Structural Equation Modeling: A Multidisciplinary Journal

ISSN: 1070-5511 (Print) 1532-8007 (Online) Journal homepage: https://www.tandfonline.com/loi/hsem20

Comparison of Models for the Analysis of Intensive Longitudinal Data

Tihomir Asparouhov & Bengt Muthén

To cite this article: Tihomir Asparouhov & Bengt Muthén (2020) Comparison of Models for the Analysis of Intensive Longitudinal Data, Structural Equation Modeling: A Multidisciplinary Journal, 27:2, 275-297, DOI: 10.1080/10705511.2019.1626733

To link to this article: https://doi.org/10.1080/10705511.2019.1626733

Structural Equation Modeling: A Multidisciplinary Journal, 27: 275–297, 2020

ISSN: 1070-5511 print / 1532-8007 online
DOI: https://doi.org/10.1080/10705511.2019.1626733

Routledge
Taylor & Francis Group

Check for updates

TEACHER'S CORNER

Comparison of Models for the Analysis of Intensive Longitudinal Data

Tihomir Asparouhov and Bengt Muthén
Mplus

We discuss the differences between several intensive longitudinal data models. The dynamic structural equation model (DSEM), the residual dynamic structural equation model (RDSEM) and the repeated measures longitudinal model are compared in several simulation studies. We show that the DIC can be used to select the correct modeling framework. We discuss the consequences of incomplete or incorrect modeling for the predictors in multilevel time series models. We also illustrate the advantages of the Bayesian estimation over the REML estimation for models with categorical data, subject-specific autocorrelations, and subject-specific residual variances. Dynamic factor analysis models are discussed where autoregressive relations occur not only for the factors but also for the residuals of the measurement variables. The models are also illustrated with an empirical example.

Keywords: dynamic structural equation models, repeated measures longitudinal models, residual dynamic structural equation models, restricted maximum likelihood estimation

INTRODUCTION

This paper considers analysis of longitudinal data with many time points. Models for such analysis have been discussed in the repeated measures tradition of, e.g. Chi and Reinsel (1989) and Raudenbush and Bryk (2002) where the correlation across time is modeled via random effects and autocorrelated residuals. A second modeling tradition based on time series analysis is discussed in e.g. Molenaar (2017) and Asparouhov, Hamaker, and Muthén (2018) and focuses on the analysis of intensive longitudinal data where observations are made close in time. Such analyses are based on dynamic models where current outcomes are regressed on past outcomes. Asparouhov et al. (2018) presented a general modeling framework for such analyses, referred to as DSEM (Dynamic Structural Equation Modeling). A corresponding modeling framework for residuals correlated across time was also presented and referred to as RDSEM (Residual DSEM).

Correspondence should be addressed to Tihomir Asparouhov Mplus, Email: tihomir@statmodel.com; Bengt Muthén E-mail: bmuthen@statmodel.com

The RDSEM framework is designed to bridge the gap between the repeated measures longitudinal modeling approach and the time series approach of DSEM. The dynamic relations between the variables in DSEM are replaced by dynamic relations between their residuals in RDSEM. This way autocorrelations can be created for the residuals independently of the structural regressions in the model, much like this is done for the repeated measures models. This similarity between RDSEM and the repeated measures models are very appealing because the model results can be interpreted in the tradition of the repeated measures models. At the same time, the RDSEM framework retains the generality and flexibility of the DSEM framework. The RDSEM framework is based on the traditions of the SEM approach and can be used to study path analysis, factor analysis, mediation analysis and their evolution across time. The framework can easily accommodate VAR (vector autoregressive) residuals or multivariate autoregressive residuals or any other combination that can be formulated through structural dynamic modeling.

In this article, we illustrate the differences between the DSEM and the RDSEM models through several simulation

studies. We compare these models to the standard multilevel SEM model which ignores the autocorrelations in the data. We also compare the models to the repeated measures longitudinal model, based on the REML (restricted maximum likelihood) estimation for linear mixed models implemented in SAS and SPSS. Methods for comparing the different modeling approaches in terms of model fit are discussed as well. The DSEM model has been implemented in Mplus 8 while the RDSEM model has been implemented in Mplus 8.1 for continuous variable and in Mplus 8.2 for categorical variables.

The outline of this article is as follows. First, we present the general DSEM and RDSEM models. We then consider several basic models such as the autoregressive model and the regression model. We illustrate how the DIC (deviance information criterion) can be used to evaluate model fit and to determine which of the two modeling frameworks is the better fit for the data. We also consider the concept of exogeneity in multilevel time series models. We study the consequences of incorrect or incomplete specification of the distribution of the covariates. This is important because in cross-sectional analysis typically only the conditional model is evaluated $[Y|X]$ and the distribution of the covariate is ignored without any consequences. Such an approach may not be a feasible strategy for DSEM nd RDSEM models. Further simulation studies illustrate the effect of unevenly spaced times of observations in the DSEM and RDSEM frameworks. We then discuss the problems of the REML estimation and conduct montecarlo studies that expose the poor performance of the REML estimation for models with categorical data, subject-specific autocorrelations, or subject-specific residual variances. We conclude the article with several more advanced examples such as ARMA(1,1) models, MEAR (measurement error autoregressive) models, dynamic factor analysis models and an empirical example.

THE DSEM AND RDSEM MODELS

Here we briefly summarize the DSEM and the RDSEM models. Let Y_{it} be the vector of dependent variables and X_{it} be the vector of independent variables for individual i at time t. The DSEM model is described as follows. First, we decompose the variables into within and between components.

$$Y_{it} = Y_{w,it} + Y_{b,i} \tag{1}$$

$$X_{it} = X_{w,it} + X_{b,i}. \tag{2}$$

This decomposition can be interpreted as a random intercept only two-level regression. The variable $Y_{b,i}$ is interpreted as the time-invariant contribution to Y_{it} or as the mean $E(Y_{it}|i)$. The variable $Y_{w,it}$ is the time-specific deviation from that mean at time t. The variables $X_{b,i}$ and $X_{w,it}$ are interpreted similarly. The structural part of the model is now expressed separately for the two components. The between level model is the same as the between level model in the standard two-level SEM model.

$$Y_{b,i} = \nu_2 + \Lambda_b \eta_{b,i} + K_b X_{b,i} + \varepsilon_{b,i} \tag{3}$$

$$\eta_{b,i} = \alpha_b + B_b \eta_{b,i} + \Gamma_b X_{b,i} + \xi_{b,i} \tag{4}$$

where $\eta_{b,i}$ is the vector of between level latent variables. The within level model incorporates lagged predictors, i.e., variables from the previous L periods where L is the lag of the model.

$$Y_{w,it} = \nu_1 + \sum_{l=0}^{L} \Lambda_{w,l} \eta_{w,i,t-l} + \sum_{l=0}^{L} R_l Y_{w,i,t-l} + \sum_{l=0}^{L} K_{w,l} X_{w,i,t-l} + \varepsilon_{w,it} \tag{5}$$

$$\eta_{w,it} = \alpha_w + \sum_{l=0}^{L} B_{w,l} \eta_{w,i,t-l} + \sum_{l=0}^{L} Q_l Y_{w,i,t-l} + \sum_{l=0}^{L} \Gamma_{w,l} X_{w,i,t-l} + \xi_{w,it}. \tag{6}$$

where $\eta_{w,it}$ is the vector of within-level latent variables. The variables $\varepsilon_{b,i}$, $\xi_{b,i}$, $\varepsilon_{w,it}$, and $\xi_{w,it}$ represent the residuals in the above equations.

The RDSEM model is introduced similarly. The between level model of the RDSEM model is the same as the between level model of the DSEM model. The within level model, however, is different. The within level model is now further separated into a "structural part" and an "autoregressive part". First, the structural part explicates the contemporaneous relationship between the variables, i.e., the relationships between the variables from the same time period.

$$Y_{1,it} = \nu_1 + \Lambda_{1,0} \eta_{1,it} + R_0 Y_{1,it} + K_{1,0} X_{1,it} + \hat{Y}_{1,it} \tag{7}$$

$$\eta_{1,it} = \alpha_1 + B_{1,0} \eta_{1,it} + Q_0 Y_{1,it} + \Gamma_{1,0} X_{1,it} + \hat{\eta}_{1,it}. \tag{8}$$

The variables $\hat{Y}_{1,it}$ and $\hat{\eta}_{1,it}$ take the role of the residuals in the above structural models. Equations (7–8) can also be viewed as the definitions for $\hat{Y}_{1,it}$ and $\hat{\eta}_{1,it}$. These equations are also equivalent to the within-level model in a standard two-level SEM model, i.e., there are no autoregressive relations here. All autoregressive relations are given in the following model for $\hat{Y}_{1,it}$ and $\hat{\eta}_{1,it}$

$$\hat{Y}_{1,it} = \sum_{l=1}^{L} \Lambda_{1,l} \hat{\eta}_{1,i,t-l} + \sum_{l=1}^{L} R_l \hat{Y}_{1,i,t-l} + \varepsilon_{1,it} \tag{9}$$

$$\hat{\eta}_{1,it} = \sum_{l=1}^{L} B_{1,l}\hat{\eta}_{1,i,t-l} + \sum_{l=1}^{L} Q_l \hat{Y}_{1,i,t-l} + \xi_{1,it}. \quad (10)$$

The RDSEM framework allows us to separate the structural part of the model from the autoregressive part. The autoregressive part is entirely contained in the model for the residuals and can be thought of as being auxiliary in the following sense. The autoregressive part allows us to model the time-series nature of the data while preserving the focus of the model on the contemporaneous relationship between the variables.

The RDSEM and DSEM models address different substantive questions. In many situations, this factor alone will determine the choice between these two modeling options. The RDSEM model focuses on the contemporaneous relations between the variables. The DSEM model focuses on structural relations that transcend across periods, i.e., structural relations between current and past observations. In some situations, however, it is possible to use statistical methodology to determine the best choice. One such example was discussed in Asparouhov et al. (2018) in the context of adding a covariate to the autoregressive model. In this article we focus on the DIC, see Spiegelhalter, Best, Carlin, and Van Der Linde (2002) and Asparouhov et al. (2018), as a method for evaluating the fit of the DSEM and the RDSEM models for comparative purposes.

It is shown in Asparouhov et al. (2018) that the RDSEM model can be viewed as a special case of the DSEM model and that for estimation purposes an RDSEM model can be approximated by a DSEM model. Such an approximation, however, is based on treating the residuals as latent variables in the model and augmenting the model with new residuals with variance fixed to a small positive number close to zero. There are several drawbacks to this approach. First, the model becomes unnecessarily complicated. Second, the Bayesian estimation becomes quite inefficient due to the new small variance residuals, i.e., the convergence is much slower. That is in addition to the fact that the model is expanded and individual iterations are much slower. The third disadvantage is the fact that the DIC of this approximation setup cannot be used for model comparison. The new latent variables are treated in DSEM as model parameters and the small residual variances of the new residuals affect the DIC results making it difficult to use. In Mplus 8.2 the RDSEM model is estimated directly without the need for this approximation setup based on the additional latent variables. With this new algorithm, the RDSEM model becomes much more practical to use. The estimation is substantially easier in terms of faster convergence due to better mixing in the MCMC algorithm. Also, the DIC can now be used for model comparison. The new estimation of the RDSEM model is outlined in Appendix C in Asparouhov et al. (2018). The RDSEM model is implemented in Mplus 8.2 for single-level models as well as two-level models. It is currently not available for cross-classified models.

Didactic empirical examples for the DSEM model can be found in Hamaker, Asparouhov, Brose, Schmiedek, and Muthén (2018), McNeish and Hamaker (2018), and Öhrlund, Schultzberg, and Bartusch (2019). Didactic empirical examples for the RDSEM model can be found in Bolger and Laurenceau (2013). Data requirements for empirical applications are discussed in Schultzberg and Muthén (2018). In the following sections, we consider several basic DSEM and RDSEM examples and present some montecarlo results as well as an empirical illustration.

THE AUTOREGRESSIVE MODEL

Consider a single-level autoregressive model of lag 2, see Greene (2014), in the DSEM and RDSEM frameworks. Let Y_t be the observed variable at time t. The DSEM AR(2) model is

$$Y_t = \alpha + r_1 Y_{t-1} + r_2 Y_{t-2} + \varepsilon_t \quad (11)$$

$$\varepsilon_t \sim N(0, \sigma) \quad (12)$$

The RDSEM AR(2) model is

$$Y_t = \mu + \varepsilon_t \quad (13)$$

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \delta_t \quad (14)$$

$$\delta_t \sim N(0, v) \quad (15)$$

The two models are equivalent and are reparameterizations of each other as follows: $r_1 = \rho_1$, $r_2 = \rho_2$, $v = \sigma$ and

$$E(Y_t) = \mu = \alpha/(1 - r_1 - r_2). \quad (16)$$

We can verify this model equivalence with a simulation study. We generate data according to the DSEM model and estimate the data using the DSEM model and the RDSEM model followed by the transformation in equation (16). We generate 100 samples of size 500 using the model parameters given in Table 1. The results in this table show that both methods recover the true parameter values and the coverage is near the nominal level. The mean-squared error (MSE) of the parameter estimates, not presented here, is also identical between the two estimations. The

TABLE 1
Single Level AR(2) Model: Bias(Coverage)

Parameter	*True Value*	*DSEM*	*RDSEM*
α	1	.02(.97)	.02(.98)
r_1	0.5	.00(.95)	.00(.96)
r_2	0.2	.01(.96)	.00(.95)
σ	1	.00(.92)	.00(.95)

average parameter estimate for μ in the RDSEM model is 3.33 but after the reparameterization in (16), the correct value α is obtained. A similar table, confirming the equivalence of the two models, can be produced by generating data according to the RDSEM model and analyzing it with the DSEM model.

