Covariance of Random Effects in Multilevel Modeling

Tihomir Asparouhov and Bengt Muthén

Mplus Web Note No.24
Version 1

October 17, 2024

1 Introduction

In this note we discuss the importance of modeling the covariance parameters between the random effects in multilevel models. Complex models with a larger number of random effects and moderate sample sizes often have difficulties modeling the full variance covariance matrix for the random effects. Convergence issues arise when random effects are highly correlated or when random effect variances collapse to zero. A simpler model, which is often used initially, assumes that the random effects are independent. This model is much more likely to converge than the model estimating all covariances. The simpler model evaluates the possibility of within level parameters to be modeled as cluster specific, without the complication of necessarily specifying the correct between level model. With very small samples, it is often the only model that can be estimated. Without specifying a between level model for the random effects, Mplus will default to the simpler model of independent random effects.

In this note we explore the implications of ignoring the covariances of the random effects. We show that some biases can occur and that between models which attempt to incorporate, even incorrectly, some covariance between the random effects are successful to some extent in resolving these biases. We consider three different settings for our exploration: two-level random regression, two-level random regression with latent centering, and multilevel mediation models.

2 Two-level Random Regression

Suppose that Y_{ij} is a dependent variable and X_{ij} is a covariate for individual i in cluster j. Consider the two-level regression model

$$Y_{ij} = \alpha_j + \beta_j X_{ij} + \varepsilon_{ij} \tag{1}$$

$$\varepsilon_{ij} \sim N(0, \theta), X_{ij} \sim N(\mu_x, \sigma_x), (\alpha_j, \beta_j) \sim N(\mu, \Sigma). \tag{2}$$

The model contains two random effects (cluster specific parameters): the random intercept α_j and the random slope β_j. The vector μ is the vector of the mean parameters for the random effects and Σ is an unrestricted covariance matrix. Note that X_{ij} has a distribution that is independent of j and it does not have a cluster specific mean, i.e., the covariate X_{ij} is group-centered.

We want to investigate the implications of ignoring the between level covariance parameter $Cov(\alpha_j, \beta_j)$. We consider two models. Model 1 is the full model as specified above and Model 2 is the model where Σ is estimated as a diagonal matrix, i.e., the covariance between the random effects is fixed to 0.

We illustrate the effect of ignoring the random effect covariance with a simulation study. The between level model, i.e, the model for the random effects, acts as a prior to the cluster specific model. If the size of the cluster is substantial the prior will have no meaningful impact on the model estimation. In fact, if the size of the clusters is substantial two things happen. First a mis-specification in the between level model will have little impact on the posterior distribution of the random effects. Second, the posterior distribution of the random effects will be narrow, i.e., the random effects will be precisely determined within each cluster. We conclude that if the cluster sizes are large, ignoring the random effect covariance parameter will have no substantial impact on the rest of the model parameters. Therefore to study the impact of the random effect covariance we must consider small cluster sizes. In this simulation study we use a sample with 1000 clusters of size 5.

Figure 1 contains the Mplus input file for this simulation study and Figures 2 and 3 contain the results for the two models. The results indicate that Model 1 estimation works well while when we exclude the random effect covariance (Model 2), the variances of the random effects are overestimated, i.e., the estimates are biased and the confidence interval coverage of these parameters drops substantially. A small bias also appears for the within level variance parameter. All other parameters appear to be unbiased.

For models with random effects, Mplus also provides estimates for the random effects. This is useful in evaluating the model beyond just the model parameters. In simulation studies, the random effect estimates can be compared to their true values since those are known. There are two statistics that Mplus computes for the random effects. The mean squared error (MSE) difference between the estimates and the true values as well as the correlation between the estimates and the true values in the entire sample across the clusters. In simulation studies with multiple replications, these two statistics are averaged across the replications and can be found in the Mplus output for each random effect. In our simulation study, these results are given at the bottom of Figures 2 and 3. We see that for Model 1, the cluster specific estimates for α_j and β_j are closer to their true values and are more highly corrected with the true values. Thus, the impact of the random effect covariance goes beyond model parameter estimates. The random effect estimates are also impacted.

