Stepwise latent variable modeling: An overview of approaches

Jeroen K. Vermunt · ORCID: 0000-0001-8778-3762

[1] Department of Methodology and Statistics, School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands

---

Address for correspondence: Department of Methodology and Statistics, School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands.

E-mail: `j.k.vermunt@uvt.nl`.

Phone: (+31) 134662748.

Abstract: Stepwise approaches for the estimation of latent variable models are becoming increasingly popular, both in the context of models for continuous (factor analysis and latent trait models) and discrete (latent class and latent profile models) latent variables. Examples include two-stage path analysis, structural-after-measurement and Croon's bias-corrected estimation of structural equation models, and two- and three-step latent class and latent Markov modeling. These methods have in common that the measurement/clustering part of the model is estimated first, followed by the estimation of a — possibly complex – structural model. In this paper, we review the existing approaches, which differ in how the information on the latent variable(s) is used when estimating the structural model. We show that based on these differences, stepwise latent variable modeling approaches can be classified

into three main types: the fixed parameters, the single indicator, and the bias adjustment approach. We discuss similarities and differences between these approaches, as well as between approaches proposed specifically for either continuous or discrete latent variables. Special attention is paid to heterogeneous measurement error resulting from missing data or measurement non-invariance, standard error estimation, and software implementations.

Key words: Structural equation modeling; latent class analysis; composite scores; measurement error; classification error; reliability; Croon's bias correction

1 Introduction

Recently, we have seen a renewed interest in stepwise latent variable modeling approaches, which involve separating the estimation of the measurement and the structural parts of the model of interest. Examples for continuous latent variables include Structural After Measurement (SAM) estimation (Rosseel and Loh, 2023), two-stage path analyses (Lai and Hsiao, 2022), two-step latent trait modeling (Kuha and Bakk, 2023), and Measurement and Uncertainty Preserving ParamETric (MUPPET) modeling (Levy, 2023; Levy and McNeish, 2024). Similar approaches have been proposed for discrete latent variables, such as three-step latent class analysis (Bolck, Croon, and Hagenaars, 2004; Vermunt, 2010) and two-step latent class analysis (Bakk and Kuha, 2018).

Typically, two types of arguments are mentioned by authors contributing to the field of stepwise latent variable modeling; that is, arguments related to *model estimation*

and to *model building.* Regarding model estimation, separate estimation of the measurement and the structural parameters yields fewer convergence problems, is less affected by model misspecifications, and is computationally less demanding when dealing with large models (Bartolucci et al., 2016; Perez Alonso, 2024; Rosseel and Loh, 2023). In addition, stepwise model building, where one constructs measurements for the variables of interest and subsequently uses these to investigate their relationships, is also what most applied researchers prefer doing, among others to comply with the standard practice in their field, to prevent interpretational confounding, or to circumvent the need to remove cases with missing values on covariates (Burt, 1976; Clouth et al., 2022; Levy, 2023; Vermunt, 2010). Given that most researchers use such a strategy, stepwise latent variable modeling approaches have the potential to yield great improvement over naive use of estimated latent variable scores in subsequent analyses without accounting for their uncertainty (Rein et al., 2024).

As shown in Table 1, stepwise approaches have been developed for each of the four types of latent variables models described by Bartholomew and Knott (1999) and Lazarsfeld and Henry (1968). These are: 1) *factor analytic models* for continuous latent variables and continuous observed indicators, 2) *latent trait models* for continuous latent variables and discrete observed indicators, 3) *latent profile models* for discrete latent variables and continuous observed indicators, and 4) *latent class models* for discrete latent variables and discrete observed indicators. Latent trait models are also referred to as item response theory (IRT) models or categorical factor analytic models, and latent profile and latent class models are also referred to as mixture models.

The structural models for which stepwise approaches have been developed are some-

times simple linear or logistic regression models in which a single latent variable serves as dependent or independent variable, but can also concern more complex models, such as path models containing multiple latent variables, latent Markov models or dynamic factor models for longitudinal data, or latent variable models for multilevel data.

Stepwise latent variable approaches appear in the literature under different names such as Structural After Measurement (SAM) estimation, two-stage analysis, two-step analysis, Croon's bias-corrected estimation, Croon's method, and three-step analysis. However, the first step of all these approaches deals with the estimation of the measurement model for the latent variables without accounting for the structural model. The approaches differ from one another with respect to how the measurement model parameters are used when estimating the parameters of the structural model. Based on these differences, stepwise latent variable modeling approaches can be classified into three main types, which we will refer to as the *fixed parameters*, the *single indicator*, and the *bias adjustment* approach. Table 1 lists the most important contributions to each of these.

In the fixed parameters approach, when estimating the parameters of the structural model, the measurement model parameters are fixed to the values obtained in the first step. Examples of this approach include Burt's approach (Burt, 1976), global SAM estimation (Rosseel and Loh, 2022), two-step latent trait analysis (Kuha and Bakk, 2023), MUPPET modeling (Levy, 2023; Levy and McNeish, 2024), two-step latent class analysis (Bakk and Kuha, 2018), and three-step latent Markov modeling (Bartolucci et al., 2015).

The single indicator approach involves obtaining predictions for the latent variables

using the parameters of the measurement model. The predicted scores are used as single indicators when estimating the structural model, while accounting for their unreliability or misclassifications. This is a rather old idea within the SEM framework, and referred to as two-step SEM (Bollen, 1989). Lai and Hsiao (2021) showed how to use this approach when the measurement model is a latent trait model instead of factor analytic model, and refer to this approach as two-stage path analysis. Moreover, Lai et al. (2023) showed how to apply two-stage path analysis when there is measurement invariance. In the context of latent class and latent profile analysis, this approach was proposed by Vermunt (2010) and Gudicha and Vermunt (2012), who refer to it as three-step analysis with maximum likelihood adjustment for classification errors (see also Asparouhov and Muthén, 2014). Somewhat related single indicator approaches have been proposed by Savalei (2019) and Skrondal and Kuha (2012).

The bias adjustment approach involves obtaining predictions for the latent variables using the parameters of the measurement model, and subsequently adjusting their covariances (or associations) so that these represent the true latent variable covariances. The adjusted covariances are used as if they were observed covariances when estimating the structural model parameters. This method was originally developed by Croon (2002) who derived the required adjustment for both factor analytic and latent class models. Devlieger and Rosseel (2017) used this approach in what they refer to as factor score path analysis. Although derived and implemented in a slightly different manner, the recently proposed local SAM approach by Rosseel and Loh (2023) can be seen as a special case of Croon's method for continuous latent variables. Croon's method for discrete latent variables was extended in various ways by Bolck et al. (2004), Vermunt (2010), Gudicha and Vermunt (2011), and Bakk, Tekle and Vermunt (2013), and is typically referred to as the Bolck-Croon-Hagenaars

(BCH) approach.

The next three sections provide more details on the implementation of the three types of stepwise estimation methods with factor analytic models, latent trait and categorical factor analytic models, and latent class, latent profile, and mixture models. Subsequently, we discuss topics such as heterogeneous measurement error resulting from missing data or measurement non-invariance, standard error estimation, and software implementations. We conclude with a short discussion.

2   Continuous latent and observed variables

2.1   Background

Assuming the data are centered, and therefore ignoring the mean structure, the regression equations defining the measurement and structural parts of a structural equation model (SEM) for response vector $\mathbf{y}$ are:

$$\begin{aligned} \mathbf{y} &= \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\epsilon}_i \\ \boldsymbol{\eta} &= \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta}_i. \end{aligned} \tag{2.1}$$

The free parameters are the factor loadings, residual covariances, regression coefficients, and residual factor covariances, which are collected in the matrices $\boldsymbol{\Lambda}$, $\boldsymbol{\Theta}$, $\mathbf{B}$ and $\boldsymbol{\Psi}$. These yield the expected covariance matrix $\boldsymbol{\Sigma}$ as follows:

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda} Var(\boldsymbol{\eta})\boldsymbol{\Lambda}' + \boldsymbol{\Theta},$$

with

$$Var(\boldsymbol{\eta}) = (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B}')^{-1}.$$

The various stepwise approaches have in common that they first estimate the measurement parameters $\mathbf{\Lambda}$ and $\mathbf{\Theta}$, while the structural parameters $\mathbf{B}$ and $\mathbf{\Psi}$ are estimated in a next step of the analysis. The approaches differ in how the information on the measurement parameters is used when estimating the structural parameters. Two of the approaches make use of factor scores, either as single indicators or as an intermediate step, to obtain an adjusted factor-score covariance matrix.

