

ON STRUCTURAL EQUATION MODELING WITH DATA THAT ARE NOT MISSING COMPLETELY AT RANDOM

BENGT MUTHÉN

DAVID KAPLAN

GRADUATE SCHOOL OF EDUCATION

MICHAEL HOLLIS

GRADUATE SCHOOL OF ARCHITECTURE AND URBAN PLANNING
UNIVERSITY OF CALIFORNIA, LOS ANGELES

A general latent variable model is given which includes the specification of a missing data mechanism. This framework allows for an elucidating discussion of existing general multivariate theory bearing on maximum likelihood estimation with missing data. Here, missing completely at random is not a prerequisite for unbiased estimation in large samples, as when using the traditional listwise or pairwise present data approaches. The theory is connected with old and new results in the area of selection and factorial invariance. It is pointed out that in many applications, maximum likelihood estimation with missing data may be carried out by existing structural equation modeling software, such as LISREL and LISCOMP. Several sets of artificial data are generated within the general model framework. The proposed estimator is compared to the two traditional ones and found superior.

Key words: maximum likelihood, ignorability, selectivity, factor analysis, factorial invariance, LISREL.

1. Introduction

Confirmatory factor analysis and structural equation modeling (see e.g., Jöreskog, 1969, 1977) need often be applied in situations where data are missing on certain variables and it cannot be realistically assumed that the data are missing completely at random. Ordinary methods would in these cases give estimates that are both inefficient and have large sample bias. Existing missing data theory that provide better alternatives (see e.g., Anderson, 1957; Beale & Little, 1975; Little & Rubin, 1987; Rubin, 1974, 1976) does not seem to have been adapted in factor analysis and structural equation modeling practice. Reasons for this may include lack of familiarity with missing data theory and the fact that general maximum likelihood estimation requires special computational routines as in Finkbeiner (1979); see also Dempster, Laird, & Rubin (1977).

This paper has several aims. A formulation is given for a possible extension of a latent variable structural equation model to include missingness. Estimation will not be attempted for all the parameters of the extended model. Within this framework, however, one aim is to explicate some parts of missing data theory regarding maximum likelihood

The research of the first author was supported by grant No. SES-8312583 from the National Science Foundation and by a Spencer Foundation grant. We wish to thank Chuen-Rong Chan for drawing the path diagram.

Requests for reprints should be sent to Bengt Muthén, Graduate School of Education, University of California, Los Angeles, CA 90024.

inference. The common assumption of missing completely at random is often quite unrealistic and correct maximum likelihood inference can be obtained under the much less restrictive assumption of the missing data mechanism being "ignorable" (Rubin, 1976). We will both consider missingness being predicted by observed variables and by latent factors and make connections with classical results on inference in selective samples and factorial invariance (see e.g., Meredith, 1964; Muthén & Jöreskog, 1983). The aim of the paper is also to show that for many missing data situations the proposed likelihood approach may be carried out with existing structural equation modeling software, such as the generally available LISREL (Jöreskog & Sörbom, 1984) and LISCOMP (Muthén, 1987) programs. The aim of our study is furthermore to use the above missing data structural equation model framework to illustrate situations where missingness is strongly selective rather than completely at random, as in the factor analysis studies of Finkbeiner (1979) and Brown (1983). Hence, the loss of efficiency in the classical missing data methods of the listwise deletion and pairwise present approaches will be overshadowed by concerns about size of parameter bias. These two estimators will be compared to the alternative likelihood approach in a set of artificial data situations.

2. A Simple Model for Missingness

Consider the confirmatory factor analysis model for a vector of p continuous, latent response variables \mathbf{y}^* (c.f. e.g., Muthén, 1984) measuring a vector of m latent variables $\boldsymbol{\eta}$,

$$\mathbf{y}^* = \mathbf{v} + \Lambda \boldsymbol{\eta} + \boldsymbol{\varepsilon}. \quad (1)$$

Let $E(\boldsymbol{\eta}) = \boldsymbol{\kappa}$. We may expand this to a general structural equation model by letting

$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta}. \quad (2)$$

We may also generalize the model to include several groups.

Although the latent \mathbf{y}^* 's are potentially available for each sample unit, we consider a selection variable \mathbf{s}^* for each \mathbf{y}^* variable that decides if the latent value is actually selected for observation as \mathbf{y} , or is missing. A simple extension of the model is to specify linear relations for the vector of latent selection variables \mathbf{s}^* ,

$$\mathbf{s}^* = \Gamma_{\boldsymbol{\eta}} \boldsymbol{\eta} + \Gamma_{\mathbf{y}} \mathbf{y}^* + \boldsymbol{\delta}. \quad (3)$$

(Parenthetically, we note that we may simplify (3) by deleting the $\Gamma_{\mathbf{y}} \mathbf{y}^*$ term, since observed predictors can be handled by the $\Gamma_{\boldsymbol{\eta}} \boldsymbol{\eta}$ term, defining some of the $\boldsymbol{\eta}$'s to be identical to the \mathbf{y}^* 's; the formulation of (3) is chosen for clarity.) Let there be a threshold parameter τ_j for variable s_j^* . For sample unit i , let $s_{ij} = 1$ for $s_{ij}^* > \tau_j$. This denotes that the i -th unit's y_{ij}^* value is selected to be observed, that is, y_{ij} is not missing. If $s_{ij} = 0$ ($s_{ij}^* \leq \tau_j$), y_{ij} is missing. While the γ parameters influence the strength of selectivity of the missingness, the τ parameters influence the amount of missingness. The variables of \mathbf{y}^* and \mathbf{s}^* are taken to be multivariate normal.

We assume random sampling of the $(\mathbf{y}^* \ \mathbf{s}^*)$ vector, but will only consider the likelihood of the (not missing) observations on \mathbf{y}^* , which does not in general correspond to a random sample of observations on \mathbf{y}^* .

In the population, using ordinary assumptions, we have in the confirmatory factor analysis (CFA) formulation

$$\boldsymbol{\mu} = E(\mathbf{y}) = \mathbf{v} + \Lambda \boldsymbol{\kappa}, \quad \boldsymbol{\Sigma} = V(\mathbf{y}) = \Lambda \boldsymbol{\Psi} \Lambda' + \boldsymbol{\Theta}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}. \quad (4)$$

The parameters of the original model are (with the more general structural equation

model parameters in parentheses)

$$v, \Lambda, \Theta_{\varepsilon\varepsilon}, \Psi \text{ (or } B, \Psi), \kappa \text{ (or } \alpha, \kappa). \tag{5}$$

Let these parameters be stacked in the vector θ . The parameters related to the missingness mechanism are

$$\Gamma, \Theta_{\delta\delta}, \Theta_{\delta\varepsilon} \text{ (and } \Theta_{\delta\zeta}), \tau, \tag{6}$$

allowing δ to be correlated with ε (and ζ) residuals. Let these parameters be stacked in the parameter vector ϕ .

It should be noted that many other specifications of the missing data mechanism are possible and can lead to the same conclusions. This specification is chosen for simplicity.

3. Likelihood and Estimators

In this paper we will consider maximum likelihood estimation of the θ parameters, while estimation of the ϕ parameters will not be attempted. The likelihood of the sample of the y observations is in general a function of θ and ϕ and may be written as (Rubin, 1976; Little, 1982, 1983)

$$\log L(\theta, \phi; y) = \sum_{g=1}^G \log \phi^g(\theta; y) + \log f(\theta, \phi; y). \tag{7}$$

Here,

$$\begin{aligned} \log \phi^g(\theta; y) = & \text{constant} - \frac{1}{2} N^g \log |\Sigma^g| \\ & - \frac{1}{2} N^g \text{tr } \Sigma^{g-1} [S^g + (\bar{y}^g - \mu^g)(\bar{y}^g - \mu^g)'], \end{aligned} \tag{8}$$

giving the log of the regular multivariate normal likelihood for the group of N^g sample units that exhibit the g -th missing data pattern by deleting appropriate rows and columns in the mean vectors and covariance matrices for y . The second term of (7) is due to the missingness mechanism and will be explicated below in sections 3.1 and 3.2. The likelihood of (7) will be termed “the true likelihood,” while for the purposes of this paper the first term of (7) will be termed “the quasi-likelihood,” since it ignores the missingness mechanism. It may be noted that the term “quasi-likelihood” is not used in the same sense as in Wedderburn’s (1974) work.

Given the model of section 2, we will first explicate the missing data theory conditions of “ignorability of the missing data mechanism.” Under this condition the true likelihood is such that when maximizing the quasi-likelihood with respect to θ , correct maximum likelihood estimates of θ are obtained. The estimator that maximizes the quasi-likelihood will be termed the FQL (full, quasi-likelihood) estimator and will be of primary interest to us. It was used, for example, by Trawinski and Bargman (1964) and Hartley and Hocking (1971) for the estimation of μ and Σ , and by Finkbeiner (1979) for factor analysis estimation of θ .

The FQL estimator will be compared to the standard listwise present (deletion) approach, which does not use the full quasi-likelihood, but only the part corresponding to the group with no missingness. This will be called the LQL estimator, for listwise (present) quasi-likelihood. The LQL estimator will in general give biased estimates even in infinitely large samples, unless the s^* ’s are uncorrelated with the y^* ’s (this condition will be referred to below as MCAR, or missing completely at random). The bias is clearly realized when applying the classic selection formulas of Pearson (1912) and Lawley (1943–1944) to our model; see also Johnson & Kotz (1972, p. 70). Here, selection takes place on s^* ,

affecting the distribution of y . Due to normality, the required linear and homoskedastic regression of y on s^* holds, and

$$\begin{aligned} \mu_y^s &= \mu_y + \Sigma_{ys^*} \Sigma_{s^*s^*}^{-1} (\mu_{s^*}^s - \mu_{s^*}), \\ \Sigma_{yy}^s &= \Sigma_{yy} + \Sigma_{ys^*} \Sigma_{s^*s^*}^{-1} (\Sigma_{s^*s^*}^s - \Sigma_{s^*s^*}) \Sigma_{s^*s^*}^{-1} \Sigma_{ys^*}' \end{aligned} \tag{9}$$

Here, $\mu_{s^*}^s$ and $\Sigma_{s^*s^*}^s$ are the mean vector and covariance matrix of a truncated normal s^* distribution. In the case of the LQL estimator, the superscript s refers to the subpopulation for which all s^* 's exceed their τ 's. LQL estimation effectively assumes random sampling of y in this subpopulation, which in large samples gives a mean vector and covariance matrix that tend to μ_y^s , Σ_{yy}^s , respectively.

FQL and LQL will also be compared to the pairwise present approach (PPA) which does not in general maximize any proper likelihood. Indeed, the covariance matrix is sometimes not positive definite.

Below, we consider certain components of the true likelihood for the special case of the s^* missingness predictors being a subset of the latent response variables of y^* and the special case of the missingness predictors being a subset of the latent variables of η . Substantive interpretations of these different forms of missingness will be given in section 6. In each case we will furthermore consider the special situation where for $y' = (y'_1, y'_2)$, the variables of y_1 have probability zero of being missing, that is, the corresponding τ 's are minus infinity, and the variables of y_2 have nonzero probabilities of being missing. In studying these special cases, the more general situation should be clear.

3.1 Missingness Predicted by the Latent Response Variables

Here, $\Gamma_\eta = 0$ and (dropping the y subscript on Γ_y)

$$s^* = \Gamma y^* + \delta, \tag{10}$$

so that with ordinary assumptions using the CFA formulation,

$$\begin{aligned} E \begin{pmatrix} y^* \\ s^* \end{pmatrix} &= \begin{pmatrix} v + \Lambda \kappa \\ \Gamma(v + \Lambda \kappa) \end{pmatrix}, \\ V \begin{pmatrix} y^* \\ s^* \end{pmatrix} &= \begin{pmatrix} \Lambda \Psi \Lambda' + \Theta_{ee} & \text{symmetric} \\ \Gamma(\Lambda \Psi \Lambda' + \Theta_{ee}) + \Theta_{de} & \Gamma(\Lambda \Psi \Lambda' + \Theta_{ee}) \Gamma' + \Theta_{dd} \end{pmatrix}. \end{aligned} \tag{11}$$

Consider first the true likelihood component for a sample unit where all variables of y_2^* are observed. Let

$$\Pr (s_{i1}^* > \tau_1, s_{i2}^* > \tau_2, \dots, s_{ip}^* > \tau_p | y_i) = \pi_i. \tag{12}$$

Integrating over the latent s^* 's, we may write this component as

$$L_i = \phi(y_{1i}) \pi_i \int_{\tau_1}^\infty \dots \int_{\tau_p}^\infty \frac{\phi(y_{2i}, s^* | y_{1i})}{\pi_i} ds^*, \tag{13}$$

where the term inside the integrals is the density of the truncated normal y_2, s^* distribution conditional on y_{1i} , where the s^* 's are truncated from below at τ_1, \dots, τ_p . Then,

$$\begin{aligned} L_i &= \phi(y_{1i}) \phi(y_{2i} | y_{1i}) \int_{\tau_1}^\infty \dots \int_{\tau_p}^\infty \phi(s^* | y_{1i}, y_{2i}) ds^* \\ &= \phi(y_i) \int_{\tau_1}^\infty \dots \int_{\tau_p}^\infty \phi(s^* | y_i) ds^*. \end{aligned} \tag{14}$$

We note from (10) and (11) that only if $\Theta_{\delta e} = \mathbf{0}$ does the conditional normal density inside the integrals not involve model parameters θ but only missingness parameters ϕ . If this holds, the true likelihood component for this sample unit may be written symbolically as the product $\phi(\theta; y_i)f(\phi; y_i)$, where the first term is the ordinary normal density term when ignoring missingness (a function of θ only), and the second term is a component due to missingness which does not involve θ ; see (7). We also assume that we have “distinctness” in the Rubin (1976, p. 582) sense. “The parameter ϕ is distinct from θ if there are no a priori ties, via parameter space restrictions or prior distributions, between ϕ and θ .” We would not in general impose, for instance, equality restrictions between any θ and ϕ parameters. Consequently, when differentiating the true likelihood component of (14) to get maximum likelihood estimates of θ , the second term does not contribute.