Figure 4 shows the Mplus input file for Model 1 that can be used with real data applications. This setup is simpler than the setup for the simulation study given in Figure 1 because it does not require starting values or specifying all model parameters. In simulation studies, the starting values are used to compute the confidence interval coverage. The setup in Figure 4 can be used with the ML or the Bayes estimator.

Figure 1: Random Regression Simulation Study

```
montecarlo:
        names are y x;
        nobservations = 5000;
        ncsizes = 1;
        csizes = 1000(5);
        nreps = 100;
        within=x;

ANALYSIS:
        TYPE = TWOLEVEL RANDOM;
        estimator=bayes;

model population:

        %WITHIN%
        x*1; y*1;
        s1 | y on x;
        [x*0.8];

        %BETWEEN%
        y*1 s1*1;
        s1 with y*0.5;
        [s1*0.5 y*1.2];

model:

        %WITHIN%
        x*1; y*1;
        s1 | y on x;
        [x*0.8];

        %BETWEEN%
        y*1 s1*1;
        s1 with y*0.5;
        [s1*0.5 y*1.2];
```

Figure 2: Random Regression Results

```
MODEL RESULTS

                                ESTIMATES              S. E.     M. S. E.  95%  % Sig
                 Population   Average   Std. Dev.   Average             Cover Coeff

Within Level

 Means
  X                   0.800     0.8007     0.0139     0.0139     0.0002 0.900 1.000

 Variances
  X                   1.000     0.9987     0.0219     0.0200     0.0005 0.920 1.000

 Residual Variances
  Y                   1.000     1.0047     0.0257     0.0252     0.0007 0.900 1.000

Between Level

 S1       WITH
  Y                   0.500     0.5027     0.0457     0.0430     0.0021 0.940 1.000

 Means
  Y                   1.200     1.1984     0.0380     0.0374     0.0014 0.930 1.000
  S1                  0.500     0.4989     0.0380     0.0362     0.0014 0.920 1.000

 Variances
  Y                   1.000     0.9980     0.0575     0.0626     0.0033 0.950 1.000
  S1                  1.000     1.0157     0.0567     0.0579     0.0034 0.920 1.000


CORRELATIONS AND MEAN SQUARE ERROR OF THE TRUE FACTOR VALUES AND THE FACTOR SCORES

                          CORRELATIONS             MEAN SQUARE ERROR
                      Average     Std. Dev.          Average     Std. Dev.
  S1                  0.909        0.006              0.419        0.011
  Y                   0.886        0.007              0.464        0.012
```

Figure 3: Random Regression Results without Random Effect Covariance

```
MODEL RESULTS

                                  ESTIMATES              S. E.     M. S. E.  95%  % Sig
                 Population   Average   Std. Dev.   Average             Cover Coeff

Within Level

 Means
  X                   0.800     0.8008     0.0140     0.0142     0.0002 0.970 1.000

 Variances
  X                   1.000     0.9993     0.0221     0.0201     0.0005 0.900 1.000

 Residual Variances
  Y                   1.000     0.9773     0.0250     0.0242     0.0011 0.860 1.000

Between Level

 Means
  Y                   1.200     1.1982     0.0383     0.0398     0.0015 0.990 1.000
  S1                  0.500     0.4999     0.0395     0.0379     0.0015 0.920 1.000

 Variances
  Y                   1.000     1.1520     0.0650     0.0689     0.0273 0.300 1.000
  S1                  1.000     1.1659     0.0610     0.0645     0.0312 0.190 1.000


CORRELATIONS AND MEAN SQUARE ERROR OF THE TRUE FACTOR VALUES AND THE FACTOR SCORES

                        CORRELATIONS               MEAN SQUARE ERROR
                    Average    Std. Dev.          Average    Std. Dev.
  S1                  0.899       0.007             0.443       0.013
  Y                   0.872       0.008             0.493       0.013
```

Figure 4: Real-Data Random Regression Model Estimation

```
variable:
        names are y x c;
        within=x;
        cluster=c;

data: file=1.dat;

ANALYSIS:
        TYPE = TWOLEVEL RANDOM;
        estimator=bayes;

model:

        %WITHIN%
        s | y on x;

        %BETWEEN%
        s with y;
```