Let $\mathbf{f}$ be the factor scores which are obtained as $\mathbf{f} = \mathbf{A}\mathbf{y}$, with $\mathbf{A}$ being the factor-score matrix. The most common type of factor score is the regression factor score, with factor-score matrix

$$\mathbf{A}_R = \mathbf{\Phi}\mathbf{\Lambda}'\mathbf{\Sigma}^{-1}, \tag{2.2}$$

or, equivalently,

$$\mathbf{A}_R = (\mathbf{\Phi}^{-1} + \mathbf{\Lambda}'\mathbf{\Theta}^{-1}\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{\Theta}^{-1}.$$

Here $\mathbf{\Phi}$ denotes the factor covariance matrix obtained from the measurement model. This yields modal a posteriori (MAP) or expected a posteriori (EAP) estimates of the factor scores. The posterior variance of the regression factor scores equals:

$$Var_{\mathbf{f}} = \mathbf{\Phi} - \mathbf{\Phi}'\mathbf{\Lambda}'\mathbf{\Sigma}^{-1}\mathbf{\Lambda}\mathbf{\Phi}. \tag{2.3}$$

Maximum likelihood or Barlett factor scores are obtained with factor-score matrix

$$\mathbf{A}_B = (\mathbf{\Lambda}'\mathbf{\Theta}^{-1}\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{\Theta}^{-1}. \tag{2.4}$$

Instead of using a factor-score matrix based on the measurement model parameters, one may also use a simple sum score of the items as an estimate of the latent variable, in which case the factor-score matrix equals $\mathbf{A}_S = \mathbf{1}$.

2.2   Fixed parameters

To prevent what he refers to as interpretational confounding, Burt (1976) proposed estimating the measurement model separately for each latent factor, and subsequently estimating the structural model of interest with the measurement parameters – loadings and residual covariances – fixed. A similar but slightly more general fixed parameters approach was recently proposed by Rosseel and Loh (2023), who refer to it as global SAM estimation. An important difference is that global SAM was presented as an estimation method, whereas Burt presented his approach as a modeling method. In addition, the global SAM approach is more flexible in the sense that the measurement model may be estimated jointly for all latent variables, separately for each latent variable, or separately for blocks of latent variables. Moreover, the estimator used for the measurement model can differ from the one used for the structural model.

Another fixed measurement parameters approach called MUPPET modeling was proposed by Levy (2023), who also stressed the issue of interpretational confounding. MUPPET is a Bayesian two-step estimation approach for factor analytic models with covariates and outcome variables. The Bayesian estimation framework allows accounting for the uncertainty in the measurement parameters when estimating the structural parameters.

2.3   Single indicator

The single indicator approach has been advocated among others by Lai and Hsiao (2022) who refer to it as two-stage path modeling. Rein et al. (2024) proposed using this approach in the context of latent vector autoregressive models for intensive

longitudinal data subject to measurement error.

The single-indicator approach assumes that factor scores $\mathbf{f}$ are computed after estimating the measurement model, where $\mathbf{A}$ can be any type of factor-score matrix. When estimating the structural parameters, the factor scores are used as indicators for the latent variables $\boldsymbol{\eta}$. Using $\mathbf{f} = \mathbf{Ay}$ and substituting $\mathbf{y}$ by the measurement part of the SEM defined in equation 2.1 yields:

$$\mathbf{f} = \mathbf{A\Lambda\eta} + \mathbf{A\epsilon} = \mathbf{\Lambda}^{*}\boldsymbol{\eta} + \boldsymbol{\epsilon}^{*}, \tag{2.5}$$

which shows how the factor scores are related to the true latent variable scores. Moreover, this equation demonstrates factor scores $\mathbf{f}$ can serve as single indicators of the latent variables $\boldsymbol{\eta}$ by using $\mathbf{\Lambda}^{*} = \mathbf{A\Lambda}$ as their "loading" matrix" and $\mathbf{A\Theta A}'$ as the covariance matrix $\mathbf{\Theta}^{*}$ of their residuals $\boldsymbol{\epsilon}^{*} = \mathbf{A\epsilon}$. In other words, the structural parameters can be estimated using a SEM with a single indicator per factor, where loadings and residual variances depend on the parameters of the measurement model and the type of factor scores used.

The measurement model for the factor scores can also be written in terms of factor covariances, which yields

$$Var(\mathbf{f}) = \mathbf{A\Lambda}Var(\boldsymbol{\eta})\mathbf{\Lambda}'\mathbf{A}' + \mathbf{A\Theta A}'. \tag{2.6}$$

The reliability of the factor scores equals:

$$\boldsymbol{\rho} \;=\; \operatorname{diag}\left(\frac{\mathbf{A\Lambda}Var(\boldsymbol{\eta})\mathbf{\Lambda}'\mathbf{A}'}{Var(\mathbf{f})}\right) \tag{2.7}$$

that is, the ratio between the (appropriately weighed) true and the estimated factor variance.

A useful feature of regression factor scores obtained from the measurement model

for a single factor (or by ignoring factor correlations) is that their fixed loading and residual variance $\mathbf{a}'_R \boldsymbol{\lambda}$ and $\mathbf{a}'_R \boldsymbol{\Theta} \mathbf{a}_R$ can be computed from the individual factor scores and their variances. Moreover, the fixed loading $\mathbf{a}'_R \boldsymbol{\lambda}$ equals the factor score reliability ρ, which in turn equals the ratio between the variance of the factor scores and the true factor variance, $var(f)/\phi$. Using the definition of $\mathbf{a}_R$ from equation 2.2, the expected variance of regression factor scores can be expressed as follows:

$$var(f) = \mathbf{a}'_R \boldsymbol{\Sigma} \mathbf{a}_R = \phi \boldsymbol{\lambda}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^{-1} \boldsymbol{\lambda} \phi = \phi \boldsymbol{\lambda}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\lambda} \phi = \mathbf{a}'_R \boldsymbol{\lambda} \phi. \tag{2.8}$$

Using equation 2.7 for a single factor and noting that $\phi = var(\eta)$, we obtain

$$\rho = \frac{(\mathbf{a}'_R \boldsymbol{\lambda})^2 \phi}{var(f)} = \frac{(\mathbf{a}'_R \boldsymbol{\lambda})^2 \phi}{\mathbf{a}'_R \boldsymbol{\lambda} \phi} = \mathbf{a}'_R \boldsymbol{\lambda}.$$

Multiply by ϕ/ϕ yields

$$\rho = \frac{\mathbf{a}'_R \boldsymbol{\lambda} \phi}{\phi} = \frac{var(f)}{\phi}.$$

This shows the reliability of regression factor scores equals the ratio between the estimated and the true factor variance. Using this equation, the factor score variance $var(f)$ can be decomposed as follows:

$$var(f) = \phi \rho = \phi \rho \rho + \phi \rho (1 - \rho),$$

which is a special case of the more general equation 2.6, with $\rho = \mathbf{a}'_R \boldsymbol{\lambda}$ and the $\phi \rho (1 - \rho) = \mathbf{a}'_R \boldsymbol{\Theta} \mathbf{a}_R$. This shows that using the single indicator approach with regression factor scores involves fixing their loading to ρ and their residual variance to $\phi \rho (1 - \rho)$.

Another important characteristic of regression factor scores is that their true factor variances ϕ can be derived from the factor scores and their variances as follows:

$$\phi = var(f) + var_f. \tag{2.9}$$

where var_f represents the posterior variance of the individual's factor scores (see equation 2.3). Equation 2.9 follows from the definition of $var(f)$ in equation 2.8 and the application of the definition in equation 2.3 to a single factor, which yields var_f. Note that this equation is equivalent to the Expectation step for obtaining the expected factor covariance matrix when using the Expected-Maximization algorithm for maximum likelihood factor analysis.

In summary, stepwise SEM using regression factor scores as single indicators involves fixing their loadings and residual variances to ρ and $\phi\rho(1 - \rho)$, respectively. The quantities ρ and ϕ can be computed from the factor scores and their variances, thus without knowledge of original measurement model parameters. This is a very convenient since factor scores and their variances are standard output provided by software for factor analysis.

