Dynamic Structural Equation Models

Tihomir Asparouhov, Bengt Muthén and Ellen Hamaker
Part 4

July 6, 2017

Overview

- The general DSEM model and estimation
- DIC
- Model estimated means and variances
- DSEM output options
- DSEM plots
- Centering
- Subject-specific variances
- Unevenly spaced and individual-specific times of observations
- Three level AR(1) models: within day v.s. between day autoregressive modeling

Reference for this talk is the ”Dynamic Structural Equation Models” paper available online http://statmodel.com/download/DSEM.pdf

The general DSEM model

- Merge ”time series”, ”structural equation”, ”multilevel” and ”TVEM(time varying effect modeling)” concepts in a generalized modeling framework in Mplus V8
- Y_{it}, η_{it} and X_{it} - are the observed dependent variables, latent factors and predictors for individual i at time t
- Four distinct sources of correlation in such observed data:
  - correlation due to individual specific effects (multilevel)
  - correlation due to proximity of observations (time series)
  - correlation between different variables (SEM)
  - correlation due to the same stage of evolution (TVEM)
- DSEM finds these correlations

DSEM model

- Includes three separate models: single level, twolevel , cross-classified
- Main decomposition equation

$$Y_{it} = Y_{1,it} + Y_{2,i} + Y_{3,t}$$

- $Y_{2,i}$, $Y_{3,t}$ are the "individual" and "time" specific contribution. These are latent variables. $Y_{1,it}$ is the residual.
- Includes three separate models:
  - single level DSEM: type=general, N=1, $Y_{2,i}$, $Y_{3,t}$ are removed
  - two-level DSEM: type=twolevel, $Y_{3,t}$ is removed
  - cross-classified DSEM: type=cross, full version
- We describe the cross-classified DSEM as it is the most general model, however ....

DSEM model

- Cross-classified DSEM requires that the time scale is aligned for all individuals - not every data set is applicable, ex. observational studies. Time t specific random effect apply for all individuals so time t has to mean the same thing, ex second grade.
- The two-level DSEM much simpler formulation
- The two-level DSEM is the most common and introductory model for applications
- The two-level DSEM can be estimated with less data, fewer requirements for size of N and T as compared to cross-classified DSEM, for example unbalanced designs
- The two-level DSEM easier to estimate as compared to cross-classified DSEM: much fewer number of random effects
- Mplus 8 speed for two-level DSEM always acceptable, Mplus 8 speed for cross-classified DSEM: depends on the model, some models acceptable, models with random variances or random autoregressive parameters can be very slow
- Single level model - one individual modeled separately

DSEM model

- The within level model includes latent variables and observed variables from the previous L (lag) periods

$$Y_{1,it} = \nu_1 + \sum_{l=0}^{L} \Lambda_{1,l}\eta_{1,i,t-l} + \sum_{l=0}^{L} R_l Y_{1,i,t-l} + \sum_{l=0}^{L} K_{1,l}X_{1,i,t-l} + \varepsilon_{1,it}$$

$$\eta_{1,it} = \alpha_1 + \sum_{l=0}^{L} B_{1,l}\eta_{1,i,t-l} + \sum_{l=0}^{L} Q_l Y_{1,i,t-l} + \sum_{l=0}^{L} \Gamma_{1,l}X_{1,i,t-l} + \xi_{1,it}.$$

- Note that all predictors are centered i.e. $Y_{1,i,t-l}$ is not $Y_{i,t-l}$ (covariates X are optional)

DSEM model

- The usual structural equations at level 2 and 3.

$$Y_{2,i} = \nu_2 + \Lambda_2 \eta_{2,i} + \varepsilon_{2,i}$$

$$\eta_{2,i} = \alpha_2 + B_2 \eta_{2,i} + \Gamma_2 x_{2,i} + \xi_{2,i}$$

$$Y_{3,t} = \nu_3 + \Lambda_3 \eta_{3,t} + \varepsilon_{3,t}$$

$$\eta_{3,t} = \alpha_3 + B_3 \eta_{3,t} + \Gamma_3 x_t + \xi_{3,t}$$

- These include not just between parts of Y_{it} but also observed between level variables

DSEM model

- Random parameters on within level
  - intercepts
  - slopes
  - loadings
  - auto-regressive parameters
  - variances - new V8 feature available for DSEM and non-DSEM
  - random covariance? Only via random factor variances

- We have not found an easy to interpret, random covariance model, that is based on normally distributed random effects which can be used in linear equations as predictors or to be predicted by other variables

DSEM model

- Every within level random parameter s has an individual specific part $s_{2,i}$ and time specific part $s_{3,t}$

$$s = s_{2,i} + s_{3,t}$$

- $s_{2,i}$, $s_{3,t}$ are normally distributed random effects which are a part of the between level latent variable vectors $\eta_{2,i}$ and $\eta_{3,t}$

- Random variances are special

$$s = Exp(s_{2,i} + s_{3,t})$$

- This way we always keep these positive

DSEM model

- The general model on the within level can now also be written with indices i and t for all the possible random parameters

$$Y_{1,it} = \nu_1 + \sum_{l=0}^{L} \Lambda_{1,lit} \eta_{1,i,t-l} + \sum_{l=0}^{L} R_{lit} Y_{1,i,t-l} + \sum_{l=0}^{L} K_{1,lit} X_{1,i,t-l} + \varepsilon_{1,it}$$

$$\eta_{1,it} = \alpha_{1,it} + \sum_{l=0}^{L} B_{1,lit} \eta_{1,i,t-l} + \sum_{l=0}^{L} Q_{lit} Y_{1,i,t-l} + \sum_{l=0}^{L} \Gamma_{1,lit} X_{1,i,t-l} + \xi_{1,it}$$

DSEM model

- The above model assumes conditional normality
- Ordered polytomous and binary dependent variables using the underlying Y^* approach
- Missing data: MAR likelihood based treatment via MCMC estimation. If there is autocorrelation in the data the missing data will be imputed from the neighbouring observations rather than from the average for the person! Note that standard econometrics methodology even for single level models does not include missing data. Even for single level data with missing observations this is new.

Residual DSEM model (available in future Mplus release)

No change in the between level model. The within level model further splits the autoregressive and the structural part

$$Y_{1,it} = Y_{0,it} + \hat{Y}_{1,it}$$

$$\eta_{1,it} = \eta_{0,it} + \hat{\eta}_{1,it}$$

- The variables $Y_{0,it}$ and $\eta_{0,it}$ represent the linear predictor part (no random element)
- The variables $\hat{Y}_{1,it}$ and $\hat{\eta}_{1,it}$ represent the auto-regressive part and can be thought of as being the residuals

Residual DSEM model

The linear predictor model for $Y_{0,it}$ and $\eta_{0,it}$

$$Y_{0,it} = \nu_1 + \sum_{l=0}^{L} K_{1,lit} X_{1,i,t-l}$$

$$\eta_{0,it} = \alpha_{1,it} + \sum_{l=0}^{L} \Gamma_{1,lit} X_{1,i,t-l}$$

The auto-regressive model for $\hat{Y}_{1,it}$ and $\hat{\eta}_{1,it}$

$$\hat{Y}_{1,it} = \sum_{l=0}^{L} \Lambda_{1,lit} \hat{\eta}_{1,i,t-l} + \sum_{l=0}^{L} R_{lit} \hat{Y}_{1,i,t-l} + \varepsilon_{1,it}$$

$$\hat{\eta}_{1,it} = \sum_{l=0}^{L} B_{1,lit} \hat{\eta}_{1,i,t-l} + \sum_{l=0}^{L} Q_{lit} \hat{Y}_{1,i,t-l} + \xi_{1,it}$$

