/2$ rows and columns, which can be a fairly large matrix. For example, if just three regression parameters and one general mean have to be estimated for $t = 4$, then the covariance matrix is a (22×22) matrix. In the third step, the structural parameters are estimated using a weighted least squares approach, where the weighted distance between estimates of the first two stages and known functions of the structural parameters are minimized (for details see Küsters, 1987, or, Schepers, 1991).

The 'generalized estimating equations' or GEE approach proposed by Liang and Zeger (1986) originally was developed for the estimation of regression parameters in panel or multivariate models, where the association between the responses was treated as a nuisance. Furthermore, instead of estimating the correlations in the underlying model, the correlations in the observed responses are modeled. The regression parameters are estimated iteratively by finding a solution to the so-called 'generalized estimating equations'

$$\sum_{i=1}^{n} X_i' \mathrm{D}_i \mathrm{W}_i^{-1} \mathrm{e}_i = 0,$$

where $\mathrm{e}_i = (y_i - \Phi(X_i\beta))$, $\Phi(\cdot)$ denotes the standard normal cumulative distribution function and is a model of the expectation of the vector variable $y_i = (y_{i1}, \ldots, y_{ij})'$. The matrix D_i is a diagonal matrix with diagonal elements $\partial\Phi(x_{ij}'\beta)/\partial\eta_{ij}$, where x_{ij} is the vector of exogenous variables of cluster i at observation point j and $\eta_{ij} = x'_{ij}\beta$. The matrix W_i is a covariance matrix of the observed responses and is a function of a 'working' correlation matrix, where the association structure between the responses is modeled. At every iteration this correlation matrix is calculated using the estimated residuals $\hat{\mathrm{e}}_i = (y_i - \Phi(X_i\hat{\beta}))$ (for details see Spiess and Hamerle, 1995).

If in the underlying model an equicorrelation structure with a positive correlation is assumed, then maximum likelihood estimation of the random effects probit model leads to ML estimates of the regression and correlation parameters (e.g. Butler and Moffit, 1982; for details see Spiess and Hamerle, 1995).

In a first simulation study the following factors were variied: The type of correlation matrix (Equicorrelation structure and autoregressive correlation structure of order one (AR(1))), the value of the correlation structure parameter (0.2 and 0.8) and the sample size ($n = 50, 100, 250, 500, 1000$). According to every model specification, 500 data sets were generated. Three covariates were generated: One dichotomous, one normally and one uniformly distributed covariate. All covariates variied over the nt observations but were held constant over the 500 simulated data sets generated according to each model specification. For every subject $t = 5$ observations were simulated for every model specification.

The results of this simulation experiment suggest that the GEE as well as the ML estimation procedure is robust with respect to convergence of the estimates. On the other hand, the MECOSA estimation procedure failed to converge more often for data sets with small to medium sample sizes. Not surprisingly, the larger the sample size, the smaller the bias of all estimators. For the model specifications considered, there were no systematic and significant differences in the biases of the estimators. Concerning efficiency, there were no significant differences between ML and GEE estimators for low 'true' correlations. For moderate to high serial correlations, ML estimators were the most efficient estimators. The MECOSA estimators turned out to be the most inefficient estimators. Moreover, for sample sizes smaller than $n = 1000$ the root mean squared errors were significantly underestimated using the MECOSA approach. In general, the GEE approach seems to be superior to the MECOSA approach concerning convergence of the estimates as well as efficiency of the estimators in finite samples (see also Spiess and Hamerle, 1995).

The first part of table 1 in section 3.2 — as an example — gives the estimation results using the ML, GEE and MECOSA approach for a model with $t = 5$, $n = 500$, 500 simulated data sets, one dichotomous, one normally and one uniformly distributed covariate and corresponding 'true' values of the regression parameters $\beta_d = 1$, $\beta_n = 1$ and $\beta_{gl} = -1$. The 'true' value of the parameter weighting the constant term was $\beta_k = -.2$ and the 'true' value of the correlation was $\rho = .8$. However, to save space only the estimation results for β_{gl} and ρ are given.