3 Two-level Random Regression with Latent Centering

In this section we repeat the simulation study from the previous section but now the covariate X_{ij} is no longer assumed to be group-mean centered. We will use latent mean centering as in Asparouhov and Muthén (2019). The model is given as follows

$$X_{ij} = X_{w,ij} + X_{b,j} \tag{3}$$

$$Y_{ij} = \alpha_j + \beta_j X_{w,ij} + \varepsilon_{ij} \tag{4}$$

$$\varepsilon_{ij} \sim N(0, \theta), X_{w,ij} \sim N(\mu_x, \sigma_x), (X_{b,j}, \alpha_j, \beta_j) \sim N(\mu, \Sigma). \tag{5}$$

The model has three random effects: $X_{b,j}$, α_j, and β_j. Model 1 is again the model with unrestricted variance covariance Σ for the random effects and Model 2 is the mis-specified between level model with all covariances fixed to 0.

Figure 5 contains the Mplus input file for this simulation study and Figures 6 and 7 contain the results for the two models. Model 1 outperforms Model 2 in terms of model parameter bias, confidence interval coverage, MSE for the random effect estimates and correlation between random effect estimates and the true random effect values. Unlike in the previous section, here we see that with Model 2, not only the variances of the random intercepts are biased but also the means of the random intercepts are biased.

The latent centering regression model is somewhat more complex than the model estimated in the previous section. It involves the product of two-latent variables and currently it can only be estimated with the Bayes estimator. We are not aware of an explicit formula that can be used to approximate the biases. We can infer, however, that the biases disappear as the cluster sizes increase due to the diminishing effect of the between level model on the estimation of the random effects. The observed data log-likelihood, which is the basis for the Bayes MCMC estimation, contains the product of $X_{b,j}$ and β_j. The expected value of that product depends on the covariance parameter between of the two effects and in Model 2 this is incorrectly specified which can conceivably explain the mean effect. However, including the covariance between $X_{b,j}$ and β_j alone in the estimation is not sufficient to resolve the mean bias. The full covariance matrix is needed as in Model 1.

Figure 8 shows the Mplus input file for Model 1 estimation of the two-level regression with latent centering that can be used with real data applications.

Figure 5: Random Regression with Latent Centering Simulation Study

```
montecarlo:
        names are y x;
        nobservations = 5000;
        ncsizes = 1;
        csizes = 1000(5);
        nreps = 100;

ANALYSIS:
        TYPE = TWOLEVEL RANDOM;
        estimator=bayes;
        biter=(500);

model population:

        %WITHIN%
        x*1; y*1;
        s1 | y on x;

        %BETWEEN%
        y*1 x*1 s1*1;
        [s1*0.5 x*0.8 y*1.2];
        y with x*0.7;
        s1 with y-x*0.5;

model:

        %WITHIN%
        x*1; y*1;
        s1 | y on x;

        %BETWEEN%
        y*1 x*1 s1*1;
        [s1*0.5 x*0.8 y*1.2];
        y with x*0.7;
        s1 with y-x*0.5;
```

Figure 6: Random Regression with Latent Centering Results

	Population	ESTIMATES Average	Std. Dev.	S. E. Average	M. S. E.	95% Cover	% Sig Coeff
Within Level							
Variances							
X	1.000	0.9956	0.0205	0.0220	0.0004	0.970	1.000
Residual Variances							
Y	1.000	0.9996	0.0254	0.0256	0.0006	0.970	1.000
Between Level							
S1 WITH							
Y	0.500	0.5112	0.0514	0.0513	0.0027	0.940	1.000
X	0.500	0.5100	0.0455	0.0436	0.0021	0.940	1.000
Y WITH							
X	0.700	0.7166	0.0476	0.0493	0.0025	0.940	1.000
Means							
Y	1.200	1.2026	0.0338	0.0382	0.0011	0.980	1.000
X	0.800	0.8001	0.0314	0.0348	0.0010	1.000	1.000
S1	0.500	0.5069	0.0387	0.0372	0.0015	0.940	1.000
Variances							
Y	1.000	1.0680	0.0616	0.0667	0.0084	0.830	1.000
X	1.000	1.0250	0.0494	0.0536	0.0030	0.950	1.000
S1	1.000	1.0264	0.0560	0.0614	0.0038	0.940	1.000