With Barlett factor scores, the fixed loading and residual variance equal 1 and $\mathbf{a}'_B \mathbf{\Theta} \mathbf{a}_B$, respectively. When applied to a single factor at a time, Barlett and regression scores are proportional to one another; that is, Barlett scores are $1/\rho$ times larger than regression scores. This implies their loading is $1/\rho$ times larger and their residual variance $1/\rho^2$ times larger. This yields a loading of 1 and a residual variance of $\phi(1 - \rho)/\rho$.

Instead of using factor scores, the single indicator approach can also be applied using a simple sum score of the items. In that case, $\mathbf{a}_S = \mathbf{1}$, implying $\lambda^* = \mathbf{a}'_S \boldsymbol{\lambda}$ and $\theta^* = \mathbf{a}'_S \mathbf{\Theta} \mathbf{a}_S$ are simple sums of loadings and residual covariances from the measurement model for the factor concerned. The sum score reliability computed by applying equation 2.6 for a single factor yields the familiar omega reliability for composite scores (McDonald, 1999; Raykov, 1997). Note that this correction for the unreliability

of sum scores requires estimating the measurement model. Somewhat related is the work by Oberski and Satorra (2013) and Savalei (2019), who proposed using the single indicator approach with composite scores with a known reliability ρ. This involves fixing their loading and residual variance to 1 and $var(f)(1 - \rho)$, respectively.

Other related work concerns the improved regression calibration by Skrondal and Kuha (2012), which can be used to correct for covariate measurement error in regression models. It involves using EAP factor scores as predictors while accounting for their uncertainty using their posterior variances. In fact, these authors make use of the fact that the total variance of the true latent factor equals the sum of the variance of the estimated factor scores and their posterior variance (see equation 2.9).

2.4   Bias adjustment

The approach proposed by Croon (2002) involves obtaining factor scores and their covariance matrix $Var(\mathbf{f})$, and subsequently computing the adjusted factor covariance matrix $Var(\boldsymbol{\eta})$. This matrix is used as input when estimating the parameters of the structural model. Croon's method can be best understood by writing the relationship between the factor scores and the true factors defined in equation 2.5 as follows:

$$\mathbf{A}\boldsymbol{\Lambda}\boldsymbol{\eta} = \mathbf{f} - \mathbf{A}\boldsymbol{\epsilon},$$

or, in terms of covariances, as

$$\mathbf{A}\boldsymbol{\Lambda}Var(\boldsymbol{\eta})\boldsymbol{\Lambda}'\mathbf{A}' = Var(\mathbf{f}) - \mathbf{A}\boldsymbol{\Theta}\mathbf{A}'.$$

This shows $Var(\boldsymbol{\eta})$ can be obtained from $Var(\mathbf{f})$ as follows:

$$Var(\boldsymbol{\eta}) = (\mathbf{A}\boldsymbol{\Lambda})^{-1}(Var(\mathbf{f}) - \mathbf{A}\boldsymbol{\Theta}\mathbf{A}')(\boldsymbol{\Lambda}'\mathbf{A}')^{-1}.$$

Note that Croon (2002) derived this formula in scalar form for each of the elements of the factor covariance matrix $Var(\boldsymbol{\eta})$, and, moreover, assumed the measurement models were estimated separately for each factor involved. The more general form given here allows for the simultaneous estimation of the measurement models of all factors, which is required when there are cross-loadings or correlated residuals among items loading on different factors (Hayes and Usami, 2020). As can be seen, the terms $\mathbf{A}\boldsymbol{\Lambda}$ and $\mathbf{A}\boldsymbol{\Theta}\mathbf{A}'$ are the same as those appearing in the derivation of the single indicator approach, which shows among others that for regression factor scores, Croon's correction can be applied without knowledge of the original measurement model parameters when there are no cross-loadings or correlated errors. Note also that with Barlett factor scores, Croon's correction formula simplifies to

$$Var(\boldsymbol{\eta}) = Var(\mathbf{f}) - \mathbf{A}_B\boldsymbol{\Theta}\mathbf{A}'_B.$$

Based on this equation, Hoshino and Bentler (2013) proposed equating the diagonal elements of $Var(\boldsymbol{\eta})$ to the estimated factor variances from the first step and the off-diagonal elements to the corresponding entries in $Var(f)$. This works because the off-diagonal elements of $\mathbf{A}_B\boldsymbol{\Theta}\mathbf{A}'_B$ equal 0 as long as there are no cross-loadings or residual correlations between items loading of different factors. Somewhat related is the work by Skrondal and Laake (2001) who showed that unbiased linear regression coefficients can be obtained without any bias adjustment of $Var(f)$ by using regression factor scores for predictors and Barlett factor scores for outcome variables.

Croon's bias adjustment has been extended for dealing with various types of more complex structural models, such as for micro-macro multilevel analysis (Croon and van Veldhoven, 2007), multilevel path analysis (Devlieger and Rosseel, 2019), and path models with latent interactions (Cox and Kelcey, 2021).

Although the local SAM approach by Rosseel and Loh (2023) is similar to Croon's approach, it uses a slightly different starting point. In local SAM, we rewrite equation 2.5 as follows

$$\mathbf{A}\mathbf{\Lambda}\boldsymbol{\eta} = \mathbf{A}(\mathbf{y} - \boldsymbol{\epsilon}),$$

where $\mathbf{A}$ is referred to as the mapping matrix. It should be chosen such that $\mathbf{A}\mathbf{\Lambda} = \mathbf{I}$, which yields

$$\boldsymbol{\eta} = \mathbf{A}(\mathbf{y} - \boldsymbol{\epsilon}),$$

or, in terms of covariances,

$$Var(\boldsymbol{\eta}) = \mathbf{A}(Var(\mathbf{y}) - \mathbf{\Theta})\mathbf{A}'. \qquad (2.10)$$

The form of $\mathbf{A}$ depends on the estimation method used to find $Var(\boldsymbol{\eta})$; for example, with maximum likelihood (ML) estimation, it coincides with the Barlett factor-score matrix $\mathbf{A}_B$ (see equation 2.6). It can easily be verified that $\mathbf{A}_B\mathbf{\Lambda} = \mathbf{I}$. Note that local SAM differs from Croon's approach in that it does not obtain factor scores as an intermediate step, but that instead $Var(\boldsymbol{\eta})$ is computed directly from $\mathbf{y}$. As in Croon's approach, the structural model is estimated using $Var(\boldsymbol{\eta})$ as input data.

Recently, Perez Alonso et al. (2024) proposed a rather advanced application of the SAM approach; that is, they used it for estimation of a mixture multigroup SEM. Their measurement model consists of a multiple group factor analysis (for many groups) and their structural model uses the group-specific factor covariances in a mixture model aimed at clustering the groups based on their structural parameters.

Similar to Croon's method and local SAM, Gerbing and Anderson (1987) and Lance et al. (1988) proposed estimating the structural parameters using $Var(\boldsymbol{\eta})$ as data matrix. However, rather than using bias-adjusted factor scores or a mapping matrix,

they used the fact $Var(\boldsymbol{\eta})$ equals the step-one factor covariance matrix Φ when the measurement models for all latent variables appearing in the structural model are estimated as a single block with a free factor covariance matrix.

3   Continuous latent and categorical observed variables

3.1   Background

When the latent variables are continuous and the observed variables are ordinal categorical, the measurement model may take on the form of a categorical factor analysis (CatFA). As factor analysis with continuous observed variables, CatFA involves estimating a linear factor model, but with the observed variables' polychoric correlation matrix instead of their covariance matrix as input data. Parameters are typically estimated by diagonally weighted least squares (DWLS).