3.2 Part II: Estimation of regression and correlation structure parameters

Unfortunately, the GEE approach as proposed by Liang and Zeger (1986) does not allow the estimation of the underlying correlations. Therefore, based on an extended GEE approach proposed by Prentice (1988), in their work Qu, Williams, Beck and Medendorp (1992) and Qu, Piedmonte and Williams (1994) propose the simultaneous, iterative estimation of both sets of parameters, that is, of the regression parameters and of the correlation structure parameters of the underlying model, using generalized estimating equations (GEE$_{Qu}$ approach). The estimating equations for the regression parameters are the same as in the GEE approach described above, up to the covariance matrix W_i, the elements of which are now estimated variances and covariances of the observed responses modeled using uni- and bivariate standard normal cumulative distribution functions and estimated correlation structure parameters. The estimating equations for the correlation structure parameters, denoted as ϑ_B, are given by the equations

$$\sum_i \mathrm{E}_i' \mathrm{M}_i^{-1} \mathrm{v}_i = 0.$$

The elements of the column vector v_i are the differences between the $t \times (t-1)/2$ cross products of the residuals e_{ij} and $\mathrm{e}_{ij'}$ and the corresponding off-diagonal elements of W_i. The matrix M_i is a diagonal matrix with the variances of the cross-products of the residuals on the diagonal. The matrix E_i' is the derivative of the off-diagonal elements of W_i with respect to the structural parameter ϑ_B.

In the mixed approach the regression parameters are iteratively estimated using the generalized estimating equations as in the approach proposed Qu et al. (1992) and Qu et al. (1994). However, the correlation structure parameters are estimated simultaneously using pseudo-score equations (GEPSE approach). The pseudo-score equations are given by

$$\sum_i \mathrm{E}_i' \, \mathrm{V}_i^{-1} \, \mathrm{w}_i = 0.$$

The elements of the vector w_i are all $t(t-1)/2$ possible different products $(2y_{ij} - 1)(2y_{ij'} - 1)$, V_i is a diagonal matrix with the probabilities of the variables y_{ij} and $y_{ij'}$ taking on specific values, given the covariates, regression parameters and the correlation, on the diagonal and E_i' is the same as in the approach proposed by Qu et al. (1992) and Qu et al. (1994). Note that the above pseudo-score equations are just the vector of first derivatives of the pseudo-maximum likelihood function with respect to ϑ_B

$$l(\vartheta_B) = \sum_i l_i(\vartheta_B) = \sum_i \sum_{\substack{j,j' \\ (j'<j)}} \log P_{i(j,j')},$$

set to zero, where summation is over the probabilities of all possible $t(t-1)/2$ pairs of responses taking on specific values. In this formulation, pairs of responses are assumed to be independent (Spiess, 1998; Spiess and Keller, 1999). Both sets of parameter estimates, i.e. regression and correlation structure parameter estimates, are calculated as if the estimators were orthogonal, thereby preserving the robustness of the regression parameter estimators with respect to misspecification of the correlation matrix. In contrast to the pseudo-ML approach proposed by Gourieroux, Monfort and Trognon (1984) where the regression parameters are estimated under the assumption of independence, using the mixed approach, the regression parameters are estimated taking the associations between the responses into account.

In the second simulation study data sets were generated according to the same model specifications as in the first simulation study. However, only samples of sizes $n = 50, 100, 500$ were generated. As in the first simulation study, the ML estimator of the random effects probit model was calculated as a reference estimator in the equicorrelation case.

The results of this second simulation experiment suggest that all three estimation procedures (GEE$_{Qu}$, GEPSE and ML) are equally robust with respect to convergence of the estimates. If convergence problems occured, then they mainly occured in small samples. As expected, there were no differences for the regression parameters in the two non-ML approaches. However, there were significant differences with respect to the correlation structure parameters. Whereas for low 'true' correlations the differences between the approaches were very small, for moderate and high 'true' correlations, the GEPSE estimator was the most efficient non-ML estimator for all models considered and generally led to very efficient regression and correlation structure parameter estimators relative to the ML approach (cf. Spiess, 1998).