CORRELATIONS AND MEAN SQUARE ERROR OF THE TRUE FACTOR VALUES AND THE FACTOR SCORES

	CORRELATIONS Average	Std. Dev.	MEAN SQUARE ERROR Average	Std. Dev.
S1	0.871	0.008	0.493	0.012
Y	0.867	0.007	0.501	0.011
X	0.926	0.005	0.379	0.009

Figure 7: Random Regression with Latent Centering Results without Random Effect Covariance

	Population	ESTIMATES Average	Std. Dev.	S. E. Average	M. S. E.	95% Cover	% Sig Coeff
Within Level							
Variances							
X	1.000	1.0151	0.0211	0.0228	0.0007	0.890	1.000
Residual Variances							
Y	1.000	0.9998	0.0252	0.0257	0.0006	0.970	1.000
Between Level							
Means							
Y	1.200	1.1130	0.0325	0.0366	0.0086	0.290	1.000
X	0.800	0.7566	0.0314	0.0335	0.0029	0.720	1.000
S1	0.500	0.5073	0.0387	0.0373	0.0015	0.920	1.000
Variances							
Y	1.000	0.9351	0.0549	0.0589	0.0072	0.840	1.000
X	1.000	0.9364	0.0486	0.0504	0.0064	0.760	1.000
S1	1.000	1.0199	0.0554	0.0611	0.0034	0.960	1.000

CORRELATIONS AND MEAN SQUARE ERROR OF THE TRUE FACTOR VALUES AND THE FACTOR SCORES

	CORRELATIONS Average	Std. Dev.	MEAN SQUARE ERROR Average	Std. Dev.
S1	0.856	0.009	0.519	0.014
Y	0.822	0.011	0.578	0.014
X	0.917	0.005	0.404	0.009

Figure 8: Real-Data Random Regression with Latent Centering

```
variable:
        names are y x c;
        cluster=c;

data: file=1.dat;

ANALYSIS:
        TYPE = TWOLEVEL RANDOM;
        estimator=bayes;

model:

        %WITHIN%
        s | y on x;

        %BETWEEN%
        y x s with y x s;
```

4 Multilevel Mediation 1-1-1

In this mediation model, all observed variables vary across the lower level. For individual i in cluster j let Y_{ij} denote the dependent variable, M_{ij} the mediator and X_{ij} the covariate. The model is given by the following equations

$$Y_{ij} = Y_{w,ij} + Y_{b,j} \tag{6}$$

$$M_{ij} = M_{w,ij} + M_{b,j} \tag{7}$$

$$X_{ij} = X_{w,ij} + X_{b,j} \tag{8}$$

$$Y_{w,ij} = \beta_{1,j} M_{w,ij} + \beta_{2,j} X_{w,ij} + \varepsilon_{w,ij} \tag{9}$$

$$M_{w,ij} = \beta_{3,j} X_{w,ij} + \xi_{w,ij}. \tag{10}$$

$$\varepsilon_{w,ij} \sim N(0, \theta_w), \xi_{w,ij} \sim N(0, \psi_w), X_{w,ij} \sim N(0, \sigma_w) \tag{11}$$

There are six between level random effects in this model: the three random intercepts $Y_{b,j}$, $M_{b,j}$, $X_{b,j}$ and the three random slopes $\beta_{1,j}$, $\beta_{2,j}$, and $\beta_{3,j}$. Because this model is also based on latent centering, it can only be estimated with Bayes estimator but not with the ML estimator. With the ML estimator, it is possible to estimate somewhat similar models based on observed or hybrid centering, see Asparouhov and Muthén (2019). However, these alternative methods are typically not recommended due to biases that are associated with those types of centering.