With continuous latent variables and categorical observed variables, the measurement model may also be a latent trait or IRT model, typically estimated using marginal maximum likelihood. A unidimensional latent trait model for response pattern $\mathbf{y}_i$ can be defined as follows:

$$P(\mathbf{y}_i) = \int P(\mathbf{y}_i|\eta)f(\eta)d\eta = \int \prod_{j=1}^{J} P(y_{ij}|\eta)f(\eta)d\eta,$$

where the probability of the response of person i on item j conditional on the latent trait value η, $P(y_{ij}|\eta)$, is modeled using a logit or a probit model in which the latent trait enters as predictor. EAP estimates of the latent trait scores and their posterior

variances are defined as follows:

$$
\begin{aligned}
f_i &= E(\eta|\mathbf{y}_i) = \frac{\int \eta P(\mathbf{y}_i|\eta)f(\eta)d\eta}{P(\mathbf{y}_i)} \\
var_{f_i} &= var(\eta|\mathbf{y}_i) = \frac{\int \eta^2 P(\mathbf{y}_i|\eta)f(\eta)d\eta}{P(\mathbf{y}_i)} - E(\eta|\mathbf{y}_i)^2.
\end{aligned}
$$

Since there is no closed form expression for the integral, it has to be solved either numerically or by Monte Carlo simulation.

3.2   Fixed parameters

One particular fixed parameters approach for dealing with continuous latent variables and categorical observed variables is the global SAM approach. Although not mentioned explicitly by Rosseel and Loh (2023), their global SAM approach can also be applied in combination with a CatFA for dichotomous or ordinal items. This method is implemented in the lavaan package in R.

Kuha and Bakk (2023) proposed a two-step approach which involves fixing the latent trait measurement parameters defining the conditional response probabilities $P(y_{ij}|\eta)$ when estimating the structural part of the model. In its most general form, their structural model consisted of a path for multiple latent variables (with fixed measurement models), with in addition observed predictors affecting the latent variables and observed outcomes affected by the latent variables.

The MUPPET approach for factor analytic models has recently been extended to allow for latent trait type measurement models (Levy and Neish, 2024). MUPPET uses a two-stage Bayesian estimation method for dealing with covariates and outcome variables.

3.3   Single indicator

Lai and Hsiao (2022) showed how the single indicator approach they refer to as two-stage path analysis can be used with latent variable scores obtained from a latent trait model. They proposed using EAP trait scores, which are IRT based equivalents of regression factor scores. Related work was done by Wang et al. (2019), who applied the single indicator approach to deal with measurement error in MAP latent trait scores used as the outcome variable in a linear mixed effects model.

As explained in the context of the continuous indicators case, with regression (and thus EAP) factor scores we can fix the loading of the factor scores to their reliability ρ and their residual variance to $\varphi\rho(1-\rho)$. However, different from the factor analytic case, the reliability of IRT trait scores varies across individuals because their posterior variance of the EAP scores, var_{f_i}, varies across response patterns. More specifically, Lai and Hsiao (2022) proposed using

$$\lambda_i^* = \rho_i = 1 - \frac{var_{f_i}}{\varphi} \tag{3.1}$$

and

$$\theta_i^* = \varphi\rho_i(1-\rho_i)$$

as fixed loading and residual variance for person i. Note that, similar to the factor analysis case, φ can be obtained from the individual factor scores and their variances; that is, by using $var_f = E(var_{f_i})$ in equation 2.9.

Note that the single indicator approach using latent trait scores would be much simpler and, moreover, computationally less demanding if we replace var_{f_i} in the equation 3.1 by $var_f = E(var_{f_i})$; that is, if we ignore the heterogeneity of the measurement error when estimating the structural parameters. Lai and Hsiao (2022) indicated that

unmodeled heterogeneity may affect estimation in a path model, but at this time it is unclear how this might occur. Most likely, it has little impact as long as the measurement error heterogeneity is uncorrelated with the other variables appearing in the structural model.

3.4   Bias adjustment

Although not described in the Rosseel and Loh (2023) paper, their local SAM approach is available in the lavaan package for categorical variables. First CatFA models are estimated using DWLS and subsequently equation 2.10 is applied with a Barlett mapping matrix and $Var(\mathbf{y})$ replaced by the polychoric correlation matrix. The structural model is estimated with $Var(\boldsymbol{\eta})$ serving as the input data matrix.

Although not proposed in the literature yet, it also seems possible to implement Croon's approach with EAP trait scores from IRT measurement models. We saw that with regression factor scores $\mathbf{a}'\boldsymbol{\lambda} = \lambda^* = \rho$ and $\mathbf{a}'\boldsymbol{\Theta}\mathbf{a} = \theta^* = \phi\rho(1-\rho)$. This suggests that the trait score covariance matrix can be corrected using $\boldsymbol{\lambda}^*$ and θ^*, defined as

$$
\begin{aligned}
\lambda^* &= \rho = 1 - \frac{E(var_{f_i})}{\varphi} \\
\theta^* &= \varphi\rho(1-\rho),
\end{aligned}
$$

with

$$
\varphi = var(f) + E(var_{f_i}).
$$

This adjustment does not account for the heterogeneous measurement error. It is also possible to apply the correction with individual-specific λ_i^* and θ_i^*; that is, by applying Croon's correction formula separately to the individual contributions to the factor-score covariance matrix and taking their average. For a variance, this would

be

$$var(\eta) = E\{[f_i - E(f)]^2/\rho_i^2 - \varphi_i(1 - \rho_i)/\rho_i\},$$

and for a covariance between two factors η_1 and η_2

$$var(\eta_1, \eta_2) = E\{[f_{1i} - E(f_1)][f_{2i} - E(f_2)]/[\rho_{1i}\rho_{2i}]\}.$$

Hoshino and Bentler (2013) indicated their stepwise procedure using maximum like-lihood factor scores can also be applied with IRT type measurement models. As was explained above for factor analytic models, their approach involves equating the diagonal elements of $Var(\boldsymbol{\eta})$ to the estimated true factor variances from the first step and the off-diagonal elements to the factor score covariances.

4   Categorical latent variables

4.1   Background

Denoting the single discrete latent variable by η, a particular class by c and the number of classes by C, a latent class measurement model for $P(\mathbf{y}_i)$, the probability of the response pattern of individual i, can be defined as follows:

$$P(\mathbf{y}_i) = \sum_{c=1}^{C} P(\eta = c)P(\mathbf{y}_i|\eta = c) = \sum_{c=1}^{C} P(\eta = c)\prod_{j=1}^{J} P(y_{ij}|\eta = c).$$

Here $P(\eta = c)$, $P(\mathbf{y}_i|\eta = c)$, and $P(y_{ij}|\eta = c)$ represent the class proportion, the probability of the response pattern, and the probability of a particular response on item j for class c. A latent profile or a mixture model for continuous responses has the same structure, but with $P(\mathbf{y}_i)$, $P(\mathbf{y}_i|\eta = c)$ and $P(y_{ij}|\eta = c)$ being probability densities instead of probabilities.

The above formula is the one for a simple latent class measurement model. The most common extension containing a structural part involves including covariates $\mathbf{x}_i$ affecting class membership:

$$P(\mathbf{y}_i|\mathbf{x}_i) \;=\; \sum_{c=1}^{C} P(\eta = c|\mathbf{x}_i) \prod_{j=1}^{J} P(y_{ij}|\eta = c).$$

Other common extensions containing a structural part are models with a distal outcome z_i affected by class membership,

$$P(\mathbf{y}_i, z_i) \;=\; \sum_{c=1}^{C} P(\eta = c) \prod_{j=1}^{J} P(y_{ij}|\eta = c)P(z_{ij}|\eta = c),$$

and models with multiple latent variables affecting each other, for example, η_1 affecting η_2,

$$P(\mathbf{y}_i) \;=\; \sum_{c=1}^{C} P(\eta_1 = c_1)P(\eta_2 = c_2|\eta_1 = c_1) \prod_{j=1}^{J_1} P(y_{ij}|\eta_1 = c_1) \prod_{k=J_1+1}^{J_1+J_2} P(y_{ij}|\eta_2 = c_2),$$

The structural parameters of interest are the regression coefficients defining $P(\eta = c|\mathbf{x}_i)$, $P(z_{ij}|\eta = c)$, and $P(\eta_2 = c_2|\eta_1 = c_1)$, respectively, which as explained next can be estimated in a stepwise manner.