Consider next the true likelihood component for a sample unit where none of the variables of y_2^* are observed. Redefining π_i in an obvious way, L_i is obtained by integrating over the latent y_2^* 's,

$$\begin{aligned}
 L_i &= \phi(y_{1i})\pi_i \int_{-\infty}^{\tau_1} \cdots \int_{-\infty}^{\tau_p} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\phi(y_2^*, s^* | y_{1i})}{\pi_i} dy_2^* ds^* \\
 &= \phi(y_{1i}) \int_{-\infty}^{\tau_1} \cdots \int_{-\infty}^{\tau_p} \phi(s^* | y_{1i}) ds^*.
 \end{aligned}
 \tag{15}$$

Writing

$$s^* = \Gamma_1 y_1^* + \Gamma_2 y_2^* + \delta,
 \tag{16}$$

we note that only if $\Gamma_2 = \mathbf{0}$ and $\Theta_{\delta e} = \mathbf{0}$ does the conditional normal density in (15) not involve θ but only ϕ . If this holds, the second term of (15) does not enter into the differentiation with respect to θ .

It follows that when missingness is predicted by latent response variables of y^* , ignoring the missingness terms of the true likelihood components when differentiating with respect to model parameters θ is only correct when $\Gamma_2 = \mathbf{0}$ and $\Theta_{\delta e} = \mathbf{0}$. These two requirements explicate the Rubin (1976) “missing at random” (MAR) definition: “The missing data are missing at random if for each possible value of the parameter ϕ , the conditional probability of the observed pattern of missing data, given the missing data and the value of the observed data, is the same for all possible values of the missing data” (p. 582). In our formulation, the observed pattern of missing data is determined by s^* , which when $\Gamma_2 = \mathbf{0}$ and $\Theta_{\delta e} = \mathbf{0}$ only depends on the value of the observed data on y_1 . Conditional on y_1 , s^* and y_2^* are uncorrelated. We conclude that MAR and distinctness of θ and ϕ gives correct maximum likelihood estimation of θ when using the FQL estimator, that is, ignoring the missingness mechanism (see Rubin, sec. 7).

We may also note that if in addition to $\Gamma_2 = \mathbf{0}$ and $\Theta_{\delta e} = \mathbf{0}$, we have $\Gamma_1 = \mathbf{0}$, that is, $s^* = \delta$, then we have “missing completely at random” (MCAR). MCAR is obtained when MAR holds in addition to “observed at random” (OAR): “The observed data are observed at random if for each possible value of the missing data and the parameter ϕ , the conditional probability of the observed pattern of missing data, given the missing data and the observed data, is the same for all possible values of the observed data” (Rubin, 1976, p. 582). We may note that MCAR is what is usually meant when imprecisely using the expression “missing at random.” It is clear that MCAR is a considerably stronger requirement than MAR. MCAR implies that $\Sigma_{y_2^*} = \mathbf{0}$ in (9). It should be emphasized that only under MCAR will the LQL and PPA estimators have no large sample bias.

When $\Gamma_2 = \mathbf{0}$ and $\Theta_{\delta e} = \mathbf{0}$ does not hold, MAR does not hold and the response mechanism is not ignorable. To obtain correct maximum likelihood estimation in this

case, a model such as the present one must be specified and the maximization carried out with respect to the parameters of θ and ϕ jointly. For special cases of structural equation models without latent variables, this has been attempted in econometric modeling by, for example, Hausman and Wise (1979) and Heckman (1976). Here, s^* and y_2^* are allowed to correlate conditional on y_1 . As emphasized, however, by Little (1982, 1985), such modeling is relatively sensitive to model misspecification, since rather strong, largely unverifiable assumptions have to be made about this conditional distribution. For instance, in Hausman and Wise (1979), maximum likelihood estimation is carried out under the normality assumptions given above. In contrast, the FQL estimator merely utilizes the normality of y , an assumption to which parameter estimates have been found rather robust in complete data settings (see e.g., Boomsma, 1983; Muthén & Kaplan, 1985).

Our view is that even when MAR does not hold, it may be preferable to act as if the response mechanism is ignorable, and to employ the FQL estimator anyway. We expect it to often reduce the bias of the LQL and PPA alternatives. The study of artificial data in section 6 generally supports this proposal. If possible, one may try to include observable covariates that are likely to reduce the conditional s^* , y_2^* correlation, so that MAR is more closely approximated.

3.2 Missingness Predicted by the Latent Variables

Here, $\Gamma_y = \mathbf{0}$ and (dropping the η subscript on Γ_η)

$$s^* = \Gamma\eta + \delta, \tag{17}$$

so that with ordinary assumptions using the CFA formulation

$$\begin{aligned} E\begin{pmatrix} y^* \\ s^* \end{pmatrix} &= \begin{pmatrix} v + \Lambda\kappa \\ \Gamma\kappa \end{pmatrix}, \\ V\begin{pmatrix} y^* \\ s^* \end{pmatrix} &= \begin{pmatrix} \Lambda\Psi\Lambda' + \Theta_{\epsilon\epsilon} & \text{symmetric} \\ \Gamma\Psi\Lambda' + \Theta_{\delta\epsilon} & \Gamma\Psi\Gamma' + \Theta_{\delta\delta} \end{pmatrix}. \end{aligned} \tag{18}$$

Consider again the true likelihood component for a sample unit when all variables of y_2^* are observed. Integrating over the latent η 's and s^* 's,

$$L_i = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi(\eta, y_{1i}) \pi_i \int_{\tau_1}^{\infty} \cdots \int_{\tau_p}^{\infty} \frac{\phi(y_{2i}, s^* | \eta, y_{1i})}{\pi_i} ds^* d\eta, \tag{19}$$

where in this case

$$\pi_i = \Pr(s_{i1}^* > \tau_1, \dots, s_{ip}^* > \tau_p | \eta_i, y_{1i}). \tag{20}$$

Changing the order of integration and integrating out the latent η variables, we have

$$L_i = \phi(y_i) \int_{\tau_1}^{\infty} \cdots \int_{\tau_p}^{\infty} \phi(s^* | y_i) ds^*. \tag{21}$$

It is clear from (18) that the conditional normal s^* density in (21) will, in general, involve both θ and ϕ parameters, unless MCAR is at hand so that $\Gamma = \mathbf{0}$ and $\Theta_{\delta\epsilon} = \mathbf{0}$. MAR will not hold even if missingness is predicted by latent variables measured only by y_1 variables.

The same conclusion is reached when considering a likelihood component for a sample unit when all variables of y_2^* are missing. Having integrated out the latent η and y_2^* variables, we are left with the expression of (21), except that y_i is replaced by y_{1i} .

When missingness is predicted by the latent variables, MAR is not at hand so that the response mechanism is not ignorable and the FQL estimator will not give maximum

likelihood estimation. Both FQL and the classic missing data approaches of LQL and PPA will then exhibit large sample bias at least for some parameters. As we will see, however, the FQL estimator may in this case still be preferable.

3.3 *Testing Restricted Versus Unrestricted Models.*

At this point we may introduce the term the “unrestricted” model to refer to the model with the p parameters of μ and the $p(p + 1)/2$ parameters of Σ , while the term the “restricted” model will refer to the model where the μ and Σ elements are viewed as functions of θ parameters as in CFA. Assuming ignorability, a correct large sample chi-square test of CFA model fit is obtained for the FQL estimator via the ordinary likelihood ratio criterion, comparing the likelihood of the null hypothesis H_0 for the restricted model with the likelihood of the H_1 hypothesis for the unrestricted model (see also Rubin, 1976, section 7). Below, however, we will also refer to tests of model fit when the likelihood ratio approach is used incorrectly so that it does not yield a true chi-square variate, as with FQL when ignorability is not at hand, and with LQL and PPA when MCAR is not at hand. The phrase “perfect model fit” will refer to such a “quasi-chi-square” value being zero. With LQL and PPA, this also means that there are zero residuals between sample statistics and model predicted counterparts, but with FQL there is not a single such set of residuals, as is clear from (8).

4. Some Special Cases

Consider for simplicity the case of only two missing data patterns; sample units that have no missing data and sample units that have missing data on y_2 but no missingness on y_1 . For concreteness we may think of the y subscripts as referring to Time 1 and Time 2 measurements of the same variables, where we have missingness due to attrition at Time 2. This missing data situation is exhibited in Figure 1. Figure 2 gives a path diagram for a longitudinal structural equation model exemplifying the model to be estimated. Here, $y'_1 = (y_{11} y_{21} y_{31} y_{22} y_{32})$, $y'_2 = (y_{13} y_{23} y_{33})$. This model will be discussed in detail in section 6 under missingness case I.

4.1 *The Unrestricted Model*

Let N_{NA} be the number of sample units that do not attrit, while N is the total sample size. Then, the log likelihood of (7) can be simply written as

$$\log L = \sum_{i=1}^{N_{NA}} \log \phi(y_{1i}, y_{2i}) + \sum_{i=N_{NA}+1}^N \log \phi(y_{1i}) + \log f(y), \tag{22}$$

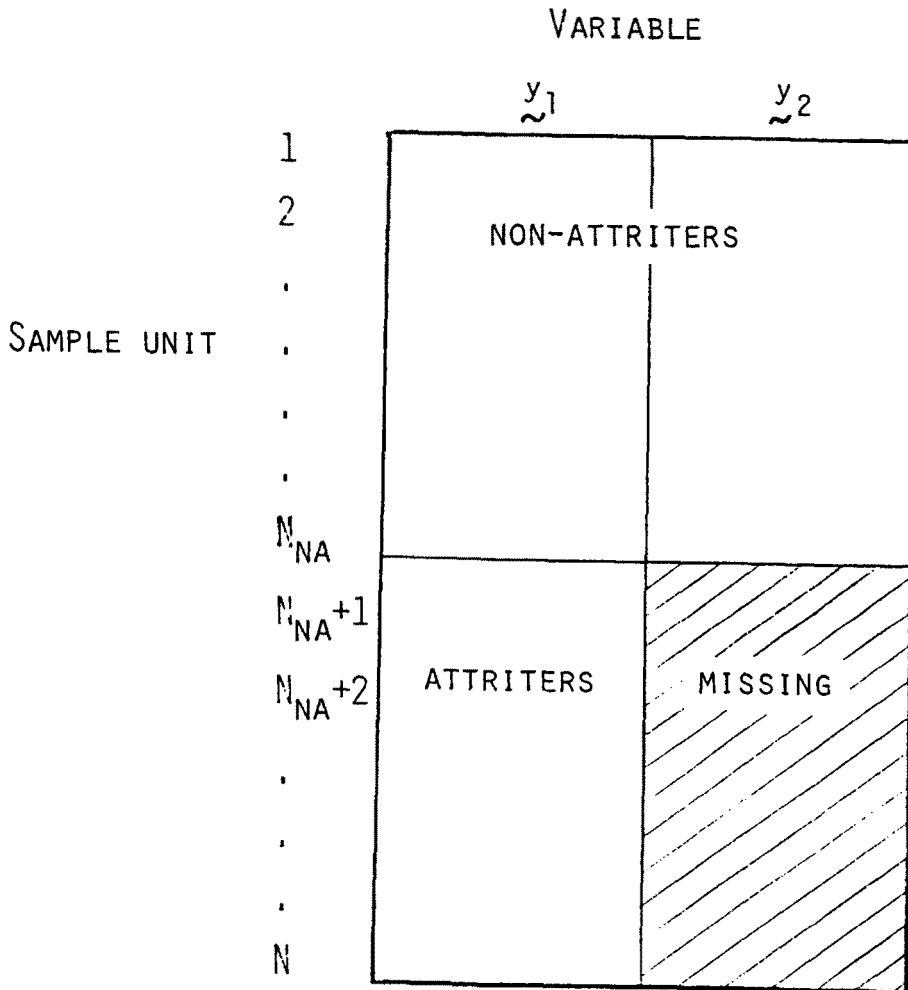
or, alternatively as

$$\log L = \sum_{i=1}^N \log \phi(y_{1i}) + \sum_{i=1}^{N_{NA}} \log \phi(y_{2i} | y_{1i}) + \log f(y). \tag{23}$$

Anderson (1957) noted that FQL estimation of the unrestricted model can be carried out via (23) in a simple, noniterative way, due to the factoring of the quasi-likelihood into two parts involving distinct sets of parameters. Rubin (1974) generalized this approach, allowing for simplified estimation with other missing data patterns and distributions as well; for an application, see, for example, Marini, Olsen, and Rubin (1980).