Note here that in the RDSEM model, the parameter v is not the residual variance of Y, although in the Mplus output it is printed as such. It is the residual variance of ε_t which in the Mplus language is denoted by $\hat{Y}$. The variance of the actual residual ε_t can be computed as in Appendix D in Asparouhov et al. (2018), and it is not the same as the parameter v.

The important difference between the two models is the fact that the variables in the RDSEM autoregressive equation (14) are centered (i.e. are with mean zero) while that is not the case for the DSEM autoregressive equation (11). The estimation algorithms for the two models are not identical. In the DSEM framework, the initial conditions Y_0 and Y_{-1} are treated as missing values. The initial conditions for ε_0 and ε_{-1} in the RDSEM model are treated as the average of the residuals ε_t, for $t > 0$.[1] These initial condition assumptions have minimal impact on the estimation when the time series is longer or the autoregressive coefficients are not large (although simulation studies not presented here indicate that the DSEM treatment is better for small sample size particularly when the posterior distribution for the autocorrelation parameters includes nonstationary models). In addition, some natural variation occurs in the MCMC estimation as different random numbers are used for the parameter updating. Nevertheless, the two estimation methods produce nearly identical results. The average DIC criterion in the above simulation is 1419 for the DSEM model and 1421 for the RDSEM model. Such a difference should be considered small and insignificant. Some of the difference is due to the different treatment of the initial conditions, and some is due to the sampling variation in the MCMC estimation.

Let us consider now the two-level autoregressive model. We use an AR(1) model for this illustration. Let Y_{it} be the observed variable for individual i at time t. The DSEM model can be written as

$$Y_{it} = Y_{b,i} + Y_{w,it} \tag{17}$$

$$Y_{w,it} = \rho Y_{w,i,t-1} + \varepsilon_{it} \tag{18}$$

$$Y_{b,i} \sim N(\mu, \sigma_b), \varepsilon_{it} \sim N(0, \sigma_w) \tag{19}$$

The RDSEM model, in this case, is written in exactly the same way. On the within level $Y_{w,it}$ is not regressed on another variable and it does not include the mean parameter. Therefore, it coincides with its residual $\hat{Y}_{w,it}$. The mean parameter μ is on the between level for both models. It is the mean of $Y_{b,i}$, which is excluded from the autoregressive equation (18). Therefore, when we estimate the above model using the DSEM or RDSEM model setup in Mplus, we can expect the parameters estimates to be the same.

[1] The RDSEM estimation can treat the initial conditions as missing variables as well using the option InitialCondition = sample.

TABLE 2
Two-Level AR(1) Model: Bias(Coverage)

Parameter	*True Value*	*DSEM*	*RDSEM*
μ	1	.00(.93)	.01(.89)
ρ	0.7	.00(.94)	.00(.93)
σ_w	1	.00(.96)	.00(.99)
σ_b	1	.01(.94)	.01(.93)

We illustrate this equivalence with a simulation study using 100 samples with $N = 200$ individuals and $T = 50$ time points generated with the DSEM model. The data are analyzed with the DSEM and the RDSEM models. The results, presented in Table 2, are nearly identical as expected. The estimates show minimal bias and the coverage is near the nominal level. Note that, unlike the single level AR model, there is no need for a reparameterization here. The mean parameter, for both models, is outside of the autoregressive part of the model. This, however, would not apply if the variable Y is a within the variable. That situation would be similar to the single level AR model.

The DIC average for the DSEM model is 28650 while for the RDSEM model it is 28722. This difference of 72 should not be interpreted as evidence that the DSEM model is better fitting than the RDSEM model. The difference is mostly due to the different treatment of the initial conditions. The average estimated number of parameters pD for the DSEM model is 265 and for the RDSEM model it is 159. This reflects the fact that in the DSEM estimation the 200 initial conditions are treated as parameters, while in the RDSEM estimation they are treated deterministically. It is possible to estimate the RDSEM model using the same initial condition treatment as the DSEM model. Using that option in Mplus the DIC and pD differences are negligible.

AUTOREGRESSIVE REGRESSION MODEL

In this section, we consider a two-level AR(1) regression model between a dependent variable Y_{it} and a covariate X_{it}. The DSEM regression model is defined as follows.

$$Y_{it} = Y_{b,i} + Y_{w,it} \tag{20}$$

$$X_{it} = X_{b,i} + X_{w,it} \tag{21}$$

$$Y_{w,it} = \beta_w X_{w,it} + r_y Y_{w,i,t-1} + \varepsilon_{it} \tag{22}$$

$$X_{w,it} = r_x Y_{w,i,t-1} + \xi_{it} \tag{23}$$

$$Y_{b,i} = \alpha + \beta_b X_{b.i} + \varepsilon_i \tag{24}$$

$$\varepsilon_{it} \sim N(0, \sigma_w), \xi_{it} \sim N(0, \psi_w), \varepsilon_i \sim N(0, \sigma_b), X_{b,i} \sim N(\mu, \psi_b) \tag{25}$$

The RDSEM model is defined similarly. The only difference is in equation (22), which for the RDSEM model is split into two equations.

$$Y_{w,it} = \beta_w X_{w,it} + \varepsilon_{it} \tag{26}$$

$$\varepsilon_{it} = r_y \varepsilon_{i,t-1} + \delta_{it} \tag{27}$$

$$\delta_{it} \sim N(0, \sigma_w) \tag{28}$$

The two models have the same number of parameters, however, the models are substantially different. The main difference is in how the dependent variable $Y_{w,it}$ relates to its predictors. In the DSEM model, the same variable from the previous period $Y_{w,i,t-1}$ affects $Y_{w,it}$ directly, while in the RDSEM model that effect occurs only through the residual from the previous period. This means that in the DSEM model, conditional on $X_{w,it}$, the covariate from the previous period $X_{w,i,t-1}$ can indirectly affect $Y_{w,it}$ through its effect on $Y_{w,i,t-1}$. In the RDSEM model the effect of the covariate is only contemporaneous, i.e., $X_{w,i,t-1}$ does not affect the value of $Y_{w,it}$, conditional on $X_{w,it}$. In that regard, the RDSEM model is more similar to a standard two-level model, which excludes the autocorrelation parameters (i.e. the model where $r_x = r_y = 0$).

We conduct a simulation study to evaluate the difference between the DSEM and the RDSEM models. First, we generate data using the DSEM model and analyze the data with the DSEM model, the RDSEM model, and the two-level SEM model that ignores the autocorrelations. We generate 100 samples with 200 individuals and 50 time points. For the data generation purposes, all residual variances are set to 1, while the intercept and mean parameters are set to 0. The results for the regression and the autoregressive parameters are presented in Table 3.

It is clear from these results that, when the data set comes from the DSEM model, the DSEM method outperforms the RDSEM and the two-level SEM methods. The DSEM method shows no bias, and the coverage is near the nominal level. The RDSEM method shows bias in the regression coefficients on both levels, and the bias for the two-level SEM method is even larger. The RDSEM model shows overestimation of the autocorrelation effect r_y. Surprisingly, the RDSEM and the two-level SEM estimates for the between level regression coefficient β_b are substantially biased, even though all the misspecifications occur only on the within level. The DIC criterion correctly picks the DSEM model as the best fitting model. The differences between the DICs for the three models are substantial.

Next, we conduct a simulation study, where the data are generated with the RDSEM model, using the same parameter values and sample size. Table 4 contains the results in this case.

The RDSEM model shows minimal bias and coverage near the nominal level as expected. The DSEM model is able to correctly estimate the between level effect β_b but yields biased estimates for the within level effect β_w and underestimates the autocorrelation coefficient r_y. We see that, regardless of the data generation method, the estimated autocorrelation coefficient r_y is higher for the RDSEM model. The two-level SEM model estimates correctly the within level effect β_w, but underestimates its standard error, which results in lower coverage. This is somewhat expected as the method does not account for the additional dependence between the data that is due to the autocorrelation. The two-level SEM model also produced a biased estimate for the between level effect β_b. This is also expected. When the autocorrelation is not accounted for, the measurement error for $X_{b,i}$ would be underestimated, under the assumption of independent observations within cluster. This results in a biased estimate for the between level regression. The DIC in this case again picks the correct model used for the data generation by a wide margin. Note that, in both simulations, the DIC for two-level SEM model is the worst.

Next, we introduce missing data in the above simulations. We generate 50% MCAR missing data for the dependent variable Y_{it}. The parameter estimates, in this case, are similar to the results in Tables 3 and 4 and we do not report these here. In this simulation, we study the ability of the DIC to distinguish between the models in the presence of

TABLE 3
Two-Level AR(1) Regression Model, DSEM Data: Bias(Coverage)

Parameter	*True Value*	*DSEM*	*RDSEM*	*Two-level SEM*
β_w	1	.00(.93)	−0.16(.00)	.81(.00)
β_b	−1	.01(.94)	.54(.09)	.68(.00)
r_y	0.7	.00(.95)	.22(.00)	–
r_x	0.7	.00(.93)	.00(.93)	–
DIC	–	57286	63242	79217

TABLE 4
Two-Level AR(1) Regression Model, RDSEM Data: Bias(Coverage)

Parameter	*True Value*	*DSEM*	*RDSEM*	*Two-level SEM*
β_w	1	−.29(.00)	.00(.92)	.00(.77)
β_b	−1	.05(.96)	−.05(.93)	.29(.20)
r_y	0.7	−.26(.00)	.00(.92)	–
r_x	0.7	.00(.91)	.00(.88)	–
DIC	–	59694	57437	68778

TABLE 5
Two-Level AR(1) Regression Model with Missing Data

True Model	*DSEM*	*RDSEM*	*Two-level SEM*	*Two-level SEM within DSEM*
DSEM	62175	69098	56854	84057
RDSEM	65790	62392	51693	73661

missing data. Table 5 contains the results of this simulation. The first thing we notice here is that the two-level SEM model has the smallest DIC values among the models. These DIC values are not comparable to the DIC values of the DSEM and RDSEM models. That is because the missing values in the DSEM and RDSEM computation of the DIC, are treated as random effects, i.e., they are treated as model parameters. In the two-level SEM model, the DIC is constructed using the marginal likelihood for the observed data only, while in the RDSEM/DSEM models we use the likelihood for the missing data as well. One can observe that difference simply by looking at the estimated number of parameters. In the two-level SEM estimation of the DSEM generated data, the average pD value is 363, which corresponds to (not precisely) the 400 random effects in the model (200 individuals times two random effects: $Y_{b,i}$ and $X_{b,i}$). On the other hand, for the DSEM and RDSEM models, pD is over 5000, which essentially adds the number of missing values to the number of random effects. In principle, the DIC is best computed with as few parameters as possible, because the value of the DIC is then more precise and it would take fewer number of MCMC iterations to estimate it well. However, in the case of the DSEM/RDSEM models, the marginalization of the likelihood is much more complicated than it is for the two-level SEM model. Since the DIC computation involves computing the likelihood in every MCMC iteration, such a complication could result in a substantially slower estimation. Therefore, in the current implementation of the DSEM/RDSEM models in Mplus, the likelihood is not marginalized and the missing values are treated as random effects. Because of that different treatments of the missing data, the DIC of the two-level model and the DIC of the DSEM/RDSEM models are not comparable. The DIC values between the DSEM and the RDSEM models are comparable since both treat the missing data the same way. Note here that this DIC analysis is specific to the Mplus MCMC implementation and may not be transferable to other software implementations.

It is possible to estimate the two-level SEM model, however, and obtain a comparable DIC value. Since the two-level SEM model is a special case of the DSEM/RDSEM models, we can estimate the two-level SEM model as a DSEM model (or an RDSEM model), where the parameters r_x and r_y are fixed to zero. We call this model the two-level SEM within the DSEM model. The DIC value for this model, also reported in Table 5, is directly comparable to the DIC values of the DSEM/RDSEM models. We can now perform the model comparisons among the three models, and we see in Table 5 that again the DIC criterion clearly identifies the correct model. That is, when we generate data with the DSEM model, the smallest DIC value is obtained by the DSEM model, and when we generate data with the RDSEM model, the smallest value is obtained with the RDSEM model. Note also that if the data set contains a large amount of missing data, the DIC may fluctuate depending on the number of iterations used in the MCMC estimation. It is recommended that these fluctuations are evaluated through selecting different random seeds (BSEED command in Mplus) and that any inference regarding model fit is based on DIC differences much larger than the DIC fluctuations based on the different random seeds. Alternatively, a very large number of iterations can be used for the estimation (using the FBITER command) so that the DIC value is estimated more precisely.