Again, as an example, the estimation results for β_{gl} and ρ are given for the model described in section 3.1 using the GEPSE approach and the approach proposed by Qu et al. (1992) and Qu et al. (1994), denoted as GEE_{Qu} approach (see the second part of table 1).

Table 1: *Mean (m), estimated standard deviation ($\widehat{sd}$) and root mean squared error (rmse) of the estimates of β_{gl} and ρ using different estimation approaches for a model with $n = 500$, $t = 5$, $\beta_{gl} = -1$ and an equicorrelation structure ($\rho = .8$) over s simulations*

m $\widehat{\text{sd}}$ rmse	ML $s = 500$	GEE $s = 500$	MECOSA $s = 497$	GEPSE $s = 500$	GEE_{Qu} $s = 500$
	-1.0067	-1.0051	-0.9986	-1.0062	-1.0062
$\hat{\beta}_{gl}$	0.0872	0.0935	0.1035	0.0884	0.0884
	0.0856	0.0901	0.1104	0.0854	0.0855
	0.8009		0.7995	0.8005	0.8012
$\hat{\rho}$	0.0244		0.0237	0.0242	0.0344
	0.0253		0.0274	0.0266	0.0356

4 The mixed approach for the estimation of the general model

Since the mixed approach was superior to the MECOSA approach and the approach proposed by Qu et al. (1992) and Qu et al. (1994) in the simulation studies, it is extended to the estimation of parameters of general mean and covariance structure models with continuous and/or ordered categorical responses, where a binary response can be considered as a special case of an ordered categorical response.

The generalized estimating equations for the estimation of the parameters of the mean structure are given by

$$\sum_{i=1}^{n} X_i' \mathrm{D}_i \mathrm{W}_i^{-1} \mathrm{e}_i = 0,$$

where $\mathrm{e}_i = (y_i - \mathrm{E}(y_i))$ and $\mathrm{E}(y_i)$ is a model of the expectation of the vector of responses y_i. The expectations now depend on the measurement level of the responses. If y_{ij} is continuous, then the expectation simply is $\gamma_j + \lambda_j' x_i$, where γ_j and λ_j are the mean structure parameters of the reduced form at observation point j. If the response is ordered categorical with $k+1$ possible different values, then the difference of a k-dimensional vector of binary variables, y_{ij}, and a vector of corresponding expectations enter into the generalized estimating equations. Note that the exogenous variables in X_i can be arranged according to the model assumptions. For example, in the general model $X_i = \mathrm{I}_t \otimes x_i'$, where I_t is the $(t \times t)$-identity matrix and $\otimes$ denotes the Kronecker product. Accordingly, $\Pi(\vartheta)$ can be vectorized, so that $\Pi(\vartheta)x_i = X_i \mathrm{Vec}(\Pi(\vartheta))$. As in the binary case, W_i is a model for the covariance matrix of the observed responses, i.e. a matrix composed of variances and covariances of continuous and/or binary variables. The elements of the matrix W_i are estimated using the estimates of the correlation structure parameters. For the estimation of the covariance part of the reduced form, (pseudo-) score equations are used: If responses are continuous, then score equations are used, if ordered categorical responses are considered, then, as in the binary case, pseudo-score equations are used.

The parameter estimates are calculated iteratively. Since the simulation results suggest that the simultaneous estimation of the parameters leads to estimates with smaller variances than their estimation in several steps, the structural parameter estimates as well as the parameter estimates of the reduced form are estimated within the same iteration step.

Like in the binary case, the asymptotic covariance matrix can be estimated by a so-called sandwich form, and is a function of the generalized estimating equations, (pseudo-) score equations and expected and exact derivatives thereof (for the binary case see Spiess, 1998, or, Spiess and Keller, 1999).