The relationship between the 6 random effects forms the between level model. Here we will explore a few different specifications for that between-level model and investigate how that specification impacts the estimates. The unstructured model (Model 1) is the first model we consider

$$(Y_{b,j}, M_{b,j}, X_{b,j}, \beta_{1,j}, \beta_{2,j}, \beta_{3,j}) \sim N(\mu, \Sigma), \tag{12}$$

where μ is an unrestricted mean parameter vector and Σ is an unrestricted variance covariance. The second model (Model 2) is the same as the first model but the variance covariance Σ is estimated as a diagonal matrix, i.e., all random effects are uncorrelated. The third model (Model 3) is given by

$$(Y_{b,j}, M_{b,j}, X_{b,j}) \sim N(\mu_1, \Sigma_1), (\beta_{1,j}, \beta_{2,j}, \beta_{3,j}) \sim N(\mu_2, \Sigma_2), \tag{13}$$

where Σ_1 is an unrestricted variance covariance matrix while Σ_2 is diagonal. The random intercepts are correlated in this model but the random slopes are uncorrelated among each other and the random intercepts as well.

The fourth model (Model 4) is based on a mediation model for the between parts of the variables. Such a model is generally of interest as it explores the same relationship between the variables but on the cluster level as well. The model is given as follows

$$Y_{b,j} = \alpha_1 + \beta_4 M_{b,j} + \beta_5 X_{b,j} + \varepsilon_{b,j} \tag{14}$$

$$M_{b,j} = \alpha_2 + \beta_6 X_{b,j} + \xi_{b,j} \tag{15}$$

$$\varepsilon_{b,j} \sim N(0, \theta_b), \xi_{b,j} \sim N(0, \psi_b) \tag{16}$$

$$(X_{b,j}, \beta_{1,j}, \beta_{2,j}, \beta_{3,j}) \sim N(\mu_3, \Sigma_3), \tag{17}$$

where μ_3 is an unrestricted mean parameter vector and Σ_3 is an unrestricted variance covariance.

We illustrate the above models with a simulation study and we compare the estimates of the four models. For this simulation study we use a more common sample size: 300 clusters of size 20. We generate data according to Model 1, i.e., using a completely unconstrained model and we want to see if Model 2, 3 and 4 will be able to produce unbiased estimates even when the between level model does not account for all the correlations between the random effects.

Figure 9 shows the input file for this simulation study using Model 1. The Model statement, not provided for brevity, is identical to the Model Population statement. The results for the four different models are reported in Figure 10. Here we report only the mean parameters for the three observed variables. Model 1 gives unbiased estimates and coverage near the nominal levels. Model 2 shows substantial bias and poor coverage. Model 3 and 4 show minor bias and coverage that is nearly acceptable. This, however, is not the case when we use bigger samples (10 times bigger). With bigger samples, Model 4 bias remains the same but the coverage drops substantially.

Model 1 estimates of all the other parameters are also unbiased. Interestingly, Model 2 random slope means and variance are also unbiased, but the estimates of the variance parameters for the random intercept are also substantially biased.

The results indicate that Model 3 and 4 perform much better than Model 2, although not as well as Model 1. This is because the mediation model on the between level allows some correlation between the variables. While the correlation is not fully unrestricted, in Model 4, the correlation between any pair of random effects is not zero. Note also that Model 4 does not estimate the means of the variables but only the intercepts. The means must be computed as new parameters and MODEL CONSTRAINT. For example, the mean of M is $\alpha_2 + \beta_6\mu_3(1)$ where $\mu_3(1)$ is the mean of X.

Next we compare the model estimates for the random effects. The correlation between the true cluster specific values and the estimated values as well as the mean squared error (MSE) averaged across replications is given in Figure 11. For all 6 random effects, Model 1 estimates are better than Model 2 estimates. These have higher correlation with the true values and smaller MSE. The differences are especially pronounced for the random intercepts of Y and M. Model 3 and 4 estimates are nearly as good as Model 1 estimates.