DSEM Initial Conditions

- At time $t = 1, ..., L$ the DSEM model uses predictors with negative time indices such as $\eta_{i,t=0}$, $\eta_{i,t=-1}$, $Y_{1,i,t=0}$, $Y_{1,i,t=-1}$, $X_{i,t=0}$, $X_{i,t=-1}$. We treat these as auxiliary parameters with their own prior.
- If sequences are long such as $T > 50$ the prior does not affect the results. For smaller time-series the priors may have minor effect.
- Mplus implements 2 options
- A. Mplus default: automatic priors, in the first 100 burnin MCMC iterations we update the priors from the sample statistics of η_{it}, $Y_{1,it}$, or $X_{i,t}$, then we discard those 100 MCMC iteration, and retain the constructed priors. Works quite well even for small T.

DSEM Initial Conditions

- B. Specify a normal prior for these auxiliary parameters in model prior. Difficult to use in practice especially when variables are not standardized.

```
model:
    f by y@1 (&1); f*0.6;
    y on f&1*0.4 y&1*0.5 y&2*0.2;
    y@0.01;

model prior: f~N(0,0.6); y~N(0,1);
```

DSEM Estimation

- MCMC with Gibbs sampler. All latent variables, missing values, initial conditions, random effects and model parameters, i.e., all unknown quantities are placed in one of 13 blocks:
  - B1: $Y_{2,i}$
  - B2: All random slopes $s_{2,i}$
  - B3: $Y_{3,t}$
  - B4: All random slopes $s_{3,t}$
  - B5: Other latent variables $\eta_{2,i}$ and $\eta_{3,t}$
  - B6: Latent variables $\eta_{1,it}$, including initial conditions where $t \leq 0$
  - B7: Missing variables Y_{it}
  - B8: Initial conditions $Y_{1,it}$ and $X_{1,it}$ for $t \leq 0$
  - B9: Threshold parameters for all categorical variables θ_3
  - B10: Underlying variables Y^*_{it} for all categorical variables
  - B11: Non-random intercepts, slope and loadings parameters θ_1
  - B12: Non-random variance, covariance and correlation parameters θ_2
  - B13: Random variance parameters

DSEM Estimation

- Determine each block conditional distribution, given all other blocks and the data
- Update (generate new values for) each block from that conditional distribution
- Repeat cycling between the blocks until convergence and use the generated values as the posterior distribution, point estimates, SE
- Mplus mini-max strategy for block formation: minimize the number of block while keeping conditional distributions explicit, i.e., maximizing the blocks. Each block is further split into the sub-blocks that are conditionally independent and update these separately. Strategy for most efficient computation and mixing. Blocks 3,6,7 sequentially updated.
- Bayes estimation inheritance: DSEM algorithm is an extension of Mplus 7.4, i.e., not developed from scratch.
- All conditional distribution are described in the DSEM paper

DIC

- DIC can be used to compare DSEM models. Implemented for models with all continuous dependent variable (no categorical).

$$D(\theta) = -2\log(p(Y|\theta))$$

$$p_D = \bar{D} - D(\bar{\theta})$$

$$DIC = D(\bar{\theta}) + 2p_D$$

- Despite the clear definition with the above formulas, there is substantial variation in what DIC actually is. The source of the variation is the definition of θ, and if it includes the latent variables or not.
- Different definitions of DIC are not comparable. You can compare only if they are using the same likelihood $[Y|\theta]$
- DIC most likely can not be used to compare models if the two models use different θ

DIC

- In DSEM the following are used in the θ vector in addition to all model parameters
  - $Y_{2,i}$ and all random effects $s_{2,i}$
  - $Y_{3,t}$ and all random effects $s_{3,t}$
  - Initial conditions
  - Latent variables $\eta_{1,it}$ if their lagged variables are used in the model
  - Missing variables Y_{it} if their lagged variable is used in the model
- To compare two models with DIC all you need to verify is that θ between the two models is "the same". Random effect with zero variance is OK.
- This list makes easy the computation of $[Y|\theta]$
- p_D - estimated number of parameters should generally be near the size of the vector θ, i.e., should be near the count of the above list
- In DSEM p_D is large and needs extra long MCMC sequence for stable estimate
- ARMA(1,1) model not comparable to AR(1) with DIC for V8.

Model fit evaluation based on comparing sample and model estimated statistics

- Assuming stationarity of the autoregressive part of the DSEM model we compute subject specific model estimated mean, variances, autocorrelations of lag L. These can be compared to their sample counterparts.
- Model fit evaluation using MSE and correlation between sample v.s. model estimated. For example, means.

$$R = Cor(\mu_i, \overline{Y_{i*}})$$

$$MSE = \sum_{i=1}^{N} (\mu_i - \overline{Y_{i*}})^2 / N.$$

- The correlation is available in the Mplus plot utilities. MSE requires saving the plot data and computing it separately.

Time-series model estimated means, variance, correlations using Yule-Walker assuming stationarity

$$Z_t = \mu + \sum_{l=1}^{L} A_l Z_{t-l} + \zeta$$

$$\Sigma = Var(\zeta)$$

$$E(Z_t) = \left(I - \sum_{l=1}^{L} A_l \right)^{-1} \mu$$

$$\Gamma_j = Cov(Z_t, Z_{t-j})$$

$$\begin{bmatrix} \Gamma_0 & \Gamma_1^T & \Gamma_2^T & \dots & \Gamma_L^T \\ \Gamma_1 & \Gamma_0 & \Gamma_1^T & \dots & \Gamma_{L-1}^T \\ \Gamma_2 & \Gamma_1 & \Gamma_0 & \dots & \Gamma_{L-2}^T \\ \dots & \dots & \dots & \dots & \dots \\ \Gamma_L & \Gamma_{L-1} & \Gamma_{L-2} & \dots & \Gamma_0 \end{bmatrix} \begin{bmatrix} I \\ -A_1^T \\ -A_2^T \\ \dots \\ -A_L^T \end{bmatrix} = \begin{bmatrix} \Sigma \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

DSEM output options

- residual option: model estimated means, variance and autocorrelations for the observed variables
- residual(cluster) option: model estimated and cluster/subject specific means, variance and autocorrelations for the observed variables
- tech4 and tech4(cluster) options: model estimated quantities for the latent variables
- stand and stand(cluster) options: standardized model estimates and standardized cluster specific model estimates
- The option with (cluster) also provides the average across cluster quantities for the cluster specific estimates - applies for residual/tech4/stand
- The (cluster) option new also for none-DSEM models
- All of the above are based on Yule-Walker and require stationarity of the autoregressive part of the model
- HTML clickable output