In summary, we see from the above simulations that the DIC criterion can be used to identify the proper modeling framework for a particular data set. In Asparouhov et al. (2018) the topic of adding covariates in time series models is also discussed in details. In particular, examples are discussed where the features of the RDSEM model and the DSEM model can be combined in one model nested above both models. The covariate can be included within the autoregressive process as in the DSEM model, and in addition, the covariate can be included outside of the autoregressive process as in the RDSEM model. The correct modeling framework can then be selected by evaluating the significance of the two regression coefficients. This method provides a valuable alternative to the DIC method, however, the combined model is generally slower to estimate and is likely to require longer time-series data.

EXOGENEITY

In most traditional statistical models, it is generally preferred that the covariate X is treated as an exogenous variable. The distribution of the covariate is not included in the model. That way, we avoid any problems that could be caused by incorrectly specifying the model for the covariate. Instead of modeling the full joint distribution for the X and Y variables, we would model only the conditional distribution of $[Y|X]$, which is also more parsimonious and is sufficient for inference in most situations. The question we want to address in this section is whether this exogeneity approach can hold up well for the DSEM and the RDSEM models. In the regression models discussed in the previous section, the covariate is not truly an exogenous variable because of the estimation of the random between level component $X_{b,i}$, the autocorrelation r_x, and potentially,

the missing data for the covariate X. Note that the missing data can be caused in the time-series framework, not just because the data is missing, but also if the times of observations are unequally spaced and missing values are inserted for the periods where observations were not taken. This is achieved in Mplus with the TINTERVAL option. Therefore, the missing data aspect is even more important in time series settings than it is with cross-sectional modeling.

Using an exogeneity approach for the random effect $X_{b,i}$ amounts to using the average $\overline{X_{i.}}$ instead of $X_{b,i}$. This approach is known as the observed centering method and is discussed in detail in Asparouhov and Muthén (2019). It is shown in that article that using such an approach in the RDSEM framework can lead to large biases in the regression parameter estimates on the between level. Thus, the exogeneity approach does not work well for the random effect $X_{b,i}$ and it is preferable that the covariate is modeled explicitly, using latent centering and the random effect $X_{b,i}$.

In this section, we evaluate the other two components of the exogeneity, namely, the missing data and the autocorrelation. We want to determine if it is necessary to include autoregressive modeling for the covariate, in addition to the autoregressive modeling for the dependent variable, and how that is affected by the presence of missing data. For completeness, we evaluate the impact on the model estimation of not just r_x, but also r_y, with and without missing data.

To do that, we generate 100 data sets with 200 individuals and 60 time points, using the DSEM and RDSEM models. The mean and intercept parameters are set to 0. The residual variance and regression parameters are set to 1. The autocorrelation parameters are set to 0.7. We consider the estimation under four different circumstance: no missing data, 50% missing data for Y, 50% missing data for X, and 25% missing data for both variables. The missing data for Y and X are generated as MCAR, where the values are missing with fixed probability. MAR missing data can be generated as well, however, in this particular study, the more advanced missing data mechanism does not reveal any new information that is not already visible with the MCAR missing data. Therefore, we restrict the discussion to the MCAR missing data. The models are estimated in four different ways: with r_x and r_y included, with only r_y included, with only r_x inclu ded, and with both autocorrelations excluded. Thus, we study four conditions for the autocorrelation modeling under the four missing data scenarios for a total of 16 different simulation studies. The performance of the DSEM and RDSEM models is studied separately. Data generated with the DSEM model are analyzed with the DSEM model. Data generated with the RDSEM model are analyzed with the RDSEM model.

TABLE 6
DSEM Regression Model: Absolute Bias for β_1/β_2. Bold Font Represents Drop in Coverage below 90%

Estimated Autocorrelations	*r_x and r_y*	*r_y only*	*r_x only*	*none*
No missing	.00/.01	.00/**.29**	**.83/.23**	**.83/.32**
Missing on Y	.00/.03	.00/**.29**	**.83/.23**	**.83/.32**
Missing on X	.00/.02	**.08/.27**	**.88/.14**	**.83/.32**
Missing on X and Y	.00/.02	**.01/.28**	**.86/.21**	**.83/.32**

TABLE 7
RDSEM Regression Model: Absolute Bias for β_1/β_2. Bold Font Represents Drop in Confidence Interval Coverage below 90%

Estimated Autocorrelations	*r_x and r_y*	*r_y only*	*r_x only*	*none*
No missing	.00/.00	.00/.00	**.00**/.00	**.00**/.00
Missing on Y	.00/.00	.00/.00	**.00**/.00	**.00**/.00
Missing on X	.00/.00	**.34**/.00	**.02**/.01	**.00**/.00
Missing on X and Y	.00/.00	**.12**/.00	**.00**/.01	**.00**/.00

Table 6 contains the results for the DSEM model. We see here that the model is not robust to any misspecifications, and it does not allow exogeneity. The distribution of both the covariate and the dependent variable has to be modeled properly. If either of the two autocorrelations is not included in the model, we can expect large biases in the parameters, even without missing data.

Table 7 contains the results for the RDSEM model. The situation here is dramatically different. The model is far more robust to misspecifications. Even if none of the autocorrelations are included in the model, the parameter estimates are unbiased. The same holds for the case when only r_x is included. We can also see, however, that if r_y is not included in the model, even though the parameter estimates are unbiased, the standard errors for β_1 are biased, which results in low confidence interval coverage. In addition, if there are missing data for X, and r_y is included in the model, it is very important to also include r_x; otherwise, the parameter estimates can be quite biased. In fact, Table 7 shows that if there are missing data for X, it is better to not include r_x and r_y at all, than to include only r_y (two wrong almost make one right if we ignore the bias in the standard error of β_1). We also see here that if there are no missing data for X, we can indeed exclude the r_x coefficient from the model and essentially treat the covariate as an exogenous variable. This conclusion unfortunately only applies to the situation when there is no contextual effect ($\beta_1=\beta_2$). To show the impact of the contextual effect on the results of Table 7, we repeat the above simulation for the RDSEM model using data generated with β_2 set to -1, which creates the contextual effect. The results of this

TABLE 8
RDSEM Regression Model with Contextual Effect: Absolute Bias for β_1/β_2. Bold Font Represents Drop in Confidence Interval Coverage below 90%

Estimated Autocorrelations	r_x *and* r_y	r_y *only*	r_x *only*	*none*
No missing	.00/.02	.00/**.26**	**.00**/.02	**.00/.26**
Missing on Y	.00/.02	.00/**.26**	**.00**/.02	**.00/.26**
Missing on X	.00/.02	**.34/.27**	**.02/.06**	**.00/.26**
Missing on X and Y	.00/.02	**.12/.26**	**.00**/.03	**.00/.26**

simulation are presented in Table 8. We see here that the prospects of treating X as an exogenous variable are further diminished and we no longer get acceptable results for the parameter estimates when r_x is excluded, even if there is no missing data. This bias is due to the fact that, when r_x is excluded, the measurement error in $X_{b,i}$ is underestimated, which results in biased estimates for β_2. As the number of time points increases, however, this bias will decrease.

In summary, the presence of missing data on X or contextual effect in the RDSEM regression requires the inclusion of autocorrelation modeling for the covariate as well. Including the autocorrelations for both X and Y in the RDSEM and DSEM models, yields satisfactory results in all situations. The prospects for treating a covariate as truly exogenous variable are very limited in the time series settings. Instead, the time-series regression should be treated as a path analysis model where both variables are treated as dependent and are properly fitted. This revelation is in stark contrast with traditional SEM models where treating a covariate as an exogenous variable is quite beneficial.

THE EFFECT OF UNEVENLY SPACED TIMES OF OBSERVATIONS ON THE CONTEMPORANEOUS EFFECTS

In this section, we study the effect of unevenly spaced time of observations on the contemporaneous relations between the variables for the DSEM and the RDSEM models. Unevenly spaced time of observations is handled in Mplus through the TINTERVAL command which inserts missing values in the data for the unobserved periods. The TINTERVAL command specifies the length of the time period δ for the analysis, and it affects the amount of missing data inserted within the original data. The estimation is more accurate with smaller δ values because the distance between the observed values is approximated more precisely. The smaller the δ value is, however, the more missing data are inserted between the observed values, which can produce slower estimation in terms of convergence. It is shown in Asparouhov et al. (2018) that up to 95% missing data can be inserted in the data, however, more complex models may allow only smaller amounts of missing data.

Here, we conduct a simulation study to evaluate the impact of the time interval δ on the contemporaneous relations in the model. Contemporaneous relations involve variables observed at the same time period. A priori, one can assume that δ may not have an effect on such relations. However, in time-series models, the variables are intra connected and this intuitive argument may not hold. For this simulation study, we use the regression model discussed in the previous section. The mean and the intercept parameters are set to 0. The residual variance parameters are set to 1. The within level regression parameter is set to 1. The between-level regression parameter is set to −1. The autocorrelation parameters are set to 0.7. For each simulation, we generate 100 samples with 200 individuals. To generate unevenly spaced times of observations, we generate 240 evenly spaced observations, where the time period is 1, and we randomly remove 75% of the observations, reducing the number of observations per individual to an average value of 60 values, which are unevenly spaced. We generate data using the DSEM model and analyze the data with the DSEM model. Similarly, we generate data with the RDSEM model and analyze the data with the RDSEM model.

To show the effect of the TINTERVAL setting δ, we analyze the data using $\delta = 1, 2$ and 100. The first setting $\delta = 1$ represents the original setting used for the data generation and we expect the model to be recovered reasonably well, including the autoregressive coefficients. The setting $\delta = 2$ represents a medium crude timescale where we can expect the autocorrelations to be lower. The setting of $\delta = 100$ represents the case where the times of observations are ignored. In that case, the observations are treated as if they come from consecutive periods. In Mplus this can be done in two ways. It can be done by not specifying the TINTERVAL option at all. Equivalently, it can be done by specifying the TINTERVAL as a very large value. In this situation, the value of 100 is sufficiently large as it is very unlikely that there are two consecutive observed values more than 100 periods apart. Here again, we can expect that the autocorrelation coefficients will be underestimated. Nevertheless, the focus of this study is on the structural effects β_1 and β_2.

The results of the simulation study are presented in Table 9. First, we see that the between level effect β_2 is not affected by the choice of δ with either the RDSEM

TABLE 9
The Effect of TINTERVAL on the DSEM and RDSEM Regressions: Absolute Bias(Coverage)

method	δ	β_1	β_2
DSEM	1	.00(.92)	.01(.97)
DSEM	2	.37(.00)	.01(.93)
DSEM	100	.65(.00)	.01(.92)
RDSEM	1	.00(.89)	.02(.96)
RDSEM	2	.00(.88)	.02(.97)
RDSEM	100	.00(.92)	.01(.96)

model or the DSEM model. The parameter estimates are unbiased and the coverage is near the nominal level. We also see that the RDSEM model produces correct estimates and coverage for β_1 as well with any δ value. While the residual autocorrelations are affected by the time frame set by δ, the structural parameters are not affected. Thus, we conclude that the δ setting with the RDSEM model is somewhat inconsequential.

On the other hand, the δ setting is important for the DSEM model. Only the correct setting of δ is able to recover the generating model. Clearly, the regression parameter on the within level depends on the δ value. This dependence is important and should be carefully considered as the δ value is chosen in the DSEM framework, not only in regards to model estimates, but also in regards to the interpretation of the model. That is because $Y_{w,i,t-1}$ is a predictor for $Y_{w,it}$. It matters what the length of the time period is and what time this predictor refers to. In fact, in the DSEM settings, the relationship between Y and X is not purely contemporaneous. It would have been, if the two predictors of $Y_{w,it}$, namely $Y_{w,i,t-1}$ and $X_{w,it}$, are independent. However, they are not independent as they are both regressed on $X_{w,i,t-1}$. The correlation between the two predictors would also affect the regression coefficients on the within level and that correlation is affected by the δ setting. Looking at this also from the perspective of the covariates, we know that the relationship between X and Y is not summarized only by the β_1 coefficient and that the previous periods covariate $X_{w,i,t-1}$ has an effect on $Y_{w,it}$ that goes through its effect on $Y_{w,i,t-1}$. Therefore, the relationship between Y and X in the DSEM model is not purely contemporaneous, and it extends beyond the current period. Subsequently, this yields model estimates for the DSEM model that depends on the δ scale.

The DSEM model dependence on the δ value can be viewed as a detriment when the δ value is difficult to set. However, model fit, as measured by the DIC criterion, should be the main method for selecting a model, rather than the stability of the estimates. While the DIC cannot be used to determine the value of δ due to the different amount of missing data used for the estimation, the DIC can be used to compare the DSEM and RDSEM models with the same δ setting. To illustrate this point we compare the DSEM and RDSEM models with DIC on the data sets generated with the different models using the $\delta = 2$ setting, i.e., a medium crude time frame. The results of this comparison are reported in Table 10. For both types of generated data, the DIC criterion identified the correct modeling framework, even though the timescale is not as precise as it should be. In particular, the DSEM model which showed different estimates for the regression coefficient on the within level is still identified as the substantially better fitting model for the DSEM generated data.