These results show that Model 1 works substantially better than Model 2 and even Model 3 and 4, although the disadvantages of Model 3 and 4 appear to be small. Accounting for the correlations between random effects is important in multilevel mediation. In some applications, Model 4 setup based on a between level mediation model is important for practical inference. If this is the case, we can either settle with Model 4 or we can use the following Model 5 which is equivalent to Model 1 but incorporates the between level mediation. Model 5 is given as follows

$$Y_{b,j} = \alpha_1 + \beta_4 M_{b,j} + \beta_5 X_{b,j} + \varepsilon_{b,j} \tag{18}$$

$$M_{b,j} = \alpha_2 + \beta_6 X_{b,j} + \xi_{b,j} \tag{19}$$

$$\varepsilon_{b,j} \sim N(0, \theta_b), \xi_{b,j} \sim N(0, \psi_b), X_{b,j} \sim N(0, \sigma_2) \tag{20}$$

$$\beta_{k,j} = \nu_k + \gamma_{1,k} Y_{b,j} + \gamma_{2,k} M_{b,j} + \gamma_{3,k} X_{b,j} + \eta_{kj} \tag{21}$$

$$(\eta_1, \eta_2, \eta_3) \sim N(\mu_3, \Sigma_3). \tag{22}$$

In this model, all random slopes are regressed on the random intercepts. This will result in a completely unconstrained variance covariance estimation on the between level for all 6 variables. Note that there are other modifications for Model 4 that can result in a Model 1 equivalent. For example, adding residual covariance between $Y_{b,j}$ and $M_{b,j}$ and all random slopes will result in a Model 1 equivalent. Such a model however is undesirable because it leads to a non-block diagonal variance covariance matrix which is more difficult to estimate with the Bayesian estimator. Model 5 is equivalent to Model 1 and thus it is guaranteed to yield best results.

Figures 12-14 give Mplus input files that can be used with real data to estimate Model 1,4, and 5.

Figure 9: Multilevel Mediation using Model 1

```
montecarlo:
        names are y m x;
        nobservations = 6000;
        ncsizes = 1;
        csizes = 300(20);
        seed = 58459;
        nreps = 100;

ANALYSIS:
    TYPE = TWOLEVEL RANDOM;
    estimator=bayes;

model population:

    %WITHIN%
    x*1; y*1; m*1;
    s1 | y on x;
    s2 | y on m;
    s3 | m on x;


    %BETWEEN%
    y*1 x*1 m*1 s1-s3*1;
    y with x*0.7;
    y with m*0.6;
    x with m*0.5;
    s1-s3 with y-x*0.5;
    s1 with s2*0.5;
    s1 with s3*0.5;
    s2 with s3*0.7;
    [s1*0.5 s2*0.8 s3*1 x*0.8 y*1.2 m*0.5];
```

Figure 10: Multilevel Mediation Results

	Population	ESTIMATES Average	Std. Dev.	S. E. Average	M. S. E.	95% Cover	% Sig Coeff
Means Model 1							
Y	1.200	1.2061	0.0585	0.0710	0.0034	0.940	1.000
M	0.500	0.5095	0.0571	0.0627	0.0033	0.970	1.000
X	0.800	0.8115	0.0544	0.0609	0.0031	0.990	1.000
Means Model 2							
Y	1.200	0.9542	0.0575	0.0600	0.0637	0.040	1.000
M	0.500	0.4123	0.0571	0.0568	0.0109	0.630	1.000
X	0.800	0.7702	0.0547	0.0571	0.0038	0.920	1.000
Means Model 3							
Y	1.200	1.1189	0.0583	0.0665	0.0099	0.800	1.000
M	0.500	0.4788	0.0583	0.0614	0.0038	0.950	1.000
X	0.800	0.8004	0.0541	0.0598	0.0029	1.000	1.000
Means Model 4							
Y	1.200	1.1480	0.0583	0.0668	0.0061	0.920	1.000
M	0.500	0.4891	0.0586	0.0614	0.0035	0.940	1.000
X	0.800	0.7977	0.0538	0.0594	0.0029	1.000	1.000

Figure 11: Multilevel Mediation Results for Estimated Random Effects

CORRELATIONS AND MEAN SQUARE ERROR OF THE TRUE FACTOR VALUES AND THE FACTOR SCORES

Model 1

	CORRELATIONS		MEAN SQUARE ERROR	
	Average	Std. Dev.	Average	Std. De
S1	0.940	0.007	0.344	0.017
S2	0.976	0.002	0.215	0.008
S3	0.974	0.003	0.228	0.010
Y	0.896	0.013	0.444	0.026
M	0.960	0.005	0.278	0.012
X	0.982	0.002	0.189	0.009