DSEM output example

```
variable:   names=y x1 x2 c;
            within=x1;
            between=x2;
            cluster=c;
            lagged=y(1);

data: file=a.dat;

analysis:  type=twolevel random;
           estimator=bayes; proc=2;

model:
 %within%
 s1 | y on x1;
 s2 | y on y&1;
 s3 | y;

 %between%
 y s1-s3 on x2;

output: standardized(cluster) residual(cluster) fscomparison;

plot: type is plot3;

savedata: stddistribution=1.dat; save=fs(200); file=2.dat;
          bparameter=3.dat;
```

DSEM output example: htm output


```
Font size: A A A A

Mplus VERSION 8
MUTHEN & MUTHEN
07/05/2017 10:58 AM

OUTPUT SECTIONS

Input Instructions
Input Warnings And Errors
Summary Of Analysis
Summary Of Data
Covariance Coverage Of Data
Univariate Sample Statistics
Model Warnings And Errors
Model Fit Information
Model Results
Standardized Model Results
R-square
Within-level Standardized Model Results For Cluster 1
Within-level R-square For Cluster 1
Within-level Standardized Model Results For Cluster 2
Within-level R-square For Cluster 2
Within-level Standardized Model Results For Cluster 3
```

DSEM output example: standardized results

Mplus - [gb00o.htm]

File Edit View Mplus Plot Diagram Window Help

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%	Significance
Within-Level Standardized Estimates Averaged Over Clusters						
S1 | Y ON						
X1	0.438	0.007	0.000	0.427	0.454	*
S2 | Y ON						
Y&1	0.247	0.007	0.000	0.232	0.260	*
S3 |						
Y	0.452	0.009	0.000	0.427	0.464	*
Between Level						
S1 ON						
X2	0.359	0.065	0.000	0.208	0.459	*
S2 ON						
X2	0.682	0.045	0.000	0.586	0.758	*

DSEM output example: cluster specific standardized results

Mplus - [gb00o.htm]

File Edit View Mplus Plot Diagram Window Help

WITHIN-LEVEL R-SQUARE FOR CLUSTER 1

Variable	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%
Y	0.607	0.097	0.000	0.460	0.871