Latent class predictions f_i based on the latent class measurement model can be obtained based on the posterior membership probabilities $P(\eta = c|\mathbf{y}_i)$, which are obtained as follows:

$$P(\eta = c|\mathbf{y}_i) = \frac{P(\eta = c)P(\mathbf{y}_i|\eta = c)}{P(\mathbf{y}_i)}.$$

The standard practice is to assign individuals to the class with the largest posterior probability, which yields MAP estimates. Other options are random assignment or proportional assignment (Bolck et al., 2004). As shown by Bolck et al. (2004), the

assignment method used defines the values of $P(f = d|\mathbf{y}_i)$, the probability of being assigned class d conditional on the response pattern. Modal assignment gives a value of 1 for the class with the largest posterior probability and 0 for the other classes, whereas random and proportional assignment give values equal to the posterior probabilities.

4.2   Fixed parameters

Bakk and Kuha (2018) proposed a two-step latent class analysis approach, which involves fixing latent class measurement parameters defining the conditional response probabilities $P(y_{ij}|\eta = c)$ when estimating the structural model. In fact, one fixes $P(\mathbf{y}_i|\eta = c)$, the probability of a full response pattern given class membership to its estimate from the first step. The structural models considered by Bakk and Kuha were regression models in which a single latent variable serves as either predictor or outcome variable. Xue and Bandeen-Roche (2002) proposed the same approach for a latent variable serving as the outcome variable, but motivated by applications where the first step was based on a much larger sample than the second.

Applications of the fixed parameters approach with more complex structural models concern the three-step latent Markov model by Bartolucci et al. (2015) and the two-step multilevel latent class model by Di Mari et al. (2023). In both applications, the first step involves estimating the measurement model parameters, ignoring the (longitudinal or multilevel) dependence structure in the data.

Although not yet mentioned in the literature, the fixed measurement parameters approach can also be used with continuous response variables, that is, with latent profile and other types of mixture models, without any modification. In that case,

the fixed parameters in the estimation of the structural model of interest are the class-specific means and (co)variances defining the multivariate densities $f(\mathbf{y}_i|\eta = c)$.

4.3   Single indicator

For latent class models with covariates, Bolck et al. (2004) derived the relationship between $P(\eta = c|\mathbf{x}_i)$ and $P(f_i = d|\mathbf{x}_i)$; that is, between the true class distribution given covariates $\mathbf{x}_i$ and the corresponding assigned class distribution given $\mathbf{x}_i$. It has the following form:

$$P(f_i = d|\mathbf{x}_i) = \sum_{c=1}^{C} P(\eta = c|\mathbf{x}_i)P(f_i = d|\eta = c). \tag{4.1}$$

As noted by Vermunt (2010), this is the equation of a latent class model with covariates $\mathbf{x}_i$, where the class assignments f_i serve as the single indicator with $P(f_i = d|\eta = c)$ defining its quality (the amount of classification errors). Vermunt (2010) proposed estimating this single indicator model by maximum likelihood treating $P(f_i = d|\eta = c)$ as fixed parameters, and called it the three-step ML method. It uses another important result from Bolck et al. (2004); that is, $P(f_i = d|\eta = c)$ can be obtained from the measurement model parameters as follows:

$$P(f = d|\eta = c) = \sum_{\mathbf{y}} P(\mathbf{y}|\eta = c)P(f = d|\mathbf{y}) = \frac{\sum_{\mathbf{y}} P(\mathbf{y})P(\eta = c|\mathbf{y})P(f = d|\mathbf{y})}{P(\eta = c)}.$$

Vermunt (2010) proposed a slight modification of this formula, which involves replacing $P(\mathbf{y})$ by its empirical distribution in the data set at hand. This yields:

$$P(f = d|\eta = c) = \frac{\sum_{i=1}^{N} \frac{1}{N} P(\eta = c|\mathbf{y}_i)P(f = d|\mathbf{y}_i)}{P(\eta = c)}. \tag{4.2}$$

Important advantages of this modification are: 1) it prevents the possibility of summing over a very large numbers of data patterns; 2) as shown by Gudicha and Vermunt (2013), it also works with latent profile and other mixture models for continuous

observed variable, where the summation over all possible **y** takes the form of a multivariate integral; 3) irrespective of the type of classification rule, it allows computing the classification error probabilities $P(f = d|\eta = c)$ using the posteriors $P(\eta = c|\mathbf{y}_i)$ of the N persons in the data file, which is standard output provided by latent class analysis software.

As shown by Bakk et al. (2013), the three-step ML approach can not only be used for latent class models with covariates, but also for latent class models with a distal outcome and SEM-like models with multiple discrete latent variables. Moreover, this stepwise modeling approach has been used with latent Markov models (Di Mari et al., 2016; Vogelsmeier et al., 2023a), multilevel latent class models (Lyrvall et al., 2024), latent class tree models (van den Bergh and Vermunt, 2019), and inverse propensity weighting for the estimation of the causal effect of a treatment on latent class membership (Clouth et al., 2022, 2023).

A similar stepwise approach for investigating the relationship between class membership and a distal outcome z_i was proposed by Lanza et al. (2013). They proposed including z_i as covariate in the step-one latent class model, which gives posterior class membership probabilities containing information on its association with the latent classes. For the structural model estimation, they use in fact a single indicator approach with proportional class assignment, but without the need to correct classification errors.

4.4 Bias adjustment

The BCH method (Bolck et al., 2004) is a bias adjustment approach for latent class and latent profile models. Similar to the single indicator approach, it is based on

equation 4.1 which shows the relationship between $P(\eta = c|\mathbf{x}_i)$ and $P(f_i = d|\mathbf{x}_i)$. What the BCH approach does is to adjust $P(f_i = d|\mathbf{x}_i)$ so that it becomes $P(\eta_i = c|\mathbf{x}_i)$, which is the same as what Croon's method does for continuous latent variables. The derivation of the adjustment by Bolck et al. (2004) starts by writing equation 4.1 in matrix form: $\mathbf{F} = \mathbf{HD}$, with $\mathbf{F}$, $\mathbf{H}$, and $\mathbf{D}$ containing the $P(f_i = d|\mathbf{x}_i)$, $P(\eta = c|\mathbf{x}_i)$, and $P(f = d|\boldsymbol{\eta} = c)$, respectively. It can now easily be seen that $\mathbf{H} = \mathbf{FD}^{-1}$. This shows the adjustment involves multiplying the $P(f_i = d|\mathbf{x}_i)$ with the inverse of the matrix with elements $P(f = d|\boldsymbol{\eta} = c)$.

As shown by Vermunt (2010), the BCH approach can be conceptualized as an analysis of an expanded data set with C records per person with weight w_{ic}. Denoting the elements of D^{-1} by g_{dc}, the weight w_{ic} is obtained as follows:

$$w_{ic} = \sum_{d=1}^{C} P(f = d|\mathbf{y}_i) g_{dc}.$$

Note that the class assignment probabilities $P(f = d|\mathbf{y}_i)$ enter not only in this formula, but also in the formula for $P(f = d|\boldsymbol{\eta} = c)$ from equation 4.2. Vermunt (2010) proposed estimating the structural parameters – the logistic regression coefficients defining $P(\eta_i = c|\mathbf{x}_i)$ – by pseudo maximum likelihood using weights w_{ic}, and using cluster robust standard errors to account for the multiple records per person and the weighting.

The three-step BCH approach can also be used with distal outcome models (Bakk et al., 2013). In fact, it has been shown to be superior to the three-step ML approach when the outcome variable is a continuous variable (Bakk and Vermunt, 2016). Bolck et al. (2004) and Bakk et al. (2013) showed how to apply this approach with multiple latent variables. Lê et al. (2024) proposed using the BCH approach together with inverse propensity weighting for the estimation of the causal effect of latent class

membership on an outcome variable.

5 Remaining topics

In this section, we discuss various remaining topics which are relevant for the practical application of stepwise latent variable modeling approaches. These include heterogeneous measurement or classification error resulting from measurement non-invariance and missing data on the latent variable indicators, correction of standard error estimates for the stepwise modeling, and software implementing stepwise modeling approaches.

5.1 Heterogeneous measurement or classification error

When discussing the implementation of the single indicator approach with IRT type measurement models (Lai and Hsiao, 2022), we already touched upon the topic of heterogeneous measurement error. Our tentative conclusion was that ignoring the heterogeneous nature of the measurement errors probably has little impact on the structural parameter estimates as long as it is uncorrelated with the other variables appearing in the structural model.