5 Discussion

The simulation studies revealed that, for the binary panel probit model, the mixed approach leads to relative more efficient estimates than the MECOSA approach or the approach proposed by Qu et al. (1992) and Qu et al. (1994). Furthermore, the mixed approach was more robust with respect to convergence of the estimates than the MECOSA approach. Although the mixed approach seems to be a promising candidate for the estimation of mean and covariance structure models, up to now no simulations have been run to evaluate this approach in the light of these more general models. However, this is the work to be done in the near future.

If as a consequence of the results of the simulation study using more general models it can be concluded that the mixed approach still seems to be superior to e.g. the MECOSA approach, then it could be generalized to estimate models with classified metric responses and one or two-sided censored responses with threshold values known a priori as well. Furthermore, the extension to mean and covariance structure panel models would be desirable. However, for panel models, the problems of initial conditions in dynamic models would have to be taken into considerations.

References

Butler, J.S. & Moffit, R. (1982). Notes and comments: A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica, 50(3)*, 761–764.

Gourieroux, C., Monfort, A. & Trognon, A. (1984). Estimation and test in probit models with serial correlation. In J.P. Florens, M. Mouchart, J.P. Raoult & L. Simar (Eds.), *Alternative approaches to time series analysis, Proceedings of the 3rd Franco-Belgian Meeting of Statisticians, november 1982* (pp. 169–209). Bruxelles: Publications des Facultès universitaires Saint-Louis, Boulevard de Jardin botanique 43.

Küsters, U. (1987). *Hierarchische Mittelwert- und Kovarianzstrukturmodelle mit nichtmetrischen endogenen Variablen.* Heidelberg: Physica-Verlag.

Lee, S.-Y., Poon, W.-Y. & Bentler, P.M. (1990). A three-stage estimation procedure for structural equation models with polytomous variables. *Psychometrika, 55*, 25–51.

Lee, S.-Y., Poon, W.-Y. & Bentler, P.M. (1992). Structural equation models with continuous and polytomous variables. *Psychometrika, 57*, 89–105.

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

Liang, K.-Y. & Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73(1)*, 13–22.

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

Muthén, B.O. & Satorra, A. (1995). Technical aspects of Muthén's LISCOMP approach to estimation of latent variable relations with a comprehensive measurement model. *Psychometrika, 60*, 489–503.

Prentice, R.L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics, 44*, 1033–1048.

Qu, Y., Williams, G.W., Beck, G.J. & Medendorp, S.V. (1992). Latent variable models for clustered dichotomous data with multiple subclusters. *Biometrics, 48*, 1095–1102.

Qu, Y., Piedmonte, M.R. & Williams, G.W. (1994). Small sample validity of latent variable models for correlated binary data. *Communications in Statistics — Simulation and Computation, 23(1)*, 243–269.

Schepers, A. (1991). *Numerische Verfahren und Implementation der Schätzung von Mittelwert- und Kovarianzstrukturmodellen mit nichtmetrischen Variablen.* Ahaus: Verlag Frank Hartmann.

Schepers, A., Arminger, G. & Küsters, U. (1991). The analysis of non-metric endogenous variables in latent variable models: The MECOSA approach. In

J. Gruber (Ed.), *Econometric Decision Models: New Methods of Modelling and Applications, Lecture Notes in Economics and Mathematical Systems: 366* (pp. 459–472). Berlin: Springer.

Spiess, M. (1998). A mixed approach for the estimation of probit models with correlated responses: Some finite sample results. *Journal of Statistical Computation and Simulation, 61*, 39–59.

Spiess, M. & Hamerle, A. (2000). Regression models with correlated binary response variables: A comparison of different methods in finite samples. To appear in: *Computational Statistics & Data Analysis, 33(4),* 439–435.

Spiess, M. & Keller, F. (1999). A mixed approach and a distribution free multiple imputation technique for the estimation of a multivariate probit model with missing values. *British Journal of Mathematical and Statistical Psychology, 52*, 1–17.