TABLE 10
DIC Comparison for DSEM and RDSEM Models with $\delta = 2$

	DSEM model	*RDSEM model*
DSEM data	180602	192537
RDSEM data	175617	171976

We emphasize here that the β_1 dependence on the length of the time interval with the DSEM model should not be interpreted as if this is a biased estimate or as if something is wrong with the DSEM estimation. Any autoregressive coefficient depends on the length of the time interval. In the DSEM model, the autoregressive part is not separated from the structural part as it is in the RDSEM model. In the DSEM model, the autoregressive and the structural part are intra-connected. Even when two variables in the regression are from the same time period in the DSEM model, the regression coefficient may depend on other coefficients in the model that are autoregressive in nature and depend on the length of the time period.

COMPARING REML AND RDSEM: THE EFFECT OF RANDOM AUTOREGRESSIVE COEFFICIENT AND RANDOM RESIDUAL VARIANCE

The REML estimation for multilevel models, see Raudenbush and Bryk (2002), has traditionally been used to estimate longitudinal models, particularly for studies with a small number of clusters. In addition to estimating random intercepts and slopes, the REML method can be used to estimate the residual autocorrelation for the dependent variable. The method is implemented in various statistical packages such as SPSS, HLM, and SAS, see Bolger and Laurenceau (2013). The RDSEM framework is more flexible than the REML estimation, and it can be used to model individual variation in the autocorrelation parameter as well as individual variation in the residual variance. In this section, we explore the consequences of ignoring these individual-level variations. This is particularly important because the REML estimation is the most commonly used approach for longitudinal modeling.

In Jongerling, Laurenceau, and Hamaker (2015) it is shown that ignoring the individually specific residual variance, in the context of the two-level AR(1) model, can result in bias in the mean of the autocorrelation parameter when the individual-specific residual variance is correlated with the random autocorrelation. In Asparouhov et al. (2018) it is shown that this also results in much larger MSE for the subject-specific autocorrelation, i.e., the distortion of the random autoregressive coefficient is not a simple shift but rather a substantial misestimation when evaluated on the individual level. In this section, we explore the consequences of ignoring the across subject variability in both the autocorrelation coefficient as well

as the residual variance in the context of the time-series regression model.

In Asparouhov and Muthén (2019) it is shown that using observed centering for the covariate with the REML estimator can result in a substantial bias in the regression coefficients. To ensure that these biases are not interfering with this study, we use a within-level covariate with no autoregressive effect. To be more specific we consider the following two-level regression model.

$$Y_{it} = \alpha_i + \beta_1 X_{it} + \varepsilon_{it} \tag{29}$$

$$\varepsilon_{it} = r_i \varepsilon_{i,t-1} + \xi_{it} \tag{30}$$

$$\xi_{it} \sim N(0, \sigma_{w,i}), X_{it} \sim N(0, \psi), \alpha_i \sim N(\alpha, \sigma_b) \tag{31}$$

$$l_i = log(\sigma_{w,i}) \sim N(\sigma_w, v_1), r_i \sim N(r, v_2), Cov(r_i, l_i) = v_{12}. \tag{32}$$

In the above model, the autoregressive parameter r_i varies across individuals. In addition, the variance of the residual ξ_{it} varies across individuals and has a log-normal distribution. As in Jongerling et al. (2015), we allow these two random effects to be correlated through the parameter v_{12}. We use a non-random regression coefficient β_1. Simulation studies with random regression coefficients did not reveal any additional information not already seen with the non-random coefficient and therefore we limit this discussion to a non-random regression coefficient β_1.

We conduct two separate simulation studies. In the first simulation study, we evaluate the effect of ignoring the random autoregressive coefficient alone. To do that, we use the above model with $v_1 = v_{12} = 0$, i.e., there is no subject-specific residual variance. In the second simulation study, we use the full model without these restrictions. We generate 100 data sets using the above model with N individuals and T time points per individual. We generate smaller sample size data sets with $N = 100$ and $T = 30$ and larger sample size data sets with $N = 500$ and $T = 100$.

First, we generate data without random residual variance using the following model parameters: $\beta_1 = 1$, $Var(\xi_{it}) = 1$, $\alpha = 0, \sigma_b = 1, r = 0.4, v_2 = 0.02$. We analyze the data using the RDSEM method, where the estimated model is the same as the data generating model, i.e., we allow for the random autocorrelation. We also analyze the data using the REML approach where the autocorrelation is assumed to identical across the individuals. Because the REML approach is asymptotically equivalent to the ML approach, see Raudenbush and Bryk (2002), which is asymptotically equivalent to the Bayesian approach, see Berger (1985), we can estimate the REML approach within the RDSEM framework by constraining the autocorrelation variance to zero or more precisely we estimate a non-random autocorrelation coefficient.

The results of this simulation study are presented in Table 11. The RDSEM approach performs well. The parameter estimates are unbiased, and the coverage is near the nominal level. The REML approach performed just as well for the β_1 parameter, however, the estimate of the autoregressive coefficient has a small bias and the coverage for that parameter is poor, particularly for the large sample where the coverage dropped to only 8%. In addition, the SMSE (square root of the mean-squared error) which measures the distance between the estimate, and the true value is much worse for the REML approach in the small and the large samples. Interestingly, the error in the REML estimate for the autocorrelation parameter increased in the large sample. In summary, ignoring the individual variation in the autocorrelation parameter in the time-series regression model, as in the REML approach, can lead to bias estimate of the autocorrelation parameter, in addition to poor coverage, and loss of efficiency. On the other hand, the regression coefficient itself is unaffected by this limitation of the REML approach.

Next, we consider the implications for the REML approach when the residual variance in the regression is subject specific. We generate data according to model (29–32) using the same parameters as in the previous simulation study and now we include the parameters $v_1 = 1$, $v_{12} = 0.08$ and $\sigma_w = 0$. Thus, the residual variance on the within level is a log-normal random effect with mean 0 and variance 1 which is also correlated with the random autoregressive coefficient. With the RDSEM method, we estimate the model (29–32) allowing for the subject-specific variation in the autoregressive parameter and the residual variance. With the REML approach, we estimate both of these parameters as non-random effects.

The results of this simulation study are reported in Table 12. The RDSEM approach performs very well in this case as well, while the REML approach shows large bias in the autocorrelation coefficient in addition to poor coverage, and a substantial increase in the mean-squared error. In addition, the REML approach, in this case, shows larger MSE for the regression parameter β_1, while the parameter estimate remains unbiased and the coverage is near the nominal level. Thus, we conclude that the REML estimation is less efficient for the regression parameter β_1 due to ignoring the individual variation of the autoregressive and residual variance parameters. In such situations, the autoregressive estimate is unreliable with the REML estimation.

TABLE 11
Time-Series Regression with Random Autocorrelation: Absolute bias/coverage/SMSE

Parameter	*N*	*T*	*RDSEM*	*REML*
r	100	30	.00/.95/.023	.02/.73/.033
β_1	100	30	.00/.96/.017	.00/.95/.017
r	500	100	.00/.95/.008	.03/.08/.071
β_1	500	100	.00/.95/.004	.00/.94/.004

TABLE 12
Time-Series Regression with Random Autocorrelation and Residual Variance: Absolute bias/coverage/SMSE

Parameter	*N*	*T*	*RDSEM*	*REML*
r	100	30	.01/.91/.029	.10/.73/.107
β_1	100	30	.00/.96/.013	.00/.95/.021
r	500	100	.00/.97/.008	.11/.00/.114
β_1	500	100	.00/.94/.003	.00/.96/.005

COMPARING REML AND RDSEM FOR CATEGORICAL DATA

The REML estimation problems for repeated measurement models can be summarized as follows. In Asparouhov and Muthén (2019) it is shown that using observed centering for the covariate with the REML estimation can lead to biased results for the contextual effect due to ignoring the sampling error of the mean, due to ignoring the additional sampling error caused by the autocorrelation in the covariate, and due to missing data in the covariate. In the previous section, we also illustrated that the REML estimator is biased and yields poor coverage results when the autocorrelation parameter varies across the level 2 units and when the residual variance on the within level varies across the level 2 units. In this section, we illustrate an additional problem with the REML estimator that is specific to the categorical data estimation, and this problem is independent of all other REML estimation problems. It turns out that REML severely underestimates the autocorrelation parameter for categorical variables.

Categorical data in the DSEM framework, based on the tetrachoric and polychoric autocorrelation are discussed in Asparouhov and Muthén (2019). In this section, we introduce categorical variables in the RDSEM framework. Consider first the standard two-level model for a binary outcome and a single covariate. Let Y_{it} be the observed binary variable for individual i at time t and X_{it} be the corresponding covariate. Assume that Y_{it} takes the values 0 and 1. The standard two-level model is given by the following equation.

$$P(Y_{it}=1)=\Phi(\alpha_i+\beta X_{it}), \tag{33}$$

where Φ is the probit distribution function, $\alpha_i \sim N(\alpha,\psi)$ represents the random intercept, and β is the probit regression slope. An alternative but an equivalent way to represent the above model is via the underlying continuous variable Y^*_{it} as follows.

$$Y_{it}=1 \Leftrightarrow Y^*_{it}>0 \tag{34}$$

$$Y^*_{it}=\alpha_i+\beta X_{it}+\varepsilon_{it}, \tag{35}$$

where $\varepsilon_{it} \sim N(0,1)$. In the standard two-level model, ε_{it} are assumed independent variables but in the repeated measures model we assume that

$$Cov(\varepsilon_{it_1},\varepsilon_{it_2})=r^{|t_2-t_1|}, \tag{36}$$

where r is the tetrachoric autocorrelation. This means that the error term ε_{it} is an AR(1) time-series process. The above model has four parameters: α, β, ψ and r.

The RDSEM model implemented in Mplus 8.2 with categorical variables aims to estimate the above model, however, it has a slightly different parameterization.

$$Y^*_{it}=a_i+bX_{it}+\hat{Y}^*_{it}, \tag{37}$$

where $a_i \sim N(a,\theta)$ and

$$\hat{Y}^*_{it}=r\hat{Y}^*_{i,t-1}+\varepsilon_{it}. \tag{38}$$

In this parameterization ε_{it} are assumed independent standard normal variables which means that $\hat{Y}^*_{it}$ is an AR(1) process with the autocorrelation parameter r. Note, however, that

$$Var(\hat{Y}^*_{it})=1/(1-r^2) \tag{39}$$

and

$$Cor(\hat{Y}^*_{it_1},\hat{Y}^*_{it_1})=r^{|t_2-t_1|}. \tag{40}$$

Models (34–36) and (37–40) are equivalent, and the relationship between the two models is given by the following equations.

$$\alpha=a\sqrt{1-r^2} \tag{41}$$

$$\beta=b\sqrt{1-r^2} \tag{42}$$

$$\psi=\theta(1-r^2). \tag{43}$$

This rescaling of the parameters is necessary due to the fact that the variance of the residual in equation (37), i.e. $Var(\hat{Y}^*_{it})$, is not 1 as it is in equation (35). In addition, note that this reparameterization is needed also when the RDSEM model (37–40) is compared with the standard two-level probit model with no serial correlation. The standard two-level probit model is essentially model (34–36) with r fixed to 0. An alternative way to compare the RDSEM binary probit model to the standard two-level model is to compare the standardized estimates obtained in Mplus via the option "OUTPUT: STDYX". Such a comparison, however, is different from the relationships given in equations (41–43). Equations (41–43) essentially put both models in the "theta" parameterization where $res-Var(Y^*_{it})$ is 1, while the standardized estimates

represent the ”delta” parameterization where $Var(Y^*_{it})$ is 1. More details on the various parameterizations used and implemented in Mplus with categorical variables can be found in Muthén and Asparouhov (2002).

Bolger and Laurenceau (2013) show how to estimate model (34–36) using the GLIMMIX procedure in SAS and the GENLINMIXED procedure in SPSS. Both methods appear to use REML estimation via a quasi-likelihood approach and yield nearly identical results. In this section, we will evaluate the performance of this method and we will use the SPSS implementation. We will also compare that method to the Bayesian estimation of the RDSEM model implemented in Mplus 8.2. Note however that the RDSEM estimates are rescaled to the “theta” parameterization using equations (41–43) which makes the model estimates comparable between Mplus and REML. In this simulation study, we are only interested in determining the quality of the point estimates. Additional simulation studies, not reported here, aimed at the standard error results did not reveal any additional information not already visible in the point estimates. Thus, in this simulation study, we focus only on the point estimates of the repeated measures probit regression. To evaluate the bias in the point estimates we can simply estimate the model once using a very large sample, instead of estimating the model with multiple smaller samples. With a large sample, we essentially obtain the asymptotic behavior of the estimator. In this simulation, we generate samples with $N = 500$ observations and $T = 50$ time points. This means that on the within level the sample size is 25000 and the autocorrelation parameter which is determined at the within level would be estimated quite precisely.