In (23), we have the usual unrestricted ML estimates $\hat{\Sigma}_{11} = S_{11}$, $\hat{\mu}_1 = \bar{y}_1$, where no superscript for \bar{y}_1 and S_{11} refer to using the whole sample at time point one. Maximization of the second term on the right-hand side of (23) gives estimates of the linear

FIGURE 1
A SPECIAL CASE OF MISSINGNESS (I)



regression parameters. Reverting to the original parameters, we have the FQL estimates

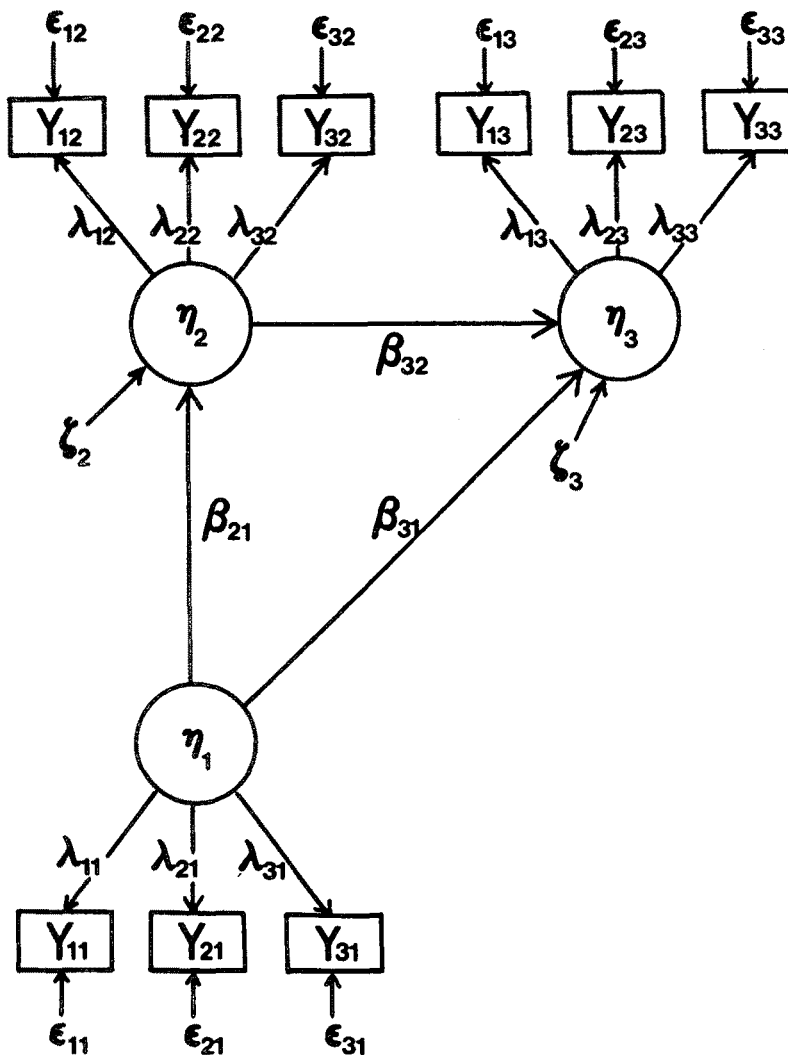
$$\begin{aligned}
 \hat{\mu}_2 &= \bar{y}_2^{NA} - S_{21}^{NA} S_{11}^{NA^{-1}} (\bar{y}_1^{NA} - \bar{y}_1), \\
 \hat{\Sigma}_{21} &= S_{21}^{NA} S_{11}^{NA^{-1}} S_{11}, \\
 \hat{\Sigma}_{22} &= S_{22}^{NA} - S_{21}^{NA} S_{11}^{NA^{-1}} (S_{11}^{NA} - S_{11}) S_{11}^{NA^{-1}} S_{21}^{NA'}.
 \end{aligned}
 \tag{24}$$

We note from (24) the well-known difference between the FQL estimator and the LQL estimator in the unrestricted case. LQL only considers the first term of (22) and does not utilize the fact that time 1 attriter observations on y_1 contribute to full likelihood estimation of both time 1 and time 2 parameters.

It is interesting to note that the equations of (24) also appear in the Pearson (1912)

Figure 2

Path Diagram for Latent Variable Model



and Lawley (1943–1944) selection context discussed in section 2. Pearson and Lawley obtained the population counterpart of (24) for the case where selection takes place on y_1 , indirectly influencing the y_2 distribution, where the NA group of (24) would correspond to the subpopulation of selected units.

4.2 The Restricted Model Under Missingness Predicted by the Latent Variables

Meredith (1964) and Olsson (1978), see also Muthén and Jöreskog (1983), used the selection formulas of (9) to study factor analysis on selected subpopulations. Their results will be used below to study the LQL and FQL estimators for some special cases of our model. Here, we will assume a CFA model where there are no restrictions placed on the parameters of κ or Ψ .

As in section 3.2 we will consider missingness predicted by the latent variables, $\mathbf{s}^* = \mathbf{\Gamma}\boldsymbol{\eta} + \boldsymbol{\delta}$. Consider a subpopulation of units obtained by truncating \mathbf{s}^* at τ . Meredith (1964), Olsson (1978), and Muthén and Jöreskog (1983) show that for this missingness case, the general selection formulas imply that the CFA model of (4) still holds in the subpopulation of selected units, with

$$\begin{aligned} \mathbf{v}^s &= \mathbf{v}, \Lambda^s = \Lambda, \Theta_{\varepsilon\varepsilon}^s = \Theta_{\varepsilon\varepsilon}, \\ \boldsymbol{\kappa}^s &= \boldsymbol{\kappa} + \boldsymbol{\Psi}\boldsymbol{\Gamma}'\boldsymbol{\Sigma}_{s^*s^*}^{-1}(\boldsymbol{\mu}_{s^*}^s - \boldsymbol{\mu}_{s^*}), \\ \boldsymbol{\Psi}^s &= \boldsymbol{\Psi} + \boldsymbol{\Psi}\boldsymbol{\Gamma}'\boldsymbol{\Sigma}_{s^*s^*}^s(\boldsymbol{\Sigma}_{s^*s^*}^s - \boldsymbol{\Sigma}_{s^*s^*})\boldsymbol{\Sigma}_{s^*s^*}^{-1}\boldsymbol{\Gamma}\boldsymbol{\Psi}. \end{aligned} \tag{25}$$

Here, $\boldsymbol{\mu}_{s^*}$ and $\boldsymbol{\Sigma}_{s^*s^*}$ are as in (18), and $\boldsymbol{\mu}_{s^*}^s$ and $\boldsymbol{\Sigma}_{s^*s^*}^s$ are given by the truncated multivariate normal distribution (with τ 's as truncation points). Hence, \mathbf{v} , Λ , and $\Theta_{\varepsilon\varepsilon}$ are invariant under this form of selection.

It is clear that in the restricted CFA model, the LQL estimator, using the NA group, would in infinitely large samples obtain perfect model fit and consistently estimate \mathbf{v} , Λ , $\Theta_{\varepsilon\varepsilon}$, $\boldsymbol{\kappa}^s$, and $\boldsymbol{\Psi}^s$, so that $\boldsymbol{\kappa}$ and $\boldsymbol{\Psi}$ estimates are, in general, biased.

To study the FQL estimator, consider again the case where there are two missing data patterns, no missingness and missingness on y_2 but not y_1 . Furthermore, assume that our CFA model may be partitioned as

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0}_2 & \Lambda_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}, \tag{26}$$

$$E \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} v_1 + \Lambda_1 \boldsymbol{\kappa}_1 \\ v_2 + \Lambda_2 \boldsymbol{\kappa}_2 \end{pmatrix},$$

$$V \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \Lambda_1 \boldsymbol{\Psi}_{11} \Lambda_1' + \Theta_{11} & \text{symmetric} \\ \Lambda_2 \boldsymbol{\Psi}_{21} \Lambda_1' & \Lambda_2 \boldsymbol{\Psi}_{22} \Lambda_2' + \Theta_{22} \end{pmatrix}. \tag{27}$$

The attrition situation of Figure 2 gives an example where this holds. In this situation, unrestricted and restricted FQL estimation will in large samples lead to the same mean vector and covariance matrix estimates. Hence, the CFA model gives perfect fit. This result is obtained as follows. Consider the population counterparts of (24) and (25). After some algebra, it can be shown that *unrestricted* estimation using FQL in infinitely large samples gives

$$\begin{aligned} \hat{\boldsymbol{\mu}}_1 &= \mathbf{v}_1 + \Lambda_1 \boldsymbol{\kappa}_1, \quad \hat{\boldsymbol{\Sigma}}_{11} = \Lambda_1 \boldsymbol{\Psi}_{11} \Lambda_1' + \Theta_{11}, \\ \hat{\boldsymbol{\mu}}_2 &= \mathbf{v}_2 + \Lambda_2 \mathbf{a}, \quad \hat{\boldsymbol{\Sigma}}_{21} = \Lambda_2 \mathbf{B} \Lambda_1', \quad \hat{\boldsymbol{\Sigma}}_{22} = \Lambda_2 \mathbf{C} \Lambda_2' + \Theta_{22}, \end{aligned} \tag{28}$$

where

$$\begin{aligned} \mathbf{a} &= \boldsymbol{\kappa}_2^{NA} - \boldsymbol{\Psi}_{21}^{NA} \Lambda_1' \boldsymbol{\Sigma}_{11}^{NA^{-1}} (\boldsymbol{\kappa}_1^{NA} - \boldsymbol{\kappa}_1), \\ \mathbf{B} &= \boldsymbol{\Psi}_{21}^{NA} (\boldsymbol{\Psi}_{11}^{NA} - \boldsymbol{\Psi}_{11})^{-1} (\Lambda_1' \boldsymbol{\Sigma}_{11}^{-1} \Lambda_1 + (\boldsymbol{\Psi}_{11}^{NA} - \boldsymbol{\Psi}_{11})^{-1})^{-1}, \\ \mathbf{C} &= \boldsymbol{\Psi}_{22}^{NA} - \boldsymbol{\Psi}_{21}^{NA} \Lambda_1' \boldsymbol{\Sigma}_{11}^{NA^{-1}} (\boldsymbol{\Sigma}_{11}^{NA} - \boldsymbol{\Sigma}_{11}) \boldsymbol{\Sigma}_{11}^{NA^{-1}} \Lambda_1 \boldsymbol{\Psi}_{21}^{NA'}. \end{aligned} \tag{29}$$

We note that time two parameter estimates exhibit large sample bias, which is due to the fact that missingness was not predicted by y_1 so that ignorability was not at hand. Equations (28) and (29) show that restricted FQL estimation will in infinitely large samples give the same likelihood value as the unrestricted analysis, that is, give perfect CFA model fit. Restricted FQL analysis will consistently estimate \mathbf{v} , Λ , $\Theta_{\varepsilon\varepsilon}$, $\boldsymbol{\kappa}_{11}$, $\boldsymbol{\Psi}_{11}$, whereas $\boldsymbol{\kappa}_2$, $\boldsymbol{\Psi}_{21}$, and $\boldsymbol{\Psi}_{22}$ will be estimated as \mathbf{a} , \mathbf{B} , and \mathbf{C} in (26). The large sample FQL bias in

the latter parameters may not always be smaller than the corresponding LQL bias, although as artificial data studies in section 6 will show, the results are encouraging.

We note that the above FQL results do not, in general, hold without the assumed missing data patterns and the model partitioning of (26) and (27). For example, when missingness occurs on one variable, but not the entire set of indicators for a certain factor, it is not possible to build on the Anderson (1957) results, since the “missingness part” of the likelihood (23), that is, the second part, involves parameters that are also involved in the full data part. This is the situation in missingness case II to be studied in section 6.

5. FQL Estimation via LISREL

The FQL estimator maximizes the quasi-likelihood part of the true likelihood of (7),

$$-\frac{1}{2} \sum_{g=1}^G N^g [\log |\Sigma^g| + \text{tr}(\Sigma^{g-1} \mathbf{T}^g)], \quad (30)$$

where

$$\mathbf{T}^g = \mathbf{S}^g + (\bar{\mathbf{y}}^g - \boldsymbol{\mu}^g)(\bar{\mathbf{y}}^g - \boldsymbol{\mu}^g)', \quad (31)$$

arranging the sample units into G groups corresponding to distinct missing data patterns. We will now show that for many cases, FQL estimation may be carried out with existing structural equation modeling software, in which mean and covariance structure modeling is allowed for in a simultaneous analysis of multiple groups. As an example, we will consider the popular LISREL program. We will also show that large sample chi-square testing of model fit can be carried out, just as in situations with no missing data.