Another situation in which measurement or classification errors are heterogeneous occurs when there is missing data on the response variables $\mathbf{y}$. More specifically, each possible missing data pattern has its own $\boldsymbol{\lambda}^*$ and $\boldsymbol{\Theta}^*$ values or its own $P(f_i = d | \eta = c)$ values, where the quality of the estimated latent variable scores decreases when the proportion of missing values increases. Alagöz and Vermunt (2022) investigated whether one can ignore the classification error heterogeneity caused by missing values

in stepwise latent class modeling. Based on a simulation study, their conclusion was that classification error heterogeneity can indeed be ignored as long as a missing at random (MAR) mechanism holds for the **y** variables concerned. An example of a situation in which the MAR assumption is violated occurs when, conditional of the observed **y** variables, covariates (which are not used in the measurement model) affect the probability of having missing values. For the continuous latent variable case, the effect of ignoring measurement error heterogeneity caused by missing data has not been investigated. But most likely the same condition holds here; that is, it can be ignored as long as we can assume the missing data is MAR when estimating the step-one model. Heterogeneous measurement and classification errors also occurs when the step-one model accounts for measurement non-invariance or differential item functioning. In that case, the terms $\boldsymbol{\lambda}^*$ and $\boldsymbol{\Theta}^*$ or the $P(f_i = d|\eta = c)$ used in the single indicator and bias adjustment approaches vary across groups with differences in their measurement model parameters. Lai et al. (2023) proposed a two-stage multiple-group path analysis in which the measurement model is a multiple-group factor analysis model and in which group-specific values are used for $\boldsymbol{\lambda}^*$ and $\boldsymbol{\Theta}^*$ when estimating the multiple-group path model. Vermunt and Magidson (2021) looked in more detail into the latent class analysis case with measurement non-invariance (see also Clouth et al., 2023). Their conclusion was that group-specific classification errors $P(f_i = d|\eta = c)$ should be accounted for when the grouping variable is itself part of the structural model. In other words, if the grouping variable is not used in the structural model, the heterogeneous classification errors can be ignored. Most likely, this same conclusion applies in the continuous latent variable case.

It should be noted that whether to account for or to ignore heterogeneous measurement or classification error is not an issue in the fixed parameters approach, since

it does not involve estimation of the latent variables scores as an intermediate step. Instead, one fixes the measurement parameters, which may differ per group if one accounted for measurement non-invariance in the first step.

5.2   Standard error computation

In the fixed parameters and the single indicator approach, we treat either the estimated parameters from the measurement model or functions of these as known when estimating the structural model parameters. As shown by Rosseel and Loh (2023) and Oberski and Satorra (2013) for the continuous latent variable case and by Bakk and Kuha (2019) and Bakk et al. (2014) for the discrete latent variable case, standard errors can be corrected for the fact that these are in fact estimates with their own sampling variation when the structural parameters are estimated by maximizing a log likelihood, which known as pseudo-maximum likelihood estimation (Gong and Samaniego, 1981).

Let $\boldsymbol{\Sigma}_S$ contain the uncorrected covariances of the parameters of the structural model, $\boldsymbol{\Sigma}_M$ the covariances of the measurement parameters or, in the single indicator case, of functions of these, $\mathbf{H}$ the second derivatives towards the free parameters, and $\mathbf{C}$ towards the free parameters with respect to the fixed parameters. The adjusted covariance matrix of the structural parameters, $\boldsymbol{\Sigma}_S^*$, is obtained as follows:

$$\boldsymbol{\Sigma}_S^* = \boldsymbol{\Sigma}_S + \mathbf{H}^{-1}\mathbf{C}\boldsymbol{\Sigma}_M\mathbf{C}'\mathbf{H}^{-1}.$$

In the single indicator approach $\boldsymbol{\Sigma}_{\boldsymbol{M}}$ can be obtained from the covariances of the measurement parameters using the delta method. Bakk et al. (2014) illustrated this for the latent class analysis case; that is, for the three-step ML procedure in which $\boldsymbol{\Sigma}_{\boldsymbol{M}}$ contains the covariances of the $\log P(f_i = d|\eta = c)/P(f_i = c|\eta = c)$ terms. It

should be noted that when dealing with dependent observations such as in multilevel latent class models (Lyrvall et al., 2024), latent Markov models (Di Mari et al., 2016; Vogelsmeier, et al., 2023a), and latent vector autoregressive models (Rein et al., 2024), $\mathbf{\Sigma_M}$ will be a cluster-robust parameter covariance matrix.

Similar types of standard error corrections are not available for the bias adjustment approach. This is why, for example, the lavaan package reports global SAM based standard errors when using the local SAM procedure. However, when applying the BCH approach for discrete latent variables, one should account for the fact that the expanded data set contains multiple records per person and, moreover, BCH weights. As proposed by Vermunt (2010), this can be dealt with using complex sampling (or cluster-robust) standard errors.

As an alternative to the frequentist approach of fixing measurement model parameters to their first-stage point estimates and subsequently adjusting the standard error estimates, one may use Bayesian multiple-stage estimation, as done in MUPPET modeling. This involves dealing with parameter uncertainty by accounting for the full first-stage posterior parameter distribution when estimating the structural parameters (Levy, 2023; Levy and McNeish, 2024). Similar Bayesian estimation methods could be developed for single indicator and bias adjustment approaches. For example, in the single indicator approach, instead of using a single set of latent variable scores obtained with point estimates of the measurement parameters, one could use multiple sets of latent variable scores drawn from their posterior distribution while accounting for the uncertainty of the measurement parameters.

5.3   Software implementations

Most of the stepwise approaches we discussed are available in latent variable modeling software. However, with some additional some effort, each of the approaches can be implemented without specific stepwise modeling routines. The fixed measurement and single indicator approaches can be implemented in SEM or latent class analysis software that allows imposing fixed value constraints on the model parameters. Slightly more tedious to implement is the heterogeneous measurement case resulting from an IRT type measurement model, for which Lai and Hsiao (2022) provided example code for OpenMx (Neale et al., 2016). The bias adjustment approach for continuous latent variables involves creating an adjusted factor covariance matrix, which can subsequently be used as input data for any structural modeling software. For categorical latent variables, one should create an expanded data set containing the BCH weights, after which any type of structural model can be estimated using a weighted analysis. The only requirement is that the routine used to estimate the structural model accepts negative weights.

LatentGOLD (Vermunt and Magidson, 2016, 2024) is one of the programs that contains options for stepwise latent variable modeling. For latent class models, it implements each of the three approaches, which it refers to as the Bakk-Kuha, ML, and BCH adjustment methods. In version 6.1, it also implements the single indicator approach for continuous latent variables, where the step-one model can be either a factor analysis or latent trait analysis. After estimation of the measurement model of interest, one saves the posterior probabilities or the logdensities of the latent class model, or the EAP factor scores and their standard errors from the factor analytic or IRT model to an output data file. This information suffices to set up the struc-

tural model, where it is possible to indicate that the measurement or classification errors are heterogeneous across levels of a grouping variable. For latent class models, the most common types of structural models (including latent Markov models) are available via the Step3 point-and-click module. Other stepwise latent class models, such as path models for multiple discrete latent variables and multilevel latent class models, can be specified via the syntax system. The structural model for continuous latent variables should always be specified with LatentGOLD's syntax system, where not only a simple regression or path model can be specified, but also a more complex model, such as a mixture multiple group SEM (Perez Alonso, 2024) or a dynamic factor model (Rein et al., 2024).

Mplus (Asparouhov and Muthén, 2014) implements the ML (single indicator) and BCH (bias adjustment) approaches for latent class models in an automated form for simple structural models (covariates and distal outcomes). However, it also allows saving BCH weights to an output file and using these weights in subsequent analysis for the estimation of more complex structural models with latent classes. The BCH approach is also available as SAS and Stata functions for dealing with covariates and distal outcomes (Dziak et al., 2017, 2022).

R packages implementing stepwise latent variable models include lavaan (Rosseel, 2012), multilevLCA (Lyrvall et al., 2023), lmfa (Vogelsmeier et al., 2023b), and tidySEM (van Lissa, 2019). Lavaan implements the fixed parameters and bias adjustment approaches for factor analysis and CatFA measurement models, and refers to these as global and local SAM. The multilevLCA package implements standard and multilevel latent class analysis with covariates for binary response variables estimated using a two-step (or fixed parameters) approach. The lmfa package implements the

single indicator approach for a continuous-time latent Markov model with covariates for intensive longitudinal data, where the first step is a mixture factor analysis. Finally, tidySEM contains a BCH function for estimating structural models in which the latent classes are treated as the grouping variable.