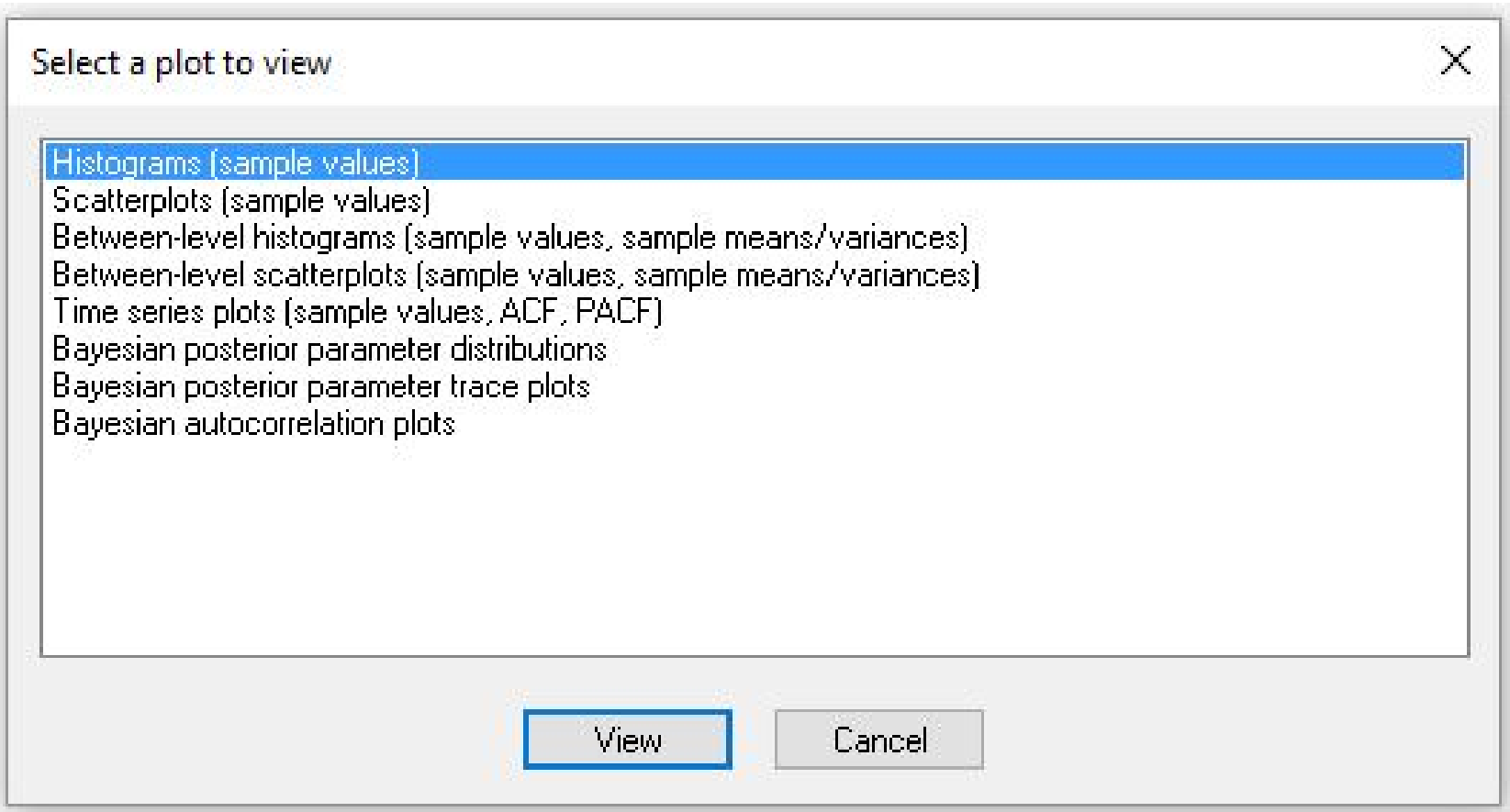
WITHIN-LEVEL STANDARDIZED MODEL RESULTS FOR CLUSTER 2

STDYX Standardization

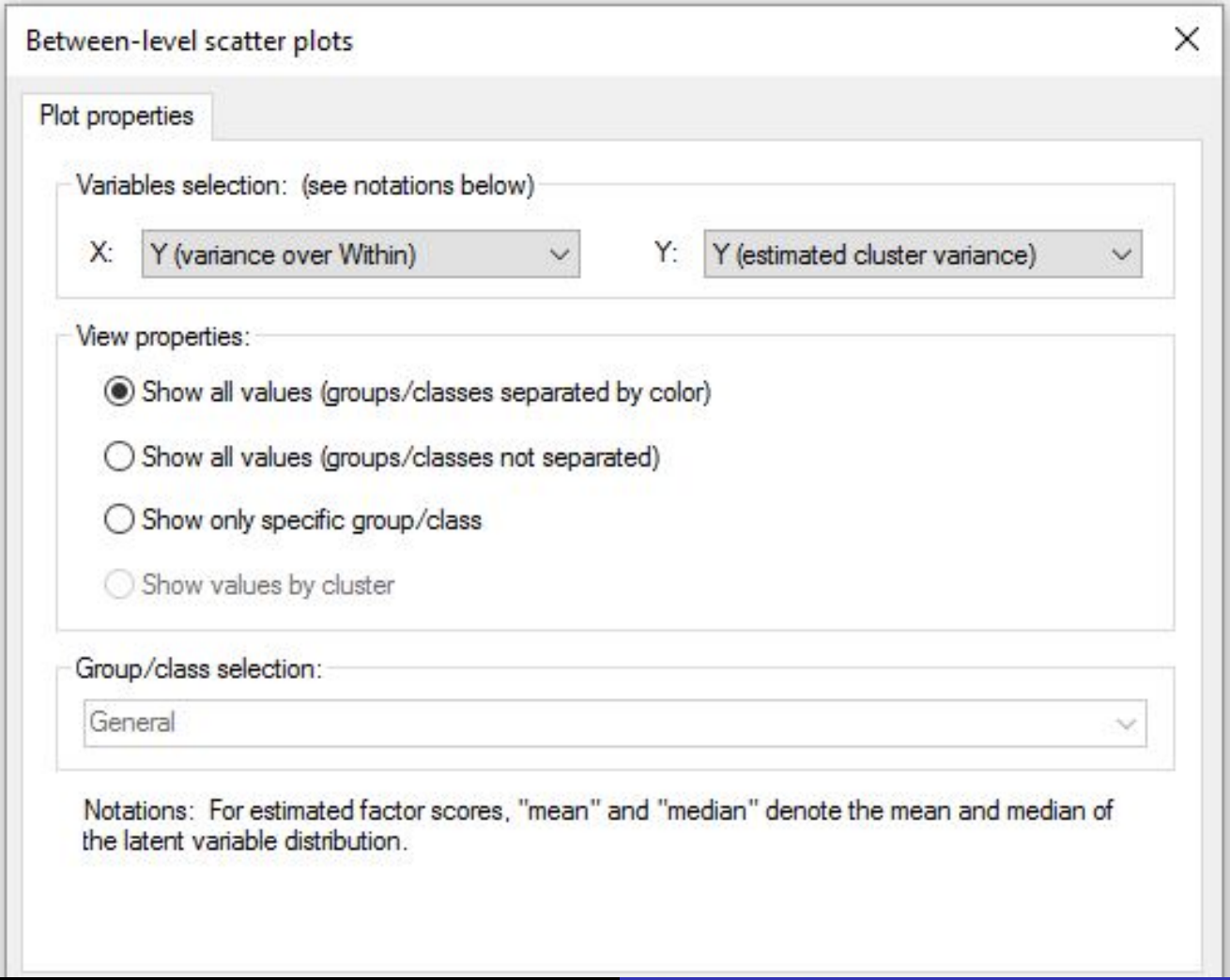
	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
S1 | Y ON						
X1	0.539	0.064	0.000	0.398	0.647	*
S2 | Y ON						
Y&1	0.250	0.070	0.000	0.108	0.388	*
S3 |						
Y	0.610	0.075	0.000	0.471	0.768	*