In the LISREL program the following function is effectively minimized

$$\sum_{g=1}^G \frac{N^g}{N} [\log |\Sigma^g| + \text{tr}(\Sigma^{g-1} \mathbf{T}^g) - \log |\mathbf{S}^g| - p], \quad (32)$$

although, equivalently (see Jöreskog & Sörbom, 1980), the fitting function (32) with \mathbf{M} instead of \mathbf{S} and $\boldsymbol{\Omega}$ instead of $\boldsymbol{\Sigma}$ is the one actually utilized, where \mathbf{M} is the sample moment matrix for moments around zero of the observations with a constant unit variable added in order to include the mean vector, and $\boldsymbol{\Omega}$ is the corresponding population matrix (Jöreskog & Sörbom, 1984).

In an ordinary LISREL analysis, the index g corresponds to different groups (sub-populations) in which independent, random sampling has been carried out. The restricted analysis usually corresponds to a set of parameter equality restrictions across the groups, which have similar model structure. The standard unrestricted analysis estimates each group's unrestricted $\boldsymbol{\mu}^g$ and Σ^g , separately. N times the minimum value of the LISREL fitting function for the restricted analysis gives the usual, large sample likelihood ratio chi-square test of model fit, since terms corresponding to the likelihood value for the unrestricted analysis does not involve any restrictions on the G $\boldsymbol{\mu}^g$ vectors, the estimation need only utilize information from the sample covariance matrices, since terms involving mean vector residuals will vanish.

Let us now return to FQL estimation in a single population with a sample involving missing data. We will assume that the number of distinct missing data patterns, G , is relatively small, so that for each pattern group g there are more than $p + 1$ sample units, yielding positive definite \mathbf{S}^g matrices. It is clear from (30) and (32) that FQL estimation can be carried out by a LISREL multiple group run with G groups. As opposed to the case in (30), the LISREL analysis will not involve reduced-sized mean vectors and covariance matrices corresponding to missing variables, but instead uses a dummy variable

approach. In terms of the special case of section 4, the LISREL approach builds on the log L version of (22) and does not utilize (23). It should be emphasized that the LISREL multiple group feature is only a technical convenience, since the model concerns a single population. This fact will affect how the H_1 analysis is carried out. If FQL estimation were to be carried out in a multiple population model, the corresponding LISREL analysis would technically involve as many groups as there are missing data patterns in all the different samples. Below, we will only consider single population analyses.

Consider first restricted, H_0 estimation by FQL using LISREL. The parameters are those of θ . This is carried out by imposing equality restrictions across the missing data pattern groups for common parameters. The fact that different numbers of variables are observed in the different groups is handled by using dummy variables in the groups with deficient numbers, the influence of which on the fitting function is made to be nil (see also Jöreskog, 1971). (The actual LISREL setup will not be given here, but is available upon request from the first author. A simple setup using LISCOMP is also available.)

It is important to note that even if H_0 does not impose any restrictions on μ , as is commonly the case in single population analyses, the FQL LISREL analysis must still include the specification of equality of mean vector parameters across missing data pattern groups for common variables. This is because the mean residuals of (31) will not vanish under the equality constraints on common elements of μ . As we will discuss in the context of the FQL H_1 , common elements of \bar{y}^g are not, in general, expected to be equal across the G groups, even in infinitely large samples.

To permit a large sample chi-square test of model fit, we also need the likelihood value for the FQL unrestricted, H_1 hypothesis. As opposed to ordinary LISREL analysis, this requires a second LISREL run. Under the FQL H_1 , the parameters are the elements of μ and Σ . In the LISREL analysis, equality constraints for these elements are again imposed across missing data groups, using the dummy variable approach. From (32), however, we note that this LISREL run will actually produce a chi-square test of H_1 against what we will call H_2 , where H_2 is the hypothesis of an unrestricted, multiple population model. Hence, H_2 is not relevant to the present analysis, but will be discussed below. For our purposes, the second LISREL run is used to get a chi-square value, which we will subtract from the chi-square of the restricted run. This gives a correct test of H_0 against H_1 , since we note that the chi-square of the restricted run is also obtained by testing against H_2 , so that the H_2 likelihood component cancels out in the subtraction. The difference in degrees of freedom on the two LISREL runs produces the correct degrees of freedom for testing H_0 against H_1 . This is the same degree of freedom that would have been used in the corresponding analysis without missing data.

It is of interest to further consider the unrestricted FQL LISREL run and the H_1 and H_2 hypotheses. Unless MCAR is at hand, the common elements of the G sample mean vectors and covariance matrices of (32) are not expected to be equal even in infinitely large samples. Hence, the test of H_1 against H_2 is a test of MCAR, where the problem has been reformulated in terms of multiple populations. When MAR, but not MCAR holds, this test is expected to lead to rejection. Our experience with real data indicates that MCAR is frequently rejected. Nevertheless, this test may be of interest to carry out as an initial analysis step. If it is not rejected and MCAR holds, the simpler LQL estimator is an alternative that gives consistent, although less efficient estimates than FQL, which is then the correct ML estimator.

We also note that unless MCAR is at hand, the elements of the G sample mean vectors and covariance matrices of (32) are not consistent estimates of the corresponding single population quantities. It seems somewhat paradoxical that given ignorability—assuming MAR and distinctness, but not MCAR—unrestricted FQL estimation gives

correct ML estimates of μ and Σ by a LISREL analysis with equality restrictions across groups, none of which uses consistent sample statistic estimators. This comes about by weighting the information from the different groups in a way that was explicated for a special two-group situation in (24) of section 3.

The idea of using multiple group LISREL analysis to analyze missing data situations is not new; see, for example, Werts, Rock, and Grandy (1979). In previous approaches, however, the original problem has been reformulated as one of multiple populations modeling. While this may also lead to useful analyses, it does not recognize the connection with the true likelihood of (7) for a single population as does the FQL LISREL approach. For instance, Werts et al. test equality for common parts of a CFA model across two populations corresponding to two missing data patterns. When MAR but not MCAR holds, we would expect this hypothesis to be rejected, whereas the original model may well not be, using our FQL approach. Also, since the FQL estimator always requires inclusion of mean parameter equalities in the LISREL run, the estimates would generally be different, unless the model common to the two groups also involves a mean structure in addition to the covariance structure (as it did in the Werts et al. case).

6. Illustrations

In this section we will use some artificial data generated by our section 2 model to exemplify the above methodology and to compare the behavior of the FQL, LQL, and PPA estimators. These estimators have been previously studied in the context of factor analysis by Finkbeiner (1979) and Brown (1983), but only under MCAR, concentrating on issues of efficiency. We will almost exclusively deal with data for which MCAR does not hold, but where there are strong correlations between the selection variables of s^* and the y^* variables. We will study cases where MAR and distinctness holds, so that FQL is correct ML, but concentrate on the perhaps most realistic case where not even MAR holds. We will be particularly interested in parameter estimate bias, but will also consider model fit and standard errors of estimates. A Monte Carlo study will not be performed, but analyses will instead be carried out on population quantities. Throughout we will use models that emanate from that of Figure 2. We will consider two basic models for the y 's, each with two variations, and two basic missing data cases, each with several variations.

The y -part models that will be studied are as follows. The first model involves all y 's of Figure 2. One version considers a CFA model, where the η_1, η_2, η_3 part is parameterized as a 3×3 covariance matrix. Another, equivalent, version considers the structural equation model (SEM) parameterization indicated. The actual (θ -) parameter values chosen for the model of Figure 2 were created on the basis of a subset of the Wheaton, Muthen, Alwin, and Summers (1977) longitudinal model, as reported in Jöreskog and Sörbom (1984), adding an extra indicator per factor, deleting correlated measurement errors, and for simplicity, standardizing the parameters to correspond to y and η variables with unit variances. Here, η_1 represented a socioeconomic background factor, whereas η_2 and η_3 were time one and time two alienation constructs. It is noteworthy that we have chosen a single population model that only imposes restrictions on the covariance matrix and not on the mean vector (mean vector restrictions could, for instance, arise by imposing measurement parameter equality restrictions across time for the time one and time two alienation indicators, estimating a factor mean change between η_2 and η_3).

The second model is the marginal part of Figure 2 that corresponds to $(y_{12} y_{22} y_{32} y_{13} y_{23} y_{33})$. One version considers this as a CFA model, where the η_2, η_3 part is parameterized as a 2×2 covariance matrix, while the other version retains the structural regression parameterization indicated in the figure.

Table 1
Full Population and MAR Non-Attriter
Subpopulation Mean Vectors and Covariance Matrices^a

Variable	Full Population Means	Covariance Matrix									Sub- Population Means
		y_{11}	y_{21}	y_{31}	y_{12}	y_{22}	y_{32}	y_{13}	y_{23}	y_{33}	
y_{11}	0.0	-15	-23	-29	-33	-34	-39	-27	-27	-27	.240
y_{21}	0.0	0.658	-16	-28	-31	-32	-37	-26	-26	-26	.248
y_{31}	0.0	0.455	0.495	-12	-39	-40	-46	-32	-32	-32	.214
y_{12}	0.0	-0.366	-0.398	-0.276	-10	-15	-17	-15	-15	-15	-.194
y_{22}	0.0	-0.340	-0.369	-0.255	0.642	-9	-17	-16	-16	-16	-.105
y_{32}	0.0	-0.239	-0.260	-0.180	0.452	0.419	-6	-18	-18	-18	-.150
y_{13}	0.0	-0.310	-0.338	-0.233	0.437	0.405	0.285	-5	-8	-8	-.134
y_{23}	0.0	-0.275	-0.299	-0.207	0.387	0.358	0.252	0.511	-4	-8	-.119
y_{33}	0.0	-0.203	-0.221	-0.153	0.286	0.265	0.187	0.378	0.335	-2	-.088

^a The full population covariance matrix is given in the lower triangular part, excluding its unit diagonal elements. The diagonal and upper triangular part contains percent bias in the non-attriter variances and covariances ((sub-population value - full population value)/full population value \times 100) at an attrition rate of 25% and an R^2 of 50%.

An interesting feature of the marginal model is that for all of the missing data situations to be considered, some variables related to the prediction of missingness are always absent in the model studied. Hence, the FQL estimator is expected to exhibit rather large bias in this case, but it will be of interest to compare its performance to that of LQL and PPA. When comparing the marginal model results with those of the full model, we can gain knowledge about the bias reduction obtained when expanding a model beyond the set of variables which are of primary interest, to include variables that are likely to explain missingness well. Background variables related to socioeconomic status is a good example, see, for example, Marini, Olsen, and Rubin (1980, p. 315-316).

The artificial data were generated as follows. Given the Wheaton et al. (1977) inspired choices of θ parameters, and the ϕ parameter values for a particular missing data situation, the population covariance matrix for the nine y 's and the s^* 's, was calculated. The corresponding mean vector was set to zero. The 9×9 population covariance matrix is given in the lower-triangular part of Table 1, where it should be noted that the unit diagonal elements are excluded. The subpopulation y mean vectors and covariance matrices were then created for the subpopulations corresponding to the missing data patterns generated by truncating the s^* 's at the threshold values. These were obtained using the selection formulas (9). The means, variances and covariances of the truncated

normal s^* distribution were obtained from Muthén (1985); see also Rosenbaum (1961) and Tallis (1961). For the marginal model, the subpopulation mean vectors and covariance matrices were reduced down to six variables, and for both the full and the marginal model variables were deleted corresponding to each pattern of missingness.

LISREL maximum likelihood estimation was then applied to carry out FQL, LQL, and PPA estimation, using population values in place of sample statistics. The FQL estimator uses the subpopulation mean vectors and covariance matrices corresponding to all missing data patterns for its H_0 and H_1 runs in a multiple group, mean and covariance structure run. LQL uses only the covariance matrix for the subpopulation corresponding to the complete data pattern in a single group, covariance structure only run. PPA uses a single group, covariance structure only run with covariance matrix elements that are obtained from the subpopulations that make for the largest possible subset of the full population.

For each case, a table will be given for the CFA and the SEM estimates, in raw and standardized form, including percentage bias. In addition, "population chi-square" values will be reported, referring to the regular chi-square measure as calculated using our population mean vectors and covariance matrices and a sample size of 600. These are then not chi-square model fit statistics in the usual sense, but do give a measure of distortion of the original model structure. For FQL, two chi-square values are given; one refers to the "test" of the model we are interested in, while the other refers to the "test" of MCAR discussed in section 5, obtained as a byproduct of the FQL H_1 run. The latter test value may be viewed as an indication of deviation from MCAR, or strength of selectivity.

The results will first be presented for missingness case I and then for missingness case II. Within each of these missingness situations we will first study the full model and then compare with the marginal model.

6.1 Results for Missingness Case I

Here, missing data may occur for y_{13} , y_{23} , y_{33} in the longitudinal framework discussed in section 4. Missingness is determined by a single s^* , reflecting the notion of attrition and not variable specific missingness.

6.1.1 MAR: Missingness Predicted by the Six Time One Observed Variables. The first missingness variation we will consider is

$$s^* = \frac{2}{3}\gamma(y_{11}^* + y_{21}^* + y_{31}^*) - \frac{1}{3}\gamma(y_{12}^* + y_{22}^* + y_{32}^*) + \delta, \quad (33)$$

such that the background indicators are twice as influential as the time one indicators. The y^* 's of (33) are identical to the corresponding y 's in this missingness case. In terms of the Wheaton et al. (1977) example, an increased value of any of the socioeconomic background indicators is taken to yield an increased propensity not to attrit (a high s^* value), while an increased value of the time one alienation indicators is taken to yield a decreased propensity not to attrit (i.e., an increase of the propensity to attrit). Hence, the observed variables at time one are assumed to be the only relevant predictors, so that alienation feelings at time two are irrelevant.