In addition to full R packages, various recent publications include R code which allows researchers to apply the newly proposed method concerned with their own data. Examples include two-step path modeling (Lai and Hsiao, 2022), MUPPET (Levy, 2023; Levy and McNeish, 2024), mixture multiple-group SEM (Perez Alonso, 2024), and stepwise latent vector autoregressive modeling (Rein et al., 2024).

6 Discussion

An overview of stepwise latent variable approaches was provided. It was shown that similar approaches have been proposed for factor analytic, latent trait, and latent class type measurement models. These involve using fixed measurement parameters, latent variable predictions as single indicators, and Croon's or similar bias adjustments of the latent variable predictions. We explained the logic underlying these approaches with the appropriate references, as well as references to applications in combination with more complex structural models. Moreover, we explained Croon's method can also be implemented with IRT based measurement models.

Special attention was paid to the issue of heterogeneous measurement and classification errors, which occurs with IRT models, with missing values on the latent variable indicators, and with measurement non-invariance. When a grouping variable causing the measurement invariance is also used in the structural model, the heterogeneity

should clearly be taken into account. In most other situations it appears that the heterogeneity can be ignored, but this is something that needs further investigation.

We showed how to obtain standard errors corrected for the stepwise estimation. However, it has also been reported that this correction is not needed or that it may even yield too conservative tests if the measurement or classification errors are not large, and the sample is not small. This issue is also a topic requiring further investigation. Moreover, work remains to be done for the bias adjustment approaches for which standard error correction is not yet available. Interesting is the Bayesian MUPPET modeling approach, which accounts for the stepwise estimation in a rather straightforward manner. Similar Bayesian estimation methods may be developed for application in conjunction with single indicator or bias adjustment approaches. Another popular approach for improved standard error estimation in stepwise modeling is bootstrapping, which surprisingly has not yet been investigated in the context of stepwise latent variable modeling.

We discussed software implementations of the stepwise approach, which include LatentGOLD, Mplus, R packages and code written in R, as well as functions written for Stata and SAS. Given that this is a lively field of research, surely, more software for stepwise latent variable modeling will become available in the near future.

Declaration of conflicting interests

Jeroen Vermunt is co-developer (together with Jay Magidson) of the LatentGOLD software.

References

[1] Alagöz, O.E.C., and Vermunt, J.K. (2022). Stepwise latent class analysis in the presence of missing values on the class indicators. Structural Equation Modeling, 29, 784-790.

[2] Asparouhov, T., and Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. Structural Equation Modeling, 21, 329341. doi:10.1080/10705511.2014.915181

[3] Bakk, Z., and Kuha, J. (2018). Two-step estimation of models between latent classes and external variables. Psychometrika, 83, 871-892. https://doi.org/10.1007/s11336-017-9592-7

[4] Bakk, Z., and Kuha, J. (2021). Relating latent class membership to external variables: An overview. British Journal of Mathematical and Statistical Psychology, 74(2), 340-362. https://doi.org/10.1111/bmsp.12227

[5] Bakk, Z., Oberski, D., and Vermunt, J.K. (2014). Relating latent class assignments to external variables:Standard errors for correct inference. Political Analysis, 22, 520-540. https://doi.org/10.1093/pan/mpu003

[6] Bakk, Z., Tekle, F.T., and Vermunt, J.K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. Sociological Methodology, 43, 272-311. https://doi.org/10.1177/0081175012470644

[7] Bakk, Z., and Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. Structural Equation Modeling, 23, 20-31.

[8] Bartholomew, D.J., and Knott, M. (1999). Latent Variable Models and Factor Analysis. London: Arnold.

[9] Bartolucci, F., Montanari, G. and Pandolfi, S. (2015). Three-step estimation of latent Markov models with covariates. Computational Statistics and Data Analysis, 83, 287-301.

[10] Bolck, A., Croon, M., and Hagenaars, J.A. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. Political Analysis 12, 3-27.

[11] Bollen, K.A. (1989). Structural equations with latent variables. Wiley.

[12] Burt, R.S. (1976). Interpretational confounding of unobserved variables in structural equation models. Sociological Methods and Research, 5, 3-52. https://doi.org/10.1177/004912417600500101

[13] Clouth, F.J., Pauws, S., Mols, F., and Vermunt, J.K. (2022). A new three-step method for using inverse propensity weighting with latent class analysis. Advances in Classification and Data Analysis, 16, 351–371.

[14] Clouth, F.J., Pauws, S., and Vermunt, J.K. (2023). Three-step latent class analysis with inverse propensity weighting in the presence of differential item functioning. Structural Equation Modeling, 30, 737-748.

[15] Cox, K., and Kelcey, B. (2021). Croon's bias corrected estimation of latent interactions. Structural Equation Modeling: A Multidisciplinary Journal, 28, 863–874.

[16] Croon, M. (2002). Using predicted latent scores in general latent structure models. In G. Marcoulides and I. Moustaki (Eds.), Latent variable and latent structure models (pp. 195-224). Lawrence Erlbaum.

[17] Croon, M.A., and van Veldhoven, M.J. (2007). Predicting group-level outcome variables from variables measured at the individual level: A latent variable multi-level model. Psychological Methods, 12, 45-57. doi:10.1037/1082-989X.12.1.45

[18] Devlieger, I., and Rosseel, Y. (2017). Factor score path analysis: An alternative for SEM? Methodology, 13, 31-38. https://doi.org/10.1027/1614-2241/a000130

[19] Devlieger, I., and Rosseel, Y. (2019). Multilevel factor score regression. Multivariate Behavioral Research, 55, 600–624. https://doi.org/10.1080/00273171.2019.1661817

[20] Dziak, J.J., Bray, B.C., and Wagner, A.T. (2017). LCA_Distal_BCH SAS macro users' guide (Version 1.1). University Park, PA: The Methodology Center, Penn State. Retrieved from http://methodology.psu.edu

[21] Dziak, J.J., Bray, B.C., and Wagner, A.T. (2020). LCA_Covariates_3Step SAS macro users' guide (Version 1.0). University Park, PA: The Methodology Center, Penn State. Retrieved from http://methodology.psu.edu

[22] Di Mari, R., Bakk, Z., Oser, J. and Kuha, J. (2023). A two-step estimator for multilevel latent class analysis with covariates. Psychometrika, 88, 1144-1170. https://doi.org/10.1007/s11336-023-09929-2

[23] Di Mari, R., Oberski, D.L. and Vermunt, J.K. (2016). Bias-adjusted three-step latent Markov modeling with covariates. Structural Equation Modeling: A Multidisciplinary Journal, 23, 649-660. https://doi.org/10.1080/10705511.2016.1191015

[24] Gerbing, D.W., and Anderson, J.C. (1987). Improper solutions in the analysis of covariance structures: Their interpretability and a comparison of alternate respecifications. Psychometrika, 52(1), 99–111. https://doi.org/10.1007/BF02293958

[25] Gong, G., and Samaniego, F.J. (1981). Pseudo maximum likelihood estimation: theory and applications. Annals of Statistics, 9, 861–869.

[26] Greene, W. H. (2003). Econometric analysis. Prentice Hall.

[27] Gudicha, D.W., and Vermunt, J.K. (2013). Mixture model clustering with covariates using adjusted three-step approaches. In: B. Lausen, D. van den Poel, and A. Ultsch (eds), Algorithms from and for Nature and Life; Studies in Classification, Data Analysis, and Knowledge Organization. 87-93. Heidelberg: Springer-Verlag GmbH.