STDY Standardization

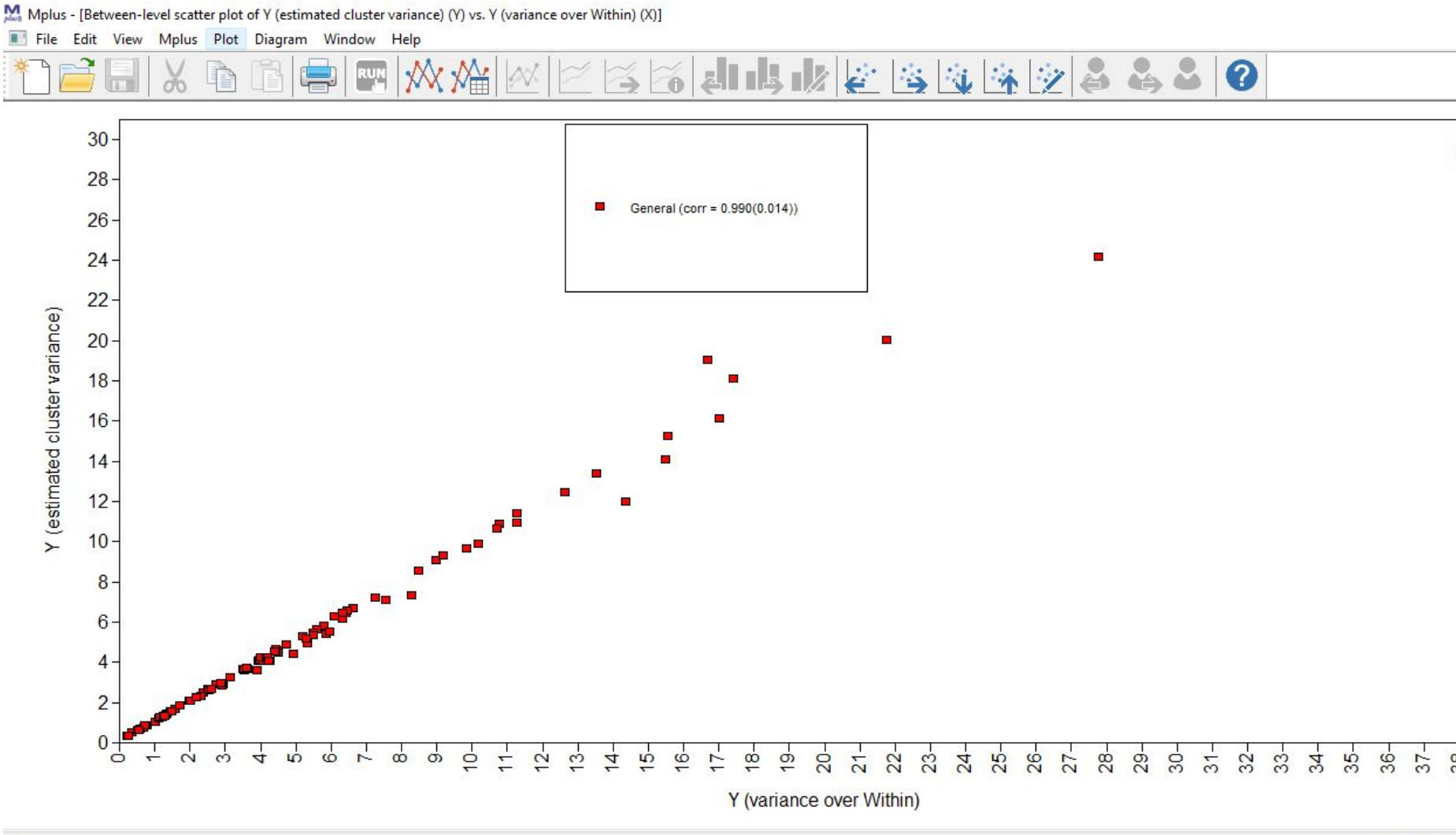
DSEM plots: plot menu



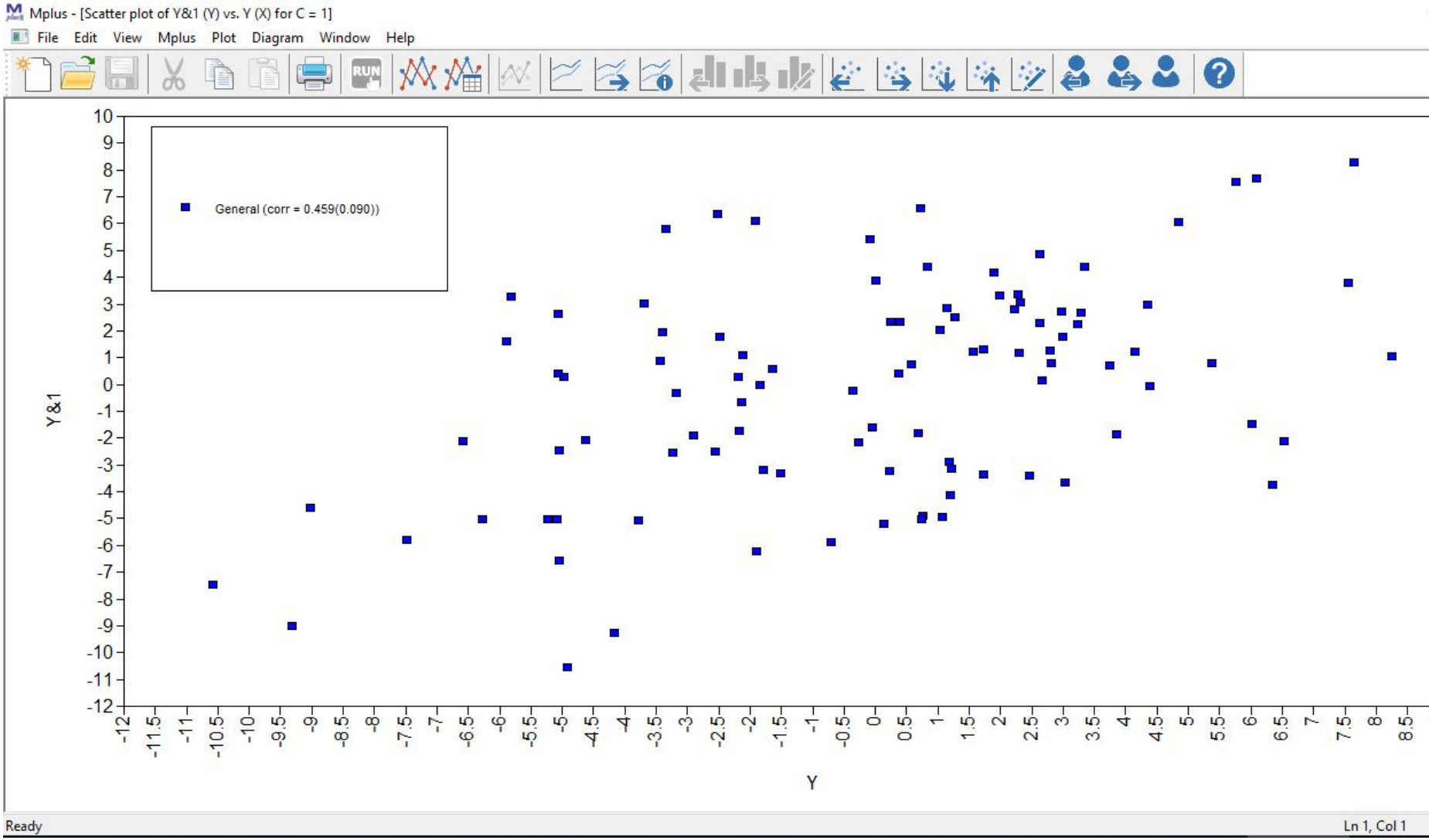
DSEM plots: plotting model estimated v.s. observed cluster specific statistic



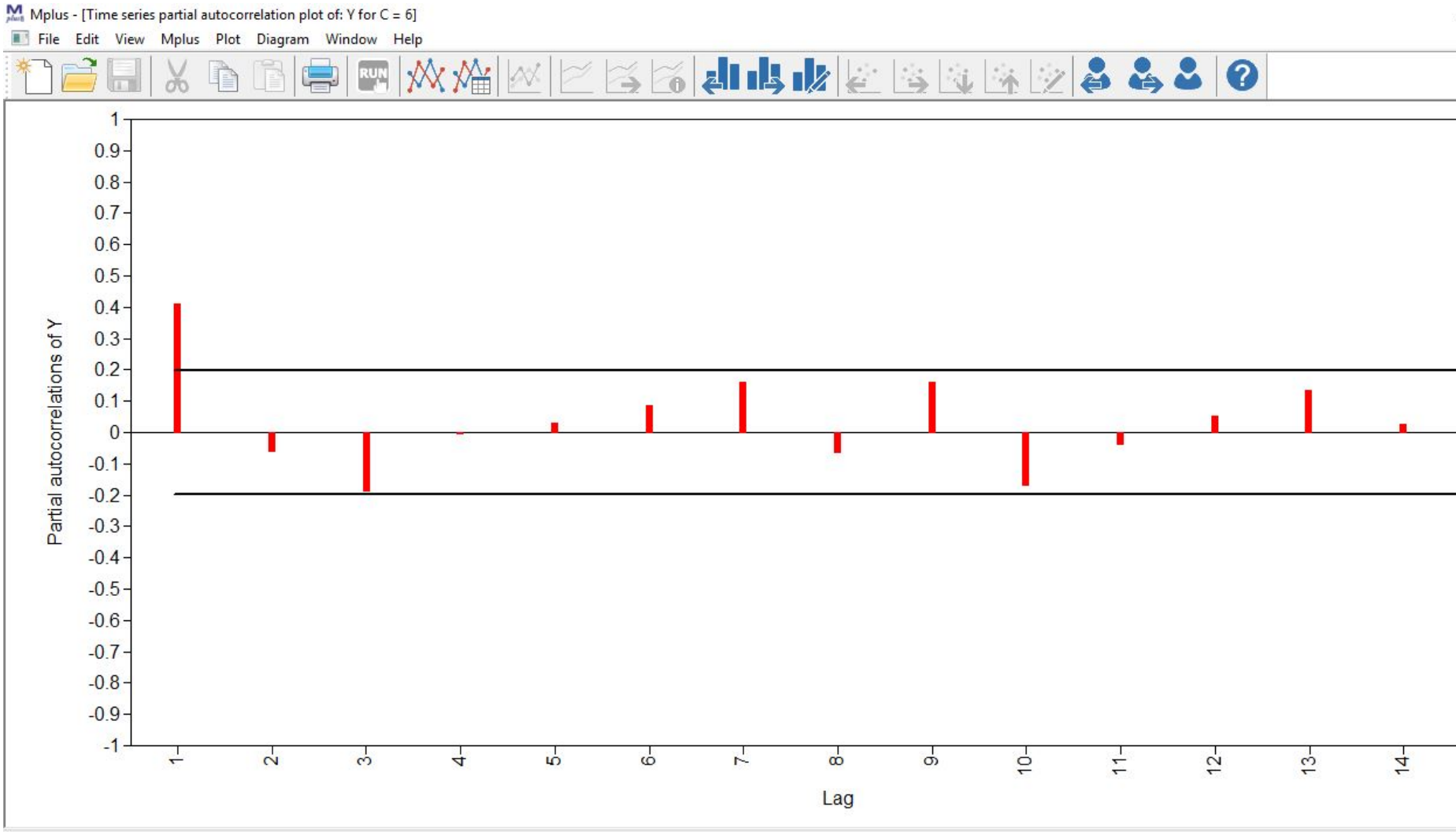
DSEM plots: plotting model estimated v.s. observed cluster specific variances