We note that this missingness situation fulfills the assumption of MAR and also ignorability, so that FQL gives correct maximum likelihood estimation. To determine the degree of selectivity in the missingness, an s^* R^2 of 50% was chosen by setting $\gamma = 0.329$ in (33). Throughout, s^* variables are taken to have zero means and unit variances. A τ value of -0.675 was chosen for s^* to give an attrition rate of 25%. To illustrate the strong selectivity of the missingness, the percentage difference between the covariance matrix analyzed by LQL, obtained for the non-attriter group by (9), and the full popu-

lation covariance matrix is given in the diagonal and upper-triangular portion of Table 1. Means are also given.

The FQL "population chi-square test" of model fit, comparing H_0 to H_1 , turns out to be zero as it should. H_0 involves 30 parameters, where the \mathbf{v} vector is included and the latent variable means are fixed at zero. The latent variable metrics are further determined by fixing one loading per factor. Since H_1 involves 9 mean vector parameters and 45 covariance matrix parameters, the model test has 24 degrees of freedom (it should be noted that the dummy variable approach in LISREL does not give the correct degrees of freedom). As a byproduct of the FQL analysis, we also get the test of MCAR discussed in section 5, comparing H_1 to H_2 with 27 degrees of freedom in a multiple population framework. This yields the value 189.68. Since MCAR does not hold, the LQL and PPA model test values are not zero, but 0.16 and 0.36, respectively.

The results of the estimation are given in Table 2. Despite the strong deviation from MCAR, the FQL estimator exhibits no bias, as was expected since MAR is at hand. (For certain CFA parameters, a very slight bias is obtained in the LISREL results; this could be avoided if it were possible to sharpen the convergence criterion.) The LQL and PPA estimators are seen to be seriously biased. (In our opinion, a bias of less than 10–15% may not be serious in most latent variable modeling contexts.) It is interesting to note that the bias is almost exclusively occurring in the structural parameters, which may be deemed the most important part of a latent variable model. Also of interest is that in terms of unstandardized values, the LQL estimator does relatively better in the SEM metric than in the CFA metric. One cause of this is perhaps that s^* is to a large portion predicted by indicators of the latent variable η_1 , which is exogenous in the structural regressions, resulting in relatively lower structural regression parameter bias (this reasoning builds on the well-known fact that in ordinary regression there is no bias if selection takes place on x). While PPA outperforms LQL in terms of overall bias in the CFA metric, the reverse is true in the SEM metric.

The standard errors of the estimates are given in parentheses. The standard errors given in connection with the true population values refer to the situation where there is no missingness, but a full sample is used. The FQL standard errors are somewhat larger than these values as is expected, although not to any great degree. For the LQL and PPA estimators, two standard error entries are given. The top one refers to the actual analysis, where it should be kept in mind that since the assumptions underlying the estimator do not hold, the values are not correct and will not represent actual sampling variation. As an added piece of information, the second standard error entry for the LQL and PPA estimators refers to what would have been obtained if MCAR held but there were missingness on the same variables with the same attrition rate (in this case, LQL and PPA would show no bias). These values may be compared to the full sample and FQL values, where the latter would remain the same under MCAR as under the present MAR situation. For instance, the SEM standard errors for LQL under MCAR show slightly higher values than FQL due to ignoring the 25% nonattriters. Such issues of estimate variability under MCAR were discussed by Brown (1983) and Finkbeiner (1979). However, we will not study this topic further since in all cases that follow none of the estimators will be consistent due to violation of MCAR and MAR, so that the sampling variability will not be adequately represented by the population standard errors. Also, the amount of large sample bias will be of primary importance.

6.1.2 Missingness Predicted by all Nine Latent Response Variables. We will now allow the partly missing time two alienation indicators to also influence the attrition.

Table 2
Missingness Case I:
Missingness Predicted by the Six Time One Observed Variables (MAR)

Parameter	Parameter Value	FOL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Measurement Parameters							
λ_{11}	0.788	0.778 ^a	0	0.788 ^a	0	0.788 ^a	0
λ_{21}	0.846 (0.051) ^b	0.846 (0.051) ^c	0	0.866 (0.077) ^c (0.058) ^d	2	0.848 (0.052) ^c (0.051) ^d	0
λ_{31}	0.585 (0.044)	0.585 (0.044)	0	0.551 (0.060) (0.051)	- 6	0.585 (0.044) (0.044)	0
λ_{12}	0.832	0.832 ^a	0	0.832 ^a	0	0.832 ^a	0
λ_{22}	0.771 (0.045)	0.771 (0.046)	0	0.765 (0.060) (0.052)	- 1	0.769 (0.046) (0.045)	0
λ_{32}	0.543 (0.044)	0.543 (0.044)	0	0.524 (0.056) (0.050)	- 3	0.541 (0.044) (0.044)	0
λ_{13}	0.759	0.759 ^a	0	0.759 ^a	0	0.759 ^a	0
λ_{23}	0.672 (0.053)	0.672 (0.061)	0	0.672 (0.066) (0.061)	0	0.672 (0.061) (0.053)	0
λ_{33}	0.497 (0.049)	0.497 (0.057)	0	0.497 (0.061) (0.057)	0	0.497 (0.054) (0.049)	0
θ_{11}	0.395 (0.037)	0.395 (0.037)	0	0.397 (0.044) (0.043)	1	0.396 (0.037) (0.037)	0
θ_{21}	0.284 (0.037)	0.284 (0.038)	0	0.280 (0.046) (0.043)	- 1	0.283 (0.038) (0.037)	0
θ_{31}	0.658 (0.043)	0.658 0.043	0	0.654 (0.048) (0.049)	- 1	0.659 (0.043) (0.043)	0
θ_{12}	0.308 (0.036)	0.308 (0.037)	0	0.305 (0.043) (0.041)	- 1	0.305 (0.037) (0.036)	- 1
θ_{22}	0.405 (0.036)	0.405 (0.037)	0	0.406 (0.043) (0.042)	0	0.406 (0.037) (0.036)	0
θ_{32}	0.705 (0.045)	0.705 (0.045)	0	0.705 (0.051) (0.052)	0	0.707 (0.045) (0.045)	0
θ_{13}	0.423 (0.044)	0.423 (0.051)	0	0.423 (0.052) (0.051)	0	0.423 (0.048) (0.044)	0
θ_{23}	0.547 (0.044)	0.547 (0.050)	0	0.547 (0.051) (0.050)	0	0.547 (0.046) (0.044)	0
θ_{33}	0.752 (0.049)	0.752 (0.056)	0	0.752 (0.056) (0.056)	0	0.752 (0.049) (0.049)	0

Table 2 (continued)

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Structural Parameters - Confirmatory Factor Analysis Model							
Unstandardized Solution							
$V(\eta_1)$	1.000 (0.100)	1.000 (0.100)	0	0.749 (0.100) (0.115)	-25	1.000 (0.100) (0.100)	0
$V(\eta_2)$	1.000 (0.091)	1.000 (0.092)	0	0.863 (0.099) (0.105)	-14	1.004 (0.092) (0.091)	0
$V(\eta_3)$	1.000 (0.111)	1.002 (0.128)	0	0.921 (0.125) (0.128)	- 8	0.921 (0.111) (0.111)	- 8
$C(\eta_2, \eta_1)$	-0.566 (0.061)	-0.566 (0.061)	0	-0.376 (0.057) (0.070)	-34	-0.566 (0.061) (0.061)	0
$C(\eta_3, \eta_1)$	-0.526 (0.063)	-0.525 (0.069)	0	-0.383 (0.062) (0.073)	-27	-0.385 (0.058) (0.063)	-27
$C(\eta_3, \eta_2)$	0.691 (0.067)	0.691 (0.073)	0	0.585 (0.070) (0.078)	-15	0.582 (0.063) (0.067)	-16
Standardized Solution							
$\rho(\eta_2, \eta_1)$	-0.566	-0.566	0	-0.467	-17	-0.566	0
$\rho(\eta_3, \eta_1)$	-0.526	-0.525	0	-0.461	-12	-0.401	-24
$\rho(\eta_3, \eta_2)$	0.691	0.690	0	0.656	- 5	0.605	-12
Structural Parameters - Structural Equation Model							
Unstandardized Solution							
β_{21}	-0.566 (0.051)	-0.566 (0.051)	0	-0.502 (0.068) (0.059)	-11	-0.597 (0.051) (0.051)	0
β_{31}	-0.198 (0.060)	-0.198 (0.069)	0	-0.219 (0.075) (0.070)	11	-0.083 (0.061) (0.060)	-58
β_{32}	0.579 (0.066)	0.579 (0.076)	0	0.582 (0.079) (0.077)	1	0.533 (0.068) (0.066)	- 8
$V(\eta_1)$	1.000 (0.100)	1.000 (0.100)	0	0.749 (0.101) (0.115)	-25	1.000 (0.100) (0.100)	0
$V(\zeta_2)$	0.680 (0.071)	0.680 (0.071)	0	0.675 (0.084) (0.082)	- 1	0.683 (0.071) (0.071)	0
$V(\zeta_3)$	0.498 (0.073)	0.498 (0.084)	0	0.497 (0.087) (0.084)	0	0.579 (0.083) (0.073)	16
Standardized Solution							
β_{21}	-0.566	-0.566	0	-0.467	-17	-0.566	0
β_{31}	-0.198	-0.198	0	-0.198	0	-0.087	-55
β_{32}	0.579	0.579	0	0.564	- 1	0.556	- 3
$R^2(\eta_2)$	0.320	0.320	0	0.218	-31	0.320	0
$R^2(\eta_3)$	0.503	0.503	0	0.461	- 7	0.372	-25

^a Fixed parameter.

^b Full sample standard error (no missingness).

^c MAR standard error.

^d MCAR standard error.

Attrition rate is 25%, R^2 is 50%.

Similar to (33), we let

$$s^* = \frac{2}{3}\gamma(y_{11}^* + y_{21}^* + y_{31}^*) - \frac{1}{3}\gamma(y_{12}^* + y_{22}^* + y_{32}^* + y_{13}^* + y_{23}^* + y_{33}^*) + \delta. \quad (34)$$

Due to the influence of the time two variables, the missingness mechanism is not ignorable in this case and the FQL estimator will show large-sample bias. In order to get a more comprehensive comparison of bias among the three estimators, both the strength of selectivity in the missingness and the attrition rate was varied. Strength of selectivity was varied by using 75%, 50%, and 25% $s^* R^2$, with γ values of 0.331, 0.270, and 0.191. Attrition rates of 25% and 33% were studied. In the table below, however, only results for 25% attrition will be given, since the 33% case yielded very similar conclusions.

For 75% R^2 , the chi-square model tests for FQL, LQL, PPA, and the MCAR test gave the values of 0.29, 0.53, 1.24, and 219.17, respectively. For 50%, the values were 0.10, 0.17, 0.53, and 175.73, while 25% gave 0.02, 0.03, 0.13, and 80.91. The degrees of freedom are as before. Table 3 gives the estimates and the percentage bias for 75%, 50%, 25%, respectively. Again, none of the estimators show a considerable amount of bias in the measurement parameters. In the CFA metric, the FQL estimator clearly outperforms both LQL and PPA and this is true for all three levels of R^2 . This is noteworthy since ignorability of the missingness is not at hand here. In most cases, FQL certainly does reduce the LQL/PPA bias as was conjectured. PPA shows less bias than LQL in the CFA metric. In terms of total percentage bias for the unstandardized SEM parameters, LQL outperforms PPA. It is noteworthy that both the FQL and PPA estimators show almost no bias for parameters involved in parts of the model where there is no missingness. The relative performance of the three estimators is not altered by the changes in R^2 . We note that at 50% and 25% R^2 , the FQL bias is in no case substantially larger than the best contender, and in most cases the bias is considerably smaller. This range of R^2 may be the most realistic in social science data. In the tables that follow, we will only consider R^2 's of 50%.

6.1.3 Missingness Predicted by the Latent Variables. In the context of latent variable models it may be relevant to consider missingness predicted by the latent variables instead of the observed ones. For instance, people may “self-select” themselves out of the sample after the time one measurement occasion as predicted by their “true” socioeconomic or alienation status rather than any combination of fallible measurements of such constructs. On the other hand, this may not be relevant if missingness is for instance related to a particular variable which involves a sensitive issue or if selection into the sample is judged by observed scores (e.g., pretest scores).

Consider the case of missingness predicted by all three latent variables

$$s^* = \frac{2}{3}\gamma\eta_1 - \frac{1}{3}\gamma(\eta_2 + \eta_3) + \delta, \quad (35)$$

and the case of missingness predicted by the time one factors η_1 and η_2 ,

$$s^* = \frac{2}{3}\gamma\eta_1 - \frac{1}{3}\gamma\eta_2 + \delta. \quad (36)$$

Throughout this section, an R^2 of 50% is chosen with an attrition rate of 25%. In (35) $\gamma = 0.619$, while in (36) $\gamma = 0.787$. The chi-square model test values for FQL, LQL, and PPA are in these two cases all zero, with zero bias in the measurement parameters. This is in line with the theory of section 4.2. This is noteworthy since with real sample data there is then no indication of a misfitting model, while at the same time quite large LQL and PPA bias may be obtained in the structural parameters. The MCAR test values are 153.46 and 160.54.