To avoid the existing known problems with the REML estimator, we use a covariate that is already centered, i.e., we generate data where the mean of the covariate is the same across all observations/clusters. In the Mplus language, this condition is obtained by specifying the covariate as a within-only variable. We also assume that the autoregressive coefficient is constant across individuals to avoid the problems reported in the previous section. We generate data according to model (37–40) using the parameter values $a = b = \theta = 0.5$ and autocorrelation parameter r values 0.1, 0.3, 0.5, and 0.7. The results of the estimation are reported in Table 13 for the RDSEM method and in Table 14 for the REML method.

TABLE 13
Repeated Measures Probit Regression: Absolute Bias for Mplus-RDSEM

Parameter	$r = .1$	$r = .3$	$r = .5$	$r = .7$
r	.01	.00	.00	.01
α	.00	.00	.00	.01
β	.00	.01	.01	.00
ψ	.01	.01	.01	.02

TABLE 14
Repeated Measures Probit Regression: Absolute Bias for SAS/SPSS-REML

Parameter	$r = .1$	$r = .3$	$r = .5$	$r = .7$
r	.05	.15	.22	.28
α	.01	.01	.00	.02
β	.00	.01	.00	.01
ψ	.00	.00	.02	.05

The RDSEM results are approximately unbiased, while the REML results show large bias for the autocorrelation parameter. In this simulation study, the autocorrelation parameter is underestimated by approximately 50%. In addition, when $r = .7$, we see that this underestimation of the autocorrelation parameter results in overestimation of the between level variance parameter ψ. This is expected because, in general, if one type of correlation is underestimated, another type of correlation will be overestimated as an attempt to compensate for the discrepancy between the data and the model. The parameters α and β appear to be unaffected, however, by the autocorrelation underestimation with the REML estimator. In Bauer and Sterba (2011), the SPSS/SAS-REML method is recommended over the maximum-likelihood estimation because of its ability to incorporate the autocorrelation parameter in two-level regression models with categorical variables. Because of its poor performance, however, as shown in Table 14, we can not affirm this recommendation and instead we recommend the RDSEM method based on the Bayesian estimation. Note, also that the RDSEM framework has numerous other advantages over the SPSS/SAS-REML method. Among these are multivariate modeling, latent variable modeling, fully structural path analysis modeling, MAR missing data modeling, multivariate autoregressive error structures, random autoregressive structures, latent centering with and without random slopes, and it is also computationally faster.

Because the autocorrelation parameter is on a correlation scale, it is difficult to argue that the REML autocorrelation parameter is from a different parameterization or that it should be interpreted differently. When multilevel models are estimated with categorical data and the probit link function, a random effect is estimated to represent a random intercept/threshold value. This modeling approach allows us to account for the correlations that exist among observations from the same cluster. The correlation is in fact modeled as a tetrachoric or polychoric correlation, i.e., this correlation is the correlation for the underlying continuous variable. In repeated measurement models, the autocorrelation should be modeled as well and the most natural approach to model that is again on the underlying continuous metric as tetrachoric or polychoric autocorrelation.

COMBINING RDSEM AND DSEM

In Mplus 8.2 it is generally not possible to combine the two frameworks. The DSEM model uses the & symbol for the lagged predictor variables and the RDSEM model uses the symbol to denote the lagged residuals. It is not possible in Mplus 8.2 to estimate a model with both of these features. Some such models, however, can be estimated via model reparameterization. In this section, we illustrate this point with a simple regression model.

Suppose that a variable Y_t is to be regressed on two covariates X_t and Z_t. Suppose that we are interested in the model where the Y_t regression on X_t is to be estimated as a DSEM regression while the Y_t regression on Z_t is to be estimated as an RDSEM regression. The regression of Y_t on X_t is a dynamic model where prior period variables are used in the regression as well. The regression of Y_t on Z_t, on the other hand, is excluded from the autoregressive process. This difference amounts to the fact that X_t and all prior period values of X_t can affect Y_t, while Z_t affects Y_t only through the current period value. To be more specific we consider the model.

$$Y_t = \beta_1 Z_t + \varepsilon_t \tag{44}$$

$$\varepsilon_t = \rho_y \varepsilon_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \zeta_t \tag{45}$$

$$X_t = \rho_x X_{t-1} + \xi_t \tag{46}$$

$$Z_t \sim N(0, \sigma_z), \zeta_t \sim N(0, \theta), \xi_t \sim N(0, \psi). \tag{47}$$

We exclude mean and intercept parameters from the above model because in the typical two-level application all intercept and mean parameters are on the between level and are not in the time series model. The model between Y_t and Z_t is the typical RDSEM regression model, while the model between ε_t and X_t is the typical SVAR (structural vector autoregressive) model, see Lütkepohl (2007), that can be estimated in the DSEM framework. The SVAR model is also equivalent to the VAR model considered in Hamaker et al. (2018) where instead of having the contemporaneous regression coefficient β_2 the residuals covariance between ξ_t and ζ_t is estimated.

The above model is not directly available in Mplus 8.2, but we can estimate the following equivalent model.

$$Y_t = b_1 Z_t + b_2 X_t + \hat{Y}_t \tag{48}$$

$$\hat{Y}_t = r_y \hat{Y}_{t-1} + b_3 X_{t-1} + \zeta_t \tag{49}$$

$$X_t = r_x X_{t-1} + \xi_t. \tag{50}$$

To see the equivalence, we set $\varepsilon_t = b_2 X_t + \hat{Y}_t$. We can now rewrite equation (48) as

$$Y_t = b_1 Z_t + \varepsilon_t. \tag{51}$$

In addition, $\hat{Y}_t = \varepsilon_t - b_2 X_t$. If we use this expression for $\hat{Y}_t$ and $\hat{Y}_{t-1}$ in equation (49), we get that

$$\varepsilon_t = r_y \varepsilon_{t-1} + b_2 X_t + (b_3 - r_y b_2) X_{t-1} + \zeta_t, \tag{52}$$

which establishes the equivalence of the two models. In fact, we see that all the parameters are unchanged except one. The reparameterization that yields the equivalence of the two models is $\rho_y = r_y$, $\rho_x = r_x$, $\beta_1 = b_1$, $\beta_2 = b_2$ and $\beta_3 = b_3 - r_y b_2$. The residual variance parameters are also unchanged. Note also that because we did not include the mean or other predictors for the covariate X_t, the variable X_t is the same as the variable $\hat{X}_t$ in the Mplus language.

In summary, if we have a set of predictors that we want to use as in an DSEM model and a set of predictors that we want to use as in an RDSEM model, the above model equivalence suggests that the RDSEM framework can be used to essentially do that. The reparameterization between the two models above could be used to present the model in whichever way is most intuitive for the particular application.

We illustrate further the above modeling technique with a simulation study based on the RDSEM representation (48–50). First, we extend the model to two-level settings as follows.

$$X_{it} = X_{b,i} + X_{w,it} \tag{53}$$

$$Y_{it} = Y_{b,i} + Y_{w,it} \tag{54}$$

$$Y_{w,it} = \beta_1 Z_{it} + \beta_2 X_{w,it} + \hat{Y}_{w,it} \tag{55}$$

$$\hat{Y}_{w,it} = \rho_y \hat{Y}_{w,i,t-1} + \beta_3 X_{w,i,t-1} + \zeta_{w,it} \tag{56}$$

$$X_{w,it} = \rho_x X_{w,i,t-1} + \xi_{it}. \tag{57}$$

$$Y_{b,i} = \alpha + \beta_4 X_{b,i} + \zeta_{b,i} \tag{58}$$

$$Z_t \sim N(\mu_z, \sigma_z), \zeta_{w,it} \sim N(0, \theta_w), \zeta_{b,i} \sim N(0, \theta_b) \tag{59}$$

$$\xi_{it} \sim N(0, \psi_w), X_{b,i} \sim N(\mu_x, \psi_b) \tag{60}$$

Using the above model, we generate 100 data sets with $N = 100$ individuals and $T = 20$ time points per individual. The following parameter values are used in this simulation study: $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 0.2$, $\beta_4 = 0.5$, $\rho_x = 0.5$, $\rho_y = 0.3$, $\alpha = 0$, $\mu_x = 0$, $\mu_z = 0$, $\sigma_z = 1$, $\theta_w = 1$, $\theta_b = 1$, $\psi_w = 1$, $\psi_b = 1$. We generate small sample size data to illustrate the fact that the model is quite easy to estimate and does not require a large number of individuals or time points. The results of this simulation study for the structural parameters in the model are presented in Table 15. The parameter estimates are nearly unbiased and the coverage is

TABLE 15
Combining RDSEM and DSEM

Parameter	*absolute bias(coverage)*
β_1	.00(.95)
β_2	.01(.90)
β_3	.00(.97)
β_4	.02(.94)
ρ_x	.01(.96)
ρ_y	.00(.93)

near the nominal level. The estimation of the model takes less than 1 s per replication.

Note that the above model can be extended to the model where $X_{w,it}$ is regressed also on $\hat{Y}_{w,i,t-1}$. This extension would make the model for $X_{w,it}$ and $\hat{Y}_{w,it}$ be equivalent to the saturated bivariate VAR(1) model. Using RDSEM and DSEM features in the same model can be useful when different parts of the model are best described by the different frameworks. In the RDSEM framework, the relationship across time between the variables is all expressed in terms of their residuals and to some extent, the dynamic relationship between the variables is lost. Adding features of the DSEM framework can be helpful in recovering the dynamic flavor of the models. Similarly, in the DSEM framework, adding features of the RDSEM framework can simplify relationships between variables when this is needed. Note that the algebraic manipulation yielding the equivalent model expressions can be performed for many RDSEM/DSEM models. In fact, every RDSEM model can be expressed as an equivalent DSEM model in more than one way. The most optimal Bayesian estimation of the model depends on how the model is specified. Most RDSEM models would best be estimated as RDSEM models rather than as converted DSEM models, particularly if the conversion involves fixing residual variances to zero.

TWO-LEVEL ARMA(1,1) REGRESSION

In this section, we illustrate some additional time-series error structures that are available in the DSEM and the RDSEM frameworks and in particular, we discuss the ARMA(1,1) model with covariates. It is shown in Schuurman, Houtveen, and Hamaker (2015) and Asparouhov et al. (2018) that the ARMA(1,1) time-series model is equivalent to the measurement error autoregressive model (MEAR), which is easy to estimate within both the DSEM and RDSEM frameworks. In this section, however, we illustrate a slightly different approach for estimating the ARMA(1,1) model, in the context of the two-level regression model.

A single level ARMA(1,1) model is defined by the following equation.

$$Y_t = \mu + aY_{t-1} + \eta_t + b\eta_{t-1}. \tag{61}$$

The model has four parameters: μ, a, b, and the variance parameter $\theta = Var(\eta_t)$. The ARMA(1,1) model can be used to fit more complex autocorrelation functions $f(k) = Corr(Y_t, Y_{t-k})$. In the AR(1) model, the autocorrelation function is exponential and thus the autocorrelation decays very rapidly. In the ARMA(1,1) model, the autocorrelation function can decay slower than exponential. In addition, for the special case when $a = 0$, the first order autocorrelation is $f(1) = b/(1 + b^2)$ while all other autocorrelations are 0, i.e., $f(k) = 0$ for $k > 1$. This can be helpful in those situations where the data exhibit significant autocorrelation for consecutive observations but zero or small autocorrelation for observations from longer periods apart.

Consider now the extended ARMA(1,1) model.

$$Y_t = \mu + aY_{t-1} + f_t + b_1 f_{t-1} + \xi_t. \tag{62}$$

Here we have added an independent error term ξ_t, which is also referred to in the time-series literature as a white noise process. The model has five parameters: μ, a, b_1, $\psi = Var(f_t)$, and $\sigma = Var(\xi_t)$. This model, however, is not identified because it is equivalent to the ARMA(1,1) model which has four parameters. That is because an MA(1) processes, represented above by the term $f_t + b_1 f_{t-1}$, plus a white noise process, represented by the term ξ_t, equals another MA(1) process, see Granger and Morris (1976). That means that models (61) and (62) are equivalent. The relationship between the parameters in the two models is as follows.

$$(1 + b^2)\theta = (1 + b_1^2)\psi + \sigma \tag{63}$$

$$b\theta = b_1\psi \tag{64}$$

The parameters μ and a remain unchanged.

To identify this extended ARMA(1,1) model we can fix the variance $\sigma = Var(\xi_t)$ to a particular value. In fact, if we fix σ to zero then obviously the two models become identical: with $b_1 = b$ and $\psi = \theta$. It is easy to show that if $\sigma > 0$ then $|b_1| > |b|$, under the regularity conditions (for invertibility) of $|b| < 1$ and $|b_1| < 1$. In addition, b_1 and b are simultaneously positive or simultaneously negative.