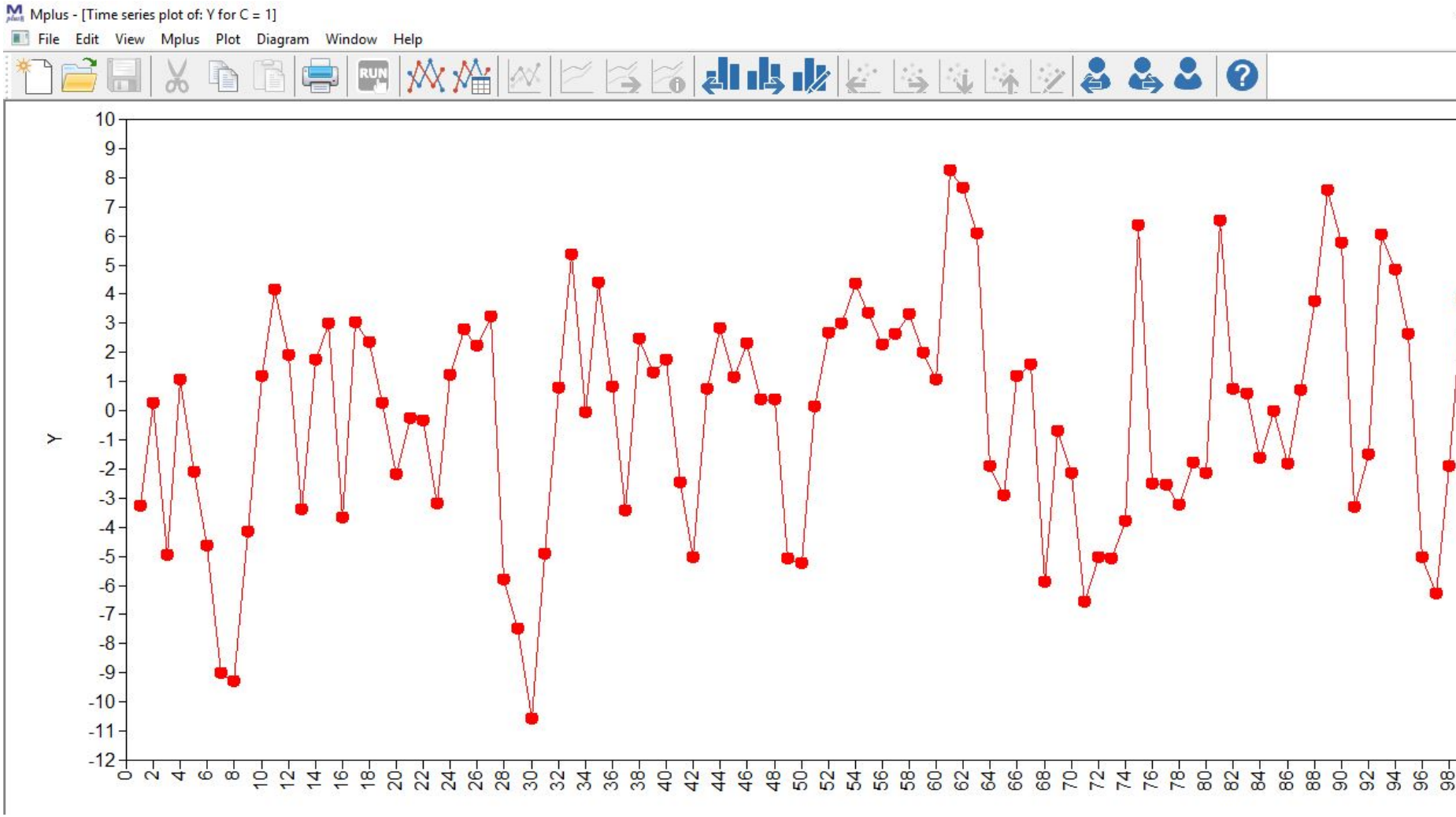
DSEM plots: cluster specific plots



DSEM plots: subject specific partial autocorrelation function



DSEM plots: subject specific time series plots



Simulation studies: Centering

- Simulation example using two-level random autoregressive AR(1) model
- Mplus latent centering

$$Y_{it} = \mu_i + r_i(Y_{i,t-1} - \mu_i) + \xi_{it}.$$

- Observed centering

$$Y_{it} = \mu_i + r_i(Y_{i,t-1} - \overline{Y_{i*}}) + \xi_{it}$$

- Uncentered

$$Y_{it} = \mu_i + r_i Y_{i,t-1} + \xi_{it}$$

Centering

- Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica: Journal of the Econometric Society, 1417-1426.
- Hamaker E.L. and Grasman R.P.P.P. (2015) To center or not to center? Investigating inertia with a multilevel autoregressive model. Front. Psychol., 5, 1492.
- Ludtke, O., Marsh, H.W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. Psychological Methods,13,203-29.
- Asparouhov, T. & Muthén, B. (2006). Constructing covariates in multilevel regression. Mplus Web Notes: No. 11.

Centering

- Ludtke bias is for two-level models, involves 2 different variables, and the bias is on the between

$$\frac{(\beta_w - \beta_b)\psi_w}{T\psi_b + \psi_w}$$

- Nickell bias is for DSEM, involves 1 variable, and the bias is on the within

$$-\frac{1+r}{T-1}$$

- Both stem from not accounting for the error in the sample mean estimate of the mean
- Both disappear when cluster sample size T increases
- Both can appear in parallel in the same example

Centering

- Note that observed centering or uncentered do not exist in case there is missing data - listwise deletion is not an option
- Hamaker and Grasman (2015) show that the uncentered method eliminates Nickell's bias. It does create other bias however, ex for σ_{11}
- Hamaker and Grasman (2015) show that using the true mean to center still creates bias

Centering - results

Table: Nickell's bias for r=0.3

T	N	Latent centering	Observed centering	Nickell's formula
10	100	0.025	-0.140	-0.144
20	50	0.006	-0.070	-0.068
30	30	0.008	-0.042	-0.045
50	50	0.000	-0.029	-0.027
100	100	-0.001	-0.014	-0.013

Nickell's formula is very accurate. Latent centering eliminates Nickell's bias.

Centering - results

Table: Nickell's bias for r=0.3

T	N	Latent centering	Observed centering	Nickell's formula
10	100	0.025	-0.140	-0.144
20	50	0.006	-0.070	-0.068
30	30	0.008	-0.042	-0.045
50	50	0.000	-0.029	-0.027
100	100	-0.001	-0.014	-0.013

Nickell's formula is very accurate. Latent centering eliminates Nickell's bias.

Centering - results

Table: Bias for $Var(\mu_i) = 3$

T	N	latent centering	Uncentered
10	100	-0.015	-1.637
20	50	0.217	-1.483
30	30	0.645	-1.256
50	50	0.378	-1.361
100	100	0.096	-1.508

For latent centering bias on $Var(\mu_i)$ as N increases (or with using weakly informative priors). For the uncentered method in will not disappear even asymptotically as the fundamentals of the model are wrong.