[28] Hayes, T., and Usami, S. (2020). Factor score regression in the presence of correlated unique factors. Educational and Psychological Measurement, 80, 5-40. https://doi.org/10.1177/0013164419854492

[29] Hoshino, T., and Bentler, P.M. (2013). Bias in factor score regression and a simple solution. In A.R. de Leon and K.C. Chough (Eds.), Analysis of mixed data: Methods and applications (pp. 43-61). Boca Raton, FL: Chapman and Hall. doi:10.1201/b14571-5

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

[31] Kuha, J. and Bakk, Z. (2023). Two-step estimation of latent trait models. https://arxiv.org/pdf/2303.16101.pdf

[32] Lai, M.H.C., and Hsiao, Y.-Y. (2022). Two-stage path analysis with definition variables: An alternative framework to account for measurement error. Psychological Methods, 27, 568-588. https://doi.org/10.1037/met0000410

[33] Lai, M.H.C., Tse, W.W.-.Y, Zhang, G, Li, Y., and Hsiao, Y.-Y. (2023) Correcting for unreliability and partial invariance: A two-stage path analysis approach, Structural Equation Modeling: A Multidisciplinary Journal, 30, 258-271, DOI: 10.1080/10705511.2022.2125397

[34] Lance, C.E., Cornwell, J.M. and Mulaik, S.A. (1988). Limited information parameter estimates for latent or mixed manifest and latent variable models. Multivariate Behavioral Research, 23, 171-187. https://doi.org/10.1207/s15327906mbr2302_3

[35] Lanza, T.S., Tan, X., and Bray, C.B. (2013). Latent class analysis with distal outcomes: A flexible model-based approach. Structural Equation Modeling, 20, 1–26.

[36] Lê, T.T., Clouth, F.J., and Vermunt, J.K. (2024). Causal latent class analysis with distal outcomes: A modified three-step method using inverse propensity weighting. Multivariate Behavioral Research. Advance online publication.

[37] Levy, R. (2023). Precluding interpretational confounding in factor analysis with a covariate or outcome via measurement and uncertainty preserving parametric modeling. Structural Equation Modeling: A Multidisciplinary Journal, 30, 719–736. https:// doi.org/10.1080/10705511.2022.2154214

[38] Levy, R., and McNeish, D. (2024). Measurement and uncertainty preserving parametric modeling for continuous latent variables with discrete indicators and external variables. Journal of Educational and Behavioral Statistics. Advance online publication.

[39] Lyrvall, J., Di Mari, R., Bakk, Z., Oser, J. and Kuha, J. (2023) multilevLCA:

An R package for single-level and multilevel latent class analysis with covariates. https://arxiv.org/abs/2305.07276

[40] Lyrvall, J., Bakk, Z., Oser, J. and Di Mari, R. (2024) Bias-adjusted three-step multilevel latent class modeling with covariates, Structural Equation Modeling: A Multidisciplinary Journal, 31, 592-603, DOI: 10.1080/10705511.2023.2300087

[41] McDonald, R.P. (1999). Test theory: A unified treatment. Psychology Press.

[42] Neale, M.C., Hunter, M.D., Pritikin, J.N., Zahery, M., Brick, T.R., Kirkpatrick, R.M., Estabrook, R., Bates, T.C., Maes, H.H., and Boker, S.M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. Psychometrika, 81, 535–549. https://doi.org/10.1007/s11336-014-9435-8

[43] Oberski, D. L., and Satorra, A. (2013). Measurement error models with uncertainty about the Error variance. Structural Equation Modeling, 20, 409–428. http://dx.doi.org/10.1080/10705511.2013.797820

[44] Perez Alonso, A.F., Rosseel, Y., Vermunt, J.K., and De Roover, K. (2024). Mixture multigroup structural equation modeling: A novel method for comparing structural relations across many groups. Psychological Methods. Advance online publication.

[45] Raykov, T. (1997). Estimation of composite reliability for congeneric measures. Applied Psychological Measurement, 21, 173-184. https://doi.org/10.1177/01466216970212006

[46] Rein, M.T., Vermunt, J.K., De Roover, K., and Vogelsmeier, L.V.D.E. (2024). Latent vector autoregressive modeling: A stepwise estimation approach. https://osf.io/preprints/psyarxiv/a2muk

[47] Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. Journal of Statistical Software, 48, 1-6. https://doi.org/10.18637/jss.v048.i02

[48] Rosseel, Y., and Loh, W.W. (2022). A structural after measurement approach to structural equation modeling. Psychological Methods. Advance online publication. https://doi.org/10.1037/met0000503

[49] Savalei, V. (2019). A comparison of several approaches for controlling measurement error in small samples. Psychological Methods, 24, 352-370

[50] Skrondal, A., and Laake, P. (2001). Regression among factor scores. Psychometrika, 66, 563-575. https://doi.org/10.1007/BF02296196

[51] Skrondal, A., and Kuha, J. (2012). Improved regression calibration. Psychometrika, 77, 649-669.

[52] Van den Bergh, M., and Vermunt, J.K. (2019). Latent class trees with the three-step approach, Structural Equation Modeling, 26, 481-492.

[53] Van Lissa, C.J., (2019). tidySEM: Tidy structural equation modeling. R package version 0.2.1. https://github.com/cjvanlissa/tidySEM/

[54] Vermunt, J.K. (2010). Latent class modeling with covariates: Two improved three-step approaches. Political Analysis 18, 450-469.

[55] Vermunt, J.K. and Magidson, J. (2016). Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax. Belmont MA: Statistical Innovations Inc.

[56] Vermunt, J.K., and J. Magidson (2021). How to perform three-step latent class analysis in the presence of measurement non-invariance or differential item functioning. Structural Equation Modeling, 28, 356-364.

[57] Vermunt, J.K., and J. Magidson (2024). Upgrade Manual for Latent GOLD Basic, Advanced, Syntax, and Choice Versions 6.0 and 6.1. Arlington, MA: Statistical Innovations Inc.

[58] Vogelsmeier, L.V.D.E., Vermunt, J.K., Bülow, A., and De Roover, K. (2023a). Evaluating covariate effects on ESM measurement model changes with latent Markov factor analysis: A three-step approach, Multivariate Behavioral Research, 58, 262-291.

[59] Vogelsmeier, L.V.D.E., Vermunt, J.K., and De Roover, K. (2023b). How to explore within-person and between-person measurement model differences in intensive longitudinal data with the R package lmfa. Behavior Research Methods, 55, 2387–2422.

[60] Wang, C., G. Xu, and X. Zhang (2019). Correction for item response theory latent trait measurement error in linear mixed effects models. Psychometrika, 84, 673-700.

[61] Xue, Q.L., and Bandeen-Roche, K. (2002). Combining complete multivariate outcomes with incomplete covariate information: A latent class approach. Biometrics, 58, 110-120.

Table 1: Types of stepwise approaches for different measurement models

Stepwise approach	Measurement model		
	Factor analysis	Latent trait and categorical factor analysis	Latent class, latent profile and mixture analysis
Fixed parameters	global SAM (Rosseel and Loh, 2023); MUPPET (Levy, 2023)	global SAM (Rosseel and Loh, 2023); two-step latent trait (Kuha and Bakk, 2024); MUPPET (Levy, 2023)	two-step latent class (Bakk and Kuha, 2018), three-step latent Markov (Bartolucci et al. (2015), two-step multilevel latent class (Di Mari et al., 2023)
Single indicator	two-step path analysis (Lai and Hsiao, 2023), improved regression calibration (Skrondal and Kuha, 2012), stepwise latent vector autoregressive modeling (Rein et al., 2024)	two-step path analysis (Lai and Hsiao, 2023) , latent trait measurement error correction (Wang et al., 2019)	three-step ML latent class (Vermunt, 2010; Bakk et al. 2013), three-step ML mixture model clustering (Gudicha and Vermunt, 2013), three-step latent Markov (Di Mari et al., 2018), three-step multilevel latent class (Lyrvall et al., 2024), three-step mixture modeling (Asparouhov and Muthén (2014)
Bias adjustment	local SAM (Rosseel and Loh, 2023), Croon's correction (Croon, 2002), factor score regression (Hayes and Usami, 2020; Hoshino and Bentler, 2013), factor score path analysis (Devlieger and Rosseel, 2017), multilevel factor score regression (Croon and van Veldhoven, 2007; Devlieger and Rosseel, 2017); Croon's bias corrected latent interactions (Cox and Kelcey, 2021), mixture multigroup SEM (Perez Alonso, 2024)	local SAM (Rosseel and Loh, 2023), factor score regression (Hoshino and Bentler, 2013), Croon's correction (this paper)	three-step BCH latent class (Bolck et al., 2004; Vermunt, 2010; Bakk and Vermunt, 2016), three-step BCH mixture model clustering (Gudicha and Vermunt, 2013)