The reason we consider the extended ARMA(1,1) model, however, is because it is faster to estimate in the RDSEM and DSEM frameworks than the ARMA(1,1) model. To estimate ARMA(1,1) model in Mplus, the variable η_t is introduced as a factor and equation (61) is simply coded in the Mplus DSEM language as **f by Y@1 (&1); Y on Y&1 f&1**. That, however, leaves the residual of Y as

a free parameter, i.e., we are essentially estimating the extended ARMA(1,1) model (62). Since that model is not identified the residual variance has to be fixed. If we fix the residual variance to 0 or to a small value such as 0.01 the mixing of the MCMC chain becomes quite slow with larger samples. Therefore, it is preferable to fix the residual variance to a value that is a bit larger, such as 0.1 or 0.2, and then use equations (63) and (64) to obtain the parameters of the ARMA(1,1) model from the parameters of the extended ARMA(1,1) model. It is important that the residual variance is not fixed to a larger value, however, because the larger σ is the closer b_1 is to 1. That, in turn, can create instability in the MCMC sequence since the posterior distribution of b_1 may include values larger than 1 that are outside of the admissible regularity space. Equations (63) and (64) can be solved in terms of b and θ as follows.

$$L = \frac{(1 + b_1^2)\psi + \sigma}{b_1 \psi} \tag{65}$$

$$b = \frac{L - \sqrt{L^2 - 4}}{2} \tag{66}$$

$$\theta = \frac{b_1 \psi}{b}. \tag{67}$$

We illustrate the above approach with a small simulation study. We generate 100 samples of size $T = 200$ using the single-level ARMA(1,1) model, and we analyze the data using the DSEM framework and the extended ARMA(1,1) model with σ fixed to 0.1. We use equations (65–67) to obtain the parameters of the ARMA(1,1) model. The results of the simulation study are presented in Table 16. The parameter estimate shows minimal bias, and the coverage is near the nominal level. It takes only 7 s to obtain these results for all 100 replications, i.e., the MCMC mixing is quite good.

Consider now the ARMA(1,1) and the extended ARMA (1,1) model in the RDSEM framework. Using the RDSEM language **f by Y@1 (&1); Y ^ on Y^1 f ^1** we estimate the following model

$$Y_t = \mu_1 + f_t + \hat{Y}_t \tag{68}$$

TABLE 16
Single Level DSEM-ARMA(1,1) Model

Parameter	*True value*	*Absolute bias(coverage)*
μ	0	.00(.91)
a	.6	.01(.92)
b	.4	.00(.92)
θ	1	.03(.95)

$$\begin{aligned}\hat{Y}_t &= a\hat{Y}_{t-1} + b_2 f_{t-1} + \xi_t \\ &= a(Y_{t-1} - \mu_1 - f_{t-1}) + b_2 f_{t-1} + \xi_t = \end{aligned} \tag{69}$$

$$aY_{t-1} - a\mu_1 + (b_2 - a)f_{t-1} + \xi_t \tag{70}$$

Combining the above equations we get

$$Y_t = \mu_1(1 - a) + aY_{t-1} + f_t + (b_2 - a)f_{t-1} + \xi_t. \tag{71}$$

This model is a reparameterization of the extended ARMA (1,1) model in the DSEM framework and

$$\mu = \mu_1(1 - a) \tag{72}$$

$$b_1 = b_2 - a. \tag{73}$$

As in the DSEM model, the variance of the residual ξ_t is an unidentified parameter and should be fixed to a particular value σ. To reduce the extended RDSEM-ARMA(1,1) model to the ARMA(1,1) model we use equation (73) followed by (65–67).

We now illustrate the RDSEM-ARMA(1,1) model with the following two-level regression.

$$X_{it} = X_{b,i} + X_{w,it} \tag{74}$$

$$Y_{it} = \alpha_i + \beta_1 X_{w,it} + \varepsilon_{w,it} \tag{75}$$

$$\alpha_i = \alpha + \beta_2 X_{b,i} + \varepsilon_{b,i} \tag{76}$$

$$X_{w,it} \sim N(0, \psi_{x,w}), X_{b,i} \sim N(\mu_x, \psi_{x,b}), \varepsilon_{b,i} \sim N(0, \theta_{y,b}) \tag{77}$$

where $\varepsilon_{w,it}$ follows an ARMA(1,1) process

$$\varepsilon_{w,it} = a\varepsilon_{w,i,t-1} + \eta_{it} + b\eta_{i,t-1} \tag{78}$$

$$\eta_{it} \sim N(0, \theta). \tag{79}$$

To estimate the above model we estimate the extended ARMA(1,1) model

$$\varepsilon_{w,it} = a\varepsilon_{w,i,t-1} + \eta_{it} + b_1\eta_{i,t-1} + \xi_{it} \tag{80}$$

and fix the variance ξ_{it} to a small but not zero and not near zero value σ and use formulas (65–67) to obtain the results for the ARMA(1,1) parameterization.

We illustrate the above process with the following simulation study. We generate 100 data sets with 200 individuals and 100 time points for each individual using the RDSEM-ARMA (1,1) model (74–79). We estimate the extended RDSEM-ARMA(1,1) model with σ fixed to 0.2 and reparameterize back to RDSEM-ARMA(1,1) scale. The results of the simulation study are reported in Table 17. The parameter estimates are unbiased, and the coverage is near the nominal level for all

TABLE 17
Two-Level RDSEM-ARMA(1,1) Regression Model

Parameter	*True value*	*Absolute bias(coverage)*
β_1	1	.00(.95)
β_2	.5	.00(.97)
α	0	.01(.96)
μ_x	0	.01(.94)
$\psi_{x,w}$	1	.00(.91)
$\psi_{x,b}$	1	.01(.96)
$\theta_{y,b}$	1	.01(.96)
a	.3	.00(.95)
b	.4	.01(.90)
θ	1	.01(.82)

but the θ variance parameter. The coverage for that parameter can be improved by increasing the number of MCMC iterations. The estimation of this model takes an average of 15 s per replication and therefore we conclude that the extended ARMA(1,1) model provides a well mixing parameterization.

Let us summarize the findings of this section. The ARMA(1,1) model can be estimated within the DSEM and the RDSEM frameworks simply by introducing a factor that models the MA part of the process. This, however, sets up the unidentified-extended ARMA(1,1) model. The simplest way to fix the identification problem is to fix the residual variance to a small but not zero and not near zero value. The parameters estimate for the ARMA (1,1) model can be obtained if desired from the parameter estimates of the extended ARMA(1,1) model. This last step is optional, however, and the reparameterization is only needed if it is necessary to obtain the parameter estimates on the standard ARMA(1,1) scale. If the error structure is not the main focus of the estimation, the reparameterization is not needed.

MEASUREMENT ERROR AUTOREGRESSIVE MODEL WITH COVARIATES

Measurement error modeling is an alternative way to specify an ARMA(1,1) error structure, however, this approach has the advantage that it can be interpreted in a more meaningful way than the approach used in the previous section, see Schuurman and Hamaker (2019). The single level MEAR model is defined as follows.

$$Y_t = \mu + f_t + \xi_t \tag{81}$$

$$f_t = rf_{t-1} + \epsilon_t. \tag{82}$$

Here the observed variable Y_t is a measurement for the latent variable f_t and the measurement error is represented by the error term ξ_t. The latent variable f_t is then modeled as an AR(1) process. In cross-sectional modeling, a latent variable typically needs more than one measurement for the model to be identified, but in time-series settings, this requirement is not needed and a single measurement is sufficient. The auto-regressive structure for f_t is the key to this identification and if $r = 0$ this would not be possible. In this model, the observed variable Y_t is a sum of an AR(1) process, represented by the term f_t, and a white noise process, represented by the term ξ_t. The sum of an AR(1) process and a white noise process is known to be equivalent to an ARMA(1,1) process, conditional on some regularity constraints, see Schuurman et al. (2015).

In this section, we illustrate the MEAR modeling with covariates in the RDSEM framework. Consider the following two-level MEAR model.

$$X_{it} = X_{b,i} + X_{w,it} \tag{83}$$

$$Y_{it} = Y_{b,i} + Y_{w,it} \tag{84}$$

$$Y_{w,it} = f_{it} + \varepsilon_{w,it} \tag{85}$$

$$f_{it} = \beta_1 X_{w,it} + \xi_{it} \tag{86}$$

$$\xi_{it} = r_1 \xi_{i,t-1} + r_2 X_{w,i,t-1} + \zeta_{it} \tag{87}$$

$$X_{w,it} = r_3 X_{w,i,t-1} + \varepsilon_{w,x,it} \tag{88}$$

$$Y_{b,i} = \alpha + \beta_2 X_{b,i} + \varepsilon_{b,i} \tag{89}$$

$$\varepsilon_{w,it} \sim N(0, \theta_{y,w}), \varepsilon_{b,i} \sim N(0, \theta_{y,b}), \zeta_{it} \sim N(0, \psi) \tag{90}$$

$$\varepsilon_{w,x,it} \sim N(0, \psi_{x,w}), X_{b,i} \sim N(\mu_x, \psi_{x,b}). \tag{91}$$

In equations (83) and (84) we separate the within and the between parts of the observed variables as in a standard two-level model. In equation (85) we specify a measurement error model for the within part of Y which measures the latent factor f_{it} with measurement error $\varepsilon_{w,it}$. In equation (86) the latent factor is regressed on the within part of the covariate. Equations (87) and (88) specify the time-series model for the residual of the latent factor and the covariate. The last equation (89) in the model is the regression equation on the between level, featuring a latent mean for the covariate and contextual effect modeling, see Asparouhov and Muthén (2019).

Without the coefficient r_2, the autoregressive part of the model reduces to an AR(1) error structure for the residual of the factor and the within part of the covariate. With the inclusion of that coefficient, however, the error structure becomes a combination of the RDSEM and DSEM models. The model for f_{it} and $X_{w,it}$ is equivalent to a bivariate SVAR model as discussed in Section 9. The relationship between the variables f_{it} and $X_{w,it}$ can be expressed equivalently without the error term ξ_{it} as follows.

TABLE 18
Two-Level RDSEM-MEAR Regression Model

Parameter	*True value*	*Absolute bias(coverage)*
β_1	1	.00(.92)
β_2	.5	.00(.98)
r_1	.5	.00(.89)
r_2	.4	.00(.95)
r_3	.5	.00(.96)
α	0	.01(.99)
μ_x	0	.01(.95)
$\psi_{x,w}$	1	.00(.93)
$\psi_{x,b}$	1	.00(.96)
$\theta_{y,w}$	.3	.02(.71)
$\theta_{y,b}$	1	.02(.94)
ψ	1	.02(.81)

$$f_{it} = \beta_1 X_{w,it} + r_1 f_{i,t-1} + (r_2 - \beta_1 r_1) X_{w,i,t-1} + \zeta_{it}. \quad (92)$$

This means that the above model can be estimated in the RDSEM framework using parameterization (83–89) or in the DSEM framework using parameterization (83–85), (88–89), (92). This alternative DSEM reparameterization can be useful in two ways. First, the mixing in the DSEM framework may be more efficient, and it may provide faster estimation in terms of the number iterations that are needed to obtain convergence. Second, the DSEM estimation is faster than the RDSEM estimation within each replication as it uses fewer number of variables and fewer updating steps.

We conduct a simulation study using the above model to evaluate the performance of the estimation method. We generate 100 samples with 200 individuals and 100 time points for each individual. The data are then analyzed with the Mplus-RDSEM estimation. The results are presented in Table 18. The parameter estimates appear to be unbiased, and the coverage is near the nominal level for almost all of the parameters, except for two of the residual variance parameters. To improve the coverage for those components as well, a much longer MCMC sequence can be used, or a bigger sample size maybe required. The estimation of this model takes 30 s per replication. Most MEAR-based models require larger number of time points T. Usually, $T = 100$ is sufficient to estimate such models well. If $T = 50$ or lower, deterioration of the estimates can be seen as well as slower convergence.

RDSEM INTERPRETATION AND THE IMPACT OF MEAR MODELING

The RDSEM model is based on the equation.

$$Y_t = [\text{Explanatory variables}] + \hat{Y}_t \quad (93)$$

where $\hat{Y}_t$ is considered the residual in the above equation which usually follows an autoregressive model. With this interpretation, $\hat{Y}_t$ is essentially treated as an auxiliary variable where the autoregressive model is there only to account for the non-independence of the observations across time and thereby to obtain proper statistical inference. In Section 9 we expanded this point of view and showed that $\hat{Y}_t$ can have meaningful dynamic modeling with other variables. From that point of view, the proper way to present the above equation is as follows.

$$Y = [\text{Non} - \text{dynamic part of } Y] + [\text{Dynamic part of } Y] \quad (94)$$

where

$$[\text{Explanatory variables}] = [\text{Non} - \text{dynamic part of } Y]. \quad (95)$$

and

$$\hat{Y} = [\text{Dynamic part of } Y]. \quad (96)$$

The non-dynamic part consists of contemporaneous relations between the variables which are not affected by variables from other periods. The dynamic part of Y is then essentially modeled as in a DSEM model where full interplay between various variables can occur across different periods. If we consider this in more detail, however, we see that with the RDSEM model the non-dynamic part of Y does not have its own residual and that is somewhat problematic from a pure modeling point of view. We should not assume that the non-dynamic part of Y is precisely predicted by explanatory variables. This problem, however, is resolved by the MEAR model. If we add a single indicator factor to equation (94) we can resolve this problem and provide a residual also for the non-dynamic part. In that case, we have

$$Y_t = [\text{Non} - \text{dynamic part of } Y] + f_t + \varepsilon_t \quad (97)$$

$$f_t = r f_{t-1} + \ldots + \xi_t \quad (98)$$

In model (97–98), $f_t = [\text{Dynamic part of } Y]$ and ε_t is the residual of the $[\text{Non} - \text{dynamic part of } Y]$. Note also that model (97–98) can be viewed as an RDSEM or DSEM model and can be estimated with both frameworks. As we saw in the previous section, once the MEAR modeling is added, the DSEM and RDSEM models usually become reparameterizations of each other.