Centering - comparison of latent centering and uncentered

- Mplus latent centering

$$Y_{it} = \mu_i + r_i(Y_{i,t-1} - \mu_i) + \xi_{it}.$$

$$Y_{it} = \mu_i(1 - r_i) + r_i Y_{i,t-1} + \xi_{it}.$$

- Uncentered

$$Y_{it} = \mu_i + r_i Y_{i,t-1} + \xi_{it}$$

- The uncentered and the latent centering are reparameterizations of each other. To obtain the correct μ_i we need to divide by $1 - r_i$
- The latent centering has the advantage of obtaining μ_i directly

Centering - comparison of latent centering and uncentered with subject specific covariate X

- Mplus latent centering

$$Y_{it} = \mu_i + \beta X_i + r_i(Y_{i,t-1} - \mu_i - \beta X_i) + \xi_{it}.$$

$$Y_{it} = \mu_i(1 - r_i) + \beta(1 - r_i)X_i + r_i Y_{i,t-1} + \xi_{it}.$$

- Uncentered

$$Y_{it} = \mu_i + \beta X_i + r_i Y_{i,t-1} + \xi_{it}$$

- The uncentered and the latent centering are NOT reparameterizations of each other as the X_i effect is random in the latent centering.

Subject specific variance

- 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 Behav Res. 50(3), 334-349.
- In this paper it is shown that if subject specific variances are ignored - the structural parameters can be slightly biased. This does not happen in regular two-level models.

$$Y_{it} = \mu_i + \varepsilon_{it}$$

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

$$v_i = Log(Var(\xi_{it}))$$

- The bias depends on how high the correlation is between r_i and v_i

Subject specific variance -results

Table: Comparing the estimation with random variance and without random variance (invariant variance): Bias(coverage)

parameter	$Cov(r_i, v_i)$	random variance	invariant variance
$E(r_i)$	high	.001(.97)	.040(.35)
$E(r_i)$	medium	.001(.98)	.028(.65)
$E(r_i)$	low	.001(.97)	.017(.83)
$E(r_i)$	none	.001(.96)	.007(.92)
$Var(r_i)$	high	.001(.97)	-.012(.47)
$Var(r_i)$	medium	.001(.93)	-.007(.78)
$Var(r_i)$	low	.001(.93)	-.004(.88)
$Var(r_i)$	none	.001(.94)	-.001(.91)

Subject specific variance - results

More detailed method for evaluation of model estimation

$$SMSE = \sqrt{(1/N)\sum_i (\hat{r}_i - r_i)^2}$$

$Cov(r_i, v_i)$	random variance	invariant variance
high	.255	.346
medium	.293	.329
low	.300	.316
none	.300	.310

Subject specific variance - conclusions

- Looking at the parameter estimates alone may not be enough when comparing estimation methods. Distortion of structural parameters due to ignoring subject specific variance is not simple shift in the autoregressive parameter. Error is actually doubled when looking at the random effects SMSE.
- Even in standard two-level models, using cluster specific variance is important if we use SMSE as a criterion
- Subject specific variance extracts more information from the data, yields more accurate estimation
- More simulations are needed to evaluate this issue in multivariate setting - study the effect of subject specific covariance.

Subject-specific times of observations

- The basic model assumes that observations are taken at equally spaced time.
- If times are subject-specific we slice the time grid in sufficiently refined grid and enter missing data for the times where observation is not taken.
- For example if several observations are taken during the day, and at different times for each individual, we slice the day in 24 hour periods and place the corresponding observations in the hour slots.
- Data from the next simulation looks like this for day 1 for individual 1.

Subject-specific times of observations: subject 1 day 1

y1	y2	y3	y4	y5	T	ID
999	999	999	999	999	1	1
999	999	999	999	999	2	1
999	999	999	999	999	3	1
999	999	999	999	999	4	1
999	999	999	999	999	5	1
999	999	999	999	999	6	1
999	999	999	999	999	7	1
999	999	999	999	999	8	1
999	999	999	999	999	9	1
999	999	999	999	999	10	1
999	999	999	999	999	11	1
999	999	999	999	999	12	1
5.026193	0.327383	1.017519	0.701296	-0.55917	13	1
999	999	999	999	999	14	1
1.628885	1.652829	2.324074	1.800932	4.013447	15	1
999	999	999	999	999	16	1
4.376545	1.652831	2.098822	6.188234	2.913506	17	1
1.534865	0.631455	-0.29779	2.798775	1.37025	18	1
0.359654	1.476764	-0.43374	0.348777	1.382437	19	1
999	999	999	999	999	20	1
999	999	999	999	999	21	1
999	999	999	999	999	22	1
999	999	999	999	999	23	1
999	999	999	999	999	24	1

Subject-specific times of observations - simulation study

```
montecarlo:
    names = y1-y5 u;
    NOBS = 30000;
    NREP = 100;
    NCSIZES = 1;
    CSIZES = 100(300);
    categorical=u;
    generate=u(1);
    within=u;
    missing=y1-y5;

model missing: [y1-y5@-15]; y1-y5 on u@30;

ANALYSIS: TYPE IS TWOLEVEL; estimator=bayes;
          biter=10000(500); proc=2;

model montecarlo:
    %within%
    [u$1*-0.83];
    f by y1-y5*1 (& 1);
    y1-y5*1;
    f@1;
    f on f&1*0.4;

    %between%
    fb by y1-y5*0.5; fb@1; y1-y5*0.2;
```

Table: Two-level DAFS AR(1) with subject-specific times - simulation study results

percentage missing values	$\hat{\phi}$ (coverage) $\phi = 0.4$	convergence rate	comp time per replication in min
.80	.39(.95)	100%	1.5
.85	.39(.90)	95%	2.5
.90	.35(.46)	55%	10
.95	.34(.55)	55%	18

Quality of the estimation deteriorates as the amount of inserted missing data exceeds 90%. Note that this missing data is imputed by the MCMC estimation, leading to large amount of imputed quantities. It works well with 80% and 85% missing data.