Table 3
Missingness Case I:
Missingness Predicted by All Nine Variables

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Measurement Parameters							
λ_{11}	0.778	0.778 ^a 0.778 0.778	- - -	0.778 ^a 0.778 0.778	- - -	0.778 ^a 0.778 0.778	- - -
λ_{21}	0.846	0.847 0.847 0.846	0 0 0	0.879 0.864 0.853	4 2 1	0.847 0.848 0.847	0 0 0
λ_{31}	0.585	0.585 0.585 0.585	0 0 0	0.539 0.558 0.573	- 8 - 5 - 2	0.585 0.585 0.585	0 0 0
λ_{12}	0.832	0.832 ^a 0.832 0.832	- - -	0.832 ^a 0.832 0.832	- - -	0.832 ^a 0.832 0.832	- - -
λ_{22}	0.771	0.770 0.770 0.771	0 0 0	0.760 0.765 0.768	- 1 - 1 0	0.768 0.769 0.770	0 0 0
λ_{32}	0.543	0.542 0.542 0.543	0 0 0	0.513 0.525 0.535	- 6 - 3 - 1	0.539 0.540 0.542	- 1 - 1 0
λ_{13}	0.759	0.759 ^a 0.759 0.759	- - -	0.759 ^a 0.759 0.759	- - -	0.759 ^a 0.759 0.759	- - -
λ_{23}	0.672	0.656 0.663 0.668	- 2 0 0	0.656 0.663 0.668	- 2 - 1 - 1	0.658 0.664 0.668	- 2 - 1 - 1
λ_{33}	0.497	0.467 0.479 0.489	- 6 - 4 - 1	0.467 0.479 0.489	- 6 - 4 - 2	0.469 0.480 0.489	- 6 - 3 - 2
θ_{11}	0.395	0.395 0.395 0.395	0 0 0	0.400 0.398 0.396	1 1 0	0.395 0.396 0.395	0 0 0
θ_{21}	0.284	0.283 0.284 0.284	0 0 0	0.274 0.279 0.282	- 4 - 2 - 1	0.283 0.283 0.283	0 0 0
θ_{31}	0.658	0.658 0.658 0.658	0 0 0	0.655 0.657 0.657	0 0 0	0.659 0.659 0.659	0 0 0
θ_{12}	0.308	0.306 0.307 0.307	- 1 0 0	0.302 0.304 0.306	- 2 - 1 - 1	0.320 0.305 0.306	- 1 - 1 - 1
θ_{22}	0.405	0.406 0.406 0.405	0 0 0	0.408 0.407 0.406	1 0 0	0.407 0.406 0.406	0 0 0
θ_{32}	0.705	0.706 0.706 0.705	0 0 0	0.706 0.706 0.706	0 0 0	0.707 0.707 0.706	0 0 0
θ_{13}	0.423	0.416 0.418 0.421	- 2 - 1 0	0.416 0.418 0.421	- 2 - 1 0	0.418 0.419 0.421	- 1 - 1 0
θ_{23}	0.547	0.550 0.549 0.548	1 0 0	0.550 0.549 0.548	1 0 0	0.549 0.549 0.548	0 0 0
θ_{33}	0.752	0.752 0.754 0.753	0 0 0	0.755 0.754 0.753	0 0 0	0.753 0.753 0.753	0 0 0

Table 3 (continued)

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Structural Parameters - Confirmatory Factor Analysis Model							
Unstandardized Solution							
$V(\eta_1)$	1.000	0.998 0.999 1.000	0 0 0	0.659 0.774 0.887	-33 -22 -10	1.000 1.000 1.000	0 0 0
$V(\eta_2)$	1.000	1.002 1.001 1.001	0 0 0	0.780 0.853 0.926	-22 -15 -7	1.006 1.004 1.002	1 0 0
$V(\eta_3)$	1.000	0.873 0.926 0.968	-13 -7 -3	0.790 0.860 0.931	-21 -14 -7	0.787 0.859 0.931	-21 -14 -7
$C(\eta_2, \eta_1)$	-0.566	-0.566 -0.566 -0.566	0 0 0	-0.286 -0.380 -0.473	-49 -33 -16	-0.567 -0.566 -0.566	0 0 0
$C(\eta_3, \eta_1)$	-0.526	-0.414 -0.462 -0.498	-21 -12 -5	-0.247 -0.340 -0.433	-53 -35 -18	-0.249 -0.342 -0.434	-53 -35 -17
$C(\eta_3, \eta_2)$	0.691	0.600 0.639 0.668	-13 -8 -3	0.465 0.540 0.616	-33 -22 -11	0.460 0.537 0.615	-33 -22 -11
Standardized Solution							
$\rho(\eta_2, \eta_1)$	-0.566	-0.565 -0.566 -0.566	0 0 0	-0.399 -0.467 -0.522	-30 -17 -8	-0.566 -0.566 -0.566	0 0 0
$\rho(\eta_3, \eta_1)$	-0.525	-0.443 -0.480 -0.506	-16 -9 -4	-0.343 -0.417 -0.477	-35 -21 -9	-0.281 -0.369 -0.450	-46 -30 -14
$\rho(\eta_3, \eta_2)$	0.691	0.642 0.664 0.679	-7 -4 -2	0.592 0.631 0.663	-14 -9 -4	0.517 0.579 0.636	-25 -16 -8
Structural Parameters - Structural Equation Model							
Unstandardized Solution							
β_{21}	-0.566	-0.566 -0.566 -0.566	0 0 0	-0.435 -0.491 -0.533	-23 -13 -6	-0.567 -0.567 -0.567	0 0 0
β_{31}	-0.198	-0.111 -0.148 -0.176	-44 -25 -11	-0.138 -0.165 -0.184	-30 -17 -7	-0.014 -0.057 -0.128	-93 -71 -35
β_{32}	0.579	0.537 0.555 0.568	-7 -4 -2	0.545 0.560 0.571	-6 -3 -1	0.465 0.503 0.541	-20 -13 -7
$V(\eta_1)$	1.000	0.999 0.999 1.000	0 0 0	0.659 0.774 0.887	-34 -23 -11	1.000 1.000 1.000	0 0 0
$V(\zeta_2)$	0.680	0.682 0.681 0.680	0 0 0	0.656 0.667 0.674	-4 -2 -1	0.684 0.683 0.682	1 0 0
$V(\zeta_3)$	0.498	0.505 0.503 0.500	1 1 0	0.503 0.502 0.500	1 1 0	0.577 0.569 0.543	16 14 9
Standardized Solution							
β_{21}	-0.566	-0.565 -0.566 -0.566	0 0 0	-0.399 -0.468 -0.522	-30 -17 -8	-0.566 -0.566 -0.566	0 0 0
β_{31}	-0.198	-0.118 -0.154 -0.179	-40 -22 -10	-0.126 -0.156 -0.180	-36 -21 -9	-0.016 -0.062 -0.132	-92 -69 -33
β_{32}	0.579	0.575 0.577 0.578	-1 0 0	0.541 0.558 0.570	-6 -3 0	0.526 0.544 0.561	-9 -6 -3
$R^2(\eta_2)$	0.320	0.320 0.320 0.320	0 0 0	0.159 0.219 0.272	-50 -32 -15	0.320 0.320 0.320	0 0 0
$R^2(\eta_3)$	0.503	0.421 0.457 0.483	-16 -9 -4	0.364 0.417 0.463	-28 -17 -8	0.267 0.337 0.417	-47 -33 -17

^a Fixed parameter.

The three entries for each parameter estimate correspond to missingness R^2 of 75%, 50%, and 25%. Attrition rate is 25%.

The bias in the structural parameters is shown in Table 4, where the top entry corresponds to (35) and the bottom entry corresponds to (36).

We may compare the biases for the case of three factors (top entries of Table 4) with that of missingness predicted by their nine observed indicators in the middle (50%) entries of Table 3, and the two factor results (bottom entries of Table 4) with that of their six observed indicators (the MAR case) in Table 2. These comparisons show that the results are affected rather little by the distinction between latent and observed variables. For instance, in the two factor case the loss of MAR when moving from missingness predicted by the six observed variables to missingness predicted by the corresponding two factors, results in a maximum FQL bias of 8%.

6.1.4 The Marginal Model. Let us now consider estimation in the marginal model for y_{12} , y_{22} , y_{23} , y_{13} , y_{23} , and y_{33} . The missingness situations to be studied were taken to be the 6.1.2 case of missingness predicted by all nine variables and the 6.1.1 case of missingness predicted by the six times one variables. An $s^* R^2$ of 50% with 25% attrition was used. This illustrates the presumably common occurrence of missingness predictors left out of the model to be estimated. The model of (1) and (3) is then misspecified, so that the δ residuals of the s^* relations are not uncorrelated with the predictors.

The test values for the nine/six predictor cases with eight degrees of freedom were 0.03/0.00, 0.02/0.00, and 0.11/0.07 for FQL, LQL, and PPA, while the MCAR test values with nine degrees of freedom were 97.27/91.17. Table 5 gives the bias values. Regarding the SEM metric it should be noted that the η_3 , η_2 structural regression is misspecified due to the left-out variable η_1 . In order not to confound this bias with the bias caused by missing data, the population parameter values are here changed to those of the misspecified regression.

FIGURE 3
A SPECIAL CASE OF MISSINGNESS (II)^a

		VARIABLE								
		y_{11}	y_{21}	y_{31}	y_{12}	y_{22}	y_{32}	y_{13}	y_{23}	y_{33}
MISSING DATA PATTERN	1	1	1	1	1	1	1	1	1	1
	2	1	1	1	1	0	1	1	1	1
	3	1	1	1	1	1	0	1	1	1
	4	1	1	1	1	0	0	1	1	1

^a1's DENOTE NOT MISSING

0's DENOTE MISSING

Table 4
Missingness Case I:
Missingness Predicted by Latent Variables

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Structural Parameters - Confirmatory Factor Analysis Model							
Unstandardized Solution							
$V(\eta_1)$	1.000	1.000 1.000	0 0	0.810 0.789	-19 -21	1.000 1.000	0 0
$V(\eta_2)$	1.000	1.000 1.000	0 0	0.843 0.855	-16 -15	1.000 1.000	0 0
$V(\eta_3)$	1.000	0.916 0.976	- 8 - 2	0.853 0.905	-15 -10	0.853 0.905	-15 -10
$C(\eta_2, \eta_1)$	-0.566	-0.566 -0.566	0 0	-0.393 -0.391	-31 -31	-0.566 -0.566	0 0
$C(\eta_3, \eta_1)$	-0.526	-0.466 -0.504	-11 - 4	-0.358 -0.383	-32 -27	-0.358 -0.382	-32 -27
$C(\eta_3, \eta_2)$	0.691	0.637 0.673	- 8 - 3	0.538 0.572	-22 -17	0.538 0.572	-22 -17
Standardized Solution							
$\rho(\eta_2, \eta_1)$	-0.566	-0.566 -0.566	0 0	-0.476 -0.476	-16 -16	-0.566 -0.566	0 0
$\rho(\eta_3, \eta_1)$	-0.525	-0.448 -0.511	-15 - 3	-0.430 -0.453	-18 -14	-0.387 -0.402	-26 -23
$\rho(\eta_3, \eta_2)$	0.690	0.666 0.682	- 3 - 1	0.634 0.651	- 8 - 6	0.582 0.601	-16 -13
Structural Parameters - Structural Equation Model							
Unstandardized Solution							
β_{21}	-0.566	-0.566 -0.566	0 0	-0.485 -0.495	-14 -13	-0.566 -0.566	0 0
β_{31}	-0.198	-0.156 -0.182	-21 - 8	-0.170 -0.198	-14 0	-0.079 -0.086	-60 -57
β_{32}	0.579	0.549 0.571	- 5 - 1	0.559 0.579	- 3 0	0.493 0.523	-15 -10
$V(\eta_1)$	1.000	1.000 1.000	0 0	0.810 0.789	-19 -21	1.000 1.000	0 0
$V(\zeta_2)$	0.680	0.680 0.680	0 0	0.652 0.661	- 4 - 3	0.680 0.680	0 0
$V(\zeta_3)$	0.498	0.493 0.500	- 1 0	0.492 0.498	- 1 0	0.560 0.573	0 15
Standardized Solution							
β_{21}	-0.566	-0.566 -0.566	0 0	-0.476 -0.476	-16 -16	-0.566 -0.566	0 0
β_{31}	-0.198	-0.163 -0.184	-18 - 7	-0.166 -0.185	-16 - 7	-0.085 -0.091	-57 -54
β_{32}	0.579	0.574 0.578	- 1 0	0.555 0.563	- 4 - 3	0.534 0.550	- 8 - 5
$R^2(\eta_2)$	0.320	0.320 0.320	0 0	0.226 0.226	-29 -29	0.320 0.320	0 0
$R^2(\eta_3)$	0.503	0.461 0.488	- 8 - 3	0.424 0.450	-16 -11	0.344 0.344	-32 -27

The two entries for each parameter estimate correspond to missingness predicted by all three latent variables, and by η_1 and η_2 . Attrition rate is 25% and R^2 is 50%.