Because of its focus on contemporaneous relationships, the RDSEM model can be viewed as basically having the same aim as regular two-level modeling of longitudinal data. In the last few sections, we have expanded that

point of view and showed that the modeling framework can tackle much more general models. The measurement error model, the ability to incorporate DSEM modeling, the ability to separate dynamic and non-dynamic relations, the ability to properly handle missing data on the dependent and the independent variables by incorporating lagged effects, and the ability to incorporate individually specific autocorrelations and residual variances, makes this framework far superior to the popular REML estimation.

DYNAMIC FACTOR ANALYSIS

The RDSEM framework offers some unique advantage for the dynamic factor analysis model as compared to the traditional DAFS (direct autoregressive factor score) models or the dynamic factor analysis model in the DSEM framework. Consider the following two-level factor analysis model. Let Y_{pit} be the $p-$th indicator variable, $p = 1, ..., P$ for individual i at time t. The indicator variables in this model measure one factor on the within level and one factor on the between level

$$Y_{pit} = Y_{b,pi} + Y_{w,pit} \tag{99}$$

$$Y_{w,pit} = \lambda_{w,p}\eta_{w,it} + \varepsilon_{w,pit} \tag{100}$$

$$Y_{b,pi} = \mu_p + \lambda_{b,p}\eta_{b,i} + \varepsilon_{b,pi} \tag{101}$$

$$\eta_{b,i} \sim N(0, \psi_b), \varepsilon_{b,pi} \sim N(0, \theta_{b,p}). \tag{102}$$

The time-series structure for the factor model is introduced through AR(1) models for the within level factor and the within level residuals

$$\eta_{w,it} = \rho\eta_{w,i,t-1} + \xi_{w,it} \tag{103}$$

$$\varepsilon_{w,pit} = r_p\varepsilon_{w,pi,t-1} + \zeta_{w,pit} \tag{104}$$

$$\xi_{w,it} \sim N(0, \psi_w), \zeta_{w,pit} \sim N(0, \theta_{w,p}). \tag{105}$$

To identify the factor scale, we fix the first loading on the within and the between levels to 1, i.e., $\lambda_{w,1} = \lambda_{b,1} = 1$. If the rest of the loadings on the within and the between levels are held equal across the two levels we can interpret the between level factor as the subject-specific factor mean across time, just as this is done in the standard two-level factor model.

The above model was suggested to us by Phil Wood and clearly is now the foundation of the RDSEM latent variable framework. The model generalizes the standard DAFS model, see Zhang, Hamaker, and Nesselroade (2008), in two ways. First, the model is extended to two-level settings where multiple individuals can be analyzed simultaneously. This extension is already available in the DSEM framework. Second, the model allows the different indicator variables to have separate and distinct autocorrelations as shown in equation (104). This additional autocorrelation is measurement specific, and it allows the measurements to correlate across time independently and in addition to the implied autocorrelation due to the factor autocorrelation. This extension of the standard DAFS model is easy to interpret. In fact, it is unrealistic to expect that all the autocorrelation in the factor model occurs via the measured latent dimension. The different measurements Y_{pit} would almost surely be subject to measurement specific autocorrelation. If that measurement specific autocorrelation is not separated from the factor model, the entire factor model may be distorted as the loading parameters would be tied up in the fitting of the measurement specific autocorrelation. We illustrate this point with the following simulation study.

We generate data according to the above model using $P = 3$ measurement variables and the following parameter values $\lambda_{w,p} = \lambda_{b,p} = 1$, $\theta_{w,p} = \theta_{b,p} = 1$, $\psi_w = \psi_b = 1$, $\mu_p = 0$, $\rho = 0.5$, $r_1 = 0.2$, $r_2 = 0.6$ and $r_3 = 0$. We generate 100 data sets using $N = 200$ individuals with two settings for the number of time points T: 30 and 100. We analyze the data using three different methods. First, we analyze the data in the RDSEM framework using the same model as the data generating model. Second, we analyze the data using the DAFS approach where the measurement specific autocorrelations are ignored, i.e., the parameters r_p are fixed to 0. Third, we analyze the data ignoring all time-series structure, i.e., as a standard two-level model with r_p and ρ fixed to 0. The results of the simulation study are presented in Table 19 for a subset of the model parameters. To compute the bias and coverage for a variance parameter when the corresponding autocorrelation is not included, we adjust the true value by dividing by $1 - r^2$, where r is the omitted autocorrelation. For example, if the residual autocorrelation for $Y_{w,2it}$ is not included, the actual residual variance should be $1/(1 - r_2^2) = 1.563$. The DIC for the two-level model is computed within the RDSEM framework so that it is comparable to the DIC of the other two models.

The results show that the RDSEM model performs well. The parameter estimates have minimal bias and the coverage is near the nominal level. The DAFS method, on the other hand, yields biased estimates for several parameters and in particular for the within level loading parameters and the factor autocorrelation parameter ρ. The two-level estimates appear to be less biased than the DAFS estimates (two wrong make one almost right), however, most of the parameters still show substantial bias particularly for the $T = 30$ case. The residual variance parameters for the indicators and for the factor on both the within and the between level appear to be biased with the two-level method. Notably, however, the within level loading parameter appears to be unbiased with

TABLE 19
Dynamic Factor Analysis: Absolute Bias(Coverage)

Parameter	*T*	*RDSEM*	*DAFS*	*Two-level*
$\lambda_{w,2}$	30	.02(.90)	.15(.07)	.00(.91)
$\theta_{w,2}$	30	.01(.86)	.43(.00)	.15(.07)
r_2	30	.01(.91)	–	–
ψ_w	30	.00(.88)	.16(.10)	.09(.56)
ρ	30	.01(.94)	.10(.00)	–
μ_2	30	.02(.88)	.02(.89)	.02(.87)
$\lambda_{b,2}$	30	.01(.95)	.01(.96)	.03(.92)
$\theta_{b,2}$	30	.03(.97)	.15(.84)	.14(.84)
ψ_b	30	.01(.93)	.04(.91)	.08(.88)
DIC	30	56055	57832	58291
$\lambda_{w,2}$	100	.01(.94)	.15(.00)	.00(.89)
$\theta_{w,2}$	100	.01(.96)	.34(.00)	.05(.37)
r_2	100	.00(.90)	–	–
ψ_w	100	.00(.94)	.17(.00)	.03(.82)
ρ	100	.00(.92)	.08(.00)	–
μ_2	100	.01(.90)	.01(.92)	.01(.95)
$\lambda_{b,2}$	100	.03(.93)	.01(.96)	.04(.89)
$\theta_{b,2}$	100	.03(.94)	.15(.84)	.09(.91)
ψ_b	100	.01(.96)	.01(.95)	.02(.92)
DIC	100	185612	193345	194970

the two-level estimation. The bias in the parameter estimates for the DAFS and the two-level models also result in a substantial drop in coverage for these parameters.

It is worth noting here that asymptotically as T and N increase to infinity, the two-level method is guaranteed to produce unbiased estimates. What is not guaranteed, however, is that the coverage for these estimates will recover. On the contrary, we can expect that the bias in the standard errors for the two-level model will remain regardless of the sample size. We demonstrated this bias earlier for simpler examples, see Table 4. The bias is due to overestimation of the number of independent observations within each cluster due to ignoring the autocorrelations in the data. The asymptotic result for the DAFS method is somewhat more dramatic. The parameter estimates of the DAFS model will remain biased even asymptotically due to the fact that the omitted measurement specific autocorrelations interfere with the estimation of the loading parameters.

The DIC criterion clearly indicates that the data is fitted best by the RDSEM model, followed by the DAFS model, followed by the two-level model. This order shows an interesting dilemma. If the choice is to be made only between the DAFS model and the two-level model, should one prefer the better fitting DAFS models or the less biased two-level model? The answer to that question probably depends on the type of inference that is needed, however, we will not discuss this further as clearly the RDSEM method outperforms both other methods in terms of bias and overall model fit.

The above model can be extended further to include autoregressive relations among the error terms across the different measurements. Instead of modeling the measurement specific autocorrelation with P univariate AR(1) models, we can use a single multivariate AR(1) model where across measurement and across time relations are included. Estimating such an extension would most likely require larger number of time points.

We conclude this section with one final simulation that compares the RDSEM and DSEM models with measurement specific autocorrelations. The RDSEM model is given in equations (99–105). The corresponding DSEM model is given by the same equations with the exception that equations (100) and (104) are replaced by the following equation

$$Y_{w,pit} = \lambda_{w,p}\eta_{w,it} + r_p Y_{w,pi,t-1} + \zeta_{w,pit}. \tag{106}$$

This model has a somewhat intricate dynamics and maybe harder to interpret in practice. For example, the RDSEM model has the simple interpretation that it estimates the same model as the two-level model but it properly accounts for the autocorrelation in the data. The RDSEM and the two-level models asymptotically would produce the same structural parameter estimates. The corresponding DSEM model does not have this simple interpretation and will not produce the same structural parameter estimates asymptotically. Nevertheless, the model can be used to explore dynamic relations between the measurements and the factor and it includes measurement specific autocorrelations. Here we conduct a simulation study to illustrate the ability of the DIC criterion to determine the proper framework for the dynamic factor model. Using the same parameters as in the previous simulation and $T = 30$ we generate data according to the DSEM model and according to the RDSEM model and we analyze it with both models, similar to the simulation study presented earlier in Table 10 for the regression model. The average DIC values across 100 replications are reported in Table 20. The results indicate that the DIC criterion can be used to properly identify the correct framework for the data. The smallest DIC value for the DSEM data is obtained by the DSEM model. The smallest DIC value for the RDSEM data is obtained by the RDSEM model.

We also see in this table that for the RDSEM data, the DSEM model with measurement specific autocorrelations

TABLE 20
DIC Comparison for DSEM and RDSEM Factor Analysis Models with Measurement Specific Autocorrelation

	DSEM model	*RDSEM model*
DSEM data	55934	57609
RDSEM data	56666	56055

fits the data better than the DAFS and the two-level models because the results of the second row in Table 20 correspond to the results in the tenth row of Table 19.

It is important to note here that the difference between the RDSEM and DSEM factor models is due only to the measurement specific autocorrelations. If these autocorrelations are not included in the model, and the only autocorrelation in the model is for the factor, then the RDSEM and DSEM factor models are identical. This, however, does not extend to the MIMIC model which includes a covariate predicting the factor. If the factor model includes a covariate predicting the factor the RDSEM and DSEM factor models would not be identical.

EMPIRICAL EXAMPLE

As an RDSEM illustration, we consider data on 10 ordinal, 5-category negative affect items, known as the PANAS scale, see Watson, Clark, and Tellegen (1988), measured on 56 consecutive days with N = 270. The data are from the older cohort of the Notre Dame Study of Health and Well-being, see Wang, Hamaker, and Bergman (2012). A one-factor model is analyzed, allowing a trend over time for the factor and letting the factor residuals have an AR(1) structure. Let Y_{ijt} denote the j-th ordinal indicator for individual i at time t. The model is described by the following equation

$$P(Y_{ijt} = k) = \Phi(\tau_{jk} - \lambda_j \eta_{it} - \xi_{ij}) - \Phi(\tau_{j,k-1} - \lambda_j \eta_{it} - \xi_{ij}) \quad (107)$$

where $k = 1, ..., 5$, τ_{jk} are the threshold parameters under the standard assumption that $\tau_{j0} = -\infty$ and $\tau_{j5} = \infty$, i.e., four threshold parameters τ_{j1}, …, τ_{j4} are estimated for each item, and λ_j are the loading parameters. The factor η_{it} represents the negative affect construct and is modeled with the following two-level linear growth autoregressive model

$$\eta_{it} = \alpha_i + \beta_i t^* + \varepsilon_{it} \quad (108)$$

where $t^* = (t - 28)/10$ is an approximately centered and rescaled version of the time variable t so that α_i can be interpreted as the average negative affect for person i. The random effects α_i and β_i are normally distributed $\alpha_i \sim N(\alpha, \sigma_\alpha)$ and $\beta_i \sim N(\beta, \sigma_\beta)$ with $Cov(\alpha_i, \beta_i) = r$. The scale of the negative affect factor η_{it} is identified by fixing α to 0 and σ_α to 1. Using the interpretation that α_i is the average negative affect for individual i across time, we identify the scale of the factor by fixing the mean and variance of the average negative affect to 0 and 1 respectively. The autoregressive part of the model is given by

$$\varepsilon_{it} = \rho \varepsilon_{i,t-1} + \zeta_{it} \quad (109)$$

where ρ is the autocorrelation and $\zeta_{it} \sim N(0, \theta)$ represent the independent/uncorrelated residuals. The random effects ξ_{ij} in equation (107) are normally distributed zero mean residuals $\xi_{ij} \sim N(0, \sigma_j)$ which allow us to fit the observed variable differences across individual that can not be explained by the factor model alone. These can also be interpreted as random threshold values, i.e., random effects for the threshold parameters, although the proper interpretation is that the random effect is a shift in the threshold parameters. This is because for each item there are 4 threshold parameters and 1 random effect shift ξ_{ij}, i.e., the random effect ξ_{ij} can be interpreted as random threshold parameter if the item is binary but if the item has more than two categories it should be interpreted as a random shift in the threshold parameters. If the indicators are continuous variables instead of categorical the interpretation of ξ_{ij} is simply the residual of the random intercept for the factor indicator. We refer to the above model as M_1.