Subject-specific times of observations

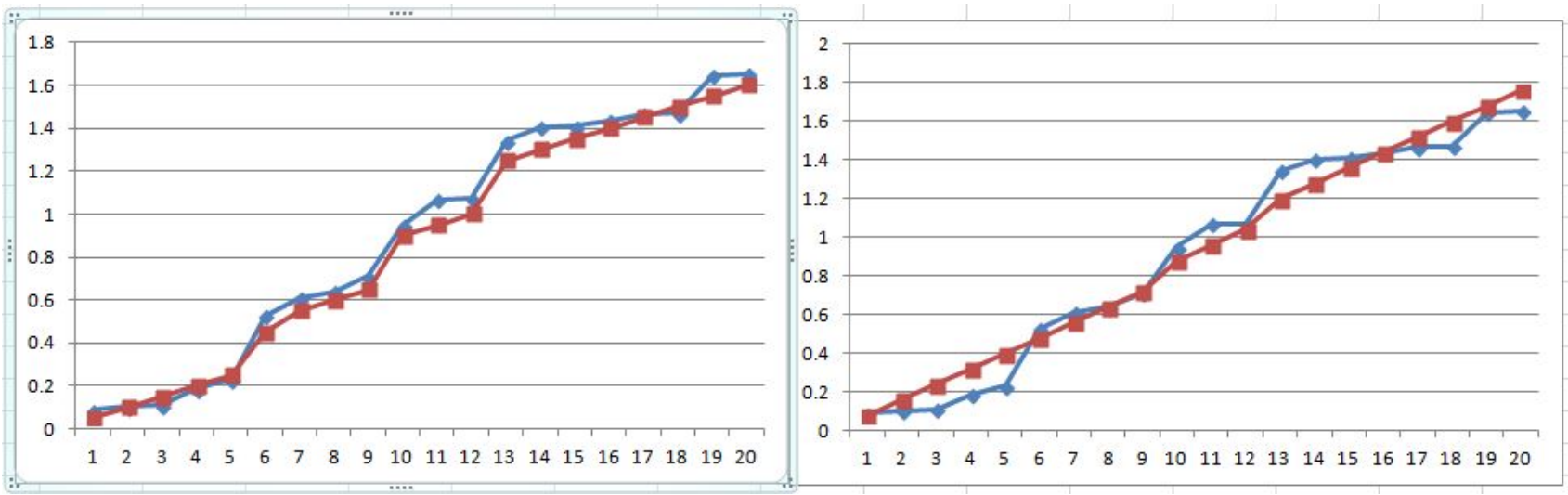
- Information contained in the unequal distances in the observations would be extracted well using the 80% to 85% missing values, eliminating the need for continuous time modeling
- Tinterval command will setup the missing data for you, given the precise times of observations and an approximation value δ
- Tinterval = $t(\delta)$ means that the continuous time variable t is replaced by the nearest integer $[t/\delta]$. There are complications if the nearest integers is the same for two or more different observations times t. Special algorithm to resolve this issue.

Subject-specific times of observations

- Split the time axis in bins of size δ. Then place each observation in the correct bin. Repeat these steps until each bin contains no more than 1 observation
  - find a bin with more than 1 observations
  - locate the nearest empty bin (look up or down)
  - move one of the extra observation to fill in the the empty bin but keep order of the observations so the extra observation bumps the rest of the observations towards the empty bin
- Mplus will warn you if the shifting process yields a discrepancy between t/δ and new time bigger than 5. Lower the δ value to resolve this.
- Fill in the remaining bins with missing values and set the time as T=1,2, ... and T is the bin number.
- Other algorithms are possible. Make your own discritzation algorithm and use Mplus with integer times.

Tinterval command comparison

Tinterval(0.05) v.s. Tinterval(0.08), Blue=true times, Red=Mplus generated times



Simulation study with varying δ

Table: Two-level AR(1) with subject-specific times. Estimates and coverage for ϕ and amount of missing data m_2 during the analysis, tinterval=δ

m	δ	$\phi = 0.8$	m_2
.80	1	.80(.91)	.80
.80	2	.81(.31)	.58
.80	3	.83(.00)	.38
.80	4	.84(.00)	.18
.80	5	.86(.00)	.05
.80	10	.92(.00)	.00
.95	1	.80(.85)	.95
.95	2	.81(.57)	.90
.95	3	.82(.20)	.85
.95	4	.83(.00)	.80
.95	5	.84(.00)	.74
.95	10	.88(.00)	.49

Subject-specific times of observations

- Tinterval command is not perfect. It is an approximate solution for the continuous process.
- The main question is how to choose δ. Three considerations:
  - Choose scale that is natural to help with interpretation of model and results - hour, day, week
  - Choose scale that does not produce more than 90% missing data, around 80% unless lower is appropriate
  - Smaller values yield better approximations but also more missing data
  - TVEM models / Cross-classified DSEM: small δ will lead to too many time specific random effects

Three-level AR(1) model

- Y_{idt} is the observed value for individual i on day d at time t

$$Y_{idt} = \mu + Y_i + E_{it} + F_{id} + G_{idt}$$

$$G_{idt} = \rho_1 G_{id,t-1} + \varepsilon_{1,idt}$$

$$F_{id} = \rho_2 F_{i,d-1} + \varepsilon_{2,id}$$

- Two type of autocorrelation parameter, ρ_1 is the autocorrelation within the day, ρ_2 is the autocorrelation between the days
- Maybe take out E_{it}?
- Model has 7 parameters: 4 variances, 2 autocorrelations, 1 intercept
- Data consists of 100 individuals, observed for 100 days, with 10 observations per day

Three-level AR(1) model - simulation study

```
montecarlo:
    names = y1-y10;
    NOBS = 10000;
    NREP = 100;
    NCSIZES = 1;
    CSIZES = 100(100);

ANALYSIS:   TYPE IS TWOLEVEL;
estimator=bayes; biter=(500); proc=2;

model montecarlo:
    %within%
    f by y1-y10@1 (& 1);
    y1-y10@0.01;
    f1 by y1@1; f2 by y2@1; f3 by y3@1;
    f4 by y4@1; f5 by y5@1; f6 by y6@1;
    f7 by y7@1; f8 by y8@1; f9 by y9@1;
    f10 by y10@1;
    f1-f10*1 (1);
    f*0.5;
    f on f&1*0.3;
    f2-f10 pon f1-f9*0.5 (2);
    f with f10@0;

    %between%
    fb by y1-y10@1; fb*0.4; y1-y10*0.1;
    [y1-y10*0] (3);
```

Three-level AR(1) model - simulation study results

	Population	ESTIMATES Average	Std. Dev.	S. E. Average	M. S. E.	95% Cover	% Sig Coeff
Within Level							
F ON							
F&1	0.300	0.2955	0.0133	0.0150	0.0002	0.970	1.000
F2 ON							
F1	0.500	0.4993	0.0042	0.0040	0.0000	0.940	1.000
Variances							
F1	1.000	1.0007	0.0049	0.0049	0.0000	0.930	1.000
Residual Variances							
F	0.500	0.4994	0.0150	0.0126	0.0002	0.880	1.000
Between Level							
Intercepts							
Y1	0.000	0.0037	0.0855	0.0572	0.0072	0.810	0.190
Variances							
FB	0.400	0.4310	0.0570	0.0658	0.0042	0.970	1.000
Residual Variances							
Y1	0.100	0.1007	0.0058	0.0051	0.0000	0.940	1.000

Three-level AR(1) model with subject-specific times of observations

- Using 50% missing data. Approximately 5 randomly spaced times of observations per day
- 5 observations a bit too low to obtain good autocorrelation parameter. Sequence is too short? Mixing estimation?
- Add the commands:
  missing=y1-y10;
  model missing: [y1-y10*0];

Three-level AR(1) model with subject-specific times of observations - simulation results

	Population	ESTIMATES Average	Std. Dev.	S. E. Average	M. S. E.	95% Cover	% Sig Coeff
Within Level							
F ON							
F&1	0.300	0.2864	0.0210	0.0166	0.0006	0.810	1.000
F2 ON							
F1	0.500	0.4428	0.0540	0.0188	0.0062	0.360	1.000
Variances							
F1	1.000	1.0444	0.0429	0.0159	0.0038	0.450	1.000
Residual Variances							
F	0.500	0.4694	0.0307	0.0189	0.0019	0.560	1.000
Between Level							
Intercepts							
Y1	0.000	0.0097	0.0665	0.0589	0.0045	0.890	0.110
Variances							
FB	0.400	0.3821	0.0640	0.0608	0.0044	0.940	1.000
Residual Variances							
Y1	0.100	0.0925	0.0098	0.0057	0.0002	0.660	1.000