Model 2

	CORRELATIONS		MEAN SQUARE ERROR	
	Average	Std. Dev.	Average	Std. Dev.
S1	0.931	0.008	0.369	0.020
S2	0.974	0.003	0.226	0.009
S3	0.972	0.003	0.235	0.010
Y	0.771	0.030	0.684	0.045
M	0.947	0.006	0.340	0.016
X	0.980	0.002	0.208	0.008

Model 3

	CORRELATIONS		MEAN SQUARE ERROR	
	Average	Std. Dev.	Average	Std. Dev.
S1	0.931	0.008	0.368	0.020
S2	0.975	0.003	0.226	0.009
S3	0.972	0.003	0.235	0.010
Y	0.885	0.014	0.475	0.030
M	0.958	0.005	0.288	0.014
X	0.981	0.002	0.193	0.009

Model 4

	CORRELATIONS		MEAN SQUARE ERROR	
	Average	Std. Dev.	Average	Std. Dev.
S1	0.940	0.007	0.345	0.018
S2	0.976	0.002	0.215	0.008
S3	0.974	0.003	0.228	0.010
Y	0.889	0.014	0.462	0.028
M	0.959	0.005	0.284	0.013
X	0.982	0.002	0.191	0.009

Figure 12: Real-Data Multilevel Mediation Model 1

```
variable:
    names are y m x c;
    cluster=c;

data: file is 1.dat;

ANALYSIS:
    TYPE = TWOLEVEL RANDOM;
    estimator=bayes;

model:

    %WITHIN%
    s1 | y on x;
    s2 | y on m;
    s3 | m on x;

    %BETWEEN%
    y-x s1-s3 with y-x s1-s3;
```

Figure 13: Real-Data Multilevel Mediation Model 4

```
variable:
    names are y m x c;
    cluster=c;

data: file is 1.dat;

ANALYSIS:
    TYPE = TWOLEVEL RANDOM;
    estimator=bayes;

model:

    %WITHIN%
    s1 | y on x;
    s2 | y on m;
    s3 | m on x;

    %BETWEEN%
    y on x m;
    m on x;
    x s1-s3 with x s1-s3;
```

Figure 14: Real-Data Multilevel Mediation Model 5

```
variable:
    names are y m x c;
    cluster=c;

data: file is 1.dat;

ANALYSIS:
    TYPE = TWOLEVEL RANDOM;
    estimator=bayes;

model:

    %WITHIN%
    s1 | y on x;
    s2 | y on m;
    s3 | m on x;

    %BETWEEN%
    y on x m;
    m on x;
    s1-s3 on x y m;
    s1-s3 with s1-s3;
```

5 Conclusion

Models with a large number of random effects lead to the need to model their variance covariance on the between level. This increases the number of model parameters substantially and as a result model parsimony is lost as well as the power to detect significance. It is tempting to ignore the between level covariances altogether but as we illustrated above there are consequences for such a strategy. Models that incorporate some kind of covariance modeling are likely to be optimal. For example, a factor analysis model for all random effects on the between level model is likely a good solution in many cases. It is often the case that the number of clusters in the sample is not large. Thus the power to reject a factor analysis model is likely to be low. Even when it is not the true between level model, and in the absence of a substantively derived model for the random effects, the factor analysis model is likely suitable. It is a parsimonious alternative to the full covariance matrix estimation and it will allow random effect covariance for every pair of random effects.

The effects of omitting or incorrectly modeling the covariances between the random effects are generally not large. The conditions that exacerbate the biases are: large number of clusters, small cluster sizes, and more complex models. DSEM analyses which are aimed at intensive longitudinal data (large cluster sizes) are likely more immune to between level model misspecifications.

This note should not be viewed as discouraging the modeling strategy where simpler models are evaluated first. In most modeling applications, using a model with independent random effects is a good first step. In that step we can see if a cluster specific parameter is needed or not. Random effects with near zero variance can be discarded at that point so that they do not cause convergence issues in later modeling stages. The note should be viewed, however, as encouraging of between level covariance modeling strategies.

References

[1] Asparouhov T, Muthén B. (2019). Latent variable centering of predictors and mediators in multilevel and timeseries models. Structural Equation Modeling: A Multidisciplinary Journal, 26, 119–42.