Next we consider the corresponding DSEM model, i.e., this is the model where we replace equations (108) and (109) with the following equation

$$\eta_{it} = \alpha_i + \beta_i t^* + \rho \eta_{i,t-1} + \zeta_{it}. \quad (110)$$

This DSEM model is equivalent to the RDSEM model because the predictor t^* is a linear function of the time t, see Asparouhov et al. (2018). The equivalence holds for any predictor which is a polynomial function of t (ex. quadratic or cubic growth) but it does not hold in general when there is a time specific predictor different from the time variable t. The reparameterization between the above DSEM and RDSEM models is fairly simple and is given in equations (65–66) in Asparouhov et al. (2018). For example, to obtain β_i in the RDSEM scale from the DSEM scale we have to divide by $1 - \rho$. We denote the DSEM model by M_2.

Model M_3 is the same as model M_1 but without the autocorrelation feature of M_1, i.e., the parameter ρ is fixed to 0. Note here that model M_3 is a standard two-level SEM model. The main difference between models M_1/M_2 and M_3 is in how the data is organized (although in Mplus this is done behind the scenes through the TINTERVAL option). Model M_3 essentially ignores the missing days in the daily diary response as these will not contribute to the likelihood of the dependent variables. Models M_1 and M_2 cannot ignore these missing days. Ignoring the missing response will alter the time distance and the correlation between the observed values. Typically, model M_3 is estimated in Mplus with the Bayes estimator because the ML estimator would require high-dimensional numerical integration and essentially become computationally prohibitive.

Model M_4 is the same as model M_3 with the constraints $\sigma_j = 0$, i.e., the random effects ξ_{ij} are eliminated from the model. The individual differences in this model are entirely explained by the differences in the factor distribution. This

model can be estimated with the ML estimator and it would require 3 dimensional numerical integration.

The final model we consider is model M_5. This model is the same as model M_3 but with the restriction $\sigma_\beta = r = 0$, i.e., the random slope β_i is a non-random slope β. Note that the random effects ξ_{ij} are included in this model. M_5 can be estimated with the two-level WLSMV limited information estimator in Mplus, see Asparouhov and Muthén (2007). Note here the natural nesting progression among the models. Model M_4 and model M_5 are not nested within each other, but are both nested in model M_3 which is nested in model M_1. Absent from this sequence is model M_2. Although in this particular case M_2 is a reparameterization of M_1 in general it is not. This feature of M_1 again emphasizes the advantage of RDSEM over DSEM. The RDSEM model is nested above the standard two-level SEM models and the parameter estimates will be on the same scale and directly comparable. The nesting also helps with model comparison. To test model M_1 v.s. model M_3 we simply need to evaluate the significance of ρ. Similarly, to test model M_3 v.s. model M_4 we test the significance of σ_j and to test model M_3 v.s. model M_5 we test the significance of σ_β.

Table 21 contains a summary of the estimation results for the above models including the computational time and the parameter estimates and standard errors for a subset of the parameters. Not surprisingly, the limited information WLSMV method is the fastest to compute, followed by the bayesian method, followed by the ML method. The autocorrelation parameter ρ is highly significant. A negative downward trend (i.e. β) is also highly significant. The results here also confirm the expected reparameterization relationship between DSEM and RDSEM as the β estimate for RDSEM can be obtained from the β estimate for DSEM by dividing it by $1 - \rho$. The variance parameter for the random slope σ_β is also significant but only marginally. The variance parameters for the random effects ξ_{ij}, i.e., σ_j are significant. The results for λ_i and σ_i for $M1$, M_2 and M_3 are very close as expected. For M_1 and M_2 these are expected to be identical since the models are equivalent. The differences are due to the random nature of the MCMC estimation. For model M_3 the estimates are similar because the omitted autocorrelation has limited or no impact on the rest of the structural parameters. Model M_4 appears to have higher loadings values. Most likely this is caused by the fact that the factor model has to compensate for the omitted random effects ξ_{ij}. Interestingly, the limited information estimation for model M_5 produces even higher loading and σ_j values. This is most likely due to omitted random slope for the time variable which would force the factor model and the random effects ξ_{ij} to absorb more of the within subject correlation. The CFI and TLI fit measures produced by the WLSMV estimation are .95 and .94 respectively which indicates that the proposed modeling framework M_5 and its time-series generalizations are reasonable approximations for these data.

DISCUSSION

In this article we provide basic insight into the differences and similarities between the DSEM and the RDSEM modeling frameworks. We showed that the DIC criterion can be used to choose the better fitting framework between the two. We also illustrated the fact that the DIC can not be used blindly across all possible models. The DIC definition changes across models and it can be used for model comparison only when the definitions are compatible.

We also considered the concept of exogeneity in multilevel time series model, i.e., the ability to limit our modeling to the endogenous variables conditional on the exogenous variables. We found that in both frameworks the exogeneity approach often leads to poor results. The RDSEM framework tends to be more suitable for such

TABLE 21
Model Comparison for Negative Affect Data: Estimate(Standard Errors)

Model	M_1	M_2	M_3	M_4	M_5
Type	RDSEM	DSEM	Two-level SEM	Two-level SEM	Two-level SEM
Estimator	Bayes	Bayes	Bayes	ML	WLSMV
Number of parameters	65	65	64	54	62
Estimation time in minutes	12.1	9.7	9.2	23.1	1.7
λ_1	1.20(.06)	1.14(.07)	1.19(.07)	1.40(.07)	1.52(.09)
λ_2	1.41(.07)	1.34(.08)	1.41(.08)	1.49(.08)	1.65(.10)
λ_3	1.40(.07)	1.33(.07)	1.40(.07)	1.43(.07)	1.42(.09)
ρ	.39(.02)	.36 (.02)	–	–	–
β	−.09(.01)	−.06(.01)	−.08(.01)	−.08(.01)	−.06(.01)
σ_β	.03 (.01)	.02(.00)	.04(.01)	.03(.00)	–
σ_1	.85(.13)	.86(.13)	.86(.14)	–	2.00(.42)
σ_2	1.12(.17)	1.13(.17)	1.12(.17)	–	2.59(.45)
σ_3	.87(.11)	.88(.11)	.88(.12)	–	1.41(.22)

a modeling approach than the DSEM framework. In some special cases, such as zero contextual effect and no missing data, the RDSEM framework worked well under the exogeneity approach. Our simulations show, however, that the most prudent approach is to fully model the covariates and to account for the multilevel and the time-series structure of the data.

We showed here that the RDSEM framework is more robust to how unevenly spaced times of observations are modeled. Even when the times of observations are incorrectly modeled the RDSEM structural parameters remain unbiased. This does not hold for the DSEM model. Nevertheless, this advantage of the RDSEM model should not be used to prefer that model when times of observations are difficult to deal with. Instead, the DIC criterion should be used to determine which model is a better fit to the data.

The RDSEM framework is a direct generalization of the repeated measures modeling framework estimated with the REML estimator and the hierarchical linear models. These modeling approaches have been staples of statistical modeling for nearly half a century, see Harville (1977), and are widely available in statistical software. Therefore, the RDSEM model has an advantage over the DSEM model because the structural part of the RDSEM model can be interpreted as in a repeated measures model. Readers familiar with the REML estimation for repeated measures model will find the interpretability of the RDSEM model very appealing, while at the same time, one can utilize the power and flexibility of the RDSEM framework. In this article we summarized the known advantages of the RDSEM estimation over the REML estimation and we added new ones. In the context of subject specific autocorrelations, subject specific residual variances, and categorical data our simulation studies demonstrated the clear advantages of the RDSEM model with the Baysian estimation over the REML estimation.

Several examples were described in this paper to illustrate how multilevel time series models can be parameterized, reparameterized or interpreted as DSEM or RDSEM models or as a combination of both. We discussed how these different parameterization options interact with the algorithms implemented in Mplus to obtain the most optimal estimation procedure. For SEM models, such as factor analysis models, we showed that the RDSEM modeling framework has a unique advantage over the classic DAFS model because it allows us to seamlessly include autoregressive models not only for the factors but also for the residual of each measurement variable. All programming scripts used in this article are available online and can be used for further research on this topic.[2]

[2] The Mplus scripts are available at statmodel.com/download/rdsem.zip.

REFERENCES

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling*, *25*, 359–388.

Asparouhov, T., & Muthén, B. (2007). Computationally efficient estimation of multilevel high-dimensional latent variable models. Proceedings of the JSM Meeting in Salt Lake City, Utah, Section on Statistics in Epidemiology, 2531–2535.

Asparouhov, T., & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling, 26*, 119–142.

Bauer, D. J., & Sterba, S. K. (2011). Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation. *Psychological Methods*, *16*, 373–390. doi:10.1037/a0025813

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis*. New York, NY: Springer-Verlag.

Bolger, N., & Laurenceau, J. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Chi, E. M., & Reinsel, G. C. (1989). Models for longitudinal data with random effects and AR(1) errors. *Journal of the American Statistical Association*, *84*, 452–459. doi:10.1080/01621459.1989.10478790

Granger, C. W. J., & Morris, M. J. (1976). Time series modelling and interpretation. *Journal of the Royal Statistical Society, Series A*, *139*, 246–257. doi:10.2307/2345178

Greene, W. H. (2014). *Econometric analysis* (7th ed.). New Jersey, USA: Prentice Hall.

Hamaker, E., Asparouhov, T., Brose, E., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, *53*, 820–841. doi:10.1080/00273171.2018.1446819

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, *72*, 320338.

Jongerling, J., Laurenceau, J. P., & Hamaker, E. (2015). A Multilevel AR (1) model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research*, *50*, 334–349. doi:10.1080/00273171.2014.1003772

Lütkepohl, H. (2007). *New introduction to multiple time series analysis*. Berlin, Germany: Springer-Verlag.

McNeish, D., & Hamaker, E. L. (2018). A primer on two-level dynamic structural equation models for intensive longitudinal data. PsyArXiv. November 28. doi:10.31234/osf.io/j56bm.

Molenaar, P. C. M. (2017). Equivalent dynamic models. *Multivariate Behavioral Research*, *52*, 242–258. doi:10.1080/00273171.2016.1277681

Muthén, B., & Asparouhov, T. (2002) Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in mplus. Mplus web note 4, 1–22. https://www.statmodel.com/download/webnotes/CatMGLong.pdf

Öhrlund, I., Schultzberg, M., & Bartusch, C. (2019). Identifying and estimating the effects of a mandatory billing demand charge. *Applied Energy*, *237*, 885–895. doi:10.1016/j.apenergy.2019.01.028

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Schultzberg, M., & Muthén, B. (2018). Number of subjects and time points needed for multilevel time series analysis: A simulation study of dynamic structural equation modeling. *Structural Equation Modeling*, *25*, 495–515.

Schuurman, N., Houtveen, J., & Hamaker, E. (2015). Incorporating measurement error in n = 1 psychological autoregressive modeling. *Frontiers in Psychology*, *6*, 1038. doi:10.3389/fpsyg.2015.01038

Schuurman, N. K., & Hamaker, E. L. (2019). Measurement error and person-specific reliability in multilevel autoregressive models. *Psychological Methods, 24*, 70–91.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*, 583–639. doi:10.1111/rssb.2002.64.issue-4

Wang, L., Hamaker, E. L., & Bergman, C. S. (2012). Investigating inter-individual difference in short-term intraindividual variability. *Psychological Methods*, *17*, 567581. doi:10.1037/a0029317

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measure of positive and negative affect: The panas scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Zhang, Z., Hamaker, E., & Nesselroade, J. (2008). Comparisons of four methods for estimating a dynamic factor model. *Structural Equation Modeling*, *15*, 377–402. doi:10.1080/10705510802154281