Table 5
Missingness Case I:
Marginal Model

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Measurement Parameters							
λ_{12}	0.832	0.832 ^a 0.832	- -	0.832 ^a 0.832	- -	0.832 ^a 0.832	- -
λ_{22}	0.771	0.770 0.770	0 0	0.766 0.766	- 1 - 1	0.768 0.768	0 0
λ_{32}	0.543	0.541 0.542	0 0	0.527 0.526	- 3 - 3	0.540 0.540	- 1 - 1
λ_{13}	0.759	0.759 ^a 0.759	- -	0.759 ^a 0.759	- -	0.759 ^a 0.759	- -
λ_{23}	0.672	0.664 0.672	- 1 0	0.664 0.672	- 1 0	0.664 0.672	- 1 0
λ_{33}	0.497	0.480 0.497	- 3 - 0	0.481 0.497	- 3 0	0.481 0.497	- 3 0
θ_{12}	0.308	0.306 0.307	- 1 - 0	0.306 0.307	- 1 0	0.304 0.304	- 1 - 1
θ_{22}	0.405	0.406 0.406	0 0	0.406 0.406	0 0	0.407 0.407	0 0
θ_{32}	0.705	0.706 0.706	0 0	0.705 0.704	0 0	0.707 0.707	0 0
θ_{13}	0.423	0.420 0.423	- 1 - 0	0.420 0.423	- 1 0	0.420 0.423	- 1 0
θ_{23}	0.547	0.548 0.547	0 0	0.548 0.547	0 0	0.548 0.547	0 0
θ_{33}	0.752	0.753 0.752	0 0	0.753 0.752	0 0	0.753 0.752	0 0
Structural Parameters - Confirmatory Factor Analysis Model							
Unstandardized Solution							
$V(\eta_2)$	1.000	1.003 1.001	0 0	0.851 0.861	-15 -14	1.006 1.005	1 0
$V(\eta_3)$	1.000	0.897 0.964	-10 - 4	0.858 0.921	-14 - 8	0.857 0.921	-14 - 8
$C(\eta_3, \eta_2)$	0.691	0.616 0.661	-11 - 4	0.539 0.584	-22 -15	0.537 0.582	-22 -16
Standardized Solution							
$\rho(\eta_3, \eta_2)$	0.691	0.749 0.776	9 13	0.631 0.656	- 9 - 5	0.579 0.605	-16 -12
Structural Parameters - Structural Equation Model							
Unstandardized Solution							
β_{32}	0.691	0.614 0.660	-11 - 5	0.634 0.679	- 8 - 2	0.534 0.579	-23 -16
$V(\eta_2)$	1.000	1.003 1.001	0 0	0.851 0.861	-15 -14	1.006 1.005	1 0
$V(\zeta_3)$	0.525	0.519 0.527	- 1 0	0.516 0.524	- 2 0	0.570 0.583	9 11
Standardized Solution							
β_{32}	0.691	0.749 0.776	9 13	0.631 0.656	- 9 - 5	0.579 0.605	-16 -12
$R^2(\eta_3)$	0.477	0.421 0.453	-12 - 5	0.398 0.431	-17 -10	0.335 0.366	-30 -23

^a Fixed parameter.

The two entries for each parameter estimate correspond to missingness predicted by all nine variables and by the six time one observed variables. In both cases, attrition rate is 25% and R^2 is 50%.

We may compare the top entries of Table 5 with the middle, 50% R^2 , entries of Table 3, and the bottom entries with those of Table 2. It is noteworthy that the bias increase in FQL is not that large when going from the full to the marginal model. In the two cases there is a maximum bias increase of 3% in the CFA structural parameters, and 7% in the SEM structural parameters. The bias increase for LQL and PPA is somewhat smaller. Overall, however, it is still true that FQL performs best. PPA is slightly better than LQL in the CFA metric, while the reverse is true in the SEM metric.

6.2 Results for Missingness Case II

In this case missingness occurs for y_{22} and y_{32} . This may, for instance, be thought of as a cross-sectional situation, where these variables involve particularly sensitive aspects of η_2 . The two variables are taken to have different s^* missingness relations, so that we have four possible data patterns, as depicted in Figure 3. We note that these patterns are not monotone (nested), see e.g., Little (1983, p. 357), as are the data patterns under missingness case I. For each of the variables we assume that the y^* value that would have been observed influences the s^* missingness propensity, so that the missingness mechanism is not ignorable for this reason alone. For y_{22} we assume the s_{22}^* relation

$$s_{22}^* = a\gamma_{22} \eta_1 - b\gamma_{22} y_{22}^* + \delta_{22}, \tag{37}$$

while for y_{32} we assume

$$s_{32}^* = a\gamma_{32} \eta_1 - b\gamma_{32} y_{32}^* + \delta_{32}. \tag{38}$$

We will study both $(a, b) = (2/3, 1/3)$ and $(a, b) = (1/3, 2/3)$, varying the relative importance of background and the specific variable. In each s^* relation an R^2 of 50% was chosen by setting $\gamma_{22} = 0.817$ and $\gamma_{32} = 0.850$. The univariate rates of missingness were set at 15% for both variables. A medium sized s_{22}^*, s_{32}^* correlation of 0.5 was chosen by allowing δ_{22} and δ_{32} to correlate 0.044 and 0.178, respectively for the two (a, b) choices. This results in the expectation of about 76% of the sample having no missingness, that is, a number comparable to that of missingness case I. The exact probabilities of each of the data patterns as numbered in Figure 3 are: 0.758, 0.092, 0.092, 0.058.

6.2.1 The Full Model. For $(a, b) = (2/3, 1/3)$ the full model test values with 24 degrees of freedom for FQL, LQL, and PPA were 0.95, 0.39, 10.56, while for $(a, b) = (1/3, 2/3)$ they were 1.04, 0.45, 11.37. In these two cases the MCAR test values were 189.47 and 126.69 with 123 degrees of freedom. The entries of Table 6 give results for parameter estimates in the same order. Compared to missingness case I, the bias in the measurement parameters is somewhat larger here. Overall, FQL is clearly the best estimator. Regarding the LQL and PPA comparison, LQL performs slightly better overall for measurement parameters, while PPA is clearly better in terms of CFA structural parameters, and in this case also slightly better overall in terms of SEM structural parameters. Again, PPA and FQL share the property of giving almost zero bias for parameters in parts of the model where there is no missingness, but in other parts FQL dominates PPA. There is little influence in the results by the change of (a, b) values. We furthermore replaced the η_1 variable as a predictor of missingness by its three indicators, but this also had little effect on the overall picture.

6.2.2 The Marginal Model. The marginal model for $y_{12}, y_{22}, y_{32}, y_{13}, y_{23}$, and y_{33} was estimated under both of the missingness conditions above. The results will be given in the order $(a, b) = (2/3, 1/3), (a, b) = (1/3, 2/3)$. The model test values were 0.06 and 0.13 for FQL, 0.02 and 0.04 for LQL, and 3.72 and 6.32 for PPA. The degrees of freedom

Table 6
Missingness Case II:
Full Model

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Measurement Parameters							
λ_{11}	0.778	0.778 ^a 0.778	- -	0.778 ^a 0.778	- -	0.778 ^a 0.778	- -
λ_{21}	0.846	0.846 0.846	0 0	0.846 0.846	0 0	0.846 0.846	0 0
λ_{31}	0.585	0.585 0.585	0 0	0.585 0.585	0 0	0.585 0.585	0 0
λ_{12}	0.832	0.832 ^a 0.832	- -	0.832 ^a 0.832	- -	0.832 ^a 0.832	- -
λ_{22}	0.771	0.721 0.687	- 6 -11	0.744 0.722	- 4 - 6	0.667 0.657	-13 -15
λ_{32}	0.543	0.489 0.454	-10 -16	0.507 0.482	- 7 -11	0.471 0.467	-13 -14
λ_{13}	0.759	0.759 ^a 0.759	- -	0.759 ^a 0.759	- -	0.759 ^a 0.759	- -
λ_{23}	0.672	0.672 0.672	0 0	0.672 0.672	0 0	0.672 0.672	0 0
λ_{33}	0.497	0.497 0.497	0 0	0.497 0.497	0 0	0.497 0.497	0 0
θ_{11}	0.395	0.395 0.395	0 0	0.395 0.395	0 0	0.395 0.395	0 0
θ_{21}	0.284	0.284 0.284	0 0	0.284 0.284	0 0	0.284 0.284	0 0
θ_{31}	0.658	0.658 0.658	0 0	0.658 0.658	0 0	0.658 0.658	0 0
θ_{12}	0.308	0.290 0.292	- 6 - 5	0.299 0.300	- 3 - 3	0.292 0.323	- 5 5
θ_{22}	0.405	0.412 0.396	2 - 2	0.406 0.391	0 - 3	0.452 0.420	12 - 4
θ_{32}	0.705	0.695 0.646	- 1 - 8	0.694 0.650	- 2 - 8	0.699 0.634	- 1 -10
θ_{13}	0.423	0.423 0.423	0 0	0.423 0.423	0 0	0.423 0.423	0 0
θ_{23}	0.547	0.547 0.547	0 0	0.547 0.547	0 0	0.547 0.547	0 0
θ_{33}	0.752	0.752 0.752	0 0	0.752 0.752	0 0	0.752 0.752	0 0

Table 6
(continued)

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Structural Parameters - Confirmatory Factor Analysis Model							
$V(\eta_1)$	1.000	1.000 1.000	0 0	0.764 0.872	-24 -13	1.000 1.000	0 0
$V(\eta_2)$	1.000	1.025 1.022	3 2	0.878 0.861	-12 -14	1.023 0.978	- 2 - 2
$V(\eta_3)$	1.000	1.002 1.002	0 0	0.907 0.915	- 9 - 9	1.002 1.002	0 0
$C(\eta_2, \eta_1)$	0.566	-0.555 -0.555	- 2 - 2	-0.498 -0.480	-12 -15	-0.529 -0.527	- 7 - 7
$C(\eta_3, \eta_1)$	-0.526	-0.525 -0.525	0 0	-0.492 -0.482	- 6 - 8	-0.526 0.525	0 0
$C(\eta_3, \eta_2)$	0.691	0.680 0.680	- 2 - 2	0.673 0.669	- 3 - 3	0.673 0.674	- 3 - 2
Standardized Solution							
$\rho(\eta_2, \eta_1)$	-0.566	-0.549 -0.549	- 3 - 3	-0.464 -0.484	-18 -14	-0.523 -0.527	- 8 - 7
$\rho(\eta_3, \eta_1)$	-0.526	-0.525 -0.525	0 0	-0.452 -0.471	-14 -10	-0.525 -0.526	0 0
$\rho(\eta_3, \eta_2)$	0.691	0.684 0.684	- 1 - 1	0.649 0.650	- 6 - 6	0.677 0.687	- 2 - 1
Structural Parameters - Structural Equation Model							
Unstandardized Solution							
β_{21}	-0.566	-0.555 -0.555	- 2 - 2	-0.498 -0.480	-12 -15	-0.529 -0.531	- 7 - 6
β_{31}	-0.198	-0.214 -0.214	8 8	-0.209 -0.209	6 6	-0.236 -0.226	19 14
β_{32}	0.579	0.561 0.561	- 3 - 3	0.569 0.569	- 2 - 2	0.548 0.564	- 5 - 3
$V(\eta_1)$	1.000	1.000 1.000	0 0	0.764 0.872	-24 -13	1.000 1.000	0 0
$V(\zeta_2)$	0.680	0.717 0.714	5 5	0.689 0.660	1 - 3	0.744 0.700	9 3
$V(\zeta_3)$	0.498	0.501 0.501	1 1	0.499 0.499	0 0	0.503 0.501	1 1
Standardized Solution							
β_{21}	-0.566	-0.549 -0.549	- 3 - 3	-0.464 -0.484	-18 -14	-0.523 -0.527	- 8 - 7
β_{31}	-0.198	-0.214 -0.214	8 8	-0.192 -0.204	- 3 3	-0.236 -0.227	19 15
β_{32}	0.579	0.567 0.567	- 2 - 2	0.560 0.552	- 3 - 5	0.554 0.567	- 4 - 2
$R^2(\eta_2)$	0.320	0.301 0.302	- 6 - 6	0.215 0.234	-32 -27	0.273 0.284	-15 -11
$R^2(\eta_3)$	0.503	0.500 0.500	0 0	0.450 0.455	-10 - 9	0.499 0.500	- 1 0

^a Fixed parameter.

The two entries for each parameter estimate correspond to the (a,b) choice ($2/3$, $1/3$) and ($1/3$, $2/3$); see text. Both variables have a missingness rate of 15% with an R^2 of 50%.

Table 7
Missingness Case II:
Marginal Model

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Measurement Parameters							
λ_{22}	0.832	0.832 ^a 0.832	- -	0.832 ^a 0.832	- -	0.832 ^a 0.832	- -
λ_{22}	0.771	0.726 0.695	- 6 -10	0.751 0.730	- 3 - 5	0.706 0.695	- 8 -10
λ_{32}	0.543	0.494 0.462	- 9 -15	0.513 0.488	- 6 -10	0.498 0.494	- 8 - 9
λ_{13}	0.759	0.759 ^a 0.759	- -	0.759 ^a 0.759	- -	0.759 ^a 0.759	- -
λ_{23}	0.672	0.672 0.672	0 0	0.672 0.672	0 0	0.672 0.672	0 0
λ_{33}	0.497	0.497 0.497	0 0	0.497 0.497	0 0	0.497 0.497	0 0
θ_{12}	0.308	0.303 0.307	- 2 0	0.306 0.308	- 1 0	0.334 0.362	8 18
θ_{22}	0.405	0.406 0.389	0 - 4	0.402 0.386	- 1 - 5	0.426 0.397	5 - 2
θ_{32}	0.705	0.691 0.642	- 2 - 9	0.691 0.647	- 2 - 8	0.686 0.621	- 3 -12
θ_{13}	0.423	0.423 0.423	0 0	0.423 0.423	0 0	0.423 0.423	0 0
θ_{23}	0.547	0.547 0.547	0 0	0.547 0.547	0 0	0.547 0.547	0 0
θ_{33}	0.752	0.752 0.752	0 0	0.752 0.752	0 0	0.752 0.752	0 0
Structural Parameters - Confirmatory Factor Analysis Model							
Unstandardized Solution							
$V(\eta_2)$	1.000	1.006 1.001	1 0	0.869 0.850	-13 -15	0.963 0.921	- 4 - 8
$V(\eta_3)$	1.000	1.002 1.002	0 0	0.907 0.907	- 9 - 9	1.002 1.002	0 0
$C(\eta_3, \eta_2)$	0.691	0.685 0.687	1 1	0.664 0.651	- 4 - 1	0.694 0.710	0 3
Standardized Solution							
$\rho(\eta_3, \eta_2)$	0.691	0.686 0.687	- 1 - 1	0.650 0.676	- 6 - 2	0.680 0.681	- 2 - 1
Structural Parameters - Structural Equation Model							
Unstandardized Solution							
β_{32}	0.691	0.685 0.687	- 1 - 1	0.664 0.676	- 4 - 2	0.694 0.710	0 3
$V(\eta_2)$	1.000	1.006 1.001	1 0	0.869 0.850	-13 -15	0.963 0.921	- 4 - 8
$V(\zeta_3)$	0.522	0.531 0.530	2 2	0.524 0.527	0 1	0.539 0.538	3 3
Standardized Solution							
β_{32}	0.691	0.686 0.687	- 1 - 1	0.650 0.651	- 6 - 6	0.680 0.681	- 2 - 1
$R^2(\eta_3)$	0.477	0.471 0.470	- 1 - 2	0.422 0.424	-12 -11	0.462 0.464	- 3 - 3

^a Fixed parameter.

The two entries for each parameter estimate correspond to the (a,b) choice ($^2/3$, $^1/3$) and ($^1/3$, $^2/3$); see text. Both variables have a missingness rate of 15% with an R^2 of 50%.

is eight. The MCAR test values were 90.04 and 95.13 with 54 degrees of freedom. Table 7 gives the bias values.

The general conclusion is that for all three estimators the bias increases rather little when going from the full to the marginal model. FQL still performs clearly best overall. PPA outperforms LQL in the CFA metric, while in this case the reverse is true in the SEM metric.

As a final analysis, let each of the two s^* 's instead be a function of η_2 only, with an R^2 of 50% ($\gamma^2 = 0.707$) and the same amount of missingness as earlier. This situation illustrates the important case where missingness is predicted by a factor and there is missingness for some of the factor's indicators. The FQL, LQL, and PPA test values with eight degrees of freedom are 0.00, 0.00, and 0.29, respectively, while the MCAR test value with 54 degrees of freedom is 62.05. The estimates are given in Table 8. In accordance with the theory of section 4.2, LQL gives perfect model fit with no measurement parameter bias. The outcome for the FQL estimator can not be predicted by the theory of that section since the assumed partitioning of (26) and (27) does not hold for missingness case II. It is, therefore, interesting to note that FQL has almost no bias, not even in the structural part.

7. Conclusions

In the artificial data studied the traditional estimation methods have been shown inferior to a likelihood estimator that uses all available information in terms of large-sample bias. We studied situations involving data that were not missing completely at random. The FQL likelihood estimator was superior even in situations that did not fulfill the prerequisites for it to be maximum likelihood. Since it is often possible to carry out such estimation using existing structural equation modeling packages, this finding has important practical implications.

The usual caution about limited generalizability from any artificial data study should, however, be kept in mind. The generalizability of the results may not only be limited by the latent variable models and the missingness situations studied, but also the particular specification of the missingness mechanism. From our experience, however, the models and mechanisms chosen seem to be common and plausible representations in latent variable contexts.

The approach of using a general structural equation modeling software package, such as LISREL, opens up the possibility for a variety of analyses with missing data. For instance, not only is confirmatory factor analysis possible, but also exploratory factor analysis (by imposing the appropriate number of restrictions), where the rotations would have to be done in a separate program. A multitude of common mean vector and covariance structure problems can also be handled, such as variance component estimation, testing of equality of covariance matrices, and so on (see e.g., Joreskog & Sorbom, 1984). A limitation to this computational approach is, however, the fact that in real data there are usually a very large number of missing data patterns, where the number of observations in each group frequently falls below the number of variables. In practice it may then be necessary to delete some parts of the data, much like the approach of listwise deletion, so that a feasible number of groups with large sample sizes remain (see also Marini, et al., 1980). Unless this loss of data introduces strongly selective missingness that is not predictable by observed variables for which there are no missingness, not much bias may result. In any case, using information from a few major missing data pattern groups will most likely yield better results than using only the complete data group. A possible alternative is to first estimate the mean vector and covariance matrix by maximum

Table 8

Missingness Case II:

Missingness Predicted by η_2 , Marginal Model

Parameter	Parameter Value	FQL		LQL		PPA	
		Estimate	% Bias	Estimate	% Bias	Estimate	% Bias
Measurement Parameters							
λ_{12}	0.832	0.832 ^a	-	0.832 ^a	-	0.832 ^a	-
λ_{22}	0.771	0.760	- 1	0.771	0	0.725	- 6
λ_{32}	0.543	0.536	- 1	0.543	0	0.514	- 5
λ_{13}	0.759	0.759 ^a	-	0.759 ^a	-	0.759 ^a	-
λ_{23}	0.672	0.672	0	0.672	0	0.672	0
λ_{33}	0.497	0.497	0	0.497	0	0.497	0
θ_{12}	0.308	0.306	- 1	0.307	0	0.305	- 1
θ_{22}	0.405	0.408	1	0.405	0	0.438	8
θ_{32}	0.705	0.706	0	0.705	0	0.718	2
θ_{13}	0.423	0.423	0	0.423	0	0.423	0
θ_{23}	0.547	0.547	0	0.547	0	0.547	0
θ_{33}	0.752	0.752	0	0.752	0	0.752	0
Structural Parameters - Confirmatory Factor Analysis Model							
Unstandardized Solution							
$V(\eta_2)$	1.000	1.003	0	0.914	- 9	1.003	0
$V(\eta_3)$	1.000	1.002	0	0.927	- 7	1.002	0
$C(\eta_3, \eta_2)$	0.691	0.689	0	0.668	- 3	0.683	- 1
Standardized Solution							
$\rho(\eta_3, \eta_2)$	0.691	0.689	0	0.663	- 4	0.683	- 1
Structural Parameters - Structural Equation Model							
Unstandardized Solution							
β_{32}	0.691	0.689	0	0.668	- 3	0.683	- 1
$V(\eta_2)$	1.000	1.003	0	0.914	- 9	1.003	0
$V(\zeta_3)$	0.522	0.527	1	0.519	- 1	0.534	2
Standardized Solution							
β_{32}	0.691	0.689	0	0.663	- 4	0.683	- 1
$R^2(\eta_3)$	0.477	0.475	0	0.440	- 8	0.467	- 2

^a Fixed parameter.

Both variables have a missingness rate of 15% with an R^2 of 50%.

likelihood, allowing any number of patterns. This may be carried out by BMDP's PAM program, Dixon (1983). A more optimal approach is, however, to carry out FQL estimation by a procedure that allows for such general patterns, as was done in Finkbeiner (1979).

It should be noted that missing data theory extends beyond the situation of continuous variables and multivariate normal distributions to, for instance, multinomial distributions (see e.g., Rubin, 1974; Little, 1983). Hence, the demonstrated strength of the FQL estimation approach may also be of interest in psychometric methods for categorical data.

References

- Anderson, T. W. (1957). Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *Journal of the American Statistical Association*, *52*, 200–203.
- Beale, E. L., & Little, R. J. A. (1975). Missing values in multivariate analysis. *Journal of the Royal Statistical Society, Series B*, *37*, 129–146.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, *46*, 443–459.
- Boomsma, A. (1983). *On the robustness of the LISREL (maximum likelihood estimation) against small sample size and nonnormality*. Unpublished doctoral, dissertation, University of Groningen, Groningen, The Netherlands.
- Brown, C. H. (1983). Asymptotic comparison of missing data procedures for estimating factor loadings. *Psychometrika*, *48*, 269–291.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, *39*, 1–38.
- Dixon, W. J. (1983). BMDP Statistical Software. Berkeley: University of California Press.
- Finkbeiner, C. (1979). Estimation for the multiple factor model when data are missing. *Psychometrika*, *44*, 409–420.
- Hartley, H. O., & Hocking, R. R. (1971). The analysis of incomplete data. *Biometrics*, *14*, 174–194.
- Hausman, J. A., & Wise, D. A. (1979). Attrition bias in experimental and panel data: The Gary income maintenance experiment. *Econometrica*, *47*, 455–474.
- Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, *5*, 475–492.
- Johnson, N. L., & Kotz, S. (1972). *Distributions in statistics: Continuous multivariate distributions*. New York: John Wiley & Sons.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*, 183–202.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*, 409–426.
- Jöreskog, 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.
- Jöreskog, K. G., & Sörbom, D. (1980). Simultaneous analysis of longitudinal data from several cohorts. Research Report 80-5, Department of Statistics, University of Uppsala, Sweden.
- Jöreskog, K. G., & Sörbom, D. (1984). LISREL VI; Analysis of linear structural relationships by maximum likelihood and least squares methods. Scientific Software.
- Lawley, D. N. (1943–1944). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society Edinburgh, Section A (Mathematics and Physics Section)*, *62*(1), 28–30.
- Little, R. J. A. (1982). Models for nonresponse in sample surveys. *Journal of the American Statistical Association*, *77*, 237–250.
- Little, R. J. A. (1983). The ignorable case. W. G. Modon, I. Olkin, & D. R. Rubin (Eds.), In *Incomplete data in sample surveys, Vol. 2: Theory and bibliographies*. New York: Academic Press.
- Little, R. J. A. (1985). A note about models for selectivity bias. *Econometrica*, *53*(6), 1469–1474.
- Little, R. J. A., & Rubin, D. R. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons.
- Marini, M. M., Olsen, A. R., & Rubin, D. B. (1980). Maximum likelihood estimation in panel studies with missing data. In K. F. Schuessler (Ed.), *Sociological Methodology*. San Francisco: Jossey Bass.
- Meredith, W. (1964). Notes on factorial invariance. *Psychometrika*, *29*, 177–185.
- Muthén, B., 1984. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, *49*, 115–132.
- Muthén, B. (1985). Moments of the censored and truncated bivariate normal distribution. Submitted for publication.

- Muthén, B. (1987). *LISCOMP. Analysis of linear structural equations using a comprehensive measurement model. User's guide*. Scientific Software.
- Muthén, B., & Jöreskog, K. (1983). Selectivity problems in quasi-experimental studies. *Evaluation Review*, 7, 139-173.
- Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, 33, 171-189.
- Olsson, U. (1978). Selection bias in confirmatory factor analysis (Department of Statistics Research, Report No. 78-4). Uppsala, Sweden: University of Uppsala.
- Pearson, K. (1912). On the general theory of the influence of selection on correlation and variation. *Biometrika*, 8, 437-443.
- Rosenbaum, S. (1961). Moments of a truncated bivariate normal distribution. *Journal of the Royal Statistical Society, Series B*, 23, 405-408.
- Rubin, D. B. (1974). Characterizing the estimation of parameters in incomplete data problems. *Journal of the American Statistical Association*, 69, 456-474.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581-592.
- Tallis, G. M. (1961). The moment generating function of the truncated multi-normal distribution. *Journal of the Royal Statistical Society, Series B*, 23, 223-229.
- Trawinski, I. M., & Bargmann, R. E. (1964). Maximum likelihood estimation with incomplete multivariate data. *Annals of Mathematical Statistics*, 35, 647-657.
- Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and the Gauss-Newton method. *Biometrika*, 61, 439-447.
- Werts, C. E., Rock, D. A., & Grandy, J. (1979). Confirmatory factor analysis applications: Missing data problems and comparison of path models between populations. *Multivariate Behavioral Research*, 14, 199-213.
- Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology*, San Francisco: Jossey Bass.

Manuscript received 12/16/85

Final version received 